OPERATIONS RESEARCH
Vol. 60, No. 4, July–August 2012, pp. 813–832
ISSN 0030-364X (print) | ISSN 1526-5463 (online)
http://dx.doi.org/10.1287/opre.1120.1075

Priority Assignment in Emergency Response

Evin Uzun Jacobson
Imperial College Business School, London SW7 2AZ, United Kingdom, e.uzun@imperial.ac.uk

Nilay Tanık Argon, Serhan Ziya
Department of Statistics and Operations Research, University of North Carolina, Chapel Hill, North Carolina 27599
{nilay@unc.edu, ziya@unc.edu}

In the aftermath of mass-casualty events, key resources (such as ambulances and operating rooms) can be overwhelmed by the sudden jump in patient demand. To ration these resources, patients are assigned different priority levels, a process that is called triage. According to triage protocols in place, each patient's priority level is determined based on that patient's injuries only. However, recent work from the emergency medicine literature suggests that when determining priorities, resource limitations and the scale of the event should also be taken into account in order to do *the greatest good for the greatest number*. This article investigates how this can be done and what the potential benefits would be. We formulate the problem as a priority assignment problem in a clearing system with multiple classes of impatient jobs. Jobs are classified based on their lifetime (i.e., their tolerance for wait), service time, and reward distributions. Our objective is to maximize the expected total reward, e.g., the expected total number of survivors. Using sample-path methods and stochastic dynamic programming, we identify conditions under which the state information is not needed for prioritization decisions. In the absence of these conditions, we partially characterize the optimal policy, which is possibly state dependent, and we propose a number of heuristic policies. By means of a numerical study, we demonstrate that simple state-dependent policies that prioritize less urgent jobs when the total number of jobs is large perform well, especially when jobs are time-critical.

Subject classifications: triage; emergency response; stochastic scheduling; stochastic orders; dynamic programming.
Area of review: Manufacturing, Service, and Supply Chain Operations.
History: Received March 2010; revisions received July 2011, November 2011; accepted January 2012. Published online in *Articles in Advance* August 2, 2012.

1. Introduction

In the aftermath of mass-casualty events and disasters, critical resources such as ambulances, rescue vehicles, operating rooms, and physicians typically are overwhelmed by the sudden jump in demand for their services. In a matter of minutes to hours, these resources become insufficient in number to provide immediate relief to all that are in need, which makes their efficient allocation essential for the eventual success of the emergency response effort. However, making good allocation decisions is not an easy task because it requires simultaneous consideration of multiple factors. Furthermore, one needs to act fast because usually there is only a short period of time during which lives can be saved. Typically, the first step of a response effort is to determine (at least roughly) the urgency of different "jobs" to which the resources need to be assigned. (Here, a job could be a single patient, a group of patients, or a rescue mission involving a large number of individuals.) Once that is done, one reasonable policy could be to start from the most urgent jobs and move onto those that are less urgent as resources become available. Unfortunately, such a simple and natural policy might not be the best policy to save the most lives. There are two main complicating factors. First, the expected "payoff" from jobs at different urgency levels are generally different from each other. For example, in mass-casualty incidents with traumatic injuries, most patients with shorter life expectancies have lower chances of going through successful operations, i.e., they have lower expected payoffs. By choosing to serve a more urgent job first, we might be forgoing the one with a higher payoff, which might not be available at the time the service of the more urgent job is over. Hence, there is a trade-off between urgency and payoff. Second, the service time requirements might be different for jobs at different urgency levels as well. For example, in mass-casualty incidents with traumatic injuries, the treatment durations for patients with shorter life expectancies are generally longer due to the complexity of their health conditions. This brings another trade-off between prioritizing urgent jobs and prioritizing those with shorter service. The objective of this article is to investigate these trade-offs among urgency, payoff, and service times, and identify classes of priority assignment policies that are both efficient and simple.

Triage, the practice of rationing medical resources depending on the severity of the patients' conditions, dates back to Napoleonic Wars. Since then, triage has been widely adopted not only in wars but also in civilian life in case of mass-casualty events or even in daily emergencies. There are several proposed and adopted triage systems in the emergency medicine literature, simple triage and rapid treatment (START) being one of the most common; see, e.g., Nocera and Garner (1999). However, to our knowledge, there has not appeared any comprehensive study on whether or not using these systems improves the outcome of emergency response efforts. In fact, more recently, adopted practices have been criticized for being too short-sighted. Several researchers from the emergency medicine community have argued that when making prioritization decisions, unlike the current practice, scarcity of the resources should be taken into account and called for more research on how that should be done (see, e.g., Frykberg 2002, Sacco 2005). One of the objectives of this article is to provide insights into this discussion as well.

In general, in the aftermath of mass-casualty incidents or disasters, there are potentially many other types of jobs that require prioritization decisions (other than patients with traumatic injuries) and for many of those decisions, there are no established rules and principles for the emergency responders to follow. For example, there are no triage rules adopted for determining priorities when evacuating already hospitalized patients before disasters strike or in their aftermath (see, e.g., Okie 2008). Similarly, there are no adopted policies for prioritizing rescue missions (e.g., for trapped victims in floods) before or after disasters. As a result, when disasters strike those who find themselves having to make such decisions do so in an ad hoc manner. Perhaps this is no more apparent than it was in the aftermath of Hurricane Katrina that hit the Gulf of Mexico coast of the United States in August 2005 (see, e.g., Darr 2006, Bernard and Mathews 2008, Fink 2009). Many physicians who did not have any training in triage and prioritization decisions had to determine in what order patients should be evacuated from hospitals during the days that followed Katrina. Part of the reason for the lack of clear prioritization guidelines is that there are unresolved ethical concerns surrounding the decision of providing help to a certain group of patients while sacrificing others for the greatest good for the greatest number; see, e.g., Larkin and Arnold (2003) and Holt (2008). What is clear, however, is that lack of guidelines and training on the subject only leads to confusion when disasters strike and can put the whole responsibility on the shoulders of a few individuals who take the initiative to make such difficult decisions. Fortunately, it appears that within the medical community discussions on these issues have intensified in recent years (see, e.g., Okie 2008, Fink 2009).

With this article, we aim to contribute to this discussion by *providing insights on how resource limitations can be taken into account when determining patient priorities in mass-casualty events and the potential benefits of resource-based decisions*. The article does not attempt to develop a decision support tool that can readily be used in real time. Our goal, rather, is to develop a relatively simple model that captures the most essential components of the decision problem, identify basic principles and rules of thumb that work well, and provide some guidance to the emergency response community in their efforts to devise practical and efficient policies.

Our formulation is similar to that of a traditional job-scheduling problem although with some important differences. Very broadly, the problem can be described as follows. There are different types of jobs each having a stochastic due date, which is unknown to the decision-maker, and an associated expected reward that will be earned if the job is taken into service before its due date. Each job has a stochastic processing time, and its distribution might depend on the type of that job. The objective is to maximize the total expected reward by dynamically determining the order according to which jobs will be processed. In this model, jobs can refer to any group of tasks that require the same set of scarce resources during an emergency response effort. For example, in the case of a bombing, jobs can be injured patients who are waiting to be transported to a hospital; or in the case of a natural disaster, they can be already hospitalized patients who are waiting to be transferred to safer locations from areas affected by the disaster. In these two examples, the scarce resource would be ground or air transportation vehicles. Similarly, jobs can be patients with traumatic injuries that are brought to a hospital following an emergency event and the scarce resource can be the operating rooms of the hospital. In each one of these cases, there is a random due date for each job because patients can die before they are safely transported and/or provided with the required medical care. The reward associated with each job can have various interpretations. If the objective of the emergency response effort is to save as many patients as possible, then the reward for a patient can be seen as the probability that the patient will survive when the required resource is provided. If the objective is to maximize the total QALY (quality adjusted life year) score, the reward can be seen as the expected QALY that would be gained by allocating the resource to the patient. In the case of prioritizing rescue missions, if the objective is to maximize the number of survivors, then the reward can be the number of disaster victims who would survive as a result of the associated rescue mission.

Throughout the article we adopt a general terminology, which allows us to emphasize the relevance of our findings to the classical job-scheduling literature. For example, we use "jobs" that are impatient instead of patients with finite lifetimes and "servers" that provide service to these jobs instead of ambulances or operating rooms. However, whenever appropriate, we interpret our results and provide insights within the context of prioritization decisions during emergency response to a disaster or a mass-casualty incident.

2. Relevant Literature

Even though patient triage has long been practiced, interestingly, there has not appeared any comprehensive study on how useful existing triage systems are or in fact whether or not triage is useful at all (Jenkins et al. 2008, Lerner et al. 2008). More recently, a number of authors (e.g., Frykberg 2002) discussed the limitations of existing practices and argued in support of making triage and priority decisions while taking into account resource limitations. However, to the best of our knowledge, there is only one line of work from the emergency medicine literature (Sacco et al. 2005, 2007) that proposes a prioritization method (called the *Sacco Triage Method* (STM)) that takes into account system conditions. More specifically, Sacco and his coauthors propose a linear-programming-based method for determining priorities when dispatching patients to hospitals. The idea is to solve a linear program at the beginning of the response effort and perhaps repeatedly thereafter as the conditions change. In addition to the fact that STM largely ignores the randomness inherent in the actual system, the method has been criticized as being impractical because it suggests using a real-time solution. Such a solution might differ drastically from one event to the other, and it highly relies on perfect system information and communication within the disaster area; see Cone and MacMillan (2005). In contrast to this line of work, which proposes a real-time solution method like STM, our objective is to identify basic rules and principles that the emergency response community can use in the development of simple and effective prioritization policies.

Within the operations literature, interest in patient triage has been relatively recent. There are four articles from this literature that are closely related to our work, namely, Glazebrook et al. (2004), Argon et al. (2008), Childers et al. (2009), and Li and Glazebrook (2010). As does this article, these four articles seek a solution to the problem of allocating service capacity to impatient jobs in a setting where all jobs are present at time zero and no additional jobs are expected to arrive. However, our work differs from these four articles in a number of ways, as we discuss next.

Glazebrook et al. (2004) consider a general job-scheduling formulation with impatient jobs having exponential lifetimes under the objective of maximizing the expected total reward. Their work is not particularly motivated by priority decisions during emergency response. The authors propose a simple state-independent policy and prove that this policy is asymptotically optimal as the mean lifetimes go to infinity. However, as Argon et al. (2008) and Li and Glazebrook (2010) demonstrate later, this simple policy does not perform well when lifetimes are sufficiently short. We discuss this policy in more detail in §5.

Argon et al. (2008) consider a formulation where jobs that belong to one of two different types (with possibly distinct lifetime and service time distributions) receive service from a single server. The objective is to determine the optimal policy that maximizes the expected number of survivors. Along with a number of analytical results that characterize the optimal policy, the authors propose two state-dependent heuristic policies that give priority to jobs with smaller mean service times but longer mean lifetimes when the system is heavily congested. In this article, we consider a formulation that generalizes the model of Argon et al. (2008) in a number of ways, making it a much more realistic representation of the actual system. First, in our work, the rewards might depend on the type of patient; i.e., unlike in the model of Argon et al. (2008), not all patients who receive service bring the same reward. This generalization significantly enriches the model. For example, it allows us to incorporate survival probabilities that differ across patient types. Several articles from the emergency response literature provide supporting evidence that the probability of survival depends highly on the type of injury and should be taken into account while giving prioritization decisions (see, e.g., Sacco et al. 2005, 2007). Furthermore, the heuristics developed in this article are general enough to allow multiple servers, which is an important generalization because in many emergency settings there is usually more than one resource available (e.g., when the resources are ambulances). Finally, in this article, we develop a simple heuristic method that requires information only on the total number of patients, unlike the heuristics proposed by Argon et al. (2008), which need information on the number of patients from each type.

Li and Glazebrook (2010) consider a formulation that is very similar to that of Argon et al. (2008) with the objective of developing a heuristic method that could be executed in real time to produce a near-optimal solution. The authors use the idea of applying a single step of the policy improvement algorithm for Markov decision processes on the state-independent policy proposed by Glazebrook et al. (2004) to develop their heuristic method. A numerical study shows that this method produces a solution that is close to the optimal performance. In addition to some minor differences in modeling assumptions, Li and Glazebrook's work also differs from our work in the main approach that they take. While their work aims at developing a near-optimal real-time solution, our aim is to analyze a stylized model to get insights about "good" policies that are simple and readily available for execution at the time of the event.

Based on a numerical study, Childers et al. (2009) analyze a similar job-scheduling problem with the motivation of ordering patients for transport in case of a health-care facility evacuation. In their model, patients are classified into two types (critical and noncritical), and there is a final due date common to all patients. They study the problem under two objectives: maximizing the number of lives saved and minimizing the holding cost of patients. Consistent with the results by Argon et al. (2008), Childers et al. (2009) conclude that when resources are severely limited, the evacuation should start with noncritical patients first and switch to critical patients as the number of patients in need decreases.

Finally, we review two relevant articles from the traditional job-scheduling literature, namely, Boxma and Forst (1986) and Emmons and Pinedo (1990), in which the performance measure is the weighted number of tardy jobs.

(Here, "weights" can be seen as "rewards" in our formulation.) The work by Boxma and Frost (1986) differs from ours in that it considers only static policies under the assumption that the due dates are independent and identically distributed (i.i.d.). Emmons and Pinedo (1990), on the other hand, consider dynamic scheduling policies, as we do in this article. One of their results states that if the processing times are i.i.d. and the due dates are either i.i.d. or have the same value, then the optimal nonpreemptive dynamic policy is to process the job with the largest weight. In §3, we prove a similar result but without the assumptions on i.i.d. due dates and deterministic weights. Furthermore, in our model, the due dates do not include the service times as in the model of Emmons and Pinedo (1990). For a more detailed review of the related literature on stochastic job scheduling and operations research methods applied to patient triage, see Argon et al. (2011).

3. General Service Time and Lifetime Distributions

In our base model, we assume that at time zero there are N jobs in need of receiving service from a single server, where $N > 1$. Jobs are impatient in the sense that if a job's waiting time in the queue exceeds its "lifetime," then it reneges, i.e., it leaves the system without receiving any service. Jobs do not renege while in service. We assume that there will not be any future job arrivals so that the problem is over as soon as all jobs in the system are cleared either after they receive service or after their lifetimes expire. A job that is taken into service brings a random reward. In the context of a mass-casualty event, this reward can be seen as the patient's QALY or the probability that the patient will survive after the service is given. In the case of prioritization of rescue missions, where a limited resource needs to be allocated among several rescue missions, the reward can be seen as the number of potential survivors associated with each mission. Finally, in our model, the service is performed in a nonpreemptive manner, i.e., once the server starts processing a job, it cannot start working on another job before completing the processing of the job that is already in service.

Each job in this system is characterized by its lifetime, service time, and reward distributions. Let Y_i be the lifetime of job i at time zero, Z_i be the nonnegative reward earned when job i is taken into service, and S_i be the service time for job i, where $i = 1, \dots, N$. We assume that $\{Y_i\}_{i=1}^N$, $\{Z_i\}_{i=1}^N$, and $\{S_i\}_{i=1}^N$ are sequences of independent random variables and that these three sequences are independent from each other. We let Π be the set of all dynamic and nonpreemptive scheduling (prioritization) policies. Here, a dynamic prioritization policy is a collection of rules that determine which job is taken into service at any given decision epoch based on the state of the system, i.e., the time of the decision epoch and the collection of jobs in the system. We also define $C_\pi(t)$ to be the total reward earned by time $t \geqslant 0$ when policy $\pi \in \Pi$ is applied. In this section, we identify characteristics of policies that maximize $C_\pi(t)$ stochastically for all $t \geqslant 0$, and hence, the total expected reward after all jobs are cleared.

