European Journal of Operational Research 255 (2016) 531–544

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

Optimizing emergency preparedness and resource utilization in mass-casualty incidents

Panagiotis P. Repoussis[a,b,*], Dimitris C. Paraskevopoulos[c], Alkiviadis Vazacopoulos[b,d], Nathaniel Hupert[e]

[a] Department of Marketing & Communication, School of Business, Athens University of Economics & Business, Athens 11362, Greece
[b] School of Business, Stevens Institute of Technology, Hoboken, NJ 07030, USA
[c] School of Management, University of Bath, Claverton Down, Bath BA2 7AY, UK
[d] Optimization Direct Inc, Harrington Park, NJ 07640, USA
[e] Departments of Healthcare Policy and Research and of Medicine, Weill Cornell College of Cornell University, New York, NY 10065, USA

ARTICLE INFO

Article history:
Received 26 August 2015
Accepted 27 May 2016
Available online 2 June 2016

Keywords:
Emergency medical services
Mass-casualty incident
Scheduling
Ambulance dispatching
Local search

ABSTRACT

This paper presents a response model for the aftermath of a Mass-Casualty Incident (MCI) that can be used to provide operational guidance for regional emergency planning as well as to evaluate strategic preparedness plans. A mixed integer programming (MIP) formulation is proposed for the combined ambulance dispatching, patient-to-hospital assignment, and treatment ordering problem. The goal is to allocate effectively the limited resources during the response so as to improve patient outcomes, while the objectives are to minimize the overall response time and the total flow time required to treat all patients, in a hierarchical fashion. The model is solved via exact and MIP-based heuristic solution methods. The applicability of the model and the performance of the new methods are challenged on realistic MCI scenarios. We consider the hypothetical case of a terror attack at the New York Stock Exchange in Lower Manhattan with up to 150 trauma patients. We quantify the impact of capacity-based bottlenecks for both ambulances and available hospital beds. We also explore the trade-off between accessing remote hospitals for demand smoothing versus reduced ambulance transportation times.

1. Introduction

Any medical incident in which casualties actually or potentially overwhelm local emergency response and hospital treatment capability may be termed a mass-casualty incident (MCI). These are typically major events, such as transportation accidents and terrorist bombings, with many casualties, though many jurisdictions define an MCI using a relatively small numerical threshold (e.g., 5 casualties from one incident in New York City while in South Korea 6 such casualties) (Arnold, Halpern, Tsai, & Smithline, 2004; Park, Shin, Song, Hong, & Kim, 2016). MCIs may overwhelm local treatment capability either due to sheer numbers of injured patients all needing treatment at the same time, or a potentially smaller number of patients who require advanced care (e.g., neurosurgical care) that is in relatively short supply locally. The International Institute for Counter-Terrorism has recorded over 33,000 terrorist incidents in the world since 1975, while lately the potential for terrorist activity is on the rise. Furthermore, the increasing frequency and severity of megastorms, such as Hurricane Sandy, has made natural-origin MCIs more likely. This paper is concerned with the development of a response model for the aftermath of an MCI that can be used to optimize resource utilization, to provide operational guidance for regional emergency planning, and to evaluate strategic preparedness plans.

As described by Mills, Argon, and Ziya (2014) an MCI creates a sudden spike in demand for the emergency response resources within an area, and as a result, even patients who are in critical condition may not have timely access to these resources that are essential for their survival. During an MCI, it sometimes happens that only a limited number of ambulances are available to transport patients, requiring ambulances to make multiple trips from the MCI site(s) to the hospitals or forcing reliance on self-transportation (der Heide, 2006). Dispatching software systems typically retrieve the locations and contact information for the hospitals nearest to the event, but this prioritizes the travel time over other factors required for optimal response, such as the availability

* Corresponding author at: School of Business, Stevens Institute of Technology, Hoboken, NJ, 07030, USA. Tel.: +1 2012163342.
E-mail addresses: prepousi@aueb.gr, prepouss@stevens.edu (P.P. Repoussis), d.paraskevopoulos@bath.ac.uk (D.C. Paraskevopoulos), alkis@optimizationdirect.com (A. Vazacopoulos), nah2005@med.cornell.edu (N. Hupert).

http://dx.doi.org/10.1016/j.ejor.2016.05.047
0377-2217/

of existing medical resources at these facilities. Emergency medical service (EMS) systems are thus challenged during MCI response to allocate effectively a set of limited resources to the patients awaiting treatment and transportation to hospitals.

From the operational perspective, we need to determine which of the available hospitals should be included in the response according to their category, trauma level, capacity and proximity to the MCI site(s), how many ambulances should be utilized, where ambulances should transport each subsequent patient, and how many patients should be transported to each hospital. The arrival times of patients at the hospitals, the hospitals' throughput capability, and the patient treatment times will dictate the treatment order (scheduling) of the transported patients at each hospital and the time required to treat all patients. During the ambulance dispatching and patient-to-hospital assignment processes, we also need to follow a triage protocol as well as to match the specific treatment needs of the patients. On the other hand, given the potential location and size of an MCI we need to measure the simulated response efficiency and to identify the bottlenecks.

