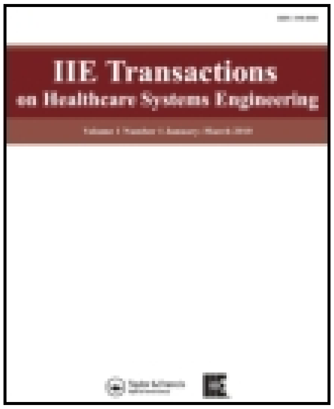

IIE Transactions on Healthcare Systems Engineering
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/uhse20

Operations research/management contributions to emergency department patient flow optimization: Review and research prospects

Soroush Saghafian[a], Garrett Austin[a] & Stephen J. Traub[b]

[a] Industrial Engineering, School of Computing, Informatics and Decision Systems Engineering, Arizona State University, Tempe, AZ, 85281, USA

[b] Department of Emergency Medicine, Mayo Clinic, Phoenix, AZ, 85054, USA

Published online: 09 Jun 2015.

To cite this article: Soroush Saghafian, Garrett Austin & Stephen J. Traub (2015) Operations research/management contributions to emergency department patient flow optimization: Review and research prospects, IIE Transactions on Healthcare Systems Engineering, 5:2, 101-123

To link to this article: http://dx.doi.org/10.1080/19488300.2015.1017676

IIE Transactions on Healthcare Systems Engineering (2015) **5**, 101–123

ISSN: 1948-8300 print / 1948-8319 online
DOI: 10.1080/19488300.2015.1017676

Operations research/management contributions to emergency department patient flow optimization: Review and research prospects

SOROUSH SAGHAFIAN[1,*], GARRETT AUSTIN[1] and STEPHEN J. TRAUB[2]

[1]*Industrial Engineering, School of Computing, Informatics and Decision Systems Engineering, Arizona State University, Tempe, AZ 85281, USA*
E-mail: soroush.saghafian@asu.edu
[2]*Department of Emergency Medicine, Mayo Clinic, Phoenix, AZ 85054, USA*

Received August 2014 and accepted February 2015

In recent years, Operations Research/Management (OR/OM) has had a significant impact on improving the performance of hospital Emergency Departments (EDs). This includes improving a wide range of processes involving patient flow from the initial call to the ED through disposition, discharge home, or admission to the hospital. We review approximately 350 related papers to (i) demonstrate the influence of OR/OM in EDs, and (ii) assist both researchers and practitioners with the OR/OM techniques already available to optimize ED patient flow. In addition, we elaborate on some practical challenges yet to be addressed. By shedding light on some less studied aspects that can have significant impacts on ED operations, we also discuss important possibilities for future OR/OM researchers.

Keywords: Operations research/management, emergency department, patient flow, operational efficiency, patient safety

1. Introduction

Growth in healthcare expenditures as a percentage of the United States Gross Domestic Product (GDP) has been explosive, outpacing even past estimates of exponential increases. In 2006, Hall *et al.* (2006) anticipated an increase in healthcare costs to 15.9% of the GDP by 2010. In fact, the World Health Organization reported that expenditures in 2010 were at 17.6% of GDP (WHO, 2010). This highlights an unquestioned need to improve the efficiency of healthcare delivery methods.

In the United States, Emergency Departments (EDs) are the gate to hospitals through which 50% of non-obstetrical admissions occur (Pitts *et al.*, 2008). Considering that admitted patients create about one-third of the U.S. healthcare bill each year (Abelson, 2013), improving ED operations may have a significant impact on U.S. healthcare expenditures. Indeed, while the direct aggregate spending on emergency care in the United States is estimated to be 5% to 10% of national health expenditures (Lee *et al.*, 2013), considering the fact that ED is the first point of contact for nearly half of all hospital admissions (Schuur and Venkatesh, 2012, and Pitts *et al.*, 2010), improving ED operations and related decisions can have broader impacts. Operations Research/Management (OR/OM) techniques seem able to contribute to such improvements.

