Journal of the Formosan Medical Association (2015) 114, 64–71

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.jfma-online.com

ORIGINAL ARTICLE

Managing emergency department overcrowding via ambulance diversion: A discrete event simulation model

Chih-Hao Lin [a], Chung-Yao Kao [b,*], Chong-Ye Huang [b]

[a] *Department of Emergency Medicine, National Cheng Kung University Hospital, College of Medicine, National Cheng Kung University, Tainan, Taiwan*
[b] *Department of Electrical Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan*

Received 20 May 2012; received in revised form 5 September 2012; accepted 5 September 2012

KEYWORDS
ambulance diversion;
emergency department overcrowding;
simulation

Background/Purpose: Ambulance diversion (AD) is considered one of the possible solutions to relieve emergency department (ED) overcrowding. Study of the effectiveness of various AD strategies is prerequisite for policy-making. Our aim is to develop a tool that quantitatively evaluates the effectiveness of various AD strategies.
Methods: A simulation model and a computer simulation program were developed. Three sets of simulations were executed to evaluate AD initiating criteria, patient-blocking rules, and AD intervals, respectively. The crowdedness index, the patient waiting time for service, and the percentage of adverse patients were assessed to determine the effect of various AD policies.
Results: Simulation results suggest that, in a certain setting, the best timing for implementing AD is when the crowdedness index reaches the critical value, 1.0 — an indicator that ED is operating at its maximal capacity. The strategy to divert all patients transported by ambulance is more effective than to divert either high-acuity patients only or low-acuity patients only. Given a total allowable AD duration, implementing AD multiple times with short intervals generally has better effect than having a single AD with maximal allowable duration.
Conclusion: An input–throughput–output simulation model is proposed for simulating ED operation. Effectiveness of several AD strategies on relieving ED overcrowding was assessed via computer simulations based on this model. By appropriate parameter settings, the model can represent medical resource providers of different scales. It is also feasible to expand the simulations to evaluate the effect of AD strategies on a community basis. The results may offer insights for making effective AD policies.

* Corresponding author. Department of Electrical Engineering, National Sun Yat-Sen University, 70 Lien-Hai Road, Kaohsiung 804, Taiwan.
E-mail address: cykao@mail.nsysu.edu.tw (C.-Y. Kao).

http://dx.doi.org/10.1016/j.jfma.2012.09.007

Introduction

Emergency department (ED) overcrowding deters timely delivery of health care[1] and is becoming a public crisis worldwide. Researchers proposed conceptual models to explain ED management and thus to enhance the understanding of ED overcrowding.[2–4]

Previous studies proposed solutions of ED overcrowding through managing the input, throughput, and output process of ED.[5–8] According to the National Hospital Ambulatory Medical Care Survey, patients who arrived by ambulance accounted for 15.5% of total ED visits.[9] Of ambulance-transported visits, 68% were triaged as emergency or urgency, and 37% resulted in hospital admission.[10] Because ambulance-transported patients tend to be sicker and may use more ED resources, ambulance diversion (AD) is considered one of the possible solutions to relieve ED overcrowding by reducing the input component.[11]

AD is implemented if the ED requests ambulances that would normally bring patients to the hospital go instead to the other hospitals presumed less crowded. The policy of AD affects regional health care. Different AD strategies are implemented in different communities.[12–14] Policymakers need a tool to quantitatively evaluate the effectiveness of individual AD strategies and to tailor policies for local practices.

The chaotic phenomenon of ED overcrowding increases the difficulty to evaluate AD effectiveness in the real world. The approach of computer simulation may be a means to solve the problem. The feasibility of simulation models of EDs[15–21] and emergency medical services system[22–24] have been well established. Methods such as simulations and queuing formulations were used to analyze problems such as waiting times and patients leaving without being seen. Simulations were also conducted to timely forecast ED crowding status[25,26] and to predict episodes of AD.[27]

Researchers used bicriteria to analyze the effect of AD on the ED performance considering patients' average waiting time and percentage of time spent on diversion.[28] However, other important issues such as the optimal timing to initiate AD, the rules of AD to divert different acuity levels of patients, and the proper interval of AD implementation are yet to be determined. To address these issues, a computer simulation program was developed in this study to evaluate the effectiveness of different AD strategies on ED overcrowding.

Materials and methods

Description of the simulation model for ED

The model consists of a set of theoretical probability distributions governing the flow of patients. The structure of the model conforms to the conceptual input—throughput—output framework of ED operations proposed by Asplin et al,[21] reflecting care processes that substantially contribute to ED overcrowding, which is illustrated in Fig. 1.

Patient arrivals (Fig. 1, point A) are simulated by a nonstationary Poisson process,[29] in which an exponential distribution governs the time between arrivals. The patient arrival rate may vary according to the time of day to reflect the actual patient arrivals. The simulation uses a previously reported algorithm[25] to implement the random nonstationary Poisson process. The ED receives ambulatory or ambulance-transported patients; the ratio between the two is an adjustable parameter in our model.