Ambulance dispatching decisions affect both the patient waiting times at the site and at the hospitals as well as the availability of ambulances. In practice, one commonly employed dispatch strategy is called "scoop-and-run", whereby patients are sent as quickly as possible to the closest hospital in order to minimize the dispatching times (this is standard practice in Israel, for example). However, this strategy ignores the specific needs of the patient, the triage protocol, and the current available capacity at the hospitals. For example, sending a large number of patients to the closest hospital may cause congestion resulting in long waiting times and unnecessary re-dispatching of patients in the worst case. As described by Carter, Chaiken, and Ignall (1972) dispatching the closest idle ambulance to an emergency call is not always the optimal policy, if the objective is to minimize the response times.

Regardless the rich literature on emergency response problems, there remain no generally accepted, evidence-based guidelines to advise dispatchers on fundamental questions, such as which hospitals to include in a specific MCI response and how many casualties to transport to each. Ambulance dispatching has been performed mostly on the basis of the reliability and validity of EMS experts' cognitive abilities. It is possible, however, that dispatching strategies using situational awareness information combined with knowledge of regional hospital capabilities, including destination facility-specific transportation times and treatment capability, could yield superior outcomes. For example, it is reasonable to expect that by balancing the load on the hospitals the level of care will be improved and the delays experienced by the patients will be reduced (Repoussis, Paraskevopoulos, Vazacopoulos, & Hupert, 2015). Although divergent modeling approaches appear in the literature, it is likely that computerized models will be increasingly important in providing public accountability for the resource allocation decisions that have to be made in emergency situations.

The contribution of this paper is three-fold. First, we address the combined ambulance dispatching, patient-to-hospital assignment, and treatment ordering problem. In particular, we propose a rigorous mathematical formulation that captures all critical compatibility issues and prioritization aspects according to the Simple Triage and Rapid Treatment (START) triage protocol. Furthermore, we consider the makespan (i.e., the latest completion time) and the total flow time as hierarchical objectives. Second, we present a hybrid MIP-based construction heuristic and local search improvement algorithms that allow us to solve and find high quality solutions for otherwise computationally intractable large scale problem instances. Third, we study the effectiveness and efficiency of the new algorithms via a comprehensive study on randomly generated small- and medium-scale instances. We also try to identify how the availability of resources as well as the spatial and temporal characteristics affect the response times and the allocation of resources. Additionally, we demonstrate the applicability of the new model on an example MCI with realistic data. We consider the hypothetical case of a bombing at the New York Stock Exchange in Lower Manhattan. For a given number of ambulances we examine 3 scenarios, regarding the size of the MCI, with up to 150 patients. We lastly examine trade-offs between increasing the available capacity (e.g., adding hospital beds) of the hospitals nearby the site and including relatively distant hospitals instead.

The remainder of the paper is structured as follows. Section 2 briefly discusses the related work regarding models; Section 3 presents the mathematical model and discusses the combined ambulance dispatching, patient-to-hospital assignment, and treatment ordering problem; and Section 4 introduces the MIP-based construction heuristic and local search metaheuristic algorithms. Subsequently, Section 5 describes the computational experiments and reports results based on realistic data, and finally the paper concludes in Section 6.

2. Related work

In practice, ambulance dispatching decisions are made in a dynamic environment; however, it is difficult to design and apply real-time dispatching tools because information is dynamic and often incomplete. Given this state of affairs, various strategic decisions can be evaluated a priori, such as, given what is known about regional surge capacity, which hospital should be included in regional disaster preparedness planning. Computer- or exercise-based modeling is therefore becoming increasingly important in providing a test-bed for the resource allocation decisions that have to be made in emergency situations, without the overt risk of harm to current patients.

Emergency vehicle deployment problems have been widely studied in the literature and various models and solution frameworks have been developed. A large part of this literature focuses on reducing dispatching response times in standard emergency call processes, with an emphasis on how to allocate emergency service stations and units. Toregas, Swain, ReVelle, and Bergman (1971) proposed a location set covering model that minimized the number of ambulances required to cover all demand points with a preset coverage standard. The Hypercube Queueing model introduced by Larson (1974) was the first model to embed queueing theory in location problems. A survey on deterministic, stochastic and dynamic ambulance location and allocation models is provided by Brotcorne, Laporte, and Semet (2003). More recent works in dynamic and real-time models for emergency vehicle dispatching and coverage relocation are those of Gendreau, Laporte, and Semet (2006) and Haghani and Yang (2007). Interested readers may also refer to Boldberg (2004) for an overview on dispatching emergency service vehicles and to Bektas, Repoussis, and Tarantilis (2014) for a detailed survey regarding models and algorithms for dynamic and stochastic vehicle dispatching problems.

Similarly, various models and decision support systems have been proposed for resource management in disaster response. The vast majority focuses on the location and allocation of emergency response units (Fiedrich, Gehbauer, & Rickers, 2000) as well as on the supply and distribution of relief supplies (Barbarosoglu & Arda, 2004; Mete & Zabinsky, 2010). Few papers consider the transportation of casualties and flows of patients between locations (Wilson, Hawe, Coates, & Crouch, 2013; Salman & Gul, 2014). Notably, many models assume that the same vehicles are used to distribute emergency supplies and simultaneously to transport casualties to treatment facilities (Yi & Kumar, 2007; Yi & Ozdamar, 2007; Ozdamar, 2011). In these works the main effort is to determine the flows of commodities and casualties between supply and demand locations