Before we proceed with the analysis, we first note a simple observation about this optimization problem: An idling policy, i.e., a policy under which the server may idle in the presence of jobs, is suboptimal in the sense of maximizing $C_\pi(t)$ along any given sample path. (This result can be easily proved by using a coupling argument.) Therefore, in the rest of the article, we consider only nonidling policies and we redefine Π as the set of all dynamic, nonpreemptive, and nonidling prioritization policies. This means that the decision epochs for our dynamic control problem are time zero and service completion instants.

We first study settings where jobs with shorter lifetimes (and thus are more urgent) have higher rewards and shorter service times. We need to define three stochastic orders to state our main result for this case. Suppose that X and Y are two random variables that are either discrete or continuous. If $\Pr\{X > u\} \leqslant \Pr\{Y > u\}$, for all $u \in (-\infty, \infty)$, then X is said to be smaller than Y in the sense of *usual stochastic orders* (denoted by $X \leqslant_{\text{st}} Y$). On the other hand, if $\Pr\{X - v > u \mid X > v\} \leqslant \Pr\{Y - v > u \mid Y > v\}$, for all $u, v \geqslant 0$, then X is said to be smaller than Y in the sense of *hazard rate orders* (denoted by $X \leqslant_{\text{hr}} Y$). Finally, let $f(t)$ and $g(t)$ be the densities or probability mass functions of X and Y, respectively. If $f(t)/g(t)$ is decreasing in t over the union of the supports of X and Y, then X is said to be smaller than Y in the sense of *likelihood ratio orders* (denoted by $X \leqslant_{\text{lr}} Y$). Note that $X \leqslant_{\text{lr}} Y \Rightarrow X \leqslant_{\text{hr}} Y \Rightarrow X \leqslant_{\text{st}} Y$. For more on these stochastic orders, see Shaked and Shanthikumar (2007).

PROPOSITION 1. *Consider a decision epoch $t_0 \geqslant 0$ at which jobs i and j are available for service. If $Y_i \leqslant_{\text{hr}} Y_j$, $S_i \leqslant_{\text{lr}} S_j$, and $Z_i \geqslant_{\text{lr}} Z_j$, then a policy that prioritizes job j over job i at time t_0 cannot be optimal in the sense of maximizing $C_\pi(t)$ stochastically at every $t \geqslant 0$.*

Proposition 1, which is proved in the appendix, states that if the lifetimes, service times, and rewards of any two jobs can be ordered according to the specified stochastic orders, then giving priority to the job with a shorter lifetime, shorter service time, and larger reward increases the total reward stochastically, and as a result, also in expectation. Thus, when determining which job to serve next, a job can be eliminated from consideration if it is "dominated" by another job whose lifetime is shorter in the sense of hazard rate ordering, and whose reward is larger and service time is shorter in the sense of likelihood ratio ordering. If these three orderings hold for any pair of jobs, then the optimal policy can be completely characterized. Hence, Proposition 1 directly leads to the following corollary, which generalizes Theorem 1 in Argon et al. (2008).

COROLLARY 1. *If $Y_1 \leqslant_{\text{hr}} Y_2 \leqslant_{\text{hr}} \cdots \leqslant_{\text{hr}} Y_N$, $S_1 \leqslant_{\text{lr}} S_2 \leqslant_{\text{lr}} \cdots \leqslant_{\text{lr}} S_N$, and $Z_1 \geqslant_{\text{lr}} Z_2 \geqslant_{\text{lr}} \cdots \geqslant_{\text{lr}} Z_N$, then a nonidling* *policy that prioritizes the job with the smallest index at every decision epoch maximizes $C_\pi(t)$ in the sense of usual stochastic orders at every $t \geqslant 0$.*

Both Proposition 1 and Corollary 1 make intuitive sense because it is reasonable to believe that high-reward jobs with short patience times and service times should get higher priority. These results are important in that they provide specific ordering conditions under which this intuition holds. In the context of emergency response, the results imply that if a group of patients have a higher chance of survival but a shorter life expectancy and treatment duration in terms of the stochastic orders given in Proposition 1 and Corollary 1, then that group should receive priority no matter how many resources (e.g., ambulances) are available and how many patients there are at any point in time during the response effort. However, at least in the case of mass-casualty incidents, the situations where these conditions hold are not common because for patients with longer life expectancies, survival probabilities are typically higher (see, e.g., Sacco et al. 2005) and treatment durations are shorter. Therefore, in the following sections, we focus on cases where jobs with higher rewards and shorter service times are not necessarily more urgent.

4. The Markovian Case

When the conditions of Proposition 1 do not hold, it appears to be difficult, if not impossible, to even partially characterize the optimal policy under general service time and lifetime distributions. Hence, in this section, we assume that service times and lifetimes are exponentially distributed, which helps us obtain partial characterizations of optimal policies and gain insights into policies that perform well under conditions that do not satisfy the *agreeability conditions* of Proposition 1. These characterizations also lead to simple heuristic policies that can be used in nonexponential settings as we discuss in §§5 and 6.

We do not claim that in reality (at least in scheduling problems that arise during emergency response efforts) service times or lifetimes are exponentially distributed. Although, to the best of our knowledge, no prior work has studied what particular distributions would be good fits, there is also no reason to expect that the exponential distribution would be a good choice neither for lifetimes nor for service times. However, the assumption of exponentially distributed lifetimes and service times (which we refer to as the Markovian assumption) allows some mathematical analysis and helps us develop insights into what kind of policies are likely to work well in practice. In fact, as we demonstrate in §6, the heuristic methods that are developed based on our analysis of the Markovian case perform well even under settings when the exponential assumptions do not hold. Thus, the main insights that come out of our analysis could be valid under conditions that are more general than the Markovian setup we assume here.

In the following, we assume that jobs are classified into two types. These job types can be seen as triage classes for patients with different injury characteristics. This simplification helps us push the analytical results further, get a better understanding of the optimal policies, and develop heuristic methods of assigning priorities. More importantly, priority decisions during emergency response mainly concern two groups of patients. For example, according to START, the casualties are categorized into four groups but the most important decision concerns the priority ordering between critically injured patients who need to be taken care of as soon as possible (classified as *immediate*) and those who also have serious injuries but can wait a little longer (classified as *delayed*). Other patients, i.e., those with minor injuries (classified as minor) and those with injuries that are so severe that chances of survival are almost zero (classified as expectant), have the lowest priority. It is clear that as long as patients are correctly classified, there is no point in giving priority to either minor patients or expectant patients. However, the priority decision between the immediate and delayed patients is not always clear. Even though the general understanding (and the current practice) is that immediate patients should have a higher priority than delayed patients, some in the emergency response community (e.g., Frykberg 2002) have suggested that this decision should ideally depend on the number of casualties and the scarcity of the available resources.

In our Markovian model, the two types of jobs are characterized by their mean lifetimes, service times, and rewards. Let $r_i > 0$ be the abandonment rate (i.e., the reciprocal of the mean lifetime), $\mu_i > 0$ the service rate (i.e., the reciprocal of the mean service time), and $\alpha_i > 0$ the expected reward for type $i \in \{1, 2\}$. We let $D_\pi(m_1, m_2)$ be the expected total reward accumulated after all jobs are cleared when prioritization policy $\pi \in \Pi$ is applied and there are initially $m_i \geqslant 1$ jobs from type $i \in \{1, 2\}$ in the system, where $m_1 + m_2 = N$. Then, the optimization problem that we would like to study is given by

$$\max_{\pi\in\Pi} D_\pi(m_1, m_2). \tag{1}$$

We first state a corollary to Proposition 1 that provides conditions under which a solution to (1) is an *index policy*, i.e., a set of state-independent decision rules that assign priorities based only on job types at any given state. The proof is provided in the appendix.

Corollary 2. *If $r_1 \leqslant r_2$, $\mu_1 \leqslant \mu_2$, and $\alpha_1 \leqslant \alpha_2$, then it is optimal to give priority to type 2 jobs at every decision epoch.*

Similar to Proposition 1 and Corollary 1, Corollary 2 concludes that jobs with the shortest mean service time and lifetime as well as the highest expected reward should be given priority. Because this is not the most interesting case (at least in the patient triage setting), in the remainder of this section, we use a dynamic programming formulation to study the structure of the optimal policy under other conditions.

In our formulation, the state of the system is defined by the vector $(q_1, q_2; Q)$, where q_i is the number of type $i \in \{1, 2\}$ jobs waiting for service, $Q \in \{P_1, P_2\}$ is the status of the server, and $Q = P_i$ indicates that the server is busy processing a job of type $i \in \{1, 2\}$. The decision epochs are time zero and service completion times. At a decision epoch, the possible actions are allocating the server to a job from type $i \in \{1, 2\}$ such that $q_i > 0$. We next present the dynamic programming equations.

Let $V(q_1, q_2; Q)$ be the value function at state $(q_1, q_2; Q)$, i.e., the maximum expected reward earned starting from state $(q_1, q_2; Q)$. Also let $\mathbb{I}_A$ denote the indicator function of event A, i.e., $\mathbb{I}_A = 1$ if A is true, and $\mathbb{I}_A = 0$ otherwise. Then, for all $q_1 = 0, 1, \ldots, m_1$ and $q_2 = 0, 1, \ldots, m_2$, we have

$$V(q_1, q_2; P_i) = \frac{\mu_i \max\{\mathbb{I}_{\{q_1>0\}}\alpha_1 + V(q_1-1, q_2; P_1), \mathbb{I}_{\{q_2>0\}}\alpha_2 + V(q_1, q_2-1; P_2)\}}{\mu_i + q_1 r_1 + q_2 r_2} + \frac{q_1 r_1 V(q_1-1, q_2; P_i) + q_2 r_2 V(q_1, q_2-1; P_i)}{\mu_i + q_1 r_1 + q_2 r_2}, \quad \text{for } i = 1, 2, \qquad (2)$$

where we let $V(q_1, q_2; Q) = 0$ if $q_1 = q_2 = 0$ or $\min\{q_1, q_2\} < 0$ for all $Q \in \{P_1, P_2\}$. We also have

$$D_{\pi^*}(m_1, m_2) = \max\{\alpha_1 + V(m_1 - 1, m_2; P_1), \alpha_2 + V(m_1, m_2 - 1; P_2)\},$$

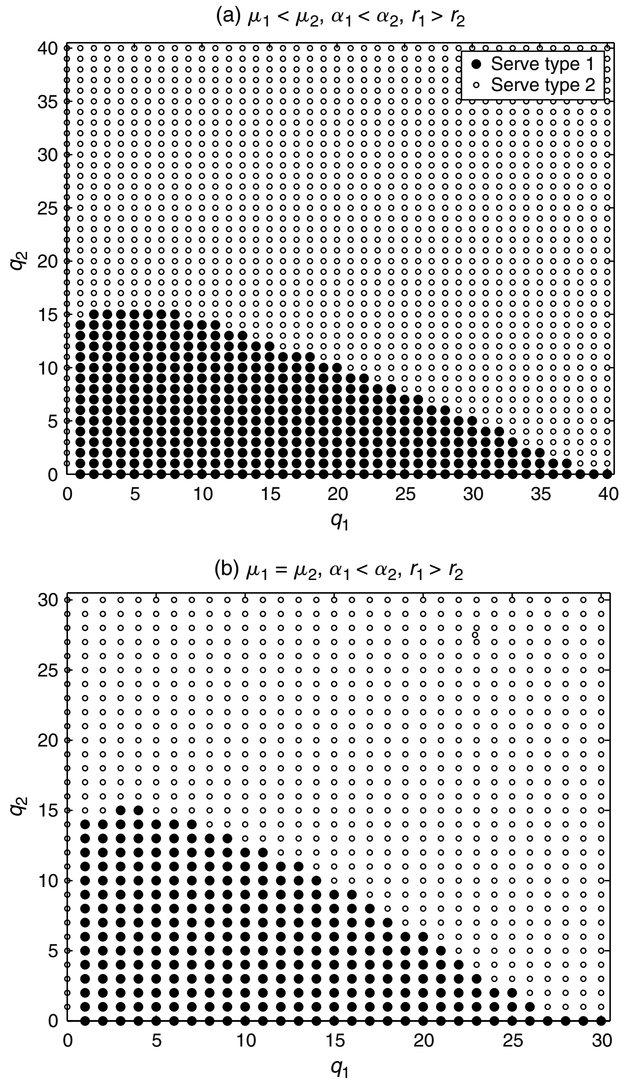
where π^* is a solution to (1).

We divided the remainder of this section into two parts. In §4.1, we provide examples that demonstrate the general structure of the optimal policy and show that under certain conditions the optimal policy can be partially characterized with thresholds. In §4.2, we provide a number of results on the optimality of index policies. Proofs of all propositions and corollaries presented in this section are given in the appendix.

4.1. Structure of the Optimal Policy

To give the reader some idea about how the optimal policies look like in general we provide two examples. In the first example, we consider a case where $\mu_1 < \mu_2$, $\alpha_1 < \alpha_2$, and $r_1 > r_2$, i.e., type 2 jobs have shorter mean service times, higher expected rewards, and longer mean lifetimes. As we discussed earlier, this is the most interesting case for the patient prioritization problem. Figure 1(a) presents the optimal actions for different values of q_1 and q_2 under this example. (In particular, Figure 1 illustrates the optimal action when the system is in state $(q_1, q_2; P_1)$ or $(q_1, q_2; P_2)$ and the next event is a service completion.) We selected this particular example because it demonstrates the most general structure for the optimal policy that we observed from several numerical examples.

Figure 1. The optimal policy for the case where (a) $\alpha_1 = 1.000$, $\alpha_2 = 1.010$, $\mu_1 = 0.901$, $\mu_2 = 0.909$, $r_1 = 1.000$, and $r_2 = 0.833$; and (b) $\alpha_1 = 1.000$, $\alpha_2 = 1.001$, $\mu_1 = \mu_2 = 0.901$, $r_1 = 0.909$, and $r_2 = 0.9009$.



As can be seen from Figure 1(a), the optimal policy gives priority to less time-critical jobs that bring a higher expected reward and that are faster to serve (i.e., type 2 jobs) when the number of jobs waiting is sufficiently large. Argon et al. (2008) observed a similar structure for the special case where the rewards are equal for both types. To demonstrate that this structure for the optimal policy is not entirely due to differences in service rates, we also studied examples where service rates are identical but expected rewards are distinct. Figure 1(b) illustrates the most general case that we observed from several examples under the condition that $\mu_1 = \mu_2$, $\alpha_1 < \alpha_2$, and $r_2 < r_1$. This example shows that the optimal policy again gives priority to less time-critical jobs that bring a higher expected reward when the number of jobs is sufficiently large. In summary, we observe that *if the type with longer mean lifetimes have at least the same expected reward and at most the same mean service time as the other type, then it should receive priority when there are many jobs in the system.*

To better understand the reason behind this structure for the optimal policy, it might help to think of the extreme hypothetical case where there is an infinite supply of types 1 and 2 jobs. In this case, one can see that it is always preferable to serve type 2 jobs with a higher reward because there is simply no advantage in serving a type 1 job instead if the mean service time of type 2 jobs is less than or equal to that for type 1 jobs. When there are fewer jobs, however, delaying service to type 2 jobs becomes a better strategy because one can "afford" to serve at least some of the type 1 jobs before switching to type 2 jobs, which have longer lifetimes. In the context of emergency response, this observation suggests that giving priority to less time-critical patients with a higher survival probability and/or shorter service time might be better when there are many patients in need of treatment. When the number of casualties is significantly high and it is clear that a large percentage of them is likely to die, then it makes more sense to use the limited resources to serve those with higher expected rewards, e.g., those who are more likely to go through a successful service. However, when there are few patients, then it makes more sense to give priority to those with shorter life expectancies (even though the chances of saving them are smaller and it takes more time to treat them) because there is enough time to get back to less time-critical patients later.