To better understand the need for improving ED operations, we note that a considerable percentage of patients report experiencing a delay in the ED, with more than half of those citing long waiting times as a cause (Kennedy *et al.*, 2004). Long waiting times are partially caused by a mismatch between "supply" and "capacity": the annual number of ED visits increased from 90.3 million to 119.2 million visits between 1996 and 2006, while the number of hospital EDs has decreased from 4019 to 3833 (Pitts *et al.*, 2008). This increasing strain has placed EDs in a state of overcapacity approximately 50% of the time (Geer and Smith, 2004). In a 2009 report to the U.S. Senate, the U.S. Government Accountability Office (GAO) emphasized that crowding continued to occur in EDs, and some patients waited longer than recommended time frames (GAO, 2009). Zilm *et al.* (2010, p. 296) note that "The 'ripple' effect of high inpatient occupancies (particularly by day of week), and delays in discharges, has extended lengths of stays in the ED, frequently resulting in 'grid lock' with few ED treatment stations available to maintain patient flow."

*Corresponding author

These issues affect not only the timeliness of serving patients, but the ability to serve altogether: Burt and McCaig (2005) found that 44.9% of all U.S. EDs experienced a period of diversion over the course of each year with as many as 1886 ambulances diverted each day. These are only a few indicators of a tremendous need for focus on improving ED operations. Such improvements are essential for increasing profit (Falvo *et al.*, 2007; McConnell *et al.*, 2006), improving patient satisfaction (Thompson *et al.*, 1996; Boudreaux and O'Hea, 2004), and – most importantly – improving patient safety (Mayer, 1979; Trzeciak and Rivers, 2003).

On a larger scale, the problem is more significant than simply managing increasing volumes of patients; some overcrowding issues can be attributed to potentially modifiable use of the ED. A high proportion of patients incorrectly seek out the ED as their first source of care (Burnett and Grover, 1996). Inappropriate ambulance use also puts increasing strain on ED resources (Richards and Ferrall, 1999). Baer *et al.* (2001) reports that a notable percentage of ED patients are recently discharged ones: "frequent flyers," considered as patients with five or more visits per year, constitute 14% of total ED visits (Huang *et al.*, 2003). Boarding patients (patients who cannot be moved to inpatient units due to lack of inpatient bed availability) represent up to 22% of the total ED patient census (Schneider *et al.*, 2003).

Historically, a common method of dealing with the inability to serve patients is to "close the doors" (e.g., through ambulance diversion) and focus on patients already in the system. Beyond the negative impact of diverting patients on overall care, however, there are financial considerations: given that 84% of a hospital's costs are fixed (Roberts *et al.*, 1999), there is little financial incentive to be on diversion. Moreover, some hospitals do not use mechanisms such as ambulance diversion, and in some states diversion is not legal. As a practical matter, barring an internal disaster no ED can truly close completely, as the Emergency Medical Treatment and Labor Act (EMTALA) passed in 1986 requires EDs to serve all patients who present to the facility, regardless of their insurance or financial status. Thus, EDs must focus on improving their patient flow process as a main mechanism for combating the aforementioned issues.

Welch *et al.* (2006) and Welch *et al.* (2011) list various metrics by which ED operations can be measured. Among the widely used metrics are LOS (length of stay), LWBS (% of patients who leave without being seen), door to diagnostic evaluation by a qualified medical professional (arrival time to provider contact time, also known as "door-to-doc" time) and ambulance diversion (amount of time ambulances are diverted away from the ED). Olshaker and Rathlev (2006, p. 354) provide a valuable, cautionary perspective regarding LOS and diversion, though the message holds true with all metrics: "[No] measure is universally applicable as a marker of overcrowding and should be used with caution when comparing performance between institutions. Diversion is not an option in some EMS systems and throughput time is ED specific and dependent on the complexity of the case mix. In spite of this, the measures have specific value in tracking individual institutional performance over time."

It is worth noting that ED operations metrics are also intertwined with hospital quality measures. For instance, the Centers for Medicare & Medicaid Services (CMS) offer an Electronic Health Record (EHR) incentive program for hospitals maintaining quality measures, and the proposed clinical quality measures for the 2014 CMS EHR Incentive Program have a strong focus on boarding time and ED LOS (CMS, 2013). We believe it is important to note, however, that there are challenges to OR/OM in the ED when it comes to assessing tradeoffs between operational improvement and quality. While ED operations have well-defined metrics, ED quality – particularly as it pertains to counting "defects" – does not. Most of the quality events brought to the attention of ED managers are rare, significant adverse events identified through physician self-report, or notification from other services. This ad-hoc approach, while useful for identifying major issues, is insufficiently sensitive to capture small changes in quality that may result from day-to-day changes in ED operations.