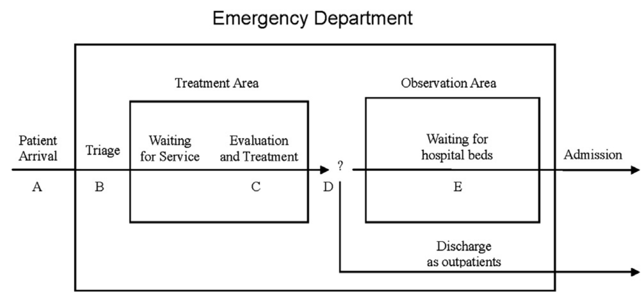

Figure 1 A simulation model for evaluating the effect of ambulance diversion strategies on emergency department overcrowding.

Upon arrival patients are triaged with an ordinal ranking scheme (Fig. 1, point B) with the numbers one to five representing the most to the least acute levels. The chance of belonging to each acuity level is governed by a multinomial distribution.[29] The simulation prioritizes patients for treatments by the most emergent acuity level and resolving ties according to waiting time. A maximal allowable waiting time is designated to patients of each acuity level. The simulation keeps track of waiting time of every patient and records any violation of exceeding the designated allowable waiting time.

We assume that more extensive ED care is required for sicker patients (Fig. 1, point C). The service time patterns are governed by the gamma distributions[29] of similar shapes that reflect the previously mentioned assumption. The shapes of the distributions can also be adjusted during the simulation according to the total number of patients in ED. This is to simulate the effect of ED overcrowding on ED operation conditions. In our model, we assume that ED overcrowding would reduce the efficiency of ED operation and prolong the service time of all patients.

Patients of different acuity levels may require different amounts of medical resources. Instead of keeping track of various kinds of resources, medical resources consumed by an acuity level i patient is simply represented by a positive number κ_i. These parameters essentially capture the average amount of medical resources consumed by patients of one acuity level relative to those by another.

After receiving emergency treatment, some patients might be admitted to the hospital (Fig. 1, point D). On completion of treatment for each patient, a random Bernoulli trial[29] is used to determine whether the patient should be admitted. The simulation uses a separate admission probability for each level of acuity so that patients with more acute conditions have higher probability to be admitted. Outpatients are immediately discharged, whereas the to-be-admitted patients are retained in the ED pending hospital bed availability.

It has been suggested that boarding of admitted patients in the ED is a major contributor to ED overcrowding (Fig. 1, point E).[21] While waiting in the observation area for an inpatient bed, an acuity level i patient is assumed to consume γ_i units of medical resources, which is less than the amount consumed by patients of the same acuity level in the treatment area. We assume that availability of inpatient beds is affected by some hospital processes with a fixed pattern and represent the process of hospital bed openings using a deterministic process, where one inpatient bed becomes available periodically and the period is a user-defined parameter.

Description of Crowdedness Index

The full capacity of ED is defined as the maximal loading of that ED, given the assumptions that: (1) all beds in the treatment and observation areas are fully occupied, and (2) patient acuity mix is the average.

The current loading of ED is defined as the medical resources currently occupied by all patients currently present in the ED, which is calculated as

$$\text{Loading of ED} = \sum_i \kappa_i N_i^T + \gamma_i N_i^0$$

where i denotes acuity levels, N_i^T is the number of patients of acuity level i in the treatment area, and N_i^O is the number of patients of acuity level i in the observation area. Recall that each acuity level i patient in the treatment area consumes κ_i units of medical resources, whereas that same patient in the observation area consumes γ_i units.

We define the Crowdedness Index (CI) as

$$\text{CI} = \frac{\text{current loading of ED}}{\text{full capacity of ED}}$$

Thus, a CI larger than 1.0 implies that: (1) additional beds are being used or (2) a full or nearly-full department with patients of greater than average acuity on the whole.

To reduce the complexity of the simulations, the current loading and the full capacity of ED in our model were conceptual. For example, if the patient with acuity 1 to 5 needs 20, 10, 5, 2, or 1 unit(s) of medical resources, respectively, and the ED can only serve 10 patients with each acuity at the same time, then the capacity of the ED would be $(20 + 10 + 5 + 2 + 1) \times 10 = 380$ units. The capacity of different EDs could be derived by this rule for comparison.

Accurate estimate of the average medical resources that each patient of a given acuity consumes is an interesting topic beyond the scope of this study. A rough estimate might be obtained from comparing the average medical expenses paid by patients of different acuity levels.

The processes of the patient flow and the essential parameters are summarized in Table 1.

Description of the AD strategies

Each AD strategy studied in this article uses CI as the only indicator for going on diversion; that is, AD is implemented when CI exceeds a certain threshold level, which is an adjustable parameter in our simulation model.

Three patient-blocking rules, that is, rules for diverting patients, could be considered in our simulation program: all AD (A-AD), high-acuity AD (H-AD), and low-acuity AD (L-AD). With implementation of A-AD, all patients transported by ambulance are diverted. By contrast, with implementation of H-AD, only patients with acuity levels 1 and 2 are diverted. Likewise, with L-AD, only patients with acuity levels 3 to 5 are diverted.

Once AD is implemented, patients will be diverted for a certain period of time, which is referred to as AD segment. Upon the end of the AD segment, the criterion for going on AD will be reassessed, and if the criterion is still satisfied, another AD segment will be initiated unless the total allowable AD duration (per day) is reached. The length of AD segment and the total allowable AD duration per day, regulated by a community-based consensus in reality, are both adjustable parameters in our simulations.

Outcome measures

We measure CI, average patient waiting time for service, and average percentage of adverse patients along the time to evaluate the effectiveness of each AD strategy. A patient is referred to as an adverse patient when the patient's waiting time for service exceeds the upper limit of waiting time for his or her acuity level. In our simulation model, the upper limits of waiting time for acuity levels 1 to 5 are specified by parameters μ_1 to μ_5, respectively. The upper limit of waiting time for each acuity level is regulated by legislation in some communities. The percentage of

Table 1 The processes and the essential parameters of the patient flow in the emergency department model.

Event of the Flow	Governed by	Essential Parameters
Patient arrivals	Non-stationary Poisson Process	λ: arrival rate
Triage	Multinomial distribution	T_i: probability of being acuity level i patient; i = 1,...,5
Treatment service	Gamma distribution	η_i: mean service time of acuity level i patient; i = 1,...,5
		ζ_i: variation of service time of acuity level i patient; i = 1,..,5
Admission/Discharge	Bermoulli trials	β_i: probability of acuity level i patient being admitted; i = 1,...,5
Hospital beds	Deterministic process	T_a: time interval between two available beds
Other Parameters		
κ_i: medical resources consumed by an acuity level i patient in the treatment area; i = 1,...,5		
γ_i: medical resources consumed by an acuity level i patient in the observation area; i = 1,...,5		
μ_i: upper limit of waiting time for acuity level i patient; i = 1,...,5		

adverse patients is defined as the ratio of the number of adverse patients to the number of total patients.

Discrete event simulations

This model is implemented in MATLAB language. For each ED strategy, every scenario was simulated 1,000 times to collect data for statistical analysis. The simulator, which follows the usual discrete event simulation engine logic,[30] consists of a main program, an initialization subroutine, a clock subroutine, several random number generators, several event programs, and a report generator. The events and the associate random variables are described in the previous section.

The main program controls the flow of the simulation as well as maintains an event list. When the simulation starts, the main program calls the initialization subroutine to initialize state variables and the clock, as well as to set up the initial event list. Then the main program calls the clock subroutine to advance the clock to the next scheduled event time, which in turn triggers the main program to call the corresponding event program.

The execution of the event program changes the status of the simulation and results in new events, and the state variables, statistics, and the event list will be updated before returning to the main program. The main program then calls the clock subroutine again and repeats the loop until the ending condition is reached. When the ending condition is reached, the main program stops the loop and calls the report generator to generate the required statistical report. Fig. 2 illustrates the flow diagram of the discrete event simulation procedure.

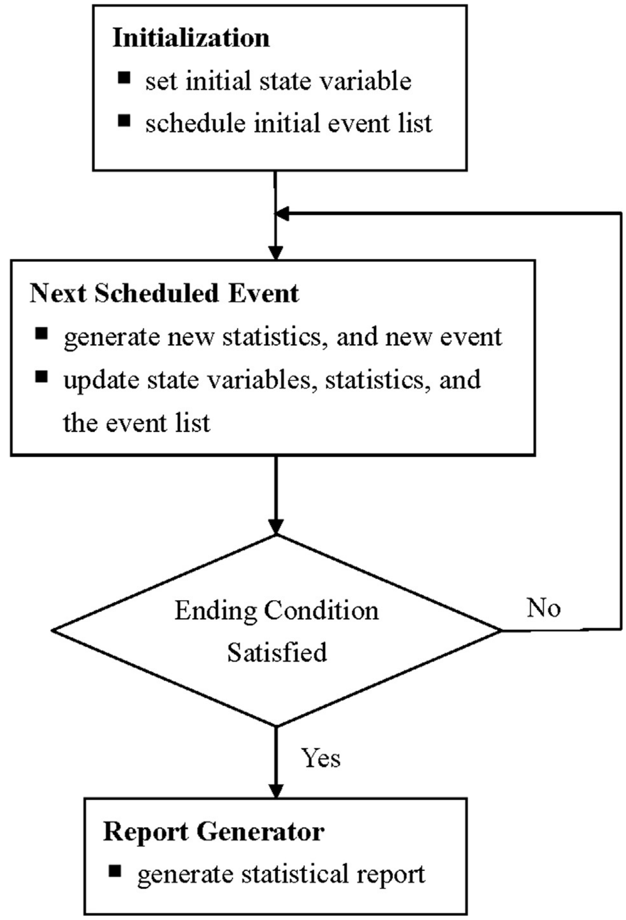

Figure 2 Flow diagram of discrete event simulation.

Assumption of the simulations

We assume the full capacity of ED is 100 units in our simulations. We assume a patient in the observation area only consumes half of the medical resources that a patient with the same acuity in the treatment area does. We assume all ED patients are brought in by ambulance in order to simplify the simulation. In reality, the percentage of ED patients transported by ambulance can be calculated by clinical observation and our simulation program can be easily amended to take walk-in patients into account if necessary.

All the simulation results are based on the parameter values shown in Table 2, many of which are derived from the observation of our ED. The parameter values used in our simulations could be obtained from clinical and administrative observations; in particular, the patient arrival rate, percentage of acuity in triage, average patient waiting time for service, the probability of hospital admission or discharge for each acuity level, the rate of hospital admission, etc., are readily available in many EDs. By adjusting these parameter values, one can easily generate models that represent many EDs of different scales.

Results

The following scenario is provided for better understanding of the simulation results described in this section and how to apply the proposed model to the real ED situations.

The ED has an average census of 10 patients per hour and is able to admit one patient to the wards every 10 minutes. The ED only accepts patients who were transferred by ambulance.

According to the registry data, 3% of patients were triaged as acuity level 1, whereas 20% were level 2, 40% were level 3, 27% were level 4, and 10% were level 5. Ninety percent of patients with acuity level 1 need to be admitted, whereas the admission rate for patients with acuity levels 2, 3, 4, and 5 are 70%, 50%, 30%, and 10%, respectively.

The local regulation requires ED physicians to assess patients with acuity level 1 within 1 minute of their ED arrival. For patients with acuity levels 2, 3, 4, and 5, the initial assess time since their ED arrival should not exceed 10 minutes, 30 minutes, 60 minutes, and 120 minutes, respectively.

Based on the internal consensus, the ED has a capacity of 100 units. Taking good care of each patient with acuity level 1 requires 20 units of resources. The ED loading for each patient with acuity levels 2, 3, 4, and 5 is 10 units, 5 units, 2 units, and 1 unit, respectively.

Question 1: Once the ED initiates AD, ambulances will be diverted for 2 hours. After that, the ED can accept incoming ambulances or initiate another round of AD. However, by the community consensus, the ED should not divert ambulances more than 8 hours per day. To effectively relieve ED overcrowding, when is the best timing for the ED to initiate

Table 2 Values of parameters assumed in the simulations.

λ = 10 patients/hour T_a = 1 bed/10 min

$\tau_1 = 0.03$	$\eta_1 = 300$ (min)	$\zeta_1 = 900$ (min)	$\beta_1 = 0.9$	$\kappa_1 = 20$ (units)	$\gamma_1 = 10$ (units)	$\mu_1 = 1$ (min)
$\tau_2 = 0.20$	$\eta_2 = 240$ (min)	$\zeta_2 = 576$ (min)	$\beta_2 = 0.7$	$\kappa_2 = 10$ (units)	$\gamma_2 = 5$ (units)	$\mu_2 = 10$ (min)
$\tau_3 = 0.40$	$\eta_3 = 180$ (min)	$\zeta_3 = 324$ (min)	$\beta_3 = 0.5$	$\kappa_3 = 5$ (units)	$\gamma_3 = 2.5$ (units)	$\mu_3 = 30$ (min)
$\tau_4 = 0.27$	$\eta_4 = 90$ (min)	$\zeta_4 = 81$ (min)	$\beta_4 = 0.3$	$\kappa_4 = 2$ (units)	$\gamma_4 = 1$ (unit)	$\mu_4 = 60$ (min)
$\tau_5 = 0.10$	$\eta_5 = 30$ (min)	$\zeta_5 = 9$ (min)	$\beta_5 = 0.1$	$\kappa_5 = 1$ (unit)	$\gamma_5 = 0.5$ (unit)	$\mu_5 = 120$ (min)

AD? That is, when the ED is very overcrowded, mildly overcrowded, or far less than overcrowded?

Question 2: To use the medical resources optimally, the ED should divert all patients when AD is initiated, or only divert patients with severe acuities, or only divert patients with mild acuities?

Question 3: By the community consensus, AD should not exceed 8 hours per day. The ED should initiate AD for 8 hours at once, or initiate AD for a shorter duration and reinitiate AD when necessary?

The answers to these three questions may be inferred from the results of the following AD strategy studies based on the proposed simulation model.

AD Strategy Study 1

This set of simulations is to assess the proper timing to initiate AD. The duration of each AD segment is set to be 2 hours, whereas the total allowable duration for AD implementation is set to be 8 hours in a day. The ED must open to accept patients once the total allowable AD duration is reached. Four AD initiating criteria are considered: AD is initiated only when the CI exceeds 0.5, 1.0, 1.5, or 2.0, respectively. The A-AD rule is implemented in this set of simulations.

The evolution of average CIs over 1 day are shown in Fig. 3. All curves exhibit the feature that the CI started to grow rapidly once the total allowable AD duration is reached. The 0.5 criterion was able to maintain the CI below 0.5 (that is, ED operates at half of its full capacity) for approximately 10 hours but then the CI quickly grows to 2.0 at approximately 15.5 hours. The 1.0 criterion, 1.5 criterion, or 2.0 criterion keeps the CI below 2.0 for approximately 20 hours.

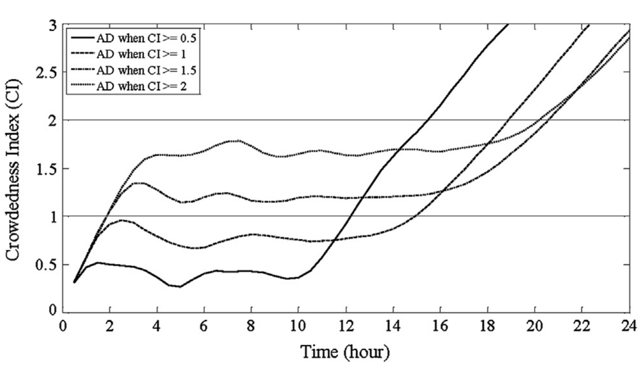

Figure 3 Ambulance Diversion Strategy Study 1: crowdedness index over 24 hours.

AD Strategy Study 2

This is to evaluate the effect of different patient-blocking rules, A-AD, H-AD, and L-AD, on relieving ED overcrowding. AD is initiated when the CI exceeds 1.0, and each AD segment is set to be 2 hours.

The average patient waiting times for service versus total allowable AD duration per day is shown in Fig. 4A. The figure shows that, regardless of the total allowable AD duration, the A-AD strategy results in the shortest average waiting time.

The average percentage of adverse patients versus total allowable AD duration per day is shown in Fig. 4B. The A-AD strategy results in almost no adverse patients if the total allowable AD duration is 14 hours or more per day. The H-AD and L-AD strategies have much higher numbers of adverse patients than using the A-AD strategy.

Provided the total allowable AD duration per day to be 8 hours, the CIs over 1 day are shown in Fig. 4C. CI is kept below 1.0 for almost 15 hours using the A-AD strategy. By contrast, neither the H-AD nor the L-AD strategy keeps CI below 1.0 after the second hour of the day. With the L-AD, the H-AD, and the A-AD strategies, the CIs increase rapidly after the 5th, 10th, and 15th hours, respectively.

Our findings suggest that, in certain settings, the A-AD strategy results in lower average patient waiting time for service, lower percentage of adverse patients, and lower CI, compared with the H-AD or the L-AD strategies.

AD Strategy Study 3

This is to compare the effectiveness of different AD segments. AD is initiated when the CI exceeds 1.0 and the A-AD strategy is implemented. The total allowable duration for AD implementation is set to be 8 hours per day. Three AD segment durations are compared: 2, 4, and 8 hours. That is, the ED can initiate a 2-hour AD for four times per day maximally, a 4-hour AD twice, or an 8-hour AD once.

The evolution of average CIs over a day is shown in Fig. 5A. Implementing the 2-hour AD multiple times can keep CI below 1.0 for the longest period compared with the other strategies. The 2-hour AD strategy also results in a steady CI curve, which implies that ED workload stays in a relatively constant status until the total allowable AD duration is reached.

Fig. 5B shows the average percentages of adverse patients. Implementing the 2-hour AD multiple times results in a smaller percentage of adverse patients in all acuity levels compared with the other strategies.

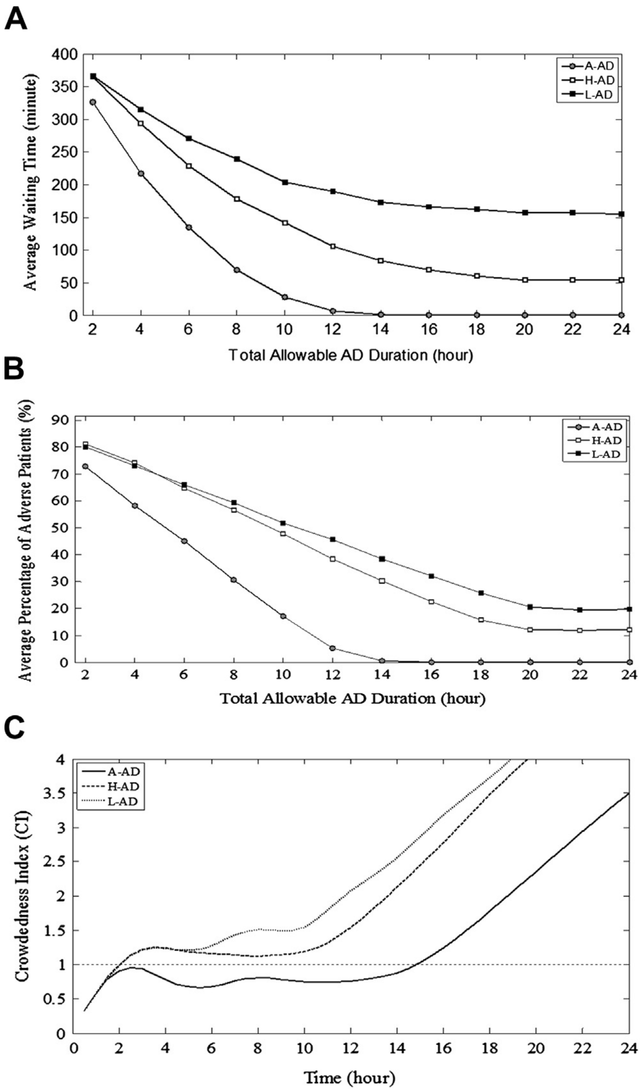

Figure 4 Ambulance Diversion Strategy Study 2. (A) Average patient waiting time versus total allowable ambulance diversion duration per day. (B) Average percentage of adverse patients versus total allowable ambulance diversion duration per day. (C) Crowdedness index over 24 hours.

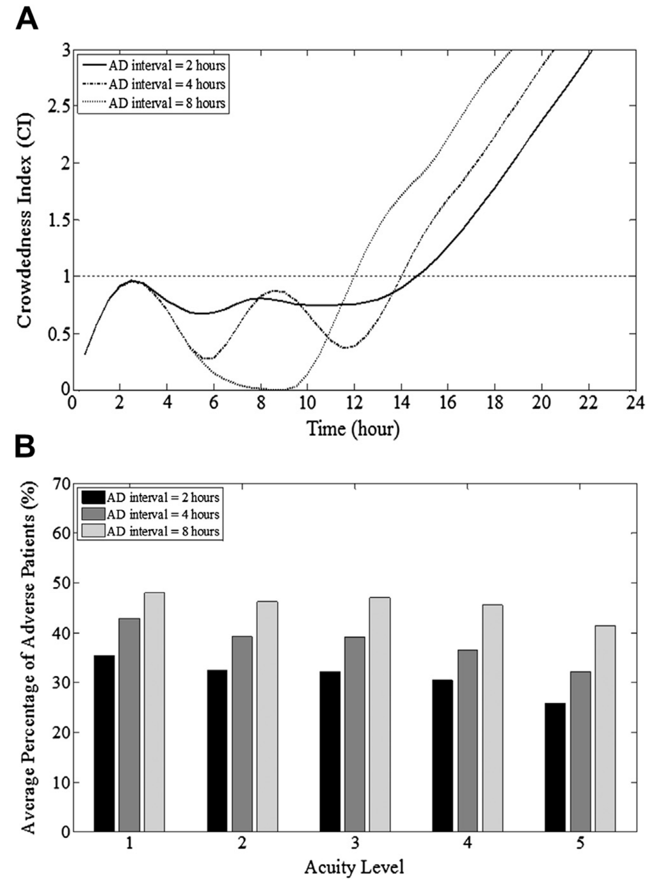

Figure 5 Ambulance Diversion Strategy Study 3. (A) Crowdedness index over 24 hours. (B) Average percentage of adverse patients in each acuity level.

Discussion

Three sets of AD strategy studies are conducted via computer simulations based on a simplified ED model where various ED operations are represented by statistical processes. Although the simplified ED model and the assumptions made on the statistical processes may not realistically reflect real ED operations, these processes can nevertheless capture the average dynamics of patient flows in EDs and are widely acceptable.[25–27] The simulation results we obtained, to some extent, do match the observation of our ED.

Study 1 examines the proper timing to initiate AD. AD is unlikely to be implemented unlimitedly in real practice. Many communities have regulations of the total allowable AD duration for individual hospitals. Under the condition that AD is initiated when CI exceeds 0.5 (i.e., the 0.5 criterion), we observe that CI increases dramatically once the total allowable AD duration is reached. The CI is more than 1.0 after the 12th hour, and is more than 2.0 after the 16th hour. This finding implies that if AD is implemented prematurely, the ED will operate at low capacity, and the overcrowding status follows soon after reaching the total allowable AD duration.

The CIs increased relatively slower using the 1.0 criterion, the 1.5 criterion, and the 2.0 criterion. Although the CIs in these scenarios are more than 2.0 eventually, the CI is significantly higher before the 20th hour using either 1.5 or 2.0 criteria, compared with that of using 1.0 criterion. This implies that, if AD is implemented in excessively overcrowded EDs, AD will not contain the problem sufficiently.

Our findings imply that the timing to initiate AD is important. In this parameter setting, initiating AD when CI just exceeds 1.0 will maintain the average CI slightly below 1.0 for most of the day and thus have the optimal ED performance. Having a real-time indicator of ED overcrowding is also essential to initiating AD at best timing.

Study 2 evaluates the effect of different patient-blocking rules on relieving ED overcrowding. The A-AD strategy has the lowest average percentage of adverse patients and the greatest effect on the average waiting time. The A-AD strategy also keeps CI below 1.0 during most of the day and thus achieves the best outcome.

Although the H-AD strategy has better outcome than the L-AD strategy on the average waiting time, their effects on the average percentage of adverse patients and the CIs appear to be similar according to the figures.

EDs may not be able to accept patients with severe injury in some circumstances; for example, due to temporary unavailability of space and essential equipment. The H-AD strategy is the choice in this situation because the ED is still able to continue services for low-acuity patients. By contrast, a tertiary trauma center may initiate the L-AD strategy and save resources for most critical patients in case of mass casualty incidents.

Field triage is more practical than ED triage when using H-AD or L-AD strategies in real practice. Patients may be diverted soon after being triaged by emergency medical technicians in the field rather than being triaged in the ED. The accuracy of field triage by emergency medical technicians seems promising[31]; however, the protocol compliance and quality insurance of local emergency medical services system should be justified.

Study 3 compares the effectiveness of different AD intervals. The 2-hour AD segment strategy has the lowest average percentage of adverse patients and thus may avoid preventable delay of essential medical treatment. This strategy also keeps the CI slightly below 1.0 for most of the day, which implies that the ED efficiently operates in a sustainable fashion.

While using the 8-hour AD strategy, the CI is reduced significantly at first, maintained in a very low level for hours, and then grows dramatically. The hours with low CI implies that the ED operates with low productivity. The 4-hour AD strategy has a similar but less rugged pattern.

The results suggest that implementing multiple times of small AD segments is better than one single large AD segment. The 2-hour AD strategy allows the ED to remediate but not cool down too much or too long. The ED reevaluates the CI after each 2-hour session of AD and then decides whether or not to continue AD.

Limitations

To reduce the complexity of the simulations, we intentionally ignore the time required for certain ED operations, such as triaging patients, cleaning the treatment area, and any administrative process. Furthermore, many aspects of ED management were grouped together into the treatment process. These included laboratory and radiologic examinations, administering of medication, pending consultations, explanation to obtain patient consents for certain procedures and treatment, medical education prior discharge, etc.

In the simulations, we assume hospital beds open at a fixed rate. Although this assumption does not realistically reflect the real hospital operation, it nevertheless does not largely change the outcomes in our simulations. In our simulations, we assess the input as the main contributor of crowding, as we deliberately overload ED capacity by setting an excessive patient influx. To assess the output as the key source of crowding, it would be necessary to choose a more realistic statistical process to represent hospital bed availability.

Our model may simulate EDs of different sizes. This study only focuses on one single ED; however, it is feasible to expand the simulations to evaluate the effect of AD strategies on a community basis. The effect of AD strategies may vary according to the number of hospitals in a single community, as well as the treatment capability and capacity of each ED.

Conclusion

An input—throughput—output simulation model is proposed for simulating daily ED operation. By appropriate parameter settings, the model can represent medical resource providers of different scales, from regional medical centers to local hospitals. Effectiveness of several AD strategies on relieving ED overcrowding was assessed via computer simulations based on this model. It is also feasible to expand the simulations to evaluate the effect of AD strategies on a community basis. The results may offer insights for making effective AD policies.

Acknowledgments

C.-H. Lin and C.-Y. Kao conceived the study. All authors contributed substantially to the study design. C.-H. Lin and C.-Y. Kao obtained the research grant. C.-Y. Huang implemented the proposed simulation model into a software program. C.-Y. Kao and C.-Y. Huang performed the statistical analysis. C.-H. Lin and C.-Y. Kao drafted the manuscript, and all authors contributed substantially to its revision. C.-Y. Kao is the corresponding author who takes responsibility for the paper as a whole.

References

1. Kellermann AL. Crisis in the emergency department. *N Engl J Med* 2006;**355**:1300—3.
2. Richardson SK, Ardagh M, Gee P. Emergency department overcrowding: the Emergency Department Cardiac Analogy Model (EDCAM). *Accid Emerg Nurs* 2005;**13**:18—23.
3. Epstein SK, Tian L. Development of an emergency department work score to predict ambulance diversion. *Acad Emerg Med* 2006;**13**:421—6.
4. Weiss SJ, Derlet R, Arndahl J, Ernst AA, Richards J, Fernandez-Frackelton M, et al. Estimating the degree of emergency department overcrowding in academic medical centers: results of the National ED Overcrowding Study (NEDOCS). *Acad Emerg Med* 2004;**11**:38—50.
5. Espinosa G, Miro O, Sanchez M, Coll-Vinent B, Milla J. Effects of external and internal factors on emergency department overcrowding. *Ann Emerg Med* 2002;**39**:693—5.
6. Anantharaman V. Impact of health care system interventions on emergency department utilization and overcrowding in Singapore. *Int J Emerg Med* 2008;**1**:11—20.
7. Nash K, Nguyen H, Tillman M. Using medical screening examinations to reduce emergency department overcrowding. *J Emerg Nurs* 2009;**35**:109—13.
8. Siegel B. Triage for overcrowding. Hospitals should fix the emergency department problems they can control. *Mod Healthc* 2003;**33**:24.
9. Burt CW, McCaig LF, Valverde RH. Analysis of ambulance transports and diversions among US emergency departments. *Ann Emerg Med* 2006;**47**:317—26.

10. Nawar EW, Niska RW, Xu J. National Hospital Ambulatory Medical Care Survey: 2005 emergency department summary. *Adv Data* 2007;**386**:1—32.
11. Scheulen JJ, Li G, Kelen GD. Impact of ambulance diversion policies in urban, suburban, and rural areas of Central Maryland. *Acad Emerg Med* 2001;**8**:36—40.
12. Warden CR, Bangs C, Norton R, Huie J. Temporal trends in ambulance diversion in a mid-sized metropolitan area. *Prehosp Emerg Care* 2003;**7**:109—13.
13. Handel DA, Pines J, Aronsky D, Genes N, Ginde AA, Hackman J, et al. Variations in crowding and ambulance diversion in nine emergency departments. *Acad Emerg Med* 2011;**18**:941—6.
14. United States General Accountability Office. Hospital emergency departments: crowding continues to occur, and some patients wait longer than recommended time frames. GAO-09-347 Report to the Chairman, Committee on Finance, U.S. Senate. April-30-2009. Available on-line.
15. Hallstrom A, Eisenberg MS, Bergner L. Modeling the effectiveness and cost-effectiveness of an emergency service system. *Soc Sci Med Med Econ* 1981;**15C**:13—7.
16. Brenner S, Zeng Z, Liu Y, Wang J, Li J, Howard PK. Modeling and analysis of the emergency department at University of Kentucky Chandler Hospital using simulations. *J Emerg Nurs* 2010; **36**:303—10.
17. Bair AE, Song WT, Chen YC, Morris BA. The impact of inpatient boarding on ED efficiency: a discrete-event simulation study. *J Med Syst* 2010;**34**:919—29.
18. Eitel DR, Rudkin SE, Malvehy MA, Killeen JP, Pines JM. Improving service quality by understanding emergency department flow: a White Paper and position statement prepared for the American Academy of Emergency Medicine. *J Emerg Med* 2010;**38**:70—9.
19. Saunders CE, Makens PK, Leblanc LJ. Modeling emergency department operations using advanced computer simulation systems. *Ann Emerg Med* 1989;**18**:134—40.
20. de Bruin AM, van Rossum AC, Visser MC, Koole GM. Modeling the emergency cardiac in-patient flow: an application of queuing theory. *Health Care Manag Sci* 2007;**10**:125—37.
21. Asplin BR, Magid DJ, Rhodes KV, Solberg LI, Lurie N, Camargo Jr CA. A conceptual model of emergency department crowding. *Ann Emerg Med* 2003;**42**:173—80.
22. Valenzuela TD, Goldberg J, Keeley KT, Criss EA. Computer modeling of emergency medical system performance. *Ann Emerg Med* 1990;**19**:898—901.
23. Su S, Shih CL. Modeling an emergency medical services system using computer simulation. *Int J Med Inform* 2003;**72**:57—72.
24. Channouf N, L'Ecuyer P, Ingolfsson A, Avramidis AN. The application of forecasting techniques to modeling emergency medical system calls in Calgary, Alberta. *Health Care Manag Sci* 2007;**10**:25—45.
25. Hoot NR, LeBlanc LJ, Jones I, Levin SR, Zhou C, Gadd CS, et al. Forecasting emergency department crowding: a discrete event simulation. *Ann Emerg Med* 2008;**52**:116—25.
26. Hoot N, Aronsky D. An early warning system for overcrowding in the emergency department. *AMIA Annu Symp Proc* 2006:339—43.
27. Leegon J, Hoot N, Aronsky D, Storkey A. Predicting ambulance diversion in an adult Emergency Department using a Gaussian process. *AMIA Annu Symp Proc* 2007:1026.
28. Nafarrate AR, Fowler JW, Wu T. Bi-criteria analysis of ambulance diversion policies. *Proceedings of the 2010 Winter Simulation Conference* 2010:2315—26.
29. Gross D, Shortle JF, Thompson JM, Harris CM. *Fundamentals of queueing theory*. 4th ed. New York: John Wiley and Sons; 2008.
30. Banks J, Carson J, Nelson BL, Nicol D. *Discrete-event system simulation*. 5th ed. New Jersey: Prentice Hall; 2009.
31. Newgard CD, Zive D, Holmes JF, Bulger EM, Staudenmayer K, Liao M, et al. A multisite assessment of the American College of Surgeons committee on trauma field triage decision scheme for identifying seriously injured children and adults. *J Am Coll Surg* 2011;**213**:709—21.