As Figure 1 clearly demonstrates, unless agreeability conditions such as those given in Corollary 2 hold, the optimal policy might be a state-dependent policy, i.e., a policy where the prioritization decisions depend on the number of jobs from each type that are waiting to receive service. Furthermore, in the absence of such convenient conditions, one cannot even guarantee the existence of an optimal policy that possesses features (e.g., monotonicity) that would make finding and describing optimal policies easier. In particular, for a fixed number of type 2 jobs, the optimal policy is not necessarily monotone in the number of type 1 jobs. For example, in Figure 1(a), if there are 15 type 2 jobs, as the number of type 1 jobs decreases, the optimal priority switches from type 2 jobs to type 1 jobs first, and then back to type 2 jobs again.

Figure 1 suggests that in general the optimal policy divides the state space into two regions separated by a single curve. A complete characterization of this curve, i.e., a complete description of the optimal policy, does not seem to be possible under all cases. Therefore, our objective here is to identify some structural properties of the optimal policy, with the ultimate goal of developing heuristic policies that closely approximate the optimal policy, i.e., the curve that separates the two regions in Figure 1.

In the following proposition, we show that under certain conditions, there exist thresholds that partially characterize the optimal policy.

Proposition 2. *Suppose that* $\mu_1 \leqslant \mu_2$, $\alpha_1 \leqslant \alpha_2$, *and* $r_1 > r_2$.

(i) *If* $r_1 \leqslant \mu_1$, *then the optimal policy is characterized by a threshold*

$$t(q_1) = \frac{\alpha_2 r_2(r_1-\mu_1) - \alpha_1 r_1(r_2-\mu_2)}{r_2[r_1(\alpha_2-\alpha_1)+\alpha_2(\mu_2-\mu_1)]} - q_1 \frac{r_1[r_2(\alpha_2-\alpha_1)+\alpha_1(\mu_2-\mu_1)]}{r_2[r_1(\alpha_2-\alpha_1)+\alpha_2(\mu_2-\mu_1)]}, \qquad (3)$$

for all $q_1 \geqslant 1$, *such that it is optimal to serve a type* 1 *job at every state where* $q_2 \leqslant t(q_1)$.

(ii) *If* $r_2 \geqslant \mu_2$, *then there exists a threshold* $\tilde{t}(q_1)$, *which is greater than or equal to* $t(q_1)$ *and possibly infinite, such that it is optimal to serve a type* 2 *job at every state where* $q_2 \geqslant \tilde{t}(q_1)$ *and* $q_2 \geqslant 1$.

Proposition 2 provides analytical support to our observation that when the number of patients in need of treatment is large, giving priority to patients who are less urgent and who have higher chances of survival and shorter service is preferable. To be specific, for the interesting case where type 2 jobs are less urgent and faster to serve and bring a higher expected reward, it implies that if type 1 jobs have a smaller abandonment rate than their service rate, then giving priority to type 1 jobs is optimal when the number of type 2 jobs is lower than a threshold $t(q_1)$. Similarly, Proposition 2 also implies that when the number of type 2 jobs is higher than another threshold $\tilde{t}(q_1)$, then it is optimal to give priority to type 2 jobs if type 2 jobs abandon the system at faster rates compared to their service.

Proposition 2 provides partial yet simple characterizations of the optimal policy, providing insights into patient prioritization decisions. The result clearly shows that optimal priority decisions can be dependent on the scale of the mass-casualty incident, i.e., the number of patients in need of treatment. Even though these characterizations do not describe the optimal policy completely, they could be useful in practice due to their simplicity. As we demonstrate in §6, a heuristic policy based on Proposition 2 performs very well.

4.2. Optimality of Index Policies

Index policies have clear practical advantages over state-dependent policies. They are easier to implement because under such policies, priority relations among types of jobs do not change with time and system state, and also there is no need to keep track of the number of jobs. We already provided sets of conditions in Corollaries 1 and 2, where a certain index policy is optimal. In this section, we obtain more conditions under which index policies are optimal.

We start with a proposition and its corollary, which provide conditions under which it is always optimal to give priority to the type with the fastest service and largest expected reward.

Proposition 3. *If $\mu_1 \leqslant \mu_2$, $\alpha_1 \leqslant \alpha_2$, $\mu_i \leqslant r_i$ for all $i = 1, 2$, and*

$$\frac{\alpha_1 r_1}{\mu_2 + r_1} \leqslant \frac{\alpha_2 r_2}{\mu_1 + r_2}, \tag{4}$$

then it is optimal to give priority to type 2 jobs at all decision epochs.

In Proposition 3, type 2 jobs have faster service, bring a higher reward, and also have a larger "immediate opportunity cost." To be specific, note that if a type 2 job is not taken into service at a decision epoch, then the probability that it will not be available at the next decision epoch is $r_2/(\mu_1 + r_2)$. Consequently, $\alpha_2 r_2/(\mu_1 + r_2)$ can be seen as the immediate opportunity cost of not providing service to that job. Hence, Proposition 3 implies that if jobs abandon the system with relatively fast rates (with respect to their service rates), then priority should be given to the type with faster service, higher expected reward, and larger immediate opportunity cost.

Corollary 3. *If $\mu_1 \leqslant \mu_2$, $\alpha_1 \leqslant \alpha_2$, $\alpha_1 r_1 \mu_1 \leqslant \alpha_2 r_2 \mu_2$, and $\mu_2 \leqslant r_2$, then it is optimal to give priority to type 2 jobs at all decision epochs.*

Corollary 3 states that jobs with shorter mean service times and higher expected rewards should receive priority at all decision epochs regardless of the system state if they also abandon the system at a sufficiently high rate (but not necessarily at a higher rate than the other type). In the context of emergency response, this means that if the type of patients who have shorter mean service time (i.e., type 2 patients) also bring a higher expected reward, then they should get the highest priority regardless of the system state if their life expectancies are significantly short (i.e., $r_2 \geqslant \max\{\mu_2, (\alpha_1\mu_1/\alpha_2\mu_2)r_1\}$).

Proposition 4. *If $\mu_1 = \mu_2 := \mu$, $\alpha_1 \leqslant \alpha_2$, and Condition (4) holds, then it is optimal to give priority to type 2 jobs at all decision epochs.*

Proposition 4 implies that the type with the larger expected reward and immediate opportunity cost should receive higher priority regardless of the system state if the service times do not depend on the type of jobs. Identical service rates is a reasonable assumption in the emergency response context when the service constitutes transporting patients from the field to a hospital (see, e.g., Sacco 2005).

Corollary 3 and Propositions 3 and 4 provide us sets of sufficient conditions that lead to the optimality of index policies. They are not necessary conditions however, and index policies might be optimal even when none of these conditions holds. Although it does not appear to be possible to identify necessary conditions, by applying a simple argument, we can characterize the structure of the "best" index policy given that there is an index policy that is optimal among all policies in Π.

Proposition 5. *If the optimal policy among the set of all index policies is also optimal within Π, then it gives priority to the job with the largest value of $\alpha_i r_i/(\mu_{3-i} + r_i)$.*

Proposition 5 describes the optimal policy under the condition that there is an index policy that is optimal. This might not hold in general because it can be clearly observed from Figure 1. However, this index policy can still perform well and thus can be used as a heuristic policy even though it might not be optimal. One important reason for expecting a reasonably good performance from this policy is that it is "myopically" optimal. As we noted earlier in this section, $\alpha_i r_i/(\mu_{3-i} + r_i)$ can be seen as the immediate cost of not providing service to a type i job. Hence, the index policy given in Proposition 5 simply gives priority to the job with the largest "immediate cost." We test the performance of this index policy along with our other heuristic policies in §6.

Remark 1. Corollary 3 and Proposition 5 generalize Propositions 5 and 2 in Argon et al. (2008) to type-dependent rewards, respectively.

5. Heuristic Policies

In §§3 and 4, we obtained partial characterizations of the optimal policy and also identified conditions under which simple state-independent policies are optimal. For the remaining cases where the optimal policy is not characterized completely, we develop two state-dependent heuristic policies (the 2-*step* and *threshold heuristics*) that are expected to perform well under a variety of conditions. We also propose an index policy, which we call the *myopic policy*, based on Proposition 5. Finally, we discuss two other index policies, namely the $\alpha r\mu$-rule and *time-critical-first rule*, which will later serve as benchmark policies in our numerical study. All heuristic policies that we propose and define in this section are general enough to be used when there are multiple identical servers. When implementing our policies with multiple servers, we do not keep track of the number of jobs from each type that are in service for ease of implementation. In the following, we use $M \geqslant 1$ to denote the number of servers.

Below, we describe these heuristic policies under the assumption that the service times and lifetimes are exponentially distributed. However, as we explain later in §6, they can be also applied in settings with other distributional conditions. We also assume, without loss of generality, that $m_i \geqslant 1$ and $q_i \geqslant 1$ for all $i \in \{1, 2\}$ because the problem becomes trivial when the number of jobs from a certain type is zero.

1. 2-*step policy*: At every decision epoch, this heuristic maximizes the expected total reward for the current decision epoch and the next one if the next event is a service completion and not an abandonment. More precisely, to obtain this heuristic, we solve the dynamic programming equations for the case with multiple servers given

in the appendix, assuming that the problem horizon is of two periods length. This gives us the following policy. Let $\Phi = \{0, 1, \ldots, \min\{m_1, M\}\}$. At time zero, choose n_1^* that attains the following maximum:

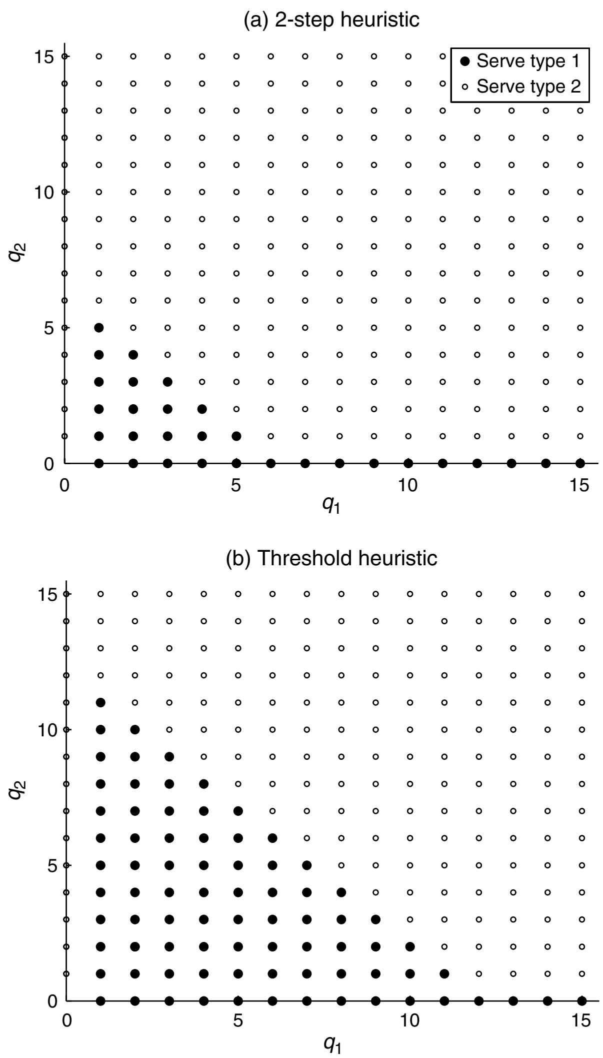
$$\max_{n_1\in\Phi}\left\{\alpha_1 n_1 + \alpha_2(M-n_1) + \frac{(n_1\mu_1+(M-n_1)\mu_2)\max\{\mathbb{1}_{\{m_1-n_1\geqslant 1\}}\alpha_1, \mathbb{1}_{\{m_2-M+n_1\geqslant 1\}}\alpha_2\}}{n_1\mu_1+(M-n_1)\mu_2+(m_1-n_1)r_1+(m_2-M+n_1)r_2}\right\},$$

and allocate n_1^* servers to type 1 jobs and $M-n_1^*$ servers to type 2 jobs. At later decision epochs, whenever there is an available server to be assigned, serve type i^* that attains

$$\max_{i\in\{1,2\}}\left\{\alpha_i + \frac{M\mu_i \max\{\mathbb{1}_{\{q_i\geqslant 2\}}\alpha_i, \alpha_{3-i}\}}{M\mu_i - r_i + q_1 r_1 + q_2 r_2}\right\}. \tag{5}$$

Figure 2(a) shows the structure of the 2-step heuristic for the same experimental settings used in Figure 1(a). From this plot, we observe that for larger numbers of types 1 and 2 jobs, the heuristic prioritizes type 2 jobs, which is consistent with the optimal policy. Note, however, that the structure of the curves separating the state space differs between the 2-step heuristic and the optimal policy.

Figure 2. A sample structure where $\alpha_1 = 1.000$, $\alpha_2 = 1.010$, $\mu_1 = 0.901$, $\mu_2 = 0.909$, $r_1 = 1.000$, and $r_2 = 0.833$.



2. *Threshold policy*: A quick examination of Figure 1(a) suggests that the optimal policy can be possibly approximated by a set of threshold values on the total number of jobs. For example, a line that passes through points ($q_1 = 0$, $q_2 = 38$) and ($q_1 = 38$, $q_2 = 0$) could be used as the boundary between the set of states in which type 1 jobs are served and those in which type 2 jobs are served. This policy is clearly not optimal, but it has the potential to perform better than state-independent policies.

The heuristic policy that we propose is described by one threshold $\mathcal{T}$. It is specifically designed for the case where $\mu_1 \leqslant \mu_2$, $\alpha_1 \leqslant \alpha_2$, and $r_1 \geqslant r_2$, although it is possible to use it in other parameter regions as well. The heuristic works as follows: At a service completion, type 1 jobs are prioritized if $q_1 + q_2 \leqslant \mathcal{T}$; otherwise, type 2 jobs are prioritized. Similarly, at time zero, if $m_1 + m_2 \leqslant \mathcal{T} + M - 1$, then the threshold policy gives priority to type 1 jobs, i.e., $\min\{m_1, M\}$ type 1 jobs are taken into service at time zero and the remaining servers are allocated to type 2 jobs. Otherwise, type 2 jobs are given priority, and as many servers as possible are allocated to them.

We use Proposition 2 in selecting the threshold $\mathcal{T}$. More precisely, we consider the equation $t(q_1) = q_2$, where $t(q_1)$ is defined by (3). First, we let $q_1 = 1$ in this equation and solve for q_2 (the solution is denoted by q_2^*), then we let $q_2 = 1$ in the same equation and solve for q_1 (the solution is denoted by q_1^*) to get

$$q_1^* = \frac{\mu_2(\alpha_1 r_1 - \alpha_2 r_2)}{r_1[r_2(\alpha_2-\alpha_1)+\alpha_1(\mu_2-\mu_1)]}, \quad \text{and}$$
$$q_2^* = \frac{\mu_1(\alpha_1 r_1 - \alpha_2 r_2)}{r_2[r_1(\alpha_2-\alpha_1)+\alpha_2(\mu_2-\mu_1)]}. \tag{6}$$

For the multiple-server problem, i.e., for $M \geqslant 2$, we replace μ_i with $M\mu_i$, for $i = 1, 2$, in the equations given by (6). Finally, we let $\mathcal{T} = \max\{q_1^*, q_2^*\}$.

One nice property of the threshold heuristic is its simple structure because it is completely characterized by a single threshold $\mathcal{T}$ that divides the state space into two regions, and the only required information is the total number of jobs in the system. Given Proposition 2, it would be reasonable to expect a relatively good performance from a policy that gives priority to more urgent jobs when the total number of jobs is below a certain threshold and to less urgent ones otherwise. This is precisely what the threshold policy does. Figure 2(b) shows the structure of the threshold heuristic for the same experimental setting used in Figure 1(a). As can be seen from this plot, the structure of the threshold policy is similar to that of the optimal policy in that the heuristic gives priority to type 2 jobs when the number of jobs in the system is large.

3. *Myopic policy*: Proposition 5 states that the policy that gives priority to the job with the largest value of $\alpha_i r_i \mu_i + \alpha_i r_1 r_2$ is optimal given that there is an optimal index policy and there is a single server. As we have discussed in §4.2, this policy can be seen as prioritizing the job with the largest immediate opportunity cost of not providing service. Hence, we call it the myopic policy. To generalize this to the multiple-server case, we replace μ_i with $M\mu_i$ in the index for a single server. More specifically, myopic policy prioritizes type i^* that attains

$$\max_{i\in\{1,2\}} \left\{ \frac{\alpha_i r_i}{M\mu_{3-i} + r_i} \right\}.$$

4. $\alpha r \mu$*-rule*: This index policy, which is proposed by Glazebrook et al. (2004), prioritizes type i^* that attains $\max_{i\in\{1,2\}}\{\alpha_i r_i \mu_i\}$. Glazebrook et al. show that when the lifetimes are exponentially distributed and there is a single server, the $\alpha r\mu$-rule is asymptotically optimal as abandonment rates approach zero. More specifically, if the abandonment rates are defined as $r_i = \theta \nu_i$ for $i = 1, 2$, then the $\alpha r\mu$-rule is asymptotically optimal as $\theta \to 0$. We can show that the $\alpha r\mu$-rule and the myopic policy behave similarly under this asymptotic condition. To see this, consider the ratio of the two indices:

$$\frac{\alpha_i r_i \mu_i}{\alpha_i r_i/(\mu_{3-i} + r_i)} = \mu_1 \mu_2 + r_i \mu_i$$

for $i = 1, 2$. This shows that the $\alpha r\mu$-rule and the myopic policy will behave similarly when r_is converge to zero for all $i \in \{1, 2\}$, everything else remaining the same. Moreover, using the asymptotic optimality of the $\alpha r\mu$-rule for small θ, we can expect that the performance of the myopic policy will be close to that of the optimal policy for small θ, under the assumption that the lifetimes are exponentially distributed and there is a single server.

5. *Time-critical-first* (*TCF*) *rule*: This index policy is based on the common practice for patient triage during daily emergencies that always gives priority to the most time-critical patients. To be more precise, this heuristic prioritizes type i^* that attains $\max_{i\in\{1,2\}}\{r_i\}$. Although this rule is expected to perform poorly in general, we still include it in our numerical analysis due to its common use in daily triage.

Among the five heuristics described in this section, the TCF rule is the easiest to implement because it simply requires an ordering of the patients with respect to their mean remaining life expectancies. The $\alpha r\mu$-rule and the myopic policy are simple policies as well, although in addition to mean life expectancies these heuristics require estimates on rewards (such as survival probabilities) and mean service times. In comparison, the 2-step and threshold policies are more sophisticated because they both prescribe state-dependent rules. However, they are also relatively easy to implement, arguably among the simplest state-dependent policies that can be expected to perform well. One of the desirable aspects of these policies is that they do not use any distributional properties other than the mean values of remaining lifetimes, service times, and rewards. The threshold policy is even simpler than the 2-step policy in that it keeps track only of the total number of jobs as opposed to the number of jobs from each type.

Finally, in this section, we present a result that shows that the proposed heuristics (i.e., the 2-step, threshold, and myopic policies) agree with the optimal policy for certain conditions under which the optimal policy can be characterized and these heuristics are well defined.

Proposition 6. *Suppose that the Markovian assumption holds and there is a single server.*

(i) *If* $\mu_1 \leqslant \mu_2$, $r_1 \leqslant r_2$, *and* $\alpha_1 \leqslant \alpha_2$, *then the* 2*-step policy, myopic policy,* $\alpha r\mu$*-rule, and TCF rule are optimal.*

(ii) *If* $\mu_1 \leqslant \mu_2 \leqslant r_2$, $\alpha_1 \leqslant \alpha_2$, *and* $\alpha_1 r_1 \mu_1 \leqslant \alpha_2 r_2 \mu_2$, *then the* 2*-step, threshold, and myopic policies, and* $\alpha r\mu$*-rule are optimal.*

(iii) *If* $\mu_1 = \mu_2 := \mu$, $\alpha_1 \leqslant \alpha_2$, *and* $\alpha_1 r_1/(\mu + r_1) \leqslant \alpha_2 r_2/(\mu + r_2)$, *then the* 2*-step, threshold, and myopic policies are optimal.*

To prove parts (i), (ii), and (iii) of Proposition 6, we show that the indicated heuristics satisfy the conditions of Corollaries 2 and 3, and Proposition 4, respectively. The details of the proof are provided in the appendix.

6. Numerical Results

In this section, we present our numerical results on the performance of the heuristics discussed in §5. Our numerical study is divided into four parts. In §6.1, we test the heuristics on the Markovian model defined in §4. In §§6.2 and 6.3, we consider two relaxations of this model, one with multiple servers and another with nonexponential service time and lifetime distributions, respectively. In §6.4, we study a case where rewards decrease with time. In all four settings, we can compute the performance of the optimal policy and hence compare the performances of the heuristics with the optimal performance. Finally, in §6.5, we provide a discussion on our numerical analysis.

6.1. The Markovian Case

In this section, we test the performance of the heuristic policies discussed in §5 under the assumption that for type $i \in \{1, 2\}$ jobs, service times, and lifetimes are exponentially distributed with rate $\mu_i > 0$ and $r_i > 0$, respectively. To cover as many different scenarios as possible, we used random samples of the system parameters. More specifically, we generated the initial numbers of jobs m_i, for $i = 1, 2$, independently and uniformly over the set $\{1, 2, \ldots, 100\}$. Moreover, we generated the expected rewards α_i and service rates μ_i, for $i = 1, 2$, independently from a uniform distribution with ranges $[0, 1]$ and $[0.5, 2.0]$, respectively. We considered five subsets of experiments depending on the range of the abandonment rates r_i, for $i = 1, 2$, which are generated independently from a uniform distribution with ranges $[2.0, 5.0]$, $[0.5, 2.0]$, $[0.1, 0.5]$, $[0.01, 0.1]$, and $[0.005, 0.01]$. (The first [last] subset corresponds to the case where jobs are most [least] time-critical.) For each subset, we generated 5,000 random scenarios where $\alpha_1 < \alpha_2$, $\mu_1 < \mu_2$, and $r_1 > r_2$. For each scenario, we calculated the expected total reward collected under each one of the five heuristic policies and the optimal policy. Then, we computed the percentage deviation of the expected total reward of each heuristic from that of the optimal policy, constructed a 95% confidence interval (CI) on the mean of these 5,000 percentage deviations, and calculated the median and maximum percentage deviation. Finally, we counted the number of scenarios in which each heuristic provided the best performance among the five heuristics. The results are presented in Table 1.

From Table 1, we observe that our two state-dependent policies perform exceptionally well when jobs are very time-critical. In this case, the difference between the state-dependent and index policies is especially pronounced in terms of the worst-case performance. On the other hand, when jobs are not time-critical, all policies except for TCF perform well, but the myopic policy and $\alpha r\mu$-rule yield a slightly better performance. This is an expected result because we already know from §5 that when the abandonment rates are small the $\alpha r\mu$-rule and the myopic policy provide near-optimal performances.

Comparing the two state-dependent heuristics, we see that the threshold policy is better than the 2-step policy in all test cases provided in Table 1. Among the three index policies considered, the myopic policy is the best across all parameter sets, and it is significantly better than the other two when jobs are time-critical.

Table 1. Performance of the heuristic policies (in terms of the percentage deviation from the optimal performance) when the service times and lifetimes are exponentially distributed and $M = 1$.

Heuristic	95% C.I.	Median	Maximum	No. of times best
		$r_i \sim$Uniform$[2.0, 5.0]$		
2-step	0.00±0.00	0.00	0.30	4,978
Threshold	0.00±0.00	0.00	0.08	4,998
Myopic	0.14±0.03	0.00	25.99	4,918
$\alpha r\mu$	1.29±0.14	0.00	43.15	4,618
TCF	49.64±0.63	50.43	98.84	93
		$r_i \sim$Uniform$[0.5, 2.0]$		
2-step	0.01±0.00	0.00	4.38	4,905
Threshold	0.00±0.00	0.00	1.68	4,979
Myopic	0.90±0.10	0.00	31.37	4,643
$\alpha r\mu$	2.27±0.18	0.00	43.02	4,328
TCF	43.33±0.58	43.28	96.12	97
		$r_i \sim$Uniform$[0.1, 0.5]$		
2-step	0.07±0.01	0.00	11.06	4,725
Threshold	0.05±0.01	0.00	6.68	4,920
Myopic	1.39±0.12	0.00	28.11	4,362
$\alpha r\mu$	1.85±0.14	0.00	30.29	4,241
TCF	33.81±0.54	32.91	90.21	150
		$r_i \sim$Uniform$[0.01, 0.1]$		
2-step	0.50±0.05	0.00	20.06	4,348
Threshold	0.47±0.05	0.00	20.04	4,500
Myopic	0.63±0.06	0.00	19.17	4,377
$\alpha r\mu$	0.68±0.06	0.00	19.87	4,352
TCF	15.41±0.37	12.53	67.69	601
		$r_i \sim$Uniform$[0.005, 0.01]$		
2-step	0.03±0.01	0.00	5.96	4,871
Threshold	0.03±0.01	0.00	5.96	4,875
Myopic	0.02±0.00	0.00	4.23	4,896
$\alpha r\mu$	0.02±0.01	0.00	4.23	4,893
TCF	10.95±0.26	8.65	51.80	219

6.2. Multiple Servers

In this section, we relax the assumption of a single server in our Markovian model and test the performance of the heuristic policies for the case with two and three identical servers. The experimental setting is the same as the one in §6.1. One difference is that we divide the service rates by the number of servers in order to partially isolate the effect of increasing the number of servers from that of increasing the total service capacity. We find the optimal performance by solving the dynamic programming equations for the multiple-server case provided in the appendix. The results are presented in Table 2.

When we compare Tables 1 and 2, we observe that the overall performance of heuristics remain similar across all parameters when the number of servers increases. More specifically, the 2-step and threshold policies perform very well, being substantially better than the index policies when jobs are time-critical and yielding a comparable performance to that of the myopic policy and $\alpha r\mu$-rule when jobs are not time-critical. Hence, this study provides supporting evidence that all heuristics under consideration are in general robust with respect to changes in the number of servers. One exception is that the worst-case performance of the 2-step heuristic degrades some with an increase in the number of servers when jobs are time-critical.

6.3. Weibull Lifetimes and Deterministic Service Times

In this section, we test the performance of the heuristics discussed in §5 under a nonexponential setting. As in Argon et al. (2008), we assume that the lifetimes come from a Weibull distribution with shape parameter $\theta_i > 0$ and scale parameter $\beta_i > 0$. (Weibull is a commonly used distribution for modeling the lifetimes of humans; see, e.g., section 2.2.2 in Hougaard 2000.) This implies that the abandonment rates are given by $r_i = \theta_i/[\beta_i \Gamma(1/\theta_i)]$ for $i = 1, 2$, where $\Gamma(\cdot)$ is the gamma function. We assume that the service times are type-dependent but deterministic, which allows us to compute the performance of the optimal policy using backward induction. In particular, we formulate the problem as a semi-Markov decision process, where the state at each decision epoch is defined by $(q_1, q_2; t)$ and

Table 2. Performance of the heuristic policies (in terms of the percentage deviation from the optimal performance) when the service times and lifetimes are exponentially distributed and $M \in \{2, 3\}$.

Heuristic	$M=2$ 95% CI	$M=2$ Median	$M=2$ Maximum	$M=2$ No. of times best	$M=3$ 95% CI	$M=3$ Median	$M=3$ Maximum	$M=3$ No. of times best
				r_i ~Uniform[2.0, 5.0]				
2-step	0.00±0.00	0.00	7.26	4,969	0.01±0.01	0.00	19.08	4,968
Threshold	0.00±0.00	0.00	0.19	5,000	0.00±0.00	0.00	0.26	5,000
Myopic	0.12±0.03	0.00	21.45	4,918	0.11±0.03	0.00	18.01	4,918
$\alpha r \mu$	1.28±0.14	0.00	46.62	4,611	1.25±0.14	0.00	48.50	4,611
TCF	50.49±0.65	50.56	98.79	0	50.08±0.68	49.88	98.75	0
				r_i ~Uniform[0.5, 2.0]				
2-step	0.03±0.02	0.00	24.96	4,846	0.03±0.01	0.00	28.02	4,850
Threshold	0.01±0.00	0.00	1.58	4,996	0.01±0.00	0.00	1.80	4,996
Myopic	0.93±0.11	0.00	31.93	4,639	0.92±0.10	0.00	31.93	4,639
$\alpha r \mu$	2.41±0.19	0.00	47.67	4,314	2.43±0.20	0.00	50.44	4,314
TCF	45.42±0.58	45.51	96.76	4	45.83±0.61	46.09	97.21	4
				r_i ~Uniform[0.1, 0.5]				
2-step	0.09±0.02	0.00	10.73	4,582	0.10±0.02	0.00	16.90	4,582
Threshold	0.05±0.01	0.00	5.57	4,949	0.06±0.01	0.00	5.84	4,947
Myopic	1.51±0.13	0.00	29.20	4,345	1.55±0.13	0.00	30.43	4,347
$\alpha r \mu$	2.00±0.15	0.00	32.08	4,221	2.07±0.16	0.00	34.48	4,223
TCF	35.54±0.54	34.75	91.06	54	36.16±0.55	35.32	91.61	56
				r_i ~Uniform[0.01, 0.1]				
2-step	0.52±0.05	0.00	20.12	4,232	0.53±0.05	0.00	20.26	4,225
Threshold	0.40±0.04	0.00	15.31	4,506	0.41±0.04	0.00	15.33	4,503
Myopic	0.67±0.06	0.00	19.81	4,358	0.69±0.06	0.00	20.30	4,359
$\alpha r \mu$	0.72±0.06	0.00	20.63	4,332	0.74±0.07	0.00	21.36	4,333
TCF	15.93±0.38	13.07	68.65	514	16.09±0.38	13.23	69.36	517
				r_i ~Uniform[0.005, 0.01]				
2-step	0.03±0.01	0.00	5.98	4,862	0.03±0.01	0.00	6.00	4,859
Threshold	0.03±0.01	0.00	5.98	4,872	0.03±0.01	0.00	6.00	4,872
Myopic	0.02±0.00	0.00	4.28	4,895	0.02±0.00	0.00	4.32	4,895
$\alpha r \mu$	0.02±0.01	0.00	4.28	4,892	0.02±0.01	0.00	4.32	4,892
TCF	11.21±0.26	8.90	52.13	129	11.24±0.26	8.93	52.41	129

$t \geqslant 0$ denotes the time. Because the service times are deterministic, there is a finite number of combinations of time points that are decision epochs, and hence we can solve the optimality equations using backward induction.

We first discuss how we adapt the heuristics described in §5 to this nonexponential setting. The main issue is that when lifetimes are not exponentially distributed, the abandonment rates change with time. Hence, when implementing the heuristics, we updated the abandonment rates at every service completion. We can show that the *updated abandonment rate* for type $i \in \{1, 2\}$ at time $t \geqslant 0$, which we denote by $r_i(t)$, is given by

$$r_i(t) = \frac{\theta_i}{\beta_i \Gamma(1/\theta_i, (t/\beta_i)^{\theta_i})} e^{-(t/\beta_i)^{\theta_i}} \tag{7}$$

when the lifetimes follow a Weibull distribution. Here, $\Gamma(a, b) := \int_b^\infty u^{a-1} e^{-u}\, du$, for $a > 0$ and $b \geqslant 0$, is the incomplete gamma function. Note that $r_i(0) = r_i$.

For the numerical experiments, we set the initial number of jobs m_i to 20 and let $\theta_i = 1.5$ for both $i = 1, 2$. (Due to the computational complexity of this nonexponential case, we could not use the same experimental setting for m_is as in §6.1.) We then generated the expected rewards α_i and service rates μ_i, for $i = 1, 2$, independently from a uniform distribution with ranges $[0, 1]$ and $[0.5, 2.0]$, respectively, and the initial abandonment rate $r_i(0)$ from a uniform distribution with five different ranges: $[2.0, 5.0]$, $[0.5, 2.0]$, $[0.1, 0.5]$, $[0.01, 0.1]$, and $[0.005, 0.01]$. For each of these five subsets of experiments, we generated 5,000 random scenarios, where $\alpha_1 < \alpha_2$, $\mu_1 < \mu_2$, and $r_1(0) > r_2(0)$. (Because the shape parameter is the same for all types of jobs, having $r_i(0) > r_j(0)$ implies that $r_i(t) \geqslant r_j(t)$ for all $t \geqslant 0$, $i, j \in \{1, 2\}$.) We computed the performance of each heuristic as in §6.1 and summarized the results in Table 3. In the same table, we also provide the results for the Markovian model under the same experimental setting for comparison purposes.

Table 3 demonstrates that the relative performances of the heuristics are similar for the exponential and nonexponential settings under the given experimental conditions. In particular, the 2-step and threshold policies perform significantly better than the index policies for the first three cases, where the jobs are time-critical; and they perform reasonably well for the last two cases, where the myopic

Table 3. Performance of the heuristic policies (in terms of the percentage deviation from the optimal performance) for the exponential and nonexponential settings when $m_1 = m_2 = 20$.

	Nonexponential setting				Exponential setting			
Heuristic	95% C.I.	Median	Maximum	No. of times best	95% C.I.	Median	Maximum	No. of times best
				$r_i \sim$Uniform[2.0, 5.0]				
2-step	0.00±0.00	0.00	0.15	4,985	0.00±0.00	0.00	0.21	4,967
Threshold	0.00±0.00	0.00	0.41	4,993	0.00±0.00	0.00	0.06	4,996
Myopic	0.21±0.06	0.00	42.91	4,905	0.14±0.04	0.00	23.09	4,910
$\alpha r\mu$	1.67±0.18	0.00	51.76	4,378	1.23±0.13	0.00	39.42	4,616
TCF	58.34±0.65	60.72	99.81	0	50.68±0.60	52.16	94.32	1
				$r_i \sim$Uniform[0.5, 2.0]				
2-step	0.00±0.00	0.00	1.05	4,913	0.01±0.00	0.00	1.86	4,837
Threshold	0.00±0.00	0.00	1.99	4,865	0.00±0.00	0.00	0.60	4,990
Myopic	1.19±0.12	0.00	30.84	4,545	0.85±0.10	0.00	30.91	4,637
$\alpha r\mu$	3.36±0.20	0.00	31.71	3,899	2.11±0.17	0.00	34.91	4,338
TCF	46.53±0.56	47.30	93.66	3	44.19±0.54	45.50	89.68	7
				$r_i \sim$Uniform[0.1, 0.5]				
2-step	0.15±0.02	0.00	10.92	4,445	0.12±0.02	0.00	10.24	4,590
Threshold	0.06±0.01	0.00	6.66	4,922	0.06±0.01	0.00	5.27	4,905
Myopic	1.60±0.11	0.00	22.13	4,148	1.01±0.09	0.00	22.09	4,422
$\alpha r\mu$	2.43±0.15	0.00	25.16	3,817	1.46±0.12	0.00	25.55	4,288
TCF	35.78±0.52	35.75	87.74	84	32.62±0.47	33.18	81.34	102
				$r_i \sim$Uniform[0.01, 0.1]				
2-step	1.74±0.11	0.00	27.06	3,664	0.78±0.07	0.00	18.75	4,169
Threshold	1.27±0.08	0.00	15.81	3,753	0.55±0.05	0.00	11.24	4,316
Myopic	0.08±0.01	0.00	9.32	4,734	0.22±0.03	0.00	11.07	4,581
$\alpha r\mu$	0.10±0.02	0.00	9.35	4,694	0.27±0.03	0.00	12.52	4,547
TCF	8.31±0.25	6.22	70.24	1,171	11.07±0.25	10.01	61.84	721
				$r_i \sim$Uniform[0.005, 0.01]				
2-step	0.07±0.01	0.00	4.53	4,600	0.03±0.01	0.00	4.19	4,818
Threshold	0.07±0.01	0.00	4.39	4,606	0.03±0.01	0.00	3.68	4,824
Myopic	0.01±0.00	0.00	0.56	4,852	0.00±0.00	0.00	0.84	4,930
$\alpha r\mu$	0.01±0.00	0.00	0.56	4,854	0.00±0.00	0.00	0.84	4,928
TCF	3.41±0.08	2.94	17.51	400	6.67±0.12	6.23	22.87	182

policy and $\alpha r\mu$-rule are the best. This shows that the state-dependent heuristics that are developed for the exponential case might also perform well in a nonexponential setting. Among all the index policies considered, myopic policy is the best, and the $\alpha r\mu$-rule performs similarly well for small abandonment rates as in the Markovian case. It appears that the only difference between the two distributional assumptions is that the performance differences among the heuristics are more pronounced in the nonexponential setting.

6.4. Diminishing Rewards

In this last part of our numerical study, we relax the assumption that the expected rewards are constant and assume that rewards diminish with time. In the context of emergency response, this models the situations where the survival probabilities of patients decrease as their delay for service increases. To the best of our knowledge, the only work in the emergency medicine literature that estimates the survival probabilities as a function of time is by Sacco et al. (2005, 2007). In these two articles, the authors categorize patients according to a scoring scheme called the RPM score, which takes value over the set $\{0, 1, \ldots, 12\}$ and is the sum of coded values for *r*espiratory rate, *p*ulse rate, and best *m*otor response. Based on the data given in Sacco et al. (2007), we plot the survival probabilities versus time for patients with initial RPM values of 6 through 12 in Figure 3(a). (All patients with initial RPM values between 0 and 5 have zero RPM after the first hour, and hence the associated curves are omitted.) In this plot, each time unit represents 30 minutes.

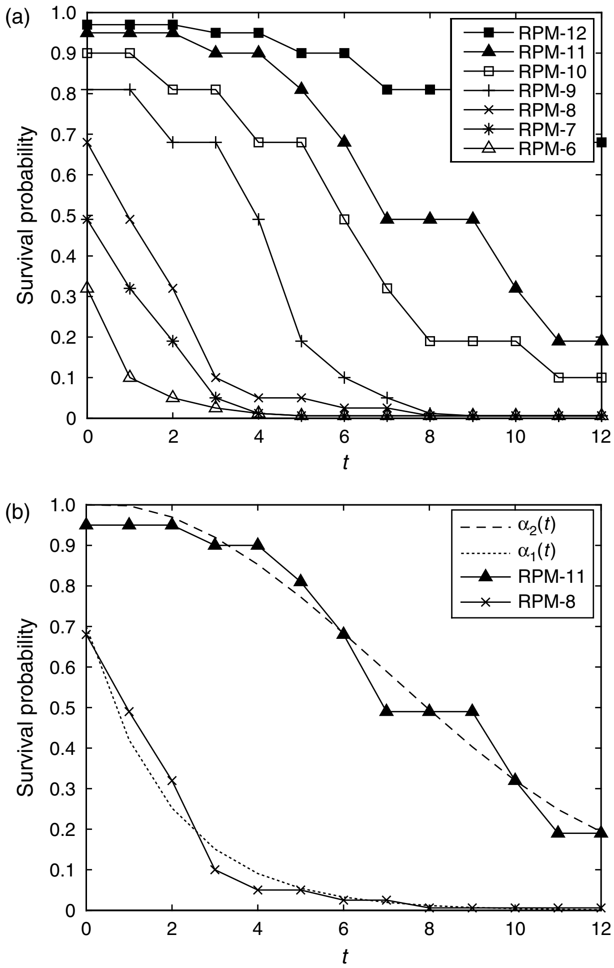
For our numerical experiment, we selected RPM-8 and RPM-11 as the reward functions for types 1 and 2 jobs, respectively, because type 1 is the more critical type. We then fitted functions to describe these reward curves by using the distribution fitting add-on EasyFitXL for MS Excel. Let $\alpha_1(t)$ and $\alpha_2(t)$ denote the fitted functions for types 1 and 2 jobs, respectively. These functions are given as

$$\alpha_1(t) = 0.7(0.6)^t \quad \text{for } t \geqslant 0,$$

$$\alpha_2(t) = \begin{cases} 0.0007t^3 - 0.015t^2 + 0.012t + 1 & \text{for } 0 \leqslant t < 12, \\ 0.2 & \text{for } t \geqslant 12 \end{cases}$$

and are plotted in Figure 3(b).

Figure 3. Survival probabilities as a function of time from Sacco et al. (2007).



In an emergency response setting, the survival probability curves can be also used to estimate the mean lifetimes of patients. Hence, in our numerical experiment, we set the initial mean lifetime for type 1 jobs, $1/r_1(0)$, to 8, which is the t-intercept of the RPM-8 function. Similarly, we set $r_2(0) = 1/12$ as $\alpha_2(t) = 0.2$ for all $t \geqslant 12$. As in §6.3, lifetimes have a Weibull distribution and its mean is updated according to Equation (7). Service times are deterministic but type-dependent, and service rates are sampled from a uniform distribution with four different ranges, specifically, $[0.1, 0.5]$, $[0.5, 1.0]$, $[1.0, 2.0]$, and $[2.0, 10.0]$. For each of these four subsets of experiments, we generated 5,000 random scenarios, where $\mu_1 < \mu_2$. We computed the performance of each heuristic and the optimal policy as in §6.3 and summarized the results in Table 4. To obtain the optimal policy, we use backward induction by updating the abandonment rates and expected rewards at each decision epoch. As explained in §6.3, this is possible because service times are deterministic.

One of the most important conclusions from Table 4 is that our heuristics, i.e., the 2-step, threshold, and myopic policies, that are developed for the exponential case with constant rewards also perform reasonably well in a nonexponential setting with diminishing rewards. To be specific, when the service is slow, all three heuristics developed in this article provide optimal performance under the given experimental conditions. In such a case, the service is so slow that only a few jobs can be saved. This means that there are essentially only a couple of decision epochs, which explains the superior performance of 2-step and myopic heuristics. For mediocre service speed, the 2-step and threshold heuristics still provide optimal performance, but the myopic policy deviates slightly from the optimal performance in a number of scenarios. Finally, for the case where the service is a lot faster than abandonments, the performance of all policies (except for TCF) degraded to about the same level. In these situations, there is enough time to save many patients, and hence more patients can be saved by policies that take into account the diminishing structure of the survival probability curves.

Table 4. Performance of the heuristic policies (in terms of the percentage deviation from the optimal performance) when service times are deterministic, lifetimes come from a Weibull distribution, expected rewards are diminishing, and $m_1 = m_2 = 20$.

Heuristic	95% CI	Median	Maximum	No. of times best
		μ_i ~Uniform[0.1, 0.5]		
2-step	0.00±0.00	0.00	0.00	5,000
Threshold	0.00±0.00	0.00	0.00	5,000
Myopic	0.00±0.00	0.00	0.00	5,000
$\alpha r \mu$	1.07±0.11	0.00	20.56	4,643
TCF	62.43±0.23	62.56	79.37	0
		μ_i ~Uniform[0.5, 1.0]		
2-step	0.00±0.00	0.00	0.00	5,000
Threshold	0.00±0.00	0.00	0.00	5,000
Myopic	0.01±0.00	0.00	3.53	4,984
$\alpha r \mu$	0.60±0.04	0.00	5.87	4,321
TCF	67.94±0.09	67.52	75.96	0
		μ_i ~Uniform[1.0, 2.0]		
2-step	0.00±0.00	0.00	0.00	5,000
Threshold	0.00±0.00	0.00	0.00	5,000
Myopic	0.10±0.01	0.00	2.67	4,705
$\alpha r \mu$	0.25±0.02	0.00	2.94	4,321
TCF	65.21±0.12	65.26	74.41	0
		μ_i ~Uniform[2.0,10.0]		
2-step	1.29±0.05	0.59	11.63	4,551
Threshold	1.29±0.05	0.59	11.63	4,551
Myopic	1.24±0.05	0.57	11.07	4,766
$\alpha r \mu$	1.22±0.05	0.55	11.07	4,835
TCF	36.85±0.37	40.55	54.40	140

6.5. Discussions

One needs to be careful about carrying over every insight from our numerical study to practice directly because the actual problem in emergency response is significantly more complicated than any mathematical model that can be analyzed. For example, without further study, it would not be reasonable to claim any one of the heuristic policies to be superior than the others for practical purposes or that their performances will actually be as close to optimality as the numerical study suggests. Nevertheless, we believe that our numerical study suggests a number of general insights that can be useful for emergency response practitioners. First, there can be significant benefits of taking resource limitations and casualty numbers into account while giving prioritization decisions, especially when patients' life expectancies are short. Second, these state-dependent policies need not be very complex; policies that simply keep track of the total number of patients and prioritize patients accordingly (as in our threshold policy) can perform well. Finally, when patients' conditions are not very critical, there exist index policies that perform well and thus can be preferred over state-dependent policies because of their simplicity. However, the choice of the index policy is important as the superiority of the myopic policy across all parameter regions, particularly over the TCF policy, clearly indicates.

7. Conclusions

Although there seems to be a consensus on the necessity of rationing resources in the aftermath of mass-casualty events, precisely how that needs to be done is a subject of an ongoing debate. This article makes a contribution to this debate by providing insights into the problem and pointing to areas where more data collection and empirical research is needed.

First and foremost, our findings provide a mathematical support to several researchers in the medical community who have recognized the potential benefits of taking into account the degree of resource limitations when determining priority levels for injured patients. One of the immediate conclusions of our work is that who gets the top priority should ideally depend on the number of patients as well as their injury characteristics. This is something that the current practice largely ignores. Typically, the priority level for a particular patient is determined based on that patient's condition only, and the most urgent patient who has a good chance to survive receives the highest priority. Although this approach might work when resources are abundant, our analysis indicates that when resources are severely limited in comparison with the demand, it might be more beneficial to give priority to less urgent patients with higher chances of survival for the overall success of the response effort.

Given that priority levels should ideally be determined in consideration with how scarce resources are, one natural question is how exactly one should define *severe scarcity*. Also, are there any conditions under which resource availabilities can be safely ignored when determining priorities? These are difficult questions to answer in practice, not only because of the complexity of the problem but also because of the lack of data needed to make more intelligent decisions. It would be naive to expect precise answers through mathematical analysis, but the insights generated from the analysis of stylized models, like the model analyzed in this article, can serve as building blocks for policies that can be used in practice.

The main contributions of this article are threefold. First, we identify several conditions under which the system-state information, i.e., the patient counts, can be ignored when determining priorities. Second, we partially characterize the optimal policy in cases where the optimal decisions could depend on the system state. Third, we demonstrate that one can develop "good" prioritization policies and rules of thumb that only consider the total number of patients as opposed to considering numbers from each type. In particular, we put patients into two different groups with respect to their time-criticality and rewards, and we show that a threshold-type policy, which gives priority to time-critical patients if the number of patients is below a threshold and to less urgent patients otherwise, can perform quite well.

We believe that with this work, we lay a foundation over which a multidisciplinary team of emergency responders (physicians as well as managers) and operations researchers can develop practical and effective prioritization rules. We also believe that any such policy needs to be tested extensively using realistic simulation models before it is put into action. Thus, one important avenue for future work is the development of a simulation test-bed for priority decisions in emergency response. It is also crucial to obtain data on lifetimes as well as survival rates at least for some of the most common types of injuries. Such data would be invaluable not only for the development of potentially effective policies but also for use in realistic simulation scenarios while evaluating these policies.

Appendix

In this appendix, we provide the proofs of our results in the order presented in the article, and we also present a dynamic programming formulation used in §5. The following lemma is needed to prove Proposition 1.

Lemma 1 (Righter 1994, Lemma 13.D.1; Among Others). *Let X and Y be two independent random variables. Then, $X \leqslant_{\rm lr} Y$ if and only if $(X \mid \{X, Y\} = \underline{m}, \max\{X, Y\} = \bar{m}) \leqslant_{\rm st} (Y \mid \min\{X, Y\} = \underline{m}, \max\{X, Y\} = \bar{m})$ for all $\underline{m} \leqslant \bar{m}$.*

Lemma 1 can equivalently be stated as follows: Given $\underline{m} = \min\{X, Y\}$ and $\bar{m} = \max\{X, Y\}$, we have that $X \leqslant_{\rm lr} Y$ if and only if $\Pr\{X = \underline{m} \mid \underline{m}, \bar{m}\} = \Pr\{Y = \bar{m} \mid \underline{m}, \bar{m}\} \geqslant \Pr\{X = \bar{m} \mid \underline{m}, \bar{m}\} = \Pr\{Y = \underline{m} \mid \underline{m}, \bar{m}\}$.

Proof of Proposition 1. We will use a coupling argument to prove the result. Suppose policy π takes job j into service at t_0 while job i is in the queue. Without loss of generality, assume that $t_0 = 0$. We will construct a policy γ which serves job i at time 0, and for which $C_\pi(t) \leqslant C_\gamma(t)$ for all $t \geqslant 0$ along any given sample path.

Let Y_l^ρ denote the remaining lifetime of job l at time 0 under policy ρ, where $l \in \{i, j\}$ and $\rho \in \{\pi, \gamma\}$. Note that by the stochastic ordering relation among the remaining lifetimes of jobs, we can couple the random variables so that $Y_i^\pi = y_i \leqslant y_j = Y_j^\gamma$. Because policy $\pi(\gamma)$ serves job $j(i)$ at time zero and the job in service will not abandon, we do not need Y_j^π and Y_i^γ. Let $Y_l^\gamma = Y_l^\pi$ for all $l \neq i, j$. Let also S_l^ρ denote the service time of job l under policy $\rho \in \{\pi, \gamma\}$, and let $S_l^\gamma = S_l^\pi$ for all $l \neq i, j$. We can couple (S_i^π, S_j^π) with (S_i^γ, S_j^γ) so that $\underline{m} := \min\{S_i^\pi, S_j^\pi\} = \min\{S_i^\gamma, S_j^\gamma\} \leqslant \bar{m} := \max\{S_i^\pi, S_j^\pi\} = \max\{S_i^\gamma, S_j^\gamma\}$ and either $S_j^\pi = S_i^\gamma := a \in \{\underline{m}, \bar{m}\}$ and $S_i^\pi = S_j^\gamma := b \in \{\underline{m}, \bar{m}\}\backslash\{a\}$ (Case I) or $S_i^\pi = S_i^\gamma = \underline{m} \leqslant S_j^\pi = S_j^\gamma = \bar{m}$ (Case II). Note that such a coupling is possible from Lemma 1 and the condition that $S_i \leqslant_{\rm lr} S_j$. Finally, let Z_l^ρ denote the reward of serving job l under policy $\rho \in \{\pi, \gamma\}$ and let $Z_l^\gamma = Z_l^\pi$ for all $l \neq i, j$. Then, we can couple (Z_i^π, Z_j^π) with (Z_i^γ, Z_j^γ) so that $\min\{Z_i^\pi, Z_j^\pi\} = \min\{Z_i^\gamma, Z_j^\gamma\} \leqslant \max\{Z_i^\pi, Z_j^\pi\} = \max\{Z_i^\gamma, Z_j^\gamma\}$, and either $Z_j^\pi = Z_i^\gamma$ and $Z_i^\pi = Z_j^\gamma$ or $Z_j^\pi = Z_j^\gamma \leqslant Z_i^\pi = Z_i^\gamma$. Such a coupling is possible from Lemma 1 and the condition that $Z_j \leqslant_{\rm lr} Z_i$. Let τ be the time instance π takes job i into service ($\tau = \infty$ if job i is not taken into service). The following cases exhaust all possibilities.

Case I:

(a) We first consider the case where $\tau < \infty$. Note that under Case I, the first decision epoch after time zero is at time a for both π and γ. γ follows π during $[a, \tau)$, and at time τ, when π takes job i into service, γ takes job j. This is possible because $y_i \leqslant y_j$. At time $\tau + b$, the states will be the same under both policies and γ follows π from then on. Hence, we have $C_\gamma(t) - C_\pi(t) = Z_i^\gamma - Z_j^\pi \geqslant 0$ for all $0 \leqslant t < \tau$, and $C_\gamma(t) - C_\pi(t) = Z_i^\gamma + Z_j^\gamma - Z_i^\pi - Z_j^\pi = 0$ for all $t \geqslant \tau$.

(b) Now suppose that $\tau = \infty$. Then, γ follows π at all decision epochs after time zero, except that it serves job j last (let the service start time be τ') if it is still available after all other jobs are cleared. Hence, we have $C_\gamma(t) - C_\pi(t) = Z_i^\gamma - Z_j^\pi \geqslant 0$ for all $0 \leqslant t < \tau'$, and if $\tau' < \infty$, $C_\gamma(t) - C_\pi(t) = Z_i^\gamma - Z_j^\pi + Z_j^\gamma \geqslant 0$ for all $t \geqslant \tau'$.

Case II:

(a) We again first consider the case where $\tau < \infty$. γ follows π at every decision epoch during $[\underline{m}, \tau - \bar{m} + \underline{m})$ and serves job j at time $\tau - \bar{m} + \underline{m}$ when π serves job i (at time τ). This is possible because $y_i \leqslant y_j$ and the service completion times under γ are $\bar{m} - \underline{m}$ units of time earlier than those under π between $\bar{m}$ and τ. The states under π and γ become the same at time $\underline{m} + \tau$, and γ follows π from then on. Hence, we have $C_\gamma(t) - C_\pi(t) = Z_i^\gamma - Z_j^\pi \geqslant 0$ for all $0 \leqslant t < \underline{m}$, $C_\gamma(t) - C_\pi(t) \geqslant Z_i^\gamma - Z_j^\pi \geqslant 0$ for all $\underline{m} \leqslant t < \tau - \bar{m} + \underline{m}$, $C_\gamma(t) - C_\pi(t) \geqslant Z_i^\gamma - Z_j^\pi + Z_j^\gamma \geqslant 0$ for all $\tau - \bar{m} + \underline{m} \leqslant t < \tau$, and $C_\gamma(t) - C_\pi(t) = Z_i^\gamma + Z_j^\gamma - Z_i^\pi - Z_j^\pi = 0$ for all $t \geqslant \tau$.

(b) We now consider the case where $\tau = \infty$. γ follows π starting at time m, except that it serves job j last (let the service start time be τ') if it is still available when all other jobs are cleared. As in Case II(a), this is possible because the service completion times under γ are $\bar{m} - \underline{m}$ units of time earlier than those under π after $\bar{m}$. Then we have $C_\gamma(t) - C_\pi(t) = Z_i^\gamma - Z_j^\pi \geqslant 0$ for all $0 \leqslant t < \underline{m}$, $C_\gamma(t) - C_\pi(t) \geqslant Z_i^\gamma - Z_j^\pi \geqslant 0$ for all $\underline{m} \leqslant t < \tau'$, and if $\tau' < \infty$, $C_\gamma(t) - C_\pi(t) = Z_i^\gamma - Z_j^\pi + Z_j^\gamma \geqslant 0$ for all $t \geqslant \tau'$.

Thus, we have shown that $C_\gamma(t) \geqslant C_\pi(t)$ for all $t \geqslant 0$ along any sample path. $\square$

Proof of Corollary 2. Let $A_\pi(t)$ be the total number of jobs taken into service by time t under policy $\pi \in \Pi$. Let also $Z_\pi(i)$ be the reward earned by the ith job to be taken into service. Then for all $t \geqslant 0$, we have

$$C_\pi(t) = \sum_{i=1}^{A_\pi(t)} Z_\pi(i). \tag{8}$$

We next take the expected value of both sides with respect to the rewards of all jobs given all service times and lifetimes of all jobs. We have

$$E_z[C_\pi(t)] = E_z\left[\sum_{i=1}^{A_\pi(t)} Z_\pi(i)\right] = \sum_{i=1}^{A_\pi(t)} E[Z_\pi(i)], \tag{9}$$

because $A_\pi(t)$ is independent of the rewards. (Here, E_z means that the expectation is taken with respect to rewards $Z_1, \ldots, Z_N$.) Then, by Equations (8) and (9), if the condition $Z_i \geqslant_{\rm lr} Z_j$ in Proposition 1 is replaced by $\alpha_i \geqslant_{\rm lr} \alpha_j$, where $\alpha_i = E[Z_i]$ for $i = 1, \ldots, N$, then the result will still hold but in the sense of stochastically increasing $E_z[C_\pi(t)]$ for all $t \geqslant t_0$.

For the Markovian model in §4, the mean rewards for both types of jobs are deterministic. Hence, if $\alpha_i \geqslant \alpha_j$, then $\alpha_i \geqslant_{\rm lr} \alpha_j$. Furthermore, because service times and lifetimes are exponentially distributed, an ordering between their means will lead to an ordering in the sense of both likelihood ratio and hazard rate orders. Therefore, Proposition 1 implies that if $\mu_1 \leqslant \mu_2$, $\alpha_1 \leqslant \alpha_2$, and $r_1 \leqslant r_2$, then the policy that gives priority to type 2 jobs will maximize the total expected reward. $\square$

We use the following two lemmas to prove Propositions 2, 3, and 4.

Lemma 2. *If* $\mu_j \leqslant \mu_i$ *for* $i, j \in \{1, 2\}$, *then* $V(q_1, q_2; P_j) \leqslant V(q_1, q_2; P_i)$.

Proof. Couple the processing times of the jobs in service for the two states such that $S_i' \leqslant S_j'$ with probability one, where S_l' denotes the processing time of the type l job in service, $l \in \{i, j\}$. Let $V_0(q_1, q_2; P_i)$ be the value function when the starting state is $(q_1, q_2; P_i)$ and we idle from time S_i' to S_j' and then follow the optimal policy. Then, from Proposition 1, we have $V(q_1, q_2; P_i) \geqslant V_0(q_1, q_2; P_i) = V(q_1, q_2; P_j)$. $\square$

Lemma 3. *Consider a job type* $j \in \{1, 2\}$ *and let* $i = 3 - j$. *Also, consider a decision epoch with* q_i *type i jobs and* q_j *type j jobs, where* $q_j \geqslant 1$. *Suppose that an optimal action at the decision epoch with* $q_i - 1$ *jobs from type i and* q_j *jobs from type j is to serve a type j job when* $q_i \geqslant 1$. *Similarly, suppose that an optimal action at the decision epoch with* q_i *jobs from type i and* $q_j - 1$ *jobs from type j is to serve a type j job when* $q_j \geqslant 2$. *If*

$$\begin{aligned} &q_1 r_1[(\alpha_j - \alpha_i) r_2 + \alpha_1(\mu_j - \mu_i)] \\ &\quad + q_2 r_2[(\alpha_j - \alpha_i) r_1 + \alpha_2(\mu_j - \mu_i)] \\ &\geqslant \alpha_j r_j (r_i - \mu_i) - \alpha_i r_i (r_j - \mu_j), \end{aligned} \tag{10}$$

and

$$\begin{aligned} &(\mu_j - \mu_i)[(r_1 - \mu_1) q_2 r_2 + (r_2 - \mu_2) q_1 r_1 \\ &\qquad - (r_1 - \mu_1)(r_2 - \mu_2)] \geqslant 0 \end{aligned} \tag{11}$$

when $q_i \geqslant 1$, *then an optimal action at the decision epoch with* q_i *type i jobs and* q_j *type j jobs is to serve a type j job.*

Proof. We will show that $\alpha_j + V(\mathbf{q} - \mathbf{e}_j; P_j) \geqslant \mathbb{1}_{\{q_i \geqslant 1\}} \alpha_i + V(\mathbf{q} - \mathbf{e}_i; P_i)$ for $i \neq j$ under the given conditions, where $\mathbf{q} = (q_1, q_2)$. When $q_i = 0$, this holds trivially. Hence, we consider only the case where $q_i \geqslant 1$.

For $q_i \geqslant 1$, we have

$$
\begin{aligned}
&V(\mathbf{q}-\mathbf{e}_j;P_j)\\
&=\frac{\mu_j \max\{\mathbb{1}_{\{q_j>1\}}\alpha_j+V(\mathbf{q}-2\mathbf{e}_j;P_j),\alpha_i+V(q_1-1,q_2-1;P_i)\}}{\mu_j+q_ir_i+(q_j-1)r_j}\\
&\quad+\frac{q_ir_iV(q_1-1,q_2-1;P_j)+(q_j-1)r_jV(\mathbf{q}-2\mathbf{e}_j;P_j)}{\mu_j+q_ir_i+(q_j-1)r_j}\\
&\geqslant\frac{\mu_j(\alpha_i+V(q_1-1,q_2-1;P_i))}{\mu_j+q_ir_i+(q_j-1)r_j}+\frac{q_ir_iV(q_1-1,q_2-1;P_j)}{\mu_j+q_ir_i+(q_j-1)r_j}\\
&\quad+\frac{(q_j-1)r_j(V(q_1-1,q_2-1;P_i)+\alpha_i-\alpha_j)}{\mu_j+q_ir_i+(q_j-1)r_j}\\
&=\frac{\alpha_j\mu_j+(\alpha_i-\alpha_j)(\mu_j+(q_j-1)r_j)}{\mu_j+q_ir_i+(q_j-1)r_j}\\
&\quad+\frac{q_ir_iV(q_1-1,q_2-1;P_j)}{\mu_j+q_ir_i+(q_j-1)r_j}\\
&\quad+\frac{(\mu_j+(q_j-1)r_j)V(q_1-1,q_2-1;P_i)}{\mu_j+q_ir_i+(q_j-1)r_j},
\end{aligned}
\tag{12}
$$

where the inequality follows because either $q_j = 1$, and hence the inequality is trivial, or by the condition that $\alpha_j + V(\mathbf{q}-2\mathbf{e}_j; P_j) \geqslant \alpha_i + V(q_1-1, q_2-1; P_i)$ for $q_j \geqslant 2$.

Similarly, for $q_i \geqslant 1$, we have

$$
\begin{aligned}
&V(\mathbf{q}-\mathbf{e}_i;P_i)\\
&=\frac{\mu_i \max\{\alpha_j+V(q_1-1,q_2-1;P_j),\mathbb{1}_{\{q_i>1\}}\alpha_i+V(\mathbf{q}-2\mathbf{e}_i;P_i)\}}{\mu_i+(q_i-1)r_i+q_jr_j}\\
&\quad+\frac{q_jr_jV(q_1-1,q_2-1;P_i)+(q_i-1)r_iV(\mathbf{q}-2\mathbf{e}_i;P_i)}{\mu_i+(q_i-1)r_i+q_jr_j}\\
&\leqslant\frac{\mu_i(\alpha_j+V(q_1-1,q_2-1;P_j))}{\mu_i+(q_i-1)r_i+q_jr_j}+\frac{q_jr_jV(q_1-1,q_2-1;P_i)}{\mu_i+(q_i-1)r_i+q_jr_j}\\
&\quad+\frac{(q_i-1)r_i(V(q_1-1,q_2-1;P_j)+\alpha_j-\alpha_i)}{\mu_i+(q_i-1)r_i+q_jr_j}\\
&=\frac{(\alpha_j-\alpha_i)(q_i-1)r_i}{\mu_i+(q_i-1)r_i+q_jr_j}+\frac{\alpha_j\mu_i+q_jr_jV(q_1-1,q_2-1;P_i)}{\mu_i+(q_i-1)r_i+q_jr_j}\\
&\quad+\frac{(\mu_i+(q_i-1)r_i)V(q_1-1,q_2-1;P_j)}{\mu_i+(q_i-1)r_i+q_jr_j},
\end{aligned}
\tag{13}
$$

where the inequality follows because $\alpha_j + V(q_1-1, q_2-1; P_j) \geqslant \mathbb{1}_{\{q_i>1\}}\alpha_i + V(\mathbf{q}-2\mathbf{e}_i; P_i)$ for the first and last terms.

Now, from (12) and (13), we get

$$
\begin{aligned}
&\alpha_j+V(\mathbf{q}-\mathbf{e}_j;P_j)-\alpha_i-V(\mathbf{q}-\mathbf{e}_i;P_i)\\
&\geqslant(\alpha_j-\alpha_i)\bigg(1-\frac{\mu_j+(q_j-1)r_j}{\mu_j+q_ir_i+(q_j-1)r_j}\\
&\qquad\qquad-\frac{q_ir_i}{\mu_i+(q_i-1)r_i+q_jr_j}\bigg)\\
&\quad+(\alpha_j\mu_j(\mu_i+(q_i-1)r_i+q_jr_j)+((\alpha_j-\alpha_i)r_i-\alpha_j\mu_i)\\
&\quad\cdot(\mu_j+q_ir_i+(q_j-1)r_j))\cdot((\mu_j+q_ir_i+(q_j-1)r_j)\\
&\quad\cdot(\mu_i+(q_i-1)r_i+q_jr_j))^{-1}\\
&\quad+\left(\frac{q_ir_i}{\mu_j+q_ir_i+(q_j-1)r_j}-\frac{\mu_i+(q_i-1)r_i}{\mu_i+(q_i-1)r_i+q_jr_j}\right)\\
&\quad\cdot V(q_1-1,q_2-1;P_j)\\
&\quad+\left(\frac{\mu_j+(q_j-1)r_j}{\mu_j+q_ir_i+(q_j-1)r_j}-\frac{q_jr_j}{\mu_i+(q_i-1)r_i+q_jr_j}\right)\\
&\quad\cdot V(q_1-1,q_2-1;P_i)
\end{aligned}
\tag{14}
$$

$$
\begin{aligned}
&=\bigg(\frac{1}{\mu_j+q_ir_i+(q_j-1)r_j}\\
&\qquad-\frac{1}{\mu_i+(q_i-1)r_i+q_jr_j}\bigg)(\alpha_j-\alpha_i)q_ir_i\\
&\quad+(((\alpha_j-\alpha_i)r_i+\alpha_j(\mu_j-\mu_i))(q_ir_i+q_jr_j)\\
&\quad+\alpha_ir_i(r_j-\mu_j)-\alpha_jr_j(r_i-\mu_i))\cdot((\mu_j+q_ir_i+(q_j-1)r_j)\\
&\quad\cdot(\mu_i+(q_i-1)r_i+q_jr_j))^{-1}\\
&\quad+\left(\frac{\mu_j+(q_j-1)r_j}{\mu_j+q_ir_i+(q_j-1)r_j}-\frac{q_jr_j}{\mu_i+(q_i-1)r_i+q_jr_j}\right)\\
&\quad\times(V(q_1-1,q_2-1;P_i)-V(q_1-1,q_2-1;P_j))\\
&=\frac{((\alpha_j-\alpha_i)r_i+\alpha_j(\mu_j-\mu_i))q_jr_j+((\alpha_j-\alpha_i)r_j+\alpha_i(\mu_j-\mu_i))q_ir_i}{(\mu_j+q_ir_i+(q_j-1)r_j)(\mu_i+(q_i-1)r_i+q_jr_j)}\\
&\quad+\frac{\alpha_ir_i(r_j-\mu_j)-\alpha_jr_j(r_i-\mu_i)}{(\mu_j+q_ir_i+(q_j-1)r_j)(\mu_i+(q_i-1)r_i+q_jr_j)}\\
&\quad+\frac{(r_i-\mu_i)q_jr_j+(r_j-\mu_j)q_ir_i-(r_j-\mu_j)(r_i-\mu_i)}{(\mu_j+q_ir_i+(q_j-1)r_j)(\mu_i+(q_i-1)r_i+q_jr_j)}\\
&\quad\times(V(q_1-1,q_2-1;P_j)-V(q_1-1,q_2-1;P_i)).
\end{aligned}
\tag{15}
$$

Finally, the right-hand side of Equation (15) is greater than or equal to zero because for the first term, Condition (10) holds, and for the second term, Condition (11) holds, and we have $V(q_1-1, q_2-1; P_i) \leqslant (\geqslant)\ V(q_1-1, q_2-1; P_j)$ if $\mu_i \leqslant (\geqslant)\ \mu_j$ by Lemma 2. ☐

Proof of Proposition 2. (i) Let $i = 2$ and $j = 1$ in Lemma 3. Then, given $r_1 > r_2$, $\alpha_1 \leqslant \alpha_2$, and $\mu_1 \leqslant \mu_2$, Condition (10) diminishes to $q_2 \leqslant t(q_1)$. Moreover, given $\mu_1 \leqslant \mu_2$, we rewrite Condition (11) as

$$r_1(r_2-\mu_2)q_1+r_2(r_1-\mu_1)q_2\leqslant(r_1-\mu_1)(r_2-\mu_2). \tag{16}$$

Note that for $r_2 < r_1 \leqslant \mu_1 < \mu_2$, Condition (16) and hence Condition (11) is satisfied for all $q_1, q_2 \geqslant 0$.

We next use induction on q_1 to prove the result. For $q_1 = q_2 = 1$, the condition that $q_2 \leqslant t(q_1)$ diminishes to $\alpha_1 r_1(\mu_1 + r_2) \geqslant \alpha_2 r_2(\mu_2 + r_1)$, which is the necessary and sufficient condition for the optimality of serving a type 1 job at the decision epoch with one job from each type; see Equation (2). Then, by Lemma 3, it is optimal to serve a type 1 job at the decision epochs with one job from type 1 and q_2 jobs from type 2 such that $q_2 \leqslant t(1)$. Now, suppose that for $q_1 = a$, where $a \geqslant 1$ is an integer, it is optimal to serve type 1 jobs at the decision epochs with a type 1 jobs and q_2 type 2 jobs such that $q_2 \leqslant t(a)$. Then, applying Lemma 3, we conclude that it is optimal to serve type 1 jobs at decision epochs with $a+1$ jobs from type 1 and q_2 jobs from type 2 such that $q_2 \leqslant t(a+1)$ because $t(q_1)$ is nonincreasing in q_1.

(ii) Let $i=1$ and $j=2$ in Lemma 3. Similar to part (i), given $r_1>r_2$, $\alpha_1\leqslant\alpha_2$, and $\mu_1\leqslant\mu_2$, Condition (10) diminishes to $q_2\geqslant t(q_1)$. Moreover, given $\mu_1\leqslant\mu_2$, we rewrite Condition (11) as

$$r_1(r_2-\mu_2)(q_1-1)+r_2(r_1-\mu_1)(q_2-1)\geqslant\mu_1\mu_2-r_1r_2. \tag{17}$$

Note that for $r_1>r_2\geqslant\mu_2\geqslant\mu_1$, Condition (17) and hence Condition (11) is satisfied for all $q_1,q_2\geqslant 1$.

We next use induction on q_1 to prove the result. We start with the case where $q_1=1$. Lemma 3 implies that if there is a decision epoch with one type 1 job and b type 2 jobs, where $b\geqslant t(1)$ and serving a type 2 job is optimal, then it is also optimal to serve type 2 jobs in all decision epochs with one type 1 job and q_2 type 2 jobs, such that $q_2\geqslant b$. Next, suppose that serving a type 2 job is optimal in decision epochs with a type 1 jobs and q_2 type 2 jobs for all $q_2\geqslant\tilde{t}(a)\geqslant t(a)$, where $a\geqslant 1$ is an integer. If there exists a decision epoch with $a+1$ type 1 jobs and d type 2 jobs, where $d\geqslant\tilde{t}(a)-1$ and serving a type 2 job is optimal, then by Lemma 3 it is also optimal to serve type 2 jobs in all decision epochs with $a+1$ type 1 jobs and q_2 type 2 jobs such that $q_2\geqslant d$. (Note that $t(q_1)$ is nonincreasing in q_1, hence Condition (10) is satisfied for decision epoch with $a+1$ type 1 jobs and $d+1$ type 2 jobs as $d+1\geqslant\tilde{t}(a)\geqslant t(a)\geqslant t(a+1)$.) This completes the proof. □

Proof of Proposition 3. Let $i=1$ and $j=2$ in Lemma 3. Then, Condition (10) becomes

$$\begin{aligned}&q_1r_1[(\alpha_2-\alpha_1)r_2+\alpha_1(\mu_2-\mu_1)]\\&\quad+q_2r_2[(\alpha_2-\alpha_1)r_1+\alpha_2(\mu_2-\mu_1)]\\&\geqslant\alpha_2r_2(r_1-\mu_1)-\alpha_1r_1(r_2-\mu_2).\end{aligned} \tag{18}$$

First note that for $q_1=q_2=1$, Condition (10) is satisfied as it diminishes to $\alpha_1r_1(\mu_1+r_2)\leqslant\alpha_2r_2(\mu_2+r_1)$. Then, note that, for $(\alpha_2-\alpha_1)r_2\geqslant\alpha_1(\mu_1-\mu_2)$ and $(\alpha_2-\alpha_1)r_1\geqslant\alpha_2(\mu_1-\mu_2)$, the left-hand side of Condition (18) is nondecreasing in q_1 and q_2, and hence Condition (10) is satisfied for all $q_1,q_2\geqslant 1$. Moreover, given $\mu_1\leqslant\mu_2$, we rewrite Condition (11) as Condition (17), which is satisfied for all $q_1,q_2\geqslant 1$ because $r_1\geqslant\mu_1$ and $r_2\geqslant\mu_2$. Finally, the necessary and sufficient condition for the optimality of serving a type 2 job at the decision epoch with one job from each type is satisfied as it diminishes to $\alpha_1r_1(\mu_1+r_2)\leqslant\alpha_2r_2(\mu_2+r_1)$; see Equation (2). Then, by Lemma 3 it is optimal to serve a type 2 job at decision epochs with q_1 type 1 jobs and one type 2 job for all $q_1\geqslant 1$. Furthermore, because it is also optimal to serve a type 2 job at decision epochs with no job from type 1 and at least one type 2 job, applying Lemma 3 multiple times completes the proof. □

Proof of Corollary 3. When $r_1\leqslant r_2$, using Corollary 2 together with the conditions $\alpha_1\leqslant\alpha_2$ and $\mu_1\leqslant\mu_2$ we conclude that serving type 2 jobs is optimal at all decision epochs. Hence, we focus on the case where $r_1>r_2$. Now, because we have $r_1>r_2\geqslant\mu_2\geqslant\mu_1$, $\alpha_1\leqslant\alpha_2$, and $\alpha_1r_1\mu_1\leqslant\alpha_2r_2\mu_2$, the conditions required in Proposition 3 are all satisfied, which completes the proof. □

Proof of Proposition 4. Let $i=1$, $j=2$, and $\mu_1=\mu_2:=\mu$ in Lemma 3. Then, Condition (10) becomes

$$(q_1+q_2-1)(\alpha_2-\alpha_1)r_1r_2\geqslant\mu(\alpha_1r_1-\alpha_2r_2). \tag{19}$$

First note that for $q_1=q_2=1$, Condition (10) is satisfied as it diminishes to $\alpha_1r_1/(\mu+r_1)\leqslant\alpha_2r_2/(\mu+r_2)$. Then note that for $\alpha_2\geqslant\alpha_1$, the left-hand side of Condition (19) is nondecreasing in q_1 and q_2, and hence Condition (10) is satisfied for all q_1, $q_2\geqslant 1$. Moreover, given $\mu_1=\mu_2$, Condition (11) is satisfied trivially. Finally, the necessary and sufficient condition for the optimality of serving a type 2 job at the decision epoch with one job from each type is satisfied as it diminishes to $\alpha_1r_1/(\mu+r_1)\leqslant\alpha_2r_2/(\mu+r_2)$; see Equation (2). Then, by Lemma 3, it is optimal to serve a type 2 job at decision epochs with one type 2 job and at least one type 1 jobs. Furthermore, because it is also optimal to serve a type 2 job at decision epochs with no type 1 job and at least one type 2 job, applying Lemma 3 multiple times completes the proof. □

Proof of Proposition 5. As the optimal policy is an index policy, it is sufficient to show that a type i job will be served under the optimal policy at a decision epoch with one job from each type if and only if the required condition holds. Using Equation (2) multiple times, we obtain

$$\begin{aligned}&\alpha_1+V(0,1;P_1)-\alpha_2+V(1,0;P_2)\\&\quad=\alpha_1+\frac{\mu_1\alpha_2}{\mu_1+r_2}-\alpha_2-\frac{\mu_2\alpha_1}{\mu_2+r_1}\\&\quad=\frac{\alpha_1r_1(\mu_1+r_2)-\alpha_2r_2(\mu_2+r_1)}{(\mu_1+r_2)(\mu_2+r_1)}.\end{aligned}$$

Hence, $\alpha_1+V(0,1;P_1)\geqslant\alpha_2+V(1,0;P_2)$ if and only if $\alpha_1r_1(\mu_1+r_2)\geqslant\alpha_2r_2(\mu_2+r_1)$, which completes the proof. □

A dynamic programming formulation generalized for the multiple-server case: Let s_i and q_i denote the number of type i jobs in service and in queue, respectively, for $i=1,2$. Then, the state of the system is denoted by $(q_1,q_2;s_1,s_2)$. At time zero, if $m_1+m_2\leqslant M$, the problem is trivial and we have $V(m_1,m_2;0,0)=\alpha_1m_1+\alpha_2m_2$. Hence suppose that $m_1+m_2>M$. Then, for all $q_1=0,1,\ldots,m_1$ and $q_2=0,1,\ldots,m_2$, we have

$$\begin{aligned}&V(m_1,m_2;0,0)\\&=\max_{n_1\in\{0,1,\ldots,\min\{m_1,M\}\}}\big\{\alpha_1n_1+\alpha_2(M-n_1)+V(m_1-n_1,m_2\\&\qquad\qquad-M+n_1;n_1,M-n_1)\big\},\end{aligned}$$

$$\begin{aligned}V(q_1,q_2;s_1,s_2)=\max\big\{&\mathbb{1}_{\{q_1\geqslant 1\}}\alpha_1+V(q_1-1,q_2;s_1+1,s_2),\\&\mathbb{1}_{\{q_2\geqslant 1\}}\alpha_2+V(q_1,q_2-1;s_1,s_2+1)\big\},\end{aligned}$$

where $s_1=0,1,\ldots,M-1$ and $s_2=M-s_1-1$, and

$$\begin{aligned}&V(q_1,q_2;s_1,s_2)\\&=\frac{s_1\mu_1V(q_1,q_2;s_1-1,s_2)+s_2\mu_2V(q_1,q_2;s_1,s_2-1)}{s_1\mu_1+s_2\mu_2+q_1r_1+q_2r_2}\\&\quad+\frac{q_1r_1V(q_1-1,q_2;s_1,s_2)+q_2r_2V(q_1,q_2-1;s_1,s_2)}{s_1\mu_1+s_2\mu_2+q_1r_1+q_2r_2},\end{aligned}$$

where $s_1=0,1,\ldots,M$ and $s_2=M-s_1$. Here, we let $V(q_1,q_2;s_1,s_2)=0$ if $q_1=q_2=0$ or $\min\{q_1,q_2,s_1,s_2\}<0$.

Proof of Proposition 6. Consider a decision epoch such that jobs from types 1 and 2 are available for service.

(i) Suppose that $\mu_1 \leqslant \mu_2$, $r_1 \leqslant r_2$, and $\alpha_1 \leqslant \alpha_2$. By Corollary 2, type 2 receives the highest priority under the optimal policy. We next show that the four heuristics under consideration all prioritize type 2 jobs.

1. 2-step policy: Let

$$\Lambda(q_1, q_2) = \alpha_2 + \frac{\mu_2 \max_2(q_1, q_2)}{\mu_2 - r_2 + q_1 r_1 + q_2 r_2} - \alpha_1 - \frac{\mu_1 \max_1(q_1, q_2)}{\mu_1 - r_1 + q_1 r_1 + q_2 r_2},$$

where $\max_1(q_1, q_2) := \max\{\mathbb{1}_{\{q_1 \geqslant 2\}}\alpha_1, \alpha_2\}$ and $\max_2(q_1, q_2) := \max\{\alpha_1, \mathbb{1}_{\{q_2 \geqslant 2\}}\alpha_2\}$. We next show that $\Lambda(q_1, q_2) \geqslant 0$, which implies that type 2 jobs are preferred over type 1 jobs under the 2-step policy at every service completion instant when $q_1, q_2 \geqslant 1$; see Equation (5). We have

$$\begin{aligned}
&\Lambda(q_1, q_2)\\
&\geqslant \alpha_2 + \frac{\mu_2\alpha_1}{\mu_2 - r_2 + q_1 r_1 + q_2 r_2} - \alpha_1 - \frac{\mu_1 \alpha_2}{\mu_1 - r_1 + q_1 r_1 + q_2 r_2}\\
&= \frac{(\alpha_2 - \alpha_1)((q_1 - 1)r_1 + q_2 r_2)(q_1 r_1 + (q_2 - 1)r_2)}{(\mu_2 - r_2 + q_1 r_1 + q_2 r_2)(\mu_1 - r_1 + q_1 r_1 + q_2 r_2)}\\
&\quad + \frac{(\alpha_2\mu_2 - \alpha_1\mu_1)((q_1 - 1)r_1 + (q_2 - 1)r_2) + \alpha_2 r_2 \mu_2 - \alpha_1 r_1 \mu_1}{(\mu_2 - r_2 + q_1 r_1 + q_2 r_2)(\mu_1 - r_1 + q_1 r_1 + q_2 r_2)}\\
&\geqslant 0, \qquad (20)
\end{aligned}$$

where the first inequality follows from the condition that $\alpha_1 \leqslant \alpha_2$, which implies that $\max_1(q_1, q_2) = \alpha_2$ and $\max_2(q_1, q_2) \geqslant \alpha_1$, and the second inequality follows from the conditions that $\alpha_1 \leqslant \alpha_2$, $\alpha_1\mu_1 \leqslant \alpha_2\mu_2$, and $\alpha_1 r_1 \mu_1 \leqslant \alpha_2 r_2 \mu_2$.

2. Myopic policy: We have

$$\frac{\alpha_2 r_2}{\mu_1 + r_2} - \frac{\alpha_1 r_1}{\mu_2 + r_1} = \frac{\alpha_2 r_2 \mu_2 - \alpha_1 r_1 \mu_1 + (\alpha_2 - \alpha_1) r_1 r_2}{(\mu_1 + r_2)(\mu_2 + r_1)} \geqslant 0,$$

because $\alpha_1 r_1 \mu_1 \leqslant \alpha_2 r_2 \mu_2$ and $\alpha_1 \leqslant \alpha_2$, which implies that the myopic policy prefers type 2 jobs over type 1 jobs.

3. The $\alpha r\mu$-rule and the TCF rule will prefer type 2 jobs over type 1 jobs because $\alpha_1 r_1 \mu_1 \leqslant \alpha_2 r_2 \mu_2$ and $r_1 \leqslant r_2$, respectively.

(ii) Suppose that $\mu_1 \leqslant \mu_2 \leqslant r_2$, $\alpha_1 \leqslant \alpha_2$, and $\alpha_1 r_1 \mu_1 \leqslant \alpha_2 r_2 \mu_2$. By Corollary 3, type 2 receives the highest priority under the optimal policy. Clearly, the TCF rule is not consistent with the optimal policy as it gives priority to the jobs with the largest abandonment rate. On the other hand, the $\alpha r\mu$-rule is consistent because it prioritizes the jobs with the largest value of $\alpha_i r_i \mu_i$. We next show that 2-step, threshold, and myopic policies give priority to type 2 jobs.

1. 2-step policy: The proof is the same as the one for part (i).
2. Threshold policy: We have

$$\begin{aligned}
\mathcal{T} &= \max\left\{\frac{\mu_2(\alpha_1 r_1 - \alpha_2 r_2)}{r_1[r_2(\alpha_2 - \alpha_1) + \alpha_1(\mu_2 - \mu_1)]}, \frac{\mu_1(\alpha_1 r_1 - \alpha_2 r_2)}{r_2[r_1(\alpha_2 - \alpha_1) + \alpha_2(\mu_2 - \mu_1)]}\right\}\\
&= \max\left\{1 - \frac{\alpha_2 r_2 \mu_2 - \alpha_1 r_1 \mu_1 + r_1 r_2(\alpha_2 - \alpha_1)}{r_1[r_2(\alpha_2 - \alpha_1) + \alpha_1(\mu_2 - \mu_1)]}, 1 - \frac{\alpha_2 r_2 \mu_2 - \alpha_1 r_1 \mu_1 + r_1 r_2(\alpha_2 - \alpha_1)}{r_2[r_1(\alpha_2 - \alpha_1) + \alpha_2(\mu_2 - \mu_1)]}\right\}\\
&\leqslant 1,
\end{aligned}$$

because $\alpha_1 r_1 \mu_1 \leqslant \alpha_2 r_2 \mu_2$, $\alpha_1 \leqslant \alpha_2$, and $\mu_1 \leqslant \mu_2$, and hence type 2 jobs are preferred over type 1 jobs at all decision epochs.

3. Myopic policy: The same proof as in the proof of part (i) also applies here.

(iii) Suppose that the conditions of Proposition 4 hold, and hence it is optimal to give priority to type 2 jobs. Under these conditions, the $\alpha r\mu$-rule [TCF rule] is not necessarily consistent with Proposition 4 because it is possible to have $\alpha_1 r_1 \geqslant \alpha_2 r_2$ [$r_1 \geqslant r_2$], in which case it gives priority to type 1 jobs. On the other hand, the myopic policy agrees with the optimal policy under the conditions of Proposition 4 by definition. We next show that the remaining two heuristics are consistent with Proposition 4 as serving a type 2 job instead of a type 1 job is the preferred action under these heuristics.

1. 2-step policy: The proof is similar to the proof for part (i). Note that (20) still holds because $\alpha_1 \leqslant \alpha_2$. Now using the conditions that $\mu_1 = \mu_2 = \mu$ and $\alpha_1 \leqslant \alpha_2$, we have

$$\begin{aligned}
&\Lambda(q_1, q_2)\\
&\geqslant \frac{(\alpha_2 - \alpha_1)((q_1 - 1)r_1 + q_2 r_2)(q_1 r_1 + (q_2 - 1)r_2) + \mu(\alpha_2 r_2 - \alpha_1 r_1)}{(\mu - r_2 + q_1 r_1 + q_2 r_2)(\mu - r_1 + q_1 r_1 + q_2 r_2)}.
\end{aligned}$$

Using this together with the fact that $\mu(\alpha_2 r_2 - \alpha_1 r_1) \geqslant (\alpha_1 - \alpha_2) r_1 r_2$, which follows from Condition (4), yields

$$\begin{aligned}
&\Lambda(q_1, q_2)\\
&\geqslant \frac{(\alpha_2 - \alpha_1)[((q_1 - 1)r_1 + q_2 r_2)(q_1 r_1 + (q_2 - 1)r_2) - r_1 r_2]}{(\mu - r_2 + q_1 r_1 + q_2 r_2)(\mu - r_1 + q_1 r_1 + q_2 r_2)},
\end{aligned}$$

which is nonnegative because $\alpha_1 \leqslant \alpha_2$ and $q_1, q_2 \geqslant 1$.

2. Threshold policy: Under the conditions of Proposition 4, we can show that

$$\mathcal{T} = \frac{\mu(\alpha_1 r_1 - \alpha_2 r_2)}{(\alpha_2 - \alpha_1) r_1 r_2} \leqslant 1.$$

Hence, type 2 jobs are preferred over type 1 jobs at all decision epochs.

Acknowledgments

The authors thank the associate editor and two anonymous referees for their comments that significantly improved the article. This work was supported by the National Science Foundation [Grants CMMI–0620736, CMMI–0715020, and CMMI–0927607].

References

Argon N, Winslow JE, Ziya S (2011) Triage in the aftermath of mass-casualty incidents. Cochran JJ, Cox LA, Keskinocak P, Kharoufeh JP, Smith JC, eds. *Wiley Encyclopedia of Operations Research and Management Science*, Vol. 8 (John Wiley & Sons, Hoboken, NJ), 5611–5620.

Argon N, Ziya S, Righter R (2008) Scheduling impatient jobs in a clearing system with insights on patient triage in mass casualty incidents. *Probab. Engrg. Informational Sci.* 22(3):301–332.

Bernard M, Mathews PR (2008) Evacuation of a maternal–newborn area during Hurricane Katrina. *MCN: Amer. J. Maternal/Child Nursing* 33(4):213–223.

Boxma OJ, Forst FG (1986) Minimizing the expected weighted number of tardy jobs in stochastic flow shops. *Oper. Res. Lett.* 5(3):119–126.

Childers AK, Visagamurthy G, Taaffe K (2009) Prioritizing patients for evacuation from a healthcare facility. *Transportation Res. Record: J. Transportation Res. Board* (2137):38–45.

Cone DC, MacMillan DS (2005) Mass-casualty triage systems: A hint of science. *Acad. Emergency Medicine* 12(8):739–741.

Darr K (2006) Katrina: Lessons from the aftermath. *Hospital Topics* 84(2):30–33.

Emmons H, Pinedo M (1990) Scheduling stochastic jobs with due dates on parallel machines. *Eur. J. Oper. Res.* 47(1):49–55.

Fink S (2009) The deadly choices at Memorial. *New York Times Magazine* (August 30):28–46.

Frykberg ER (2002) Medical management of disasters and mass casualties from terrorist bombings: How can we cope? *J. Trauma* 53(2):201–212.

Glazebrook KD, Ansell PS, Dunn RT, Lumley RR (2004) On the optimal allocation of service to impatient tasks. *J. Appl. Probab.* 41(1):51–72.

Holt GR (2008) Making difficult ethical decisions in patient care during natural disasters and other mass casualty events. *Otolaryngology—Head and Neck Surgery* 139(2):181–186.

Hougaard P (2000) *Analysis of Multivariate Survival Data* (Springer-Verlag, New York).

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 and Disaster Medicine* 23(1):3–8.

Larkin GL, Arnold J (2003) Ethical considerations in emergency planning, preparedness, and response to acts of terrorism. *Prehospital and Disaster Medicine* 18(3):170–178.

Lerner EB, Schwartz RB, Coule PL, Weinstein ES, Cone DC, Hunt RC, Sasser SM, et al. (2008) Mass casualty triage: An evaluation of the data and development of a proposed national guideline. *Disaster Medicine and Public Health Preparedness* 2(1):S25–S34.

Li D, Glazebrook KD (2010) An approximate dynamic programming approach to the development of heurtistics for the scheduling of impatient jobs in a clearing system. *Naval Res. Logist.* 57(3):225–236.

Nocera A, Garner A (1999) An Australian mass casualty incident triage system for the future based upon triage mistakes of the past: The homebush triage standard. *Australian and New Zealand J. Surgery* 69(8):603–608.

Okie S (2008) Dr. Pou and the hurricane—Implications for patient care during disasters. *New England J. Medicine* 358(1):1–5.

Righter R (1994) Scheduling. Shaked M, Shanthikumar JG, eds. *Stochastic Orders and Their Applications*, Chap. 13 (Academic Press, New York), 381–432.

Sacco WJ, Navin M, Fiedler KE, Waddell RK II, Long WB, Buckman RF Jr (2005) Precise formulation and evidence-based application of resource-constrained triage. *Acad. Emergency Medicine* 12(8):759–770.

Sacco WJ, Navin M, Waddell RK II, Fiedler KE, Long WB, Buckman RF Jr (2007) A new resource-constrained triage method applied to victims of penetrating injury. *J. Trauma* 63(2):316–325.

Shaked M, Shanthikumar JG (2007) *Stochastic Orders* (Springer, New York).

Evin Uzun Jacobson is a postdoctoral research associate in the Healthcare Management Group at Imperial College Business School. Her primary research interests are stochastic modeling and simulation, design and control of queueing systems, and statistical analysis of large-scale data sets. Her current application areas of interest include healthcare problems, especially time-critical decision making.

Nilay Tanık Argon is an assistant professor in the Department of Statistics and Operations Research at the University of North Carolina at Chapel Hill. Her research primarily focuses on the analysis, design, and control of stochastic systems with applications to healthcare and manufacturing. Her research activities include the development and analysis of simulation output analysis methods.

Serhan Ziya is an associate professor in the Department of Statistics and Operations Research at the University of North Carolina at Chapel Hill. His main research interests are management and control of service operations, particularly those from healthcare, revenue management, and pricing under capacity and inventory considerations.