Nevertheless, OR/OM techniques have significantly helped various parts of hospitals (and especially the ED) to improve their performance gauging metrics (see, e.g., Ozcan, 2009; Hopp and Lovejoy, 2013; and Green, 2012). However, to the best of our knowledge, there is no review that comprehensively describes and unifies the OR/OM contributions in EDs. To help researchers and practitioners involved in improving ED operations, we first provide such a review. We then provide some important research prospects by shedding light on some fruitful research directions for future research. We also discuss some practical challenges yet to be addressed. These require a focus on the complex details of OR/OM techniques used to model ED patient flow. With an understanding of the depth of OR/OM tools used in the ED, some of the managerial challenges in the ED can be addressed. However, we argue that although common techniques such as Mathematical Programming, Queueing Theory, Simulation Analysis, Markov models and Game Theory have already addressed many challenges, we must also look beyond traditional OR/OM methods; there are still various unanswered questions in the ED. This lends itself to our perspective regarding possibilities for future researchers on valuable but less studied aspects that can have significant impacts on the practice of ED operations.

In closing this section, we note that while the use of OR/OM to solve fundamental social problems evolved mainly from World War II (Green and Kolesar, 2004), it took more than two decades for OR/OM to be used in ED operations. An early related example of using OR/OM tools is the work of Savas (1969), where response time and round-trip time were the performance metrics studied in a computer simulation of ambulance quantity and location. In Goldman *et al.* (1968), a computer simulation is used to reallocate hospital beds and develop a modified usage policy. Time and motion studies were also among pioneers to identify issues with ED patient flow (Heckerling, 1985; Saunders, 1987). More modern techniques can be found in studies such as Green (2006), Green *et al.* (2006), Armony *et al.* (2011), Saghafian *et al.* (2012, 2014), Huang (2013), Huang *et al.* (2013), and many others that we will review. In reviewing such studies, we hope to provide a resource for both researchers and practitioners to familiarize them with past related, valuable contributions, and to invoke to plan future studies that can help EDs reach new levels of both operational efficiency and patient safety.

The remainder of the paper is organized as follows. Section 2 discusses three components of ED patient flow: flow into, within, and out of the ED. Section 3 describes various OR/OM tools for modeling and optimizing ED patient flow. Section 4 concludes and presents some important research prospects by discussing some less studied aspects of ED patient flow with high potential impact on ED operations.

2. Three components of ED patient flow: Flow into, within, and out of ED

Patient flow can be viewed from two perspectives: operational and clinical (see, e.g., Côté, 2000; Côté and Stein, 2000; and Marshall *et al.*, 2005). The former perspective refers to the physical movement of patients through a set of locations, and the latter refers to the progression of their health status. In this paper, we use the term patient flow to mean the physical movement of patients.

In what follows, we focus on patient flow into, in, and out of the ED, and discuss OR/OM contributions to each of these categories separately.

2.1. *Patient flow into the ED*

We start by considering contributions that allow altering the arrival process to the ED. Early work in this area includes studies that challenged traditional ambulance dispatch practices (e.g., Carter *et al.*, 1972). Since then, innovations on altering arrival to the ED have been ample. We categorize the related studies in this vein into subsections below.

2.1.1. *Ambulance deployment and location*

Ambulance response time is a key metric used to evaluate prehospital emergency medical services (Peleg and Pliskin, 2004). This, coupled with the understanding that shaving minutes off of response time has a great potential to save lives (Mayer, 1979; BBC, 2002), has led to a focus on optimizing this first step in the patient flow process. Policies regarding ambulance deployment and location have become commonplace. The EMS Act of 1973 mandates that 95% of requests be served within 10 minutes in rural areas, and 30 minutes in urban areas (Ball and Lin, 1993). Such stipulations on service times are not exclusive to the United States (Gendreau *et al.*, 2001; Galvão *et al.*, 2005). A survey of more than 3000 calls in Ireland showed that only 81% of calls had a response within 15 minutes (Breen *et al.*, 2000), and England and Australia struggle with response benchmarks as well (Kelly *et al.*, 2002; Stoykova *et al.*, 2004). Lowthian *et al.* (2011) report a global increase in the number of ambulance arrivals to EDs, and a 7.0–12.5% annual increase in ambulance response times, further exacerbating the problem.

Ambulance deployment and location represents some of the earliest OR/OM work not only in emergency response services, but in healthcare. Savas (1969) and Swoveland *et al.* (1973) present some of the first research on the subject of ambulance location, while Fitzsimmons (1973) tackles early work on ambulance deployment. Brandeau and Chiu (1989) summarize the original contributions to the field of ambulance location. Lee (2011) strategically organizes EMS decision-making into three categories: ambulance location, ambulance relocation and ambulance dispatching. We acknowledge contributions in these realms below, since all of them affect patient flow into the ED.

Brotcorne *et al.* (2003) and Goldberg (2004) provide a detailed review of literature on ambulance location with additional focus on relocation and dispatching. The review of Goldberg (2004) provides an in-depth summary of the Hypercube queueing model, introduced by Larson (1974), where a queueing network is characterized by a class of Markov models. Predetermined values of decision variables can be altered to reach desired performance levels for ambulance location. The robustness of the Hypercube model is evident with immediate extensions to make it less computationally rigorous (Larson, 1975), additions to address "ties" between potentially dispatched ambulances (Burwell *et al.*, 1993), through present day application where it has been used to analyze ambulance decentralization (Takeda *et al.*, 2007). Ingolfsson *et al.* (2008) analyze ambulance location optimization with a more in-depth definition of response time, incorporating often-neglected delays such as the duration of the phone call, time spent dispatching ambulances, time to contact paramedics and more. Peleg and Pliskin (2004) use a geographic decision support tool for ambulance location, improving operations as evidenced by an additional 60% of calls meeting the established 8-minute threshold. A similar effort is seen in Peters and Hall (1999). Singer and Donoso (2008) model ambulance deployment with Queueing Theory, allowing for an in-depth analysis of the trade-offs of fleet size and cycle time. Repede and Bernardo (1994) utilize a time variation coverage location approach for a 13% increase in coverage and 36% decrease in response time, without the addition of resources.

Rajagopalan *et al.* (2008) generate search algorithms that allow for exponentially quicker decision making on ambulance relocation. Gendrau *et al.* (2006) focus on relocation as a dynamic model, which relocates ambulances as they are being dispatched. Gendreau *et al.* (2001) also use a dynamic model through parallel computing to perform relocation. They suggest that the downside, the potential need to relocate the fleet at every event, is outweighed by the benefit of maximized coverage.

Andersson and Värbrand (2007) provide dynamic ambulance dispatch and relocation analysis for the most complex ambulance control situation in Sweden. This is based on a quantification of "preparedness," an effort that is continued with great detail by Lee (2011). Deo and Gurvich (2011) utilize a queueing network model to test five different decision making strategies, two of which employ the knowledge of a centralized "social planner" to aid in real-time ambulance routing decisions.

2.1.2. *Ambulance diversion*

Ambulance diversion (AD) is a technique utilized to reduce ED arrival rates by diverting incoming ambulances to other nearby hospitals (Deo and Gurvich, 2011). First reported at a New York City hospital in the 1990s, it was initially established as a flow management technique, rarely used to relieve the strain on overburdened EDs (Handel, 2011; Asplin, 2003). However, over the next two decades, U.S. hospitals would see nearly 30% more patients per year with a 12% decline in the number of EDs. Not coincidentally, diversion increased: this increase in volume resulted, for instance, in an average per ED increase in diversion hours from 57 to 190 per month in Los Angeles County (Sun *et al.*, 2006). The effect: 91% of ED directors in the United States reported overcrowding as a recurring day-to-day issue, making AD a prominent fixture in hospitals (Olshaker and Rathlev, 2006).

Despite the negative public perception surrounding the technique, AD is designed to improve patient safety and network performance. However, recent anecdotal and empirical evidence suggests that EDs across the nation are not seeing any significant improvements in their wait times while on diversion (Deo and Gurvich, 2011; Mihal and Moilanen, 2005; Kowalczyk, 2008). This discrepancy between ideal AD outcomes and actual observations has increased to the point where some EDs elect to forego diversion process altogether (Kowalczyk, 2008).

Nonetheless, large volumes of research are underway, trying to determine how to best implement AD. The American College of Emergency Physicians (ACEP) has developed guidelines by which hospitals declare, and exist in, the state of diversion in order to minimize its negative effects (Brennan *et al.*, 2000). We focus in the remainder of this section on research geared at understanding and reducing ambulance diversion to better accommodate patient arrivals.

An extensive diversion program implemented in Patel *et al.* (2006) attempts to adhere to AD guidelines to reduce diversion as a whole. The intervention resulted in a reduction in diversion hours to 7143, a 74% decrease (Patel *et al.*, 2006). Vilke *et al.* (2004) studies two EDs as they commit to eliminating AD entirely. Adding staff to the ED coupled with a strong focus on AD during a test period reduced diversion hours from 19.7 to 1.4 for one hospital and 27.7 to 0 for another. Pham *et al.* (2006) identifies at least four other successful studies that primarily enforce policies to reduce diversion.

The addition of resources is common when tackling diversion. California saw a 45% statewide decrease in AD from 2003 to 2007 (Borders *et al.*, 2009). The best practices that lead to successful diversion reduction were consistent with other findings, particularly the use of a "bed czar," where an individual is responsible for evaluating strategies focused on reducing diversion hours (Geer and Smith, 2004). Ingalls Hospital, a large urban Chicago hospital, saw its diversion hours decrease 79% in one year, largely through the introduction of an admission and discharge room (Geer and Smith, 2004).

Although the majority of evidence is positive with respect to adding resources to decrease diversion (Warden *et al.*, 2003; McConnell *et al.*, 2005; Burt and McCaig, 2006), some negative results on diversion (after the addition of beds (Han *et al.*, 2007) and staffing (Schull *et al.*, 2003) suggests that ED expansion alone is insufficient to decrease diversion, unless such an expansion is done along with removing other hospital bottlenecks. Interestingly, a significant number of diversion efforts focus on expanding ICU resources, despite evidence that the ED is the bottleneck in more than 80% of AD cases (Allon *et al.*, 2013). In addition to expansion-based approaches, some studies analyze other diversion policies, with a focus on minimizing patient waiting (Ramirez-Nafarrate *et al.*, 2012) and using a network perspective whereby AD policies are centralized (Ramirez-Nafarrate *et al.*, 2011).

2.1.3. *Ambulance alternatives*

We turn briefly to alternate roles that an ambulance can play in patient flow into the ED. Namely, we identify work on the use of EMS to perform triage, identify nonurgent patients and redirect them. With a fair amount of non-urgent patients making what might be considered "discretionary" ED visits (Sempere-Selva *et al.*, 2001; Northington *et al.*, 2005), this presents an alternative to utilizing what may be scarce ED resources.

Price *et al.* (2005) propose the use of selective diversion, whereby EMS quickly predicts patient disposition to divert critical care patients to hospitals with available resources. Two Washington-based EMS agencies offered alternate care to patients diagnosed as non-urgent (Schaefer *et al.*, 2002). The intervention reduced the percentage of patients who received ED care by 7.2%, presenting an effective mechanism to attack ED overcrowding.

Snooks *et al.* (2004), however, express skepticism towards the use of ambulance alternatives into the ED, citing a lack of established literature in non-conveyance and the ability of ambulance crews to triage with accuracy. Indeed, a scarcity of research on patients not being transported to the ED is unsurprising. EMS-initiated refusal of transport is present in only 7.0% of the largest cities (Knapp *et al.*, 2009), down from 17.0% a decade earlier (Jaslow *et al.*, 1998). We identify a common theme of a lack of reliability on EMS triage in the literature (Schaefer *et al.*, 2000; Pointer *et al.*, 2001; Hauswald, 2002; Levine *et al.*, 2006). In closing this section, we also mention that, given the high percentage of fixed cost in a hospital's total expenses, these ambulance alternatives would purely be to relieve congestion without regard to revenue (Roberts *et al.*, 1999). As a result, more research establishing ambulance alternatives is required before such a practice can be considered as an operational alternative.

2.2. *Patient flow within the ED*

In this section, we address the literature that aims at optimizing patient flow inside the ED. Miró *et al.* (2003) note that patient flow within the ED has a large effect on overcrowding, in conjunction with internal factors and external pressures. Purnell (1991) provides an early review of triage and fast track systems in the ED, touching on many important topics like patient classification and the skills of triage personnel. Wiler *et al.* (2011) review modeling approaches from a technical perspective, focusing on research on patient flow and crowding. Wiler *et al.* (2010) provide a less technical review of operations improvements in the front end of the ED – registration, triage, and fast track. Oredsson *et al.* (2011) provides a review of triage-related flow improvements, primarily recognizing fast tracks, streaming and triage as interventions.

2.2.1. *General triage interventions*

The first concepts of triage began in World War I (Keen, 1917), with widespread research and publication on triage in circulation for the greater part of the past two decades. The concept of evaluating and prioritizing patients in the ED is not a new one (Meislin *et al.*, 1988; Wright *et al.*, 1992). Iserson and Moskop (2007) note three conditions that must be satisfied to constitute triage, which we simplify:

1. At least a modest scarcity of health care resources exists.
2. A health care worker assesses each patient's medical needs, usually based on a brief examination.
3. The triage officer uses an established system or plan, usually based on an algorithm or a set of criteria, to determine specific treatment or treatment priority for each patient (pp. 275–276).

Work on triage has been remarkably extensive, making a comprehensive review well outside the scope of this work.

Triage in most U.S. EDs is typically managed through Emergency Severity Index (ESI); it is, however, not universal. For instance, the Australasian Triage Scale (ATS), the Manchester Triage Scale (MTS), and the Canadian Triage Acuity Scale (CTAS) are used in other countries (Beveridge, 1998; Richardson, 1998; Saghafian *et al.*, 2014). Despite the differences, the relative successes of these triage scales has led to a move away from the once popular three-level scale in the United States. Today, the five-level ESI system – proposed by Wuerz *et al.* (2000), which combines urgency with an estimate of resource requirements – has become the most common algorithm (Fernandes *et al.*, 2005; Chonde *et al.*, 2013). A five-level scale is better than a three-level scale according to Wang (2004), where a queueing system that models high risk patients concludes that patients should be split into as many classes as possible. FitzGerald *et al.* (2010) revisit past efforts and evaluates the future direction of triage, concluding that the five-level scale is now firmly established.

Reviews studying the effects of triage have been as extensive as the implementation of triage has been global. Outside of the aforementioned countries, we see triage employed at 97% of EDs in Switzerland in 2010 (Farrokhnia and Goransson, 2011), an innovative triage in South Africa (Gottschalk *et al.*, 2006), and modifications to established systems in Portugal (Martins *et al.*, 2009) and the Netherlands (van der Wulp *et al.*, 2008). Görannson *et al.* (2005) provides a review of the use of triage in a fair majority of Swedish EDs. Fernandes *et al.* (1999) conducts a study that finds ED triage reliable when employed under the proper circumstances, while providing several sources that suggest mixed results with triage. Harding *et al.* (2011) provides a non-technical review to study the overall effect of triage on patient flow. Results from different studies were also mixed, suggesting a triage system tailored to the patient mix may be necessary.

Although traditional triage uses a nurse to evaluate patients at triage, research has found the investment of a physician at triage to have a benefit to a combination of common performance metrics such as LOS, LWBS, and diversion levels (Partovi *et al.*, 2001; Han *et al.*, 2010; Russ *et al.*, 2010). From an OR/OM perspective, the main trade-off is between (i) using the physician (an expensive resource) at triage who might be better used to treat patients in the rooms, (ii) gaining more accurate information upfront, and (iii) issuing discharge or appropriate tests early on. Traub *et al.* (2014a) performed a mechanistic analysis of the effects of a physician in triage finding (in a single facility study) that the overall reduction in LOS was a function of rapid discharge of low-acuity patients much more so than of placing orders for patients who were ultimately seen in the main ED by another physician. Triage has been found to help reduce LWBS rates and patient LOS, even in the midst of increased patient census (Chan *et al.*, 2005; Sanchez *et al.*, 2006; Ruohonen *et al.*, 2006). One concern with this approach is that the benefit of decreased waiting times might be at the expense of quality, as significant inconsistencies in patient classification may be a byproduct of triage (Wuerz *et al.*, 1998). Batt and Terwiesch (2012) study the phenomena of state-dependent service times as seen in human-paced service systems, transportation and telecommunications; they find that the use of triage reduces service time in periods of crowding in the ED, although some thought must be given to the financial cost of triage.

A large number of Emergency Medicine studies note the high percentage of non-urgent patients who utilize the ED (see, e.g., Lowe and Bindman, 1997; Koziol-McLain *et al.*, 2000; and Carret *et al.*, 2009), and efforts have been undertaken to address this. Indeed, Derlet *et al.* (1995) assessed and referred more than 30,000 patients out of the ED, finding that nonemergency patients can be triaged out of the ED to relieve stress. Other work has been done validating the relocation of patients to this end (Washington *et al.*, 2000; Derlet and Richards, 2008). At least one study cautions against this, however, as the lower acuity patients on the five-level scale who are being deferred actually make up some portion of admissions (Vertesi, 2004). Young *et al.* (1996) note that 5.5% of patients classified as non-urgent at triage were later admitted to the hospital.

2.2.2. *Complexity-augmented triage*

With global use of triage in the ED, research into the use of variables beyond a traditional triage scale is common. From a queueing perspective, if patients can be distinguished based on measures related to their service times, then prioritization algorithms such as shortest processing time first (that are used in many industries, including manufacturing) can improve performance metrics.

Observing this, the complexity-augmented triage proposed by Saghafian *et al.* (2014) notes that an additional complexity evaluation at triage would only take a matter of seconds, but its benefit could be significant. Through simulation analysis calibrated with hospital data and various queueing models, Saghafian *et al.* (2014) show that complexity-augmented triage does indeed benefit ED performance, both for patient safety (measured by risk of adverse events) and operational efficiency (measured by LOS). Saghafian *et al.* (2014) also investigate several patient flow designs that can be utilized after the complexity-augmented triage is implemented. Ieraci *et al.* (2008) also assert that patient complexity should be factored into triage and streaming. Although only a single case is presented, their conclusions are backed with a marked improvement (a 58% reduction) in waiting time. Sprivulis (2004) takes research into complexity a step further, designating clear complexity groups for patients based on the number of procedures or investigations required, with an additional partitioning based on patient age. Evidence regarding the accuracy of classifying patients by complexity is positive, welcome news given mixed history regarding traditional acuity-based triage (Vance and Sprivulis, 2005).

2.2.3. *Patient streaming*

The innovation and implementation of patient streaming was pioneered in an Australian hospital, Flinders Medical Center. King *et al.* (2006) and Ben-Tovim *et al.* (2008) discuss restructuring patient flow in the ED of this hospital based on whether a patient will be discharged or admitted. This involved processing non-urgent patients in a First-In-First-Out (FIFO) manner and using traditional prioritization methods on admitted patients. We see an identical streaming methodology in Kinsman *et al.* (2008), resulting in continual reduction in LOS times for both patient types. Inspired by the work of King *et al.* (2006), Kelly *et al.* (2007) segregates patient flow. It is recognized that the admitted group of patients may have different barriers to overcome than the discharged group. Though it was accompanied with a resource reallocation effort, improvements were seen in some key metrics, namely ambulance turn away and waiting time (Kelly *et al.*, 2007). In at least one case, patient streaming is seen to have significant benefit only to discharged patient metrics, although inpatient care is not adversely affected (O'Brien *et al.*, 2006). As reported with patient complexity, staff are relatively accurate at predicting patient disposition (Kosowsky *et al.*, 2001; Holdgate *et al.*, 2007).

OR/OM contributions in patient streaming are similar to interventions in call centers, and more generally to resource pooling in applications where resources with noticeably different process times can be optimally partitioned to benefit a particular customer class (e.g., Whitt, 1999; Hu and Benjaafar, 2009). It is noted in Hu and Benjaafar (2009) that prioritization can be performed as an alternative to partitioning, though this would require reliable patient evaluation, which the literature suggests is difficult to achieve (Wuerz *et al.*, 1998; Fernandes *et al.*, 1999). Peck and Kim (2009) use a simulation with a fast track with nurses evaluating acuity and disposition, showing that patient waiting could be reduced by upwards of 50% with the use of streaming. Saghafian *et al.* (2012) provide detailed queueing-based analysis in the realm of disposition-based ED patient streaming where patients are sent to separate tracks based on a prediction of their disposition (admit or discharge) made by triage nurses. Comparing the two systems of physical patient streaming and traditional patient pooling in Saghafian *et al.* (2012) shows a strong advantage for patient pooling due to a low resource utilization in physical streaming flow designs known as the "anti-pooling effect." Observing this, the study of Saghafian *et al.* (2012) develops a *virtual streaming* flow design, which does not require restructuring the ED, but one which vastly outperforms pooling flow designs. This suggests caution towards past successes on patient streaming in which resources are physically separated and introduces virtual streaming as a new paradigm that can effectively achieve the advantages of both streaming and resource pooling.

2.2.4. *ED fast track*

Fast track in the ED is a dedicated stream of resources to process lower acuity patients more quickly. Welch (2009) notes that a fast track dedicated for minor injuries has been a mainstay in EDs since the 1980s. Given that about 80% of ED visits are non-urgent, the use of an ED fast track lane is a great aid in serving lower acuity patients and reducing overcrowding (Williams, 2006). Fast track implementation has found success in numerous ED environments, such as in an urban pediatric ED (Simon *et al.*, 1996; Hampers *et al.*, 1999), or a teaching hospital (Meislin *et al.*, 1988).

Roche and Cochran (2007) apply a queueing methodology to eight EDs, implementing fast tracks with different patient acuity, volume, and anticipated LOS in an effort to minimize patients LWBS. Cochran and Roche (2009) split patients into two tracks by acuity level to reduce patient LWBS and increase ED effectiveness. O'Brien *et al.* (2006) experienced a 20.3% reduction in patient wait time with the implementation of a fast track. A simulation by Gárcia *et al.* (1995) investigated redistributing resources towards a fast track, reporting the potential to reduce LOS by 25% for lower acuity patients without negatively impacting other patients. Fernandes *et al.* (1997) also reports successful results in lowering LOS and LWBS for lower acuity patients. The implementation of a fast track in Considine *et al.* (2008) saw the percentage of patients discharged within 2 hours and 4 hours increase to 53% and 92%, respectively, up from 44% and 84%. Cooke *et al.* (2002) reduced the probability of a patient waiting more than an hour by 32%, down to 59.2%.

Given the overlap between triage and fast track efforts, we see research reporting improvements in performance metrics through the implementation of both, despite an increase in patient census (Sanchez *et al.*, 2006; Kwa and Blake, 2008). Fast track benefit to LOS is ultimately aligned with a time-tested understanding in OR/OM of the benefits of processing time prioritization (Lawler and Moore, 1969; Davis and Patterson, 1975). In the rare case where performance metrics were not met (Nash *et al.*, 2007), customer satisfaction improvements legitimized the implementation of the fast track. In fact, when surveyed, there were marked improvements in categories from LOS to overall patient satisfaction (Rodi *et al.*, 2006). However, we warn that creating a fast track may also result in the "anti-pooling" effect discussed earlier for patient streaming mechanisms. Hence, careful analyses must be performed before creating a fast track to make sure resources are assigned to the fast track in an appropriate way.

2.2.5. *Bed planning*

As Saghafian *et al.* (2012, p. 1080) notes: "The most direct way to alleviate crowding and improve responsiveness [in the ED] is by adding resources. However, because this is also the most expensive approach, it is generally not the preferred option." This sheds light on an important connection between OR/OM techniques and ED operations, as OR/OM techniques are widely used for resource allocation in various industries. Resource allocation is indeed a well-known problem in OR/OM with a long history (Karchere and Hoeber, 1953; Cooper, 1963; Arrow and Hurwicz, 1977). Richardson (2003) recognizes that the discussion regarding resources is not surrounding their addition, but rather the proper allocation. In the discussion of bed planning, most research has occurred outside of the ED, namely in the ICU. We refuse to omit these papers, however, because changes outside of the ED can have significant impact within the ED. For example, an increase of 20 beds (43%) in one ICU resulted in a 66% reduction in ambulance diversion hours and 9.7% reduction in ED LOS for admitted patients (McConnell *et al.*, 2005). In fact, the lack of ICU beds seems to be a common bottleneck for patient flow (GAO, 2003; Burt and McCaig, 2006; Pham *et al.*, 2006). However, a cross-sectional study of California hospitals found that the ICU was the bottleneck in only 34 of 181 hospitals (Allon *et al.*, 2013). Thus, any conclusions relating to expanding bed capacity must be justified. Relative size of the ED and ICU must be considered prior to expanding bed capacity. A failed example of this is provided by Han *et al.* (2007), where an effort to expand the ED to improve ambulance diversion had no tangible positive impact.

First defining the issues with current bed planning practices and the necessity of proper bed management, we start with the established target occupancy level that has been in place for more than 25 years. Eighty-five percