Markov decision process model for patient admission decision at an emergency department under a surge demand

Hyun-Rok Lee¹ · Taesik Lee¹

Published online: 8 February 2017
© Springer Science+Business Media New York 2017

Abstract We study an admission control problem for patients arriving at an emergency department in the aftermath of a mass casualty incident. A finite horizon Markov decision process (MDP) model is formulated to determine patient admission decisions. In particular, our model considers the time-dependent arrival of patients and time-dependent reward function. We also consider a policy restriction that immediate-patients should be admitted as long as there is available beds. The MDP model has a continuous state space, and we solve the model by using a state discretization technique and obtain numerical solutions. Structural properties of an optimal policy are reviewed, and the structures observed in the numerical solutions are explained accordingly. Experimental results with virtual patient arrival scenarios demonstrates the performance and advantage of optimal policies obtained from the MDP model.

Keywords Disaster response · Emergency department · Admission control · Mass casualty incident · Markov decision process · Optimal policy

This manuscript is based on and extended from the authors' work presented at the 2015 Health Care Systems Engineering Conference, Lyon, France.

✉ Taesik Lee
taesik.lee@kaist.edu

Hyun-Rok Lee
hyunrok@kaist.ac.kr

¹ Industrial and Systems Engineering, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea

1 Introduction

We study an admission control problem of an emergency department (ED) facing a surge in demand in the aftermath of a mass casualty incident (MCI). Selective admission of patients has been suggested in der Heide (2006) as a potential means to cope with the demand surge. Under a selective admission policy, a decision to admit a patient is made by considering various factors including the number of remaining beds and likelihood of high-severity patient arrivals in the future. Such policies are found in the disaster preparedness plan in many hospitals, but there is a lack of rigorous analysis in the literature to support decision-making for selective admission in an ED.

When an MCI occurs, a large number of patients simultaneously require medical attention. Although there is no universally accepted definition for an MCI, literature in the field of emergency medical services (EMS) and disaster medicine seems to suggest an MCI typically involves more than dozens of patients (Arnold et al. 2003; Reilly and Markenson 2010). This surge in demand for EMS can overwhelm the local EMS system in the region. Hospitals, emergency rooms (ERs) in particular, experience a temporary shortage of medical resources, and efficient use of scarce resources is critical in saving more lives.

An effective EMS response to an MCI situation requires a high level of disaster preparedness in a few dimension, but it is difficult to attain in the real world (Institute of Medicine of the National Academies 2005). The first dimension is the ability to promptly secure extra capacity in response to the demand surge, but there is a severe limitation as to how much and how soon such extra capacity can be secured in a typical emergency department. Many hospitals and EDs are already operating with their capacity fully stretched as overcrowding in EDs is a pandemic problem all over the world. To make things worse, low-severity patients tend to arrive early at an ED and occupy its resources first, causing shortage of resources for high-severity patients that typically arrive later (der Heide 2006; Hogan et al. 1999; Waeckerle 1991). For instance, when the Oklahoma City bombing happened in 1995, there was a first wave with low-severity patients and then following the second wave with high-severity patients (Hogan et al. 1999). Another important dimension in the preparedness is coordination among the response resources, but often there is a lack of communication among the ambulance operations, incident command, and EDs in the regions (der Heide 2006; Niska and Burt 2005). This causes EDs responding to MCIs operate based on its own, partial perception of the situation. All these factors can lead to an increase in overall mortality from MCIs.

As a strategy to mitigate the demand surge, we model the selective admission decision problem of an ED under an MCI situation. Specifically, the decision problem concern whether the ED should admit a currently arriving patient or send the patient to some other ED to save its resource for future use. Our objective is to provide an optimal patient admission policy so that an ED can best utilize its scarce resources in response to MCIs. We formulate this decision problem as a Markov decision process (MDP) model with a finite planning horizon. Optimal decisions depend on the amount of remaining resources (i.e., how many more patients can we

treat) and the elapsed time from the onset of the MCI (i.e., how long has passed since the MCI).

There exist several papers in prior research that address the issue of resource allocation problem for the pre-hospital phase of MCI responses. But, to our best knowledge, the ED admission control problem under MCI situations has not been studied yet. Our work contributes to the disaster response literature by providing a rigorous decision model and rich analyses for the practically relevant yet unexplored problem in EMS system's disaster preparedness.

The remainder of this paper is structured as follows. Literature related to our problem is discussed in the Sect. 2. In Sect. 3, we present an MDP formulation to model a selective patient admission problem at an ED. Properties of an optimal policy and a numerical solution method for the MDP model are discussed. Then, experimental results for various disaster scenarios are presented in Sect. 4 to illustrate structural characteristics of the admission policies. Then we conclude our paper in Sect. 5.

2 Related literature

Our problem is a selective admission of patients into an ED, and it is closely related to the three concepts in the EMS and disaster medicine: triage, resource-constrained patient prioritization, and ambulance diversion. In this section, we review prior research in each of the three areas, and highlight how our work builds upon and contributes to the existing body of knowledge in the literature.

Triage refers to the process of prioritizing patients for resource allocation, and it is typically used in the context of ED operation and disaster responses. Triage in the context of daily ED operation is the first step of the patient flow in an ED; upon their arrival at an ED, patients are quickly assessed for their clinical condition, based on which the priority to receive treatment is determined. There are several triage systems, and 5-level triage systems, e.g., emergency severity index (ESI), are known to be reliable measures and widely used in many countries (Christ et al. 2010). In disaster medicine, triage refers to a scheme to quickly assess and classify a large number of patients at the scene of a disaster. Four-class classification is used in most widely accepted disaster triage systems, namely simple triage and rapid treatment (START) (Jenkins et al. 2008). By few quick and simple tests, victims are classified into four categories visually indicated by corresponding color codes: deceased/expectant (black), immediate (red), delayed (yellow), and minor (green). The two triage systems serves the same goal on surface—"quick sorting of patients in the setting of constrained resources", but the circumstances and philosophy behind it are quite different. In our work, we limit our attention strictly to the disaster triage.

Triage carries two meanings; first, it is a classification system to categorize patients according to the urgency of their medical condition. Second, it is a resource rationing system under the situation where available resources are less than needed. Most medical research focuses on the appropriateness of the classification part of the system. In this view, the allocation of resources is a direct, linear translation

from the classification system. While this view of triage is a standard practice, many recent studies point out that this relatively simple approach may not be the best approach to patient prioritization (Argon et al. 2008; Frykberg 2002; Jacobson et al. 2012; Jenkins et al. 2008; Mills et al. 2013; Sacco et al. 2005, 2007). These studies argue that a high priority in the classification system does not necessarily imply a high priority in the resource rationing decision under severely resource-constrained situation such as MCI cases. This is due to the possibility that the loss in the lower-class patient's chance of survival by delaying care can be larger than the gain from treating a higher-class patients first; as unfortunate as it sounds, forgoing an attempt to save one patient in the immediate-class can lead to life saving of two delayed-patients. Hick et al. (2012) suggested that emergency physicians should adjust their strategies according to the demand level and the resource availability, and several studies present triage concept that take into account the resource availability to most effectively ration the limited resources (Argon et al. 2008; Jacobson et al. 2012; Li and Glazebrook 2010; Mills et al. 2013; Sacco et al. 2005, 2007; Sung and Lee 2016).

Sacco et al. (2005) is the first to construct a mathematical formulation to frame triage as a resource allocation problem, and it is referred to as Sacco triage method (STM). They proposed a linear programming model to determine the patient priorities for ambulance transport from an MCI site. Since STM, this so-called resource-constrained triage problem has been addressed in several other studies by using more sophisticated models, primarily stochastic scheduling models and fluid models for a clearing system. Li and Glazebrook (2010) proposed a semi-Markov decision process model and solved this model with approximate dynamic programming. Their solution confirmed that patients in the highest medical severity class are not always given the highest transport priority. Jacobson et al. (2012) solved a similar scheduling problem and derived a number of characteristics of an optimal policy under various MCI conditions, and found that the optimal policy gives higher priority to a low triage class patients when there are many low triage class patients compared with high triage class patients. Mills et al. (2013) formulated a continuous optimization problem using survival probability functions for two patient classes, and propose resource-based simple triage and rapid treatment (ReSTART) protocol. Sung and Lee (2016) added the hospital component to the triage problem, and formulated a patient prioritization and a hospital selection problem as ambulance routing problem. Our model shares the theme with these studies in that it focuses on the resource allocation/rationing portion of the triage.

Development of the concept and analytic models of the resource-constrained triage has made an important contribution to the patient prioritization method. Yet, these studies primarily focus on the pre-hospital phase of EMS operations. That is, they mostly deal with patient transport priority at the disaster site. It should be noted that there exists an important prioritization problem at the hospital phase, i.e., ED operation, as well. Prioritization decision in EDs is an important component to effectively cope with demand surge from MCIs, but there exist few analytic research on the subject. Resource allocation problems in a hospital have been addressed in the context of disaster preparedness (der Heide 2006; Einav et al. 2006; Hick et al. 2012; Peleg and Kellerman 2009; Timbie et al. 2013; Waeckerle 1991),

but the nature of these studies are largely conceptual. Analytic approaches based on mathematical models will provide scientific support and augment the conceptual principles and guidelines. For example, Kilic et al. (2013) presents a queuing model for disaster situations and uses the model to determine the optimal service level in response to surge demand experienced at a hospital. In another example, Cohen et al. (2014) formulates a continuous optimization problem to determine the optimal allocation of doctors in two treatment areas. We need many more studies to address this void, and our paper makes contribution in this area.

Though not directly relevant to disaster responses, a decision problem for ambulance diversion is closely related to our problem. Ambulance diversion is one of the strategies to mitigate the overcrowding problem in an ED. It is a practice of diverting incoming ambulances to other EDs when it is experiencing severe overcrowding. In the sense that it controls the incoming demand by blocking them, it is similar to the admission control problem that we are dealing with. Since ambulance diversion can negatively affect patient outcomes, the decision to declare the diversionary state must be made with great caution. Naturally, there exist some studies which interpret, analyze and design ambulance diversion policies (Allon et al. 2013; Deo and Gurvich 2011; Ramirez-Nafarrate et al. 2014; Xu and Chan 2016). Technical approaches taken in these studies are diverse, e.g., queuing models (Allon et al. 2013; Deo and Gurvich 2011; Xu and Chan 2016), game theoretic approach (Deo and Gurvich 2011), and infinite horizon MDP (Ramirez-Nafarrate et al. 2014). While ambulance diversion problem and the selective admission problem shares a common problem of making a decision to when to block which patients, our problem differs in that it is a transient problem with a finite horizon whereas ambulance diversion problems are typically analyzed for a system under the steady state.

To our best knowledge, the resource-constrained triage at the hospital-phase, its admission control problem in particular, has not been studied. Yet, this problem is very relevant to the real world operation. For example, we have reviewed the disaster response protocol of three major teaching hospitals in Korea. These protocols explicitly specify transferring less severe patients to other hospitals when they experience surges in demand due to disaster event. In summary, we contribute to the literature by modeling and analyzing the ED admission control problem during the hospital phase of EMS operation, which is a practically relevant yet unexplored problem in the disaster response literature.

3 MDP model

In this section, we describe a selective patient admission problem, and present its formulation as an MDP model. It is a multi-class admission control problem over a finite horizon with time-dependent arrival rate and time-dependent rewards. Then we discuss the properties of optimal policy structure, followed by numerical solution method we used in this paper.

3.1 Problem description

Specifically, we consider the following problem. A large scale MCI has just been reported. Patients suffer from injury, some minor and others severe. They are triaged at the MCI site as red, yellow, green and black. Soon, injured patients are arriving at an ED. A large number of patients arrive soon after the MCI, but the number of arrivals tapers as time goes on. Patients with a less severe injury tend to arrive earlier than more severe patients. Suppose a yellow-class patient arrives at the door of the ED, and there are only a handful of beds available at the moment. Do we admit the patient to provide care or divert the patient to somewhere else so we can save a bed for more critical patient in the future?

Our problem is to determine whether to admit a patient into an ED for each arriving patient from an MCI. MDP is a suitable modeling technique to formulate these kind of sequential decision-making problems. The goal is to find the optimal action for each system state when a system state stochastically changes under the influence of the chosen action (Puterman 1994). The set of optimal actions constitute an optimal policy, and decision-making according to the optimal policy maximizes expected rewards in the decision-making problem.

It turns out that our MDP model is structurally similar to the model discussed in (Brumelle and Walczak 2003). Brumelle and Walczak (2003) solves a revenue maximization problem for a single flight through admission control of multiple fare classes with time-dependent arrivals to a single-leg flight. In fact, our model can be viewed as a special case of the generalized version of the MDP model in Brumelle and Walczak (2003). For example, Brumelle and Walczak (2003) defines the augmented states concerning batch arrivals of jobs; it also includes the reward function's dependency on the remaining inventory level.

While it bears a similarity to the model in Brumelle and Walczak (2003) in terms of its structure, our model relaxed a few key assumptions made in Brumelle and Walczak (2003) to describe ED operation under MCI situation. Specifically, Brumelle and Walczak (2003) made assumptions on the structural properties of a reward function and the action sets to maintain mathematical tractability in their analysis of the optimal policy's structural properties. In this work, we relax those assumptions to fit the plausible context of ED admission decision problem. This requires to develop a numerical solution method so that experiments under various MCI environments are conducted to explore possible insights for EMS operation.

3.2 Formulation

We make few assumptions to construct a decision-making model for ED patient admission under an MCI situation. First, for patient arrivals to an ED, we assume that patients of each severity class arrive according to independent nonhomogeneous Poisson process (NHPP) respectively. With different NHPP for each victim class, we intend to represent the following patterns (Arnold et al. 2003; der Heide 2006; Einav et al. 2006; Stratton and Tyler 2006): (1) the volume of patient arrivals at an ED peaks after some time from the onset of an MCI and decreases afterwards, (2) relatively-less injured patients arrive at an ED earlier than more-severe patients,

and (3) volume of less-injured patients is greater than more-severe patients. Note that this assumption of NHPP arrivals has potential limitations to model actual patient arrivals process in an MCI. Most critically, batch arrivals at an ED are possible in MCI situations since an MCI tends to generate bulk of patients at one point in time, for example in a building collapse. Nevertheless, in the absence of other empirically validated alternatives, we follow previous research in the literature to adopt NHPP assumptions (e.g., Kilic et al. 2013).

Next, we assume the medical condition of a patient continues to deteriorate until definitive care such as surgery is provided. For our MDP model, it means that treating patients at an ED has a time-dependent reward. The time-dependent reward in our model is derived from a survival probability function proposed in prior literature (Mills et al. 2013). This assumption of treating time as a sole determinant for patients' medical condition and their survival probability may be an oversimplification. One may wish to develop a more sophisticated reward function to incorporate other factors that influence the patient condition, but unfortunately, to our best knowledge, such a reward function is not available in the literature. Developing such reward functions that are justifiable clinically and operationally is beyond the scope of this paper. Given the limitation, we have chosen to use the survival probability function for trauma patients which was originally developed based on expert opinions (Sacco et al. 2005, 2007).

Third, we consider the number of beds as a capacity indicator of an ED for admission decision. The number of beds is a reasonable proxy to represent the capacity of an ED as the staffing levels of doctors and nurses in an ED are typically determined proportionately to the number of beds. Hence, the number of beds in our model is interpreted as the number of patients that an ED can admit, which takes staffing levels, beds, and other resources into account. Note also that it is assumed that once a bed is given to a patient, it will not become available during the planning horizon. In other words, we assume that the decision time horizon is short compared to the treatment time needed for patients and no patients leave the ED within that time horizon.

Below, we define our MDP model by five elements: decision epoch, state, action, transition probability, and reward function. Specifics of these elements for our model are as follows:

- Decision epoch

For our model, a decision epoch is the moment of a patient arrival. An admission decision is made when and only when a patient arrives at an ED. No patient arrives between decision epochs.

- System state

The state of the system is defined as vector (k, t, m) , where k is the number of remaining beds in the ED, t is the time elapsed from the occurrence of MCI, and m is the class of the current arriving patient out of the M patient classes. The number of remaining beds is non-negative integer, and we have a finite time horizon from 0 to T .

- Action

When there is an available bed in ED, there are two actions to take for a currently arriving patient at each decision epoch: admit or reject. When there is no remaining bed, the only feasible action is “reject patient”. Thus, action set A is a function of the number of remaining beds k . $A(k)$ is {admit patient, reject patient} if $k > 0$, and $A(k)$ is {reject patient} if $k = 0$.

- Reward

Reward is awarded when we take the “admit patient” action. Specifically, $r_m(t, \text{admit})$, is awarded when a class- m patient is admitted to ED at time t . If we reject a patient, no reward is given, i.e., $r_m(t, \text{reject}) = 0$. We use the patient’s survival probability function to model $r_m(t, \text{admit})$. This will be discussed in detail in Sect. 4.1.

- Transition probability

Transition of state (k, t, m) depends on the chosen action at the current state and the arrival process of patients. When there are k beds in the current state, next state can have either $k - 1$ or k beds. The number of remaining beds is reduced by one if we admit the current patient, and remains unchanged if we reject the patient. This part of transition is deterministic, depending only on the action chosen at the current state. The other two components, arrival time t and class m for a next arriving patient, are computed using the properties of NHPP arrival. Let $\lambda_m(t)$ denote an NHPP intensity function for class- m patients. $\lambda_m(t)$ is the arrival rate of class- m patients at time t . We define $\lambda(t)$ to be the sum of the arrival rates of all the patient classes, $\lambda(t) = \sum_m \lambda_m(t)$. For NHPP arrivals with $\lambda(t)$, the probability density that a next arrives at time τ given the current arrival at t is $\lambda(\tau) \exp(-\int_t^\tau \lambda(s) ds)$. The probability that a patient arriving τ is from class- m' is $\lambda_{m'}(\tau)/\lambda(\tau)$ due to superposition of NHPP. Thus, given the current time t , probability density that arrival time and class of a next arriving patient in the next state is τ and m' is computed as follows:

$$P_{\tau, m'|t} = \lambda_{m'}(\tau) \exp\left(-\int_t^\tau \lambda(s) ds\right). \quad (1)$$

Transition probability is then computed for a chosen action by using $P_{\tau, m'|t}$. For example, transition probability from state (k, t, m) to the next state $(k - 1, \tau, m')$ is $P_{\tau, m'|t} d\tau$ for infinitesimal interval $d\tau$ if the chosen action for the currently arriving patient is “admit” or 0 if “reject”.

Optimal policy for this MDP model can be obtained from an optimality equation, Eq. (2):

$$V(k, t, m) = \max_{a \in A(k)} \left[r_m(t, a) + \int_t^T \sum_{m'} \{P_{\tau, m'|t} \times V(k - 1_{\text{admit}}, \tau, m')\} d\tau \right]. \quad (2)$$

where $1_{\text{admit}} = 1$ if a is “admit patient” and 0 otherwise. In the equation, $r_m(t, a)$ is the reward function, as defined in Sect. 3.2. Argument a in Eq. (2) is an element in

$A(k)$ where $A(k)$ is {admit patient, reject patient}. $P_{\tau, m'|t}$ is a density function for state transition probability, and $V(k - 1_{admit}, \tau, m')$ denotes the value of the possible future states. Note in the future state, the number of beds decreases by 1 only when a patient is admitted in the current state. Note that the above equation is applied to states where $k > 0$. For the states where $k = 0$ (i.e., no beds available), V is zero because we cannot admit any patient in those and their future states.

Equation (2) maximizes the value of the current state by choosing an optimal action a . In doing so, it considers the values of possible future states, giving a recursive structure. This equation follows the principle of optimality, and uniqueness of this optimality equation is shown in (Brumelle and Walczak 2003). The solution to Eq. (2) is thus the optimal value function, and the set of action in each state constitutes the optimal policy to our MDP model.

3.3 Properties of policy structure

Our MDP model is structurally the same as the splittable-batches model studied in (Brumelle and Walczak 2003), hence sharing its structural properties. In this section, we briefly review some of the results in the context of our problem. See Brumelle and Walczak (2003) for detailed proofs and complete theorems.

- Monotonicity of optimal value function

The optimal value function for the model is monotonic. For each patient type, optimal value increases with the number of remaining beds k , and it decreases with the elapsed time t . Formally,

$$V(k + 1, t, m) \geq V(k, t, m) \quad \forall k \geq 0, \quad t \in (0, T], \quad m \in \{1, \dots, M\} \quad (3)$$

$$V(k, t, m) \geq V(k, s, m) \quad \forall k \geq 0, \quad t \in (0, T], \quad s \in [t, T], \quad m \in \{1, \dots, M\} \quad (4)$$

This can be proved by a sample path argument. Monotonicity of the optimal value function has an intuitive interpretation. Equation (3) suggests that we can expect greater reward when the number of remaining beds is large as we can admit more patients. Equation (4) implies that greater reward is expected at the earlier period of decision horizon due to higher future arrival of patients.

- Optimality condition of an action

A decision to reject a patient can be justified only when a bed saved for now can be used to admit a future patient with a higher reward. In other words, to admit a patient is an optimal action if the reward is greater than the reward from admitting any other patient in the future. If we know for sure $r_m(t, admit)$ for a current patient is greater than $r_m(t + \Delta, admit)$ for any future patient, we should admit the patient and it will be the optimal action to take. If, on the other hand, $r_m(t, admit)$ for a current patient is smaller than $r_m(t + \Delta, admit)$ for at least one future patient, an optimal action requires other factors, e.g., probability of those patients' arriving,

taken into account. A generalized condition to determine the optimality of an action can be derived by transforming the optimal value function:

$$\text{If } r_m(t, \text{admit}) \geq \int_t^T \sum_{m'} [P_{\tau, m'|t} \times \{V(k, \tau, m') - V(k-1, \tau, m')\}] d\tau, \quad (5)$$

then admit a patient in state (k, t, m) for $k > 0$,
otherwise reject a patient.

This condition states that a patient should be admitted if and only if the reward for admitting this patient is greater than the marginal value of a bed, which is the right-hand-side of Eq. (5). The marginal value of a bed is the expected value that can be gained in the future if a bed is kept in the given state and used later for other patient.

- Structure of optimal policy

A closer examination of the optimality condition in Eq. (5) reveals a special structure of the optimal policy. Suppose the optimal value function is discretely concave in k . Then the right-hand-side of Eq. (5) is decreasing in k . It implies that an optimal action, for a patient of type- m and at fixed t , is “admit patient” if there are more beds than some threshold number k^* . It in turn means that in the optimal policy, optimal action changes only once along the dimension of k , yielding a simple policy structure. In our model, it can be shown that the optimality equation preserves discrete concavity, and thus the optimal value function satisfies the discrete concavity condition. Another property that can be deduced from Eq. (5) is that the threshold value k^* decreases in time if $r_m(t, \text{admit})$ is non-decreasing and the marginal value of a bed is decreasing in time. It is directly evident from Eq. (5); if Eq. (5) holds for some t , then it is also true for $t + \Delta$ under the condition that r_m is non-decreasing and the right-hand side of Eq. (5) is decreasing in t . Note that for our model the marginal value of a bed decreases in time due to the positive reward function and the NHPP assumption for the arrival process.¹

3.4 Solution method: discretization and backward iteration

Difficulties in solving our MDP model stem from the continuous variable in the state definition. We need to solve for an optimal policy through optimality equation, Eq. (2), and due to the continuous state variable t it requires to compute the integration of the value function. It is usually impossible to analytically solve the integration for arbitrary arrival rate and reward functions. We use an approximation approach to numerically find the optimal policy by discretizing the continuous state variable t .

We discretize the entire time horizon into a finite set of equally-spaced time interval. The entire time horizon T is divided into N intervals, each of which has a length $\Delta t = T/N$. Then, the continuous state variable t is discretized as $t_1 = \Delta t$, $t_2 = 2 \times \Delta t, \dots, t_N = N \times \Delta t$. Admission decision in state (k, t_i, m) now concerns a patient who has arrived during time interval $(t_{i-1}, t_i]$. As a consequence of

¹ A generalized version of this statement is proved in Brumelle and Walczak (2003).

discretization, a decision epoch for an arriving patient is the nearest decision time after the moment of patient arrival. In other words, when a patient arrives during a time interval, admission decision for the patient is delayed until the end of the time interval so that the decision is made at a predefined time point. It should be noted that it is assumed here that only one patient arrives in a time interval, and thus accuracy of the obtained solution would increase as the discretization interval Δt decreases. While small Δt is preferred, making Δt small adds the computational burden, and thus the level of discretization needs to be determined such that it balances the approximation errors and computation time. We conducted preliminary experiments to test various level of discretization. It turns out that decreasing Δt below 0.1 min does not change the policy solution, only significantly increasing computation time. Thus we set $\Delta t = 0.1$ and $T = 720$ min in our experiments.

This discretization requires the transition probability Eq. (1) to be modified so that it accounts for the probability of a patient arrival during a time interval. Note that probability of a next patient's arriving in $(t_{j-1}, t_j]$, given the current arrival at t_i , is $\{exp(-\int_{t_i}^{t_{j-1}} \lambda(s)ds) - exp(-\int_{t_i}^{t_j} \lambda(s)ds)\}$. Thus, at the current elapsed time t_i , probability that arrival time and class of an arriving patient in the next state is t_j and m' is approximated as follows:

$$P_{t_j, m' | t_i} \approx \frac{\lambda_{m'}(t_j)}{\lambda(t_j)} \left\{ exp\left(-\int_{t_i}^{t_{j-1}} \lambda(s)ds\right) - exp\left(-\int_{t_i}^{t_j} \lambda(s)ds\right) \right\}. \quad (6)$$

Then the optimality equation, Eq. (2), is modified accordingly. With the discretized state space t_i , Eq. (2) is rewritten:

$$V(k, t_i, m) = \max_{a \in A(k)} \left[r_m(t_i, a) + \sum_{j=i+1}^N \sum_{m'} \left\{ P_{t_j, m' | t_i} \times V(k - 1_{admit}, t_j, m') \right\} \right]. \quad (7)$$

Now that the discretized MDP model has a finite state space, it can be easily solved by a backward iteration algorithm. The algorithm recursively computes the value of all states using the pre-calculated values of their possible next states. Note that the state transitions in our model proceed only in the direction of smaller k and larger t_i . Therefore, the value function of a state can be computed from the states close to the end of planning horizon t_N with a smaller k . We compute the optimal value starting from state $(1, t_N, m)$ for $m = 1, \dots, M$ until we get the optimal value for state $(1, 0, m)$. Then value function for state $(2, t_N, m)$ is computed for $m = 1, \dots, M$, and the process repeats until value functions for all states are computed. The boundary condition is that the value of the states with $k = 0$ is zero: $V(0, \cdot, \cdot) = 0$.

4 Computational results

4.1 Test instances

We construct hypothetical MCI scenarios to examine the optimal policies obtained from the model. To construct the test scenarios, we refer to the examples and

assumptions from previous literature (Arnold et al. 2003; Einav et al. 2006; Hogan et al. 1999; Reilly and Markenson 2010; Stratton and Tyler 2006). For the sake of simplicity in analysis, we consider only two severity classes. Assuming EMS response is practiced following the principles of disaster medicine, use of EMS resources under an MCI situation must be strictly limited to the patients under critical conditions. Therefore, a selective admission decision-making problem under MCI situations concerns only the immediate- (red) and delayed-class (yellow) patients who need definitive care for their survival. Those who suffer minor, non-life-threatening injuries (green) and those with no chance of survival (black) are excluded in resource allocation decision. It does not imply, however, that our model is limited to handle no more than two classes of patients. Additional patient-classes can be included as long as their arrival rate and reward functions are appropriately defined. It should be noted that additional patient-classes do not change the structure of the model or computational procedures.

Faster provision of care is strongly desired for both immediate- and delayed-class patients as their medical condition deteriorates with time, reducing the survival probability. But survival probability for patients in each class decreases at a different rate, which makes this decision problem non-trivial. For patient arrivals, we use NHPP with its intensity functions defined by a modified form of gamma distribution (Kilic et al. 2013). These intensity functions well depicts patient arrival patterns known to exist in MCI situations; patient arrivals rapidly increase at the beginning, reaches its peak, and gradually decreases to zero. Parameters α , β and scaling constant are varied to generate diverse MCI scenarios in terms of total number of patients and arrival patterns as shown in Table 1.

There are three main factors that we vary across the MCI scenarios: scale of an incident, peak time of patient arrival, and a ratio between immediate- and delayed-class patients. The scale factor controls the expected volume of patients from an incident, and we use 10 levels from 30 to 120. Peak time of the arrivals are controlled by shape parameter for the intensity function (α_I, α_D). For convenience,

Table 1 NHPP intensity functions and parameter values

Patient class	Arrival rate function (per min)
Immediate, $\lambda_I(t)$	$\frac{C_I}{60} \frac{(t/60)^{\alpha_I-1} \exp(-\beta_I t/60)}{\Gamma(\alpha_I) \beta_I^{\alpha_I}}$
Delayed, $\lambda_D(t)$	$\frac{C_D}{60} \frac{(t/60)^{\alpha_D-1} \exp(-\beta_D t/60)}{\Gamma(\alpha_D) \beta_D^{\alpha_D}}$
Parameter	Values
Total number of patients, $C_I + C_D$	30, 40, 50, 60 , 70, 80, 90, 100, 110, 120
Alpha for immediate, α_I	1.5, 2.5 , 3.5, 4.5, 5.5
Alpha for delayed, α_D	1.5 , 2.5, 3.5, 4.5, 5.5
Severity ratio, $C_I : C_D$	3:1, 2:1, 1:1, 1:2, 1:3 , 1:4, 1:5, 1:6, 1:7

Bold text indicates a reference scenario

we set β at 1 for all scenarios, and change α values with five levels from 1.5 to 5.5. These α values correspond to the peak arrival time of 30–270 min with an increment of 60 min. The ratio between the immediate- and delayed-class patients ranges from 3:1 to 1:7. In a reference scenario, total of 60 patients are expected to arrive at an ED. Forty-five of them are delayed-class patients, and their peak arrival will be around at 30 min after the incident. Fifteen immediate-class patients are expected with a peak arrival at 90 min after the incident. Table 1 summarizes the parameter values used for our experiments.

In addition to the arrival patterns, reward function is what makes the patient admission an interesting decision problem. A key element is that, given certain structure of reward function, the benefit from admitting a patient of one class over the other class may change over time. This time-dependence is considered in making an optimal decision along with two other factors, elapsed time (that affects future arrivals) and remaining beds. Recall, in our formulation, we define that reward $r_m(t, \text{admit})$ is awarded when a class- m patient is admitted, and no reward is given for a reject decision. That is, $r_m(t, \text{reject}) = 0$. Hereafter, we refer to $r_m(t, \text{admit})$ simply as a reward or reward function.

We test three hypothetical reward functions for our experiments: time-independent reward, time-dependent reward with delay penalty, and time-dependent reward with time-shift penalty. For convenience we use acronyms TI, TDDP, and TDTS respectively. For TI, we ignore the temporal change of the survival probability, and use a constant reward for both classes of patients. We assume the reward for admitting an immediate-patient is 4 times greater than admitting a delayed patient.² Ignoring the dynamic aspect of reward function, this establishes the most simple-minded reward structure.

For TDDP and TDTS, we define rewards based on the survival probability. We assume that patient's medical condition deteriorates in time until definitive care is provided. This has been quantitatively represented as a survival probability function in the literature (Mills et al. 2013; Sacco et al. 2005, 2007). In this study, we adopt the survival probability function for the two classes of patients, $f_I(t)$ and $f_D(t)$, proposed in (Mills et al. 2013).³ Figure 1 shows the survival probability function used in our experiment.

We define TDDP and TDTS reward by the relative gain in an admitted patient's survival probability with respect to a case where the same patient is rejected (diverted) and seeks care in another ED. For example, suppose we have an immediate-patient arriving at τ . If we admit this patient, then the patient's probability of survival is $f_I(\tau)$. If not, the patient would be diverted to another ED to receive care, and the delay in obtaining care would reduce the patient's survival probability to $\tilde{f}_I(\tau)$. We use the difference, $f_I(\tau) - \tilde{f}_I(\tau)$, as the value for the reward for admitting the patient. In TDDP, we assume that diversion of a patient reduces a patient's survival probability by a discount factor: $\tilde{f}_I(t) = \kappa_I f_I(t)$ and $\tilde{f}_D(t) = \kappa_D f_D(t)$. For TDTS, we compute $\tilde{f}(t)$ from

² Factor of 4 is adapted from Korea National Emergency Medical Center (2007), an empirical study on mortality rate for interhospital transfer between severe and non-severe patient groups.

³ Mills et al. (2013) shows five different survival probability scenarios, and we use the most optimistic scenario in this study.

Fig. 1 Survival probability for injured patients (Mills et al. 2013)

the survival probability function $f(t)$ by adding diversion time Δ_{div} . That is, $\tilde{f}_I(t) = f_I(t + \Delta_{div})$ and $\tilde{f}_D(t) = f_D(t + \Delta_{div})$. Three reward functions for the immediate- and delayed-class are shown in Fig. 2.

Each reward function has unique characteristics. For TI, immediate-patients always give a greater reward than a delayed-patient, and the difference remains constant over time. In TDDP, immediate-patient also gives greater reward for all time, but the difference between the two varies in time. On the other hand, under TDTS, it shows a more complicated pattern. This is due to the shape of the survival probability functions. The rate of decrease is higher for $f_I(t)$ upto certain time t^* , and the rate for $f_D(t)$ becomes higher after that point. It will be shown later how these features influence the shape of optimal policies.

4.2 Numerical solution

Optimal policies for the test scenarios are numerically computed. Figure 3 shows the optimal policy for the reference scenario under the three types of reward functions. Shown on x -axis is the number of remaining beds k , and on y -axis is the elapsed time t . Thus each graph in Fig. 3 is an optimal policy plotted on the state space. For ease of viewing, we present the optimal policies for delayed- and immediate-class patients on separate plots.

There are few features in Fig. 3 that can be immediately related to the properties described in Sect. 3.3. First we see in Fig. 3a that the optimal action for an immediate-class patients is to always admit the patient in all states when we use TI reward function. This observation is in line with the optimal policy structure. Specifically, the optimality condition described in Sect. 3.3 states that if admitting a

Fig. 2 Three types of reward functions. For TI, $r_I(t, admit) = 4$ and $r_D(t, admit) = 1$; for TDDP, $\kappa_I = 0.82$ and $\kappa_D = 0.96$; for TDTS, $\Delta_{div} = 50$ min

Fig. 3 Optimal policy under three types of reward functions for each patient class

current patient returns a reward no less than any other patients in the future, then it is optimal to admit the patient. From Fig. 2a, we see that the reward for an immediate patient arriving at time t is no smaller than the reward from any future patients: $r_I(t, \text{admit}) \geq r_I(t + \Delta, \text{admit})$, $r_I(t, \text{admit}) \geq r_D(t + \Delta, \text{admit})$ for $\Delta > 0$. The same explanation holds for TDDP reward function, and we also see the optimal policy under TDDP is to admit all immediate patients as shown in Fig. 3b.

Next we note in Fig. 3 that, for fixed time t , an optimal action switches from “reject” to “admit” at most once along x -axis in all cases. That is, there is a single threshold value k^* across which the optimal action switches. For example, in Fig. 3d, when a delayed-class patient arrives at time 200 min, then the optimal action is determined simply by whether the number of remaining bed is above or below 7; the switching point $k^* = 7$. As mentioned in Sect. 3.3, this property is due to the discrete concavity of the optimal value function in our MDP model.

Another observation about k^* is that it is decreasing along y (time)-axis in Fig. 3d–f. The decreasing pattern of k^* in Fig. 3d is expected from the structure of optimal policy; the reward function is $r_D(t, \text{admit}) = 1$ which is non-decreasing in t , and the marginal value of a bed in our model decreases in t . Recall from Sect. 3.3 that when these two conditions are satisfied, k^* decreases in t . For the decreasing pattern in Fig. 3e, on the other hand, the above explanation does not hold true because the TDDP reward function is a decreasing function of t . It just happens to exhibit a decreasing pattern due to the specific shape of the TDDP function and NHPP functions used in our experiment. Although not shown here, we observe that varying the location of peak of the NHPP arrival function yields an optimal policy where k^* is not monotonically decreasing in t .

Although looking apparently similar to Fig. 3d and e, the underlying mechanism behind the pattern shown in Fig. 3f is different from them. From Fig. 2c, we see that the TDTS reward function for delayed-patients increases up to 88 min, and then decreases afterwards. The increase in r_D during the first 88 min means that, per the property of optimal policies, k^* decreases in time. After 88 min, the reward from a current arriving patient (delayed-class) is greater than any future arrival. Then according to the optimality condition, it is optimal to admit delayed patients when $t > 88$ min.

Figure 3c shows a more general, irregular shape for the optimal policy. In Fig. 3c, we see that k^* decreases in t up to 16 min, increases to 48 min, and decreases afterwards. This pattern is related to the structure of TDTS reward functions and arrival functions. In Fig. 2c, r_I increases up to 16 min, and we know from the property of optimal policy that k^* decreases in time during that period. After 16 min, r_I starts to decrease, and there is a possibility that a patient with greater reward may arrive in the future. The specific shape of an optimal policy is determined by the combined effects from r_I, r_D as well as the arrival functions for the two classes of patients. A pattern shown in Fig. 3c is one instance of optimal policy given a specific choice of the reward and arrival functions.

While the above discussion relates the patterns exhibited in Fig. 3 to the properties of an optimal policy, they can be intuitively understood by a simple rationale: *it is better to save a bed when we expect a large number of patients with high reward to arrive in the future; otherwise, use the bed for a current patient*. This rationale translates to as follows. If there are many beds available, there is little

reason to reject a patient. Likewise, if a long time has passed since the outbreak of an MCI, few additional patients will arrive in the future, and thus there is little reason to reject a patient. The optimal policies shown in Fig. 3d–f show exactly such structure.

Finally we note that the optimal policy shown in Fig. 3c rejects an immediate-class patient for some portion of the state space. It turns out that for emergency care providers including emergency physicians, it is difficult to accept a policy that requires to reject an immediate-class patient and save a bed for future use. In the next section, we introduce a constraint in the model to address this problem and examine how it affects the optimal policy.

4.3 Effects of policy restriction

When optimal decisions conflict with a conventional belief and behavior, it is hard to expect they are actually embraced and implemented. Here we refer to such conflict as a behavioral constraint. These are constraints for a decision problem as they limit the policy solution space. Ideally, a decision-making model should take behavioral constraints into account in deriving an optimal policy.

For our problem, the conflict concerns a reject decision for an immediate-class patient. Under an optimal policy, an immediate-class patient can be rejected even when there is a remaining bed, as shown in Fig. 3c. Since this is not acceptable by EMS practitioners, we incorporate a rule to always admit an immediate-class patient; that is, never reject an immediate-class patient as long as there is an available bed. With this rule, our decision problem is made relevant only to delayed-class patients.

Specifically, we introduce a constraint on the action set definition. Recall in the original formulation in Sect. 3.2, action set A depends on k , and $A(k)$ is {admit patient, reject patient} if $k > 0$, and {reject patient} if $k = 0$. With the “always-admit-immediate-patient” rule, we redefine the action sets as follows:

$$A(k, m) = \begin{cases} \{\text{admit}, \text{reject}\} & \text{if } k > 0, \quad m = \text{delayed} \\ \{\text{admit}\} & \text{if } k > 0, \quad m = \text{immediate} \\ \{\text{reject}\} & \text{if } k = 0. \end{cases}$$

Modifying the action set invalidates many of the structural properties of an optimal policy. For example, optimal value function’s monotonicity with respect to time is not satisfied any more since we cannot selectively admit patients in the future. Recall that optimal policy may reject a patient and save a bed for future use when it expects a patient with higher reward in the future. But with this rule, an immediate-class patient is always admitted. A bed is used up for an immediate-class patient even if its reward is smaller than the potential future reward from admitting other patient. Thus Eq. (4) does not hold true any longer. Also, this modified MDP model is not guaranteed to satisfy the discrete concavity property or the decrease in marginal value of a bed. In the absence of mathematically-derived structural properties, we resort to numerical experiment and examine how such restriction affects the optimal policy solution.

Fig. 4 Comparison between the basis scenario policy with and without a behavioral constraint

In Fig. 4, boundaries (k^*) of optimal policies with and without the policy restriction are plotted for comparison. Since immediate-class patients are always admitted, we only show the policy solution for delayed-class patients. Recall that without the policy restriction the optimal policy under the TI and TDDP reward function is to always admit immediate-class patients (Fig. 3a, b). On the other hand, immediate-class patients are rejected under the TDTS reward function (Fig. 3c). As such, we expect that the policy restriction only affects the optimal policy for the case of TDTS reward function, and the numerical results shown in Fig. 4 agree with our expectation.

Without a mathematical proof, we have the following conjecture to explain the pattern shown in Fig. 4c. The modified MDP model has the policy restriction that constrains solution space, and thus a policy solution from the modified model is sub-optimal compared with the policy solution without the restriction. There is some portion of the state space where we would reject immediate-class patients if it were not for the policy restriction. The policy restriction forces to admit immediate-class patients in those states, and it incurs some loss in the value function. With the policy restriction, the only way to make up for this loss is to change the policy for delayed-class patients; if we are not allowed to reject immediate-class patients in those states, we may instead opt to admit delayed-class patients more aggressively and possibly induce diversion for the immediate-class patients arriving later. This conjecture leads us to expect that a new optimal policy admits some of the delayed-class patients that it would reject in the original model without the policy restriction. Indeed, Fig. 4c shows the change in the optimal policy, which is consistent to some degree with such conjecture.

It is interesting to observe that the single switching property for an optimal policy is lost due to the absence of discrete concavity in the value function. Figure 5 shows that for some fixed t there exist two switching points for optimal action. For

Fig. 5 Existence of multiple k^* under the policy restriction

example, for $t = 55$ min, we admit a delayed-class patient if there are less than 3 beds, reject the patient if we have more than 3 and less than 11 beds, and then admit if more than 11 beds are available. This does not sound intuitive, but we can certainly concoct a feasible scenario where such policy is indeed optimal strategy. Suppose that we are making an admission decision for a delayed-class patients of reward r and we expect the order of next arrivals is {immediate-class patient with smaller reward, delayed-class with larger reward, delayed-class with larger reward, and so on}. Then the optimal action would be to admit the current patient if we have only one bed. If we have two beds, on the other hand, a better action would be to reject the patient so that we can admit the delayed-class patient with larger reward.

Lastly, if we compare the outcomes from the two MDP models—with and without policy restriction-, it is easy to see that the expected number of survivors from the restricted policy model is smaller than that from the unrestricted policy, and example results from a few test instances are shown in Table 2. It shows that the unrestricted policy produces higher values for the expected number of survivors by, on average, four to nine percent in the test instances. While the unrestricted policy certainly outperforms the restricted policy, anecdotal comments from EMS

Table 2 Difference in the expected number of survivors (“gap”) between the two models for three test instances

Initial capacity	30	20	10
Minimum gap (%)	0.34	1.11	1.27
Maximum gap (%)	8.28	10.10	15.40
Average gap (%)	4.01	6.15	9.15

caregivers seem to suggest the difference is not practically significant enough to warrant major changes in their de facto practice.

4.4 Performance test

We assess the quality of policy solution obtained from the MDP model. We vary four parameters to construct test instances. Two parameters, $(C_I + C_D)$ and $(C_I : C_D)$, concern patient volume and severity, and the other parameters, α_I and α_D , characterize their arrival pattern. Refer to Table 1 in Sect. 4.1 for specific values used in the experiment. We generate 26 arrival scenarios by changing values for one of the four parameter while fixing the other three parameters at their reference value. For these 26 arrival scenarios, we use the three reward function type—TI, TDDP, and TDTS. Thus, we have a total of $26 \times 3 = 78$ test instances, and MDP model for each instance is solved to determine an optimal policy for each case. We also set three levels for the number of initially available beds at 10, 20, and 30.⁴ Then, we run a simulation where patients arrive according to NHPP with the arrival parameters of each instance, and we make admission decision as stated in the optimal policy from the MDP model.

As a performance indicator, we measure the total sum, S , of rewards earned in a simulation for the stream of patients. Then we compare S against the maximum reward sum, $\hat{S}$, for the patients. $\hat{S}$ is the reward sum when admission decision is made with all patient types and arrival times known a priori. Thus, $\hat{S}$ is the maximum possible value by admission decision for the respective patient arrivals. We refer to $S/\hat{S} = \zeta$ (%) as a relative efficiency of a decision model. Each instance is replicated 1000 times with random set of patients sampled from the arrival parameters of the test instance.

It turns out that decision-making according to MDP policy solution is very close to the maximum possible value. Even for the worst case among the test instances, the average relative efficiency $\bar{\zeta}$ is 95.41% (95% confidence interval = 0.34) in 10-bed cases, 93.74% (0.32) for 20-bed, and 94.08% (0.30) for 30-bed. Overall for all test instances, $\bar{\zeta}$ ranges from 95.41 to 100, 93.74 to 100, 94.08 to 100% in 10-, 20-, 30-bed cases respectively. It simply shows that optimal policies computed from the MDP model work very well in a wide range of MCI scenarios tested in our experiments.

We also compare the MDP policy with more simple heuristics bed management in an ED. Two heuristics are considered for comparison. “First-come-first-serve (FCFS)” policy would admit all patients as they arrive at an ED until all beds are occupied. This is the simplest response an ED can take in response to MCI. “Reject-delayed (RejD)” policy, on the other hand, diverts all delayed-class patients and only admits patients in immediate-class. Our survey on hospital disaster preparedness plans has found the RejD policy in some large hospitals in Korea. Table 3 shows the results of the comparison. For example, in 30-bed cases, MDP policy is better in 72 of 78 instances than FCFS policy, and rest of the cases, they are

⁴ Note that the number of initially available beds does not affect the optimal policy from MDP models.

Table 3 Optimal policy from MDP is compared with two ED admission policies

	Initial # of beds		
	30	20	10
Versus FCFS			
MDP is better	72/78	75/78	76/78
FCFS is better	0/78	0/78	0/78
Versus RD			
MDP is better	76/78	70/78	46/78
RD is better	0/78	0/78	0/78

Relative efficiency $\bar{\zeta}$ is compared at significance level $\alpha = 0.05$

statistically indifferent. As seen in Table 3, the MDP policy yields better results (higher $\bar{\zeta}$) than the two heuristic policies in most instances. In fact, in no instance is either FCFS or RejD policy better than the MDP policy. The results demonstrate that the MDP policy performs well to maximize the total reward sum, showing a potential to enhance life-saving capability in ED's disaster response according to the reward function considered.

4.5 Discussion

Performance of an admission policy depends on how it uses the limited resources for the right patient at the right time. To gain deeper insights, we analyze how different admission policies use the bed resource for each class of patients in comparison to the ideal case. If decision-making from a given policy is indeed identical or very similar to the ideal decision-making, then there should be zero or minimal level of difference in how the resources are spent serving each class of patients.

To do so, we first need to define a new metric $\Delta\zeta_I$. Recall that $\zeta = S/\hat{S}$ is the relative efficiency for a decision policy, and it is the ratio of total reward by using the decision policy to the maximum possible value for total reward. Also note that S can be decomposed to S_I and S_D , each representing the reward from admitting I (or D) patient. Likewise $\hat{S}$ can be decomposed to $\hat{S}_I$ and $\hat{S}_D$. Then, we define $\Delta\zeta_I = (S_I - \hat{S}_I)/\hat{S}$. $\Delta\zeta_I$ measures the difference between the reward from admitting the immediate-class patients under a proposed policy and the reward from admitting the immediate-class patients under the ideal case. $\hat{S}$ is a normalization factor. Roughly speaking, if a proposed policy admits the similar number of immediate-class patients at similar time, then $S_I \sim \hat{S}_I$, thereby $\Delta\zeta_I$ will be close to zero. If the policy admits many more (much fewer) immediate-class patients than the ideal case does, $S_I > \hat{S}_I$ ($S_I < \hat{S}_I$) and $\Delta\zeta_I$ will be positive (negative). Thus, by observing the sign of $\Delta\zeta_I$, we can tell how the policy uses the beds for immediate-class patients with respect to the ideal case. $\Delta\zeta_D$ is defined likewise, and the same interpretation applies. It should be noted that because $\Delta\zeta_I$ and $\Delta\zeta_D$ measures the amount of deviation of the policy from the ideal case, even the positive value implies inefficient admission decisions.

Fig. 6 Scatter plot of $\Delta \zeta_I$ and $\Delta \zeta_D$ value in 20-bed cases

Inefficiency in the bed usage is caused by either allocating beds to undesirable patients or leaving some beds unused. Figure 6 shows the results of $\Delta \zeta_I$ and $\Delta \zeta_D$ under the three policies—FCFS, RejD, and MDP policy.⁵ x-axis and y-axis represent $\Delta \zeta_I$ and $\Delta \zeta_D$, respectively, and each point on the plot denotes $\Delta \zeta_I$ and $\Delta \zeta_D$ from one test instance. A point on the second quadrant ($\Delta \zeta_I < 0, \Delta \zeta_D > 0$) suggests that this ED admitted too few immediate-class patients and too many delayed-class patients. In contrast, a point on the fourth quadrant ($\Delta \zeta_I > 0, \Delta \zeta_D < 0$) indicates that too many immediate-class patients were admitted while the ED admitted too few delayed-class patients. The third quadrant would suggest the ED admitted too few patients overall or it admitted patients at inappropriate times. The diagonal line indicates the Pareto frontier where $\Delta \zeta_I + \Delta \zeta_D = 0$. Note that no point will be plotted above the Pareto frontier since no policy can perform better than the decisions made under the ideal case. There are 26 points on each plot which represent the average values from one thousand patient arrival samples in each of the twenty six test instances.

⁵ Results for the TI reward cases are similar to the TDDP case, and thus are not shown to save space.

The results shown in Fig. 6 clearly show the possible problems with the two heuristic policies. For FCFS policy, because this policy does not divert any patient, it results in admitting too many delayed-class patients. Therefore, the FCFS policy loses opportunities to admit immediate-class patients to achieve higher rewards from it, and FCFS policy always yields negative $\Delta\zeta_I$ value. This problem has been pointed out in the literature (der Heide 2006; Hogan et al. 1999; Waeckerle 1991) as the potential shortcoming with the FCFS policy. Preoccupation of beds by the delayed-class patients results in low efficiency for bed usage under the FCFS policy. In contrast to the FCFS policy case, all the points from the ReJD policy cases are located on the fourth quadrant. It implies there are lost chances for using beds to effectively serve delayed-class patients. Since it rejects all delayed-class patients, it is subject to a high risk of wasting the bed resource. For test instances where there were only a few immediate-class patients, ReJD policy ended up leaving some of the beds unused. Shown in Fig. 6c, f are the results from the MDP policy cases. Under the selective admission policy computed from the MDP model, the use of beds seems to achieve the right balance as expected. Points under the MDP policy are mostly clustered near the origin, which suggests that the MDP policy tends to make decisions very close to the ideal case.

5 Conclusion

It has been well demonstrated in the literature that mathematical models and analyses can enhance disaster response capability to save more lives. Especially, they are well suited to address resource utilization problems for effective disaster responses where available EMS resources are scarce relative to the demand from patients. In this paper, we construct a MDP model that determines an optimal admission policy for patients arriving at an ED in the aftermath of an MCI.

Our model addresses a multi-class admission control problem over a finite horizon. It incorporates time-dependent arrivals of disaster victims and their time-dependent survival probabilities. Structural policies of this model are summarized and interpreted in the context of the MCI response. We recognize that optimal policy in the context of EMS operation must respect the current practice by the EMS professionals. In our problem, the concept of rejecting an immediate-class patient with a goal of maximizing the overall expected number of survivors seems to be against their standard decision-making. Solutions of the MDP model are numerically obtained with a state discretization technique that enables to handle the continuous state space of the model.

Experimental results show that the policy obtained from our model performs very close to an ideal decision-making. We also compare the MDP policy with two admission decision heuristics plausible in a real practice, and it shows that MDP policies very clearly outperform the heuristics. While specifics of an optimal policy depend on the arrival functions and reward functions, we found there is a simple, intuitive rationale behind the policy structure: when a large number of patients with high reward are expected to arrive in the future, saving a bed for a future arrival improves the expected number of survivors; otherwise, using the bed for the current

patient is better. Identifying a fundamental yet simple decision principle from a MDP model is crucial for the solution's applicability in a realistic situation.

Acknowledgements This research was supported by a Grant 'research and development of modeling and simulating the rescues, the transfer, and the treatment of disaster victims' (nema-md-2013-36) from the Man-made Disaster Prevention Research Center, National Emergency Management Agency of Korea.

References

- Allon G, Deo S, Lin W (2013) The impact of size and occupancy of hospital on the extent of ambulance diversion: theory and evidence. *Oper Res* 61(3):544–562
- Argon NT, Ziya S, Righter R (2008) Scheduling impatient jobs in a clearing system with insights on patient triage in mass casualty incidents. *Probab Eng Inf Sci* 22(3):301–332
- Arnold JL, Tsai M-C, Halpern P, Smithline H, Stok E, Ersoy G (2003) Mass-casualty, terrorist bombings: epidemiological outcomes, resource utilization, and time course of emergency needs (part I). *Prehospital Disaster Med* 18(3):220–234
- Brumelle S, Walczak D (2003) Dynamic airline revenue management with multiple semi-Markov demand. *Oper Res* 51(1):137–148
- Christ M, Grossmann F, Winter D, Bingisser R, Platz E (2010) Modern triage in the emergency. *Deutsch Ärztebl Int* 107(50):892–898
- Cohen I, Mandelbaum A, Zychlinski N (2014) Minimizing mortality in a mass casualty event: fluid networks in support of modeling and staffing. *IIE Trans* 46(7):728–741
- Deo S, Gurvich I (2011) Centralized vs. decentralized ambulance diversion: a network perspective. *Manag Sci* 57(7):1300–1319
- der Heide EA (2006) The importance of evidence-based disaster planning. *Ann Emerg Med* 47(1):34–49
- Einav S, Aharonson-Daniel L, Weissman C, Freund HR, Peleg K, Israel Trauma Group (2006) In-hospital resource utilization during multiple casualty incidents. *Ann Surg* 243(4):533–540
- Frykberg ER (2002) Medical management of disasters and mass casualties from terrorist bombings: how can we cope? *J Trauma Acute Care Surg* 53(2):201–212
- Hick JL, Hanfling D, Cantrill SV (2012) Allocating scarce resources in disasters: emergency department principles. *Ann Emerg Med* 59(3):177–187
- Hogan DE, Waeckerle JF, Dire DJ, Lillibridge SR (1999) Emergency department impact of the Oklahoma city terrorist bombing. *Ann Emerg Med* 34(2):160–167
- Institute of Medicine of the National Academies (2005) Hospital-based emergency care: at the breaking point. The National Academies Press, Washington
- Jacobson EU, Argon NT, Ziya S (2012) Priority assignment in emergency response. *Operat Res* 60(4):813–832
- Jenkins JL, McCarthy ML, Sauer LM, Green GB, Stuart S, Thomas TL, Hsu EB (2008) Mass-casualty triage: time for an evidence-based approach. *Prehospital Disaster Med* 23(1):3–8
- Kilic A, Dincer MC, Gokce MA (2013) Determining optimal treatment rate after a disaster. *J Oper Res Soc* 65(7):1053–1067
- Korea National Emergency Medical Center (2007) Development of the guidelines for inter-hospital transfer. Ewha Womans Univ., Korea National Emergency Medical Center, Seoul
- Li D, Glazebrook KD (2010) An approximate dynamic programming approach to the development of heuristics for the scheduling of impatient jobs in a clearing system. *Nav Res Logist* 57(3):225–236
- Mills AF, Argon NT, Ziya S (2013) Resource-based patient prioritization in mass-casualty incidents. *Manuf Serv Oper Manag* 15(3):361–377
- Niska RW, Burt CW (2005) Bioterrorism and mass preparedness in hospitals: United States, 2003. *Adv Data* 364:1–14
- Peleg K, Kellerman AL (2009) Enhancing hospital surge capacity for mass casualty events. *JAMA* 302(5):565–567
- Puterman ML (1994) Markov decision processes: discrete stochastic dynamic programming, vol 672. Wiley, New York
- Ramirez-Nafarrate A, Hafizoglu AB, Gel ES, Fowler JW (2014) Optimal control policies for ambulance diversion. *Eur J Oper Res* 236(1):298–312

- Reilly MJ, Markenson D (2010) Hospital referral patterns: how emergency medical care is accessed in a disaster. *Disaster Med Public Health Preparedness* 4(3):226–231
- Sacco WJ, Navin DM, Fiedler KE, Waddell RK II, Long WB, Buckman RF (2005) Precise formulation and evidence-based application of resource-constrained triage. *Acad Emerg Med* 12(8):759–770
- Sacco WJ, Navin DM, Waddell RKII, Fiedler KE, Long WB, Buckman RF (2007) A new resource-constrained triage method applied to victims of penetrating injury. *J Trauma Acute Care Surg* 63(2):316–325
- Stratton SJ, Tyler RD (2006) Characteristics of medical surge capacity demand for sudden-impact disasters. *Acad Emerg Med* 13(11):1193–1197
- Sung I, Lee T (2016) Optimal allocation of emergency medical resources in a mass casualty incident: patient prioritization by column generation. *Eur J Oper Res* 252(2):623–634
- Timbie JW, Ringel JS, Fox DS, Pillemer F, Waxman DA, Moore M, Hansen CY, Knebel AR, Ricciardi R, Kellermann AL (2013) Systematic review of strategies to manage and allocate scarce resources during mass casualty events. *Ann Emerg Med* 61(6):677–689
- Waeckerle JF (1991) Disaster planning and response. *N Engl J Med* 324(12):815–821
- Xu K, Chan CW (2016) Using future information to reduce waiting times in the emergency department via diversion. *Manuf Serv Oper Manag* 18(3):314–331

Hyun-Rok Lee is a Ph.D. candidate in the industrial and Systems Engineering Department at KAIST. His research interests include modeling, simulation and optimization in healthcare applications.

Taesik Lee is an associate professor in the Industrial and Systems Engineering Department at KAIST. His primary research area is modeling, simulation, and optimization for applications in healthcare delivery systems and disaster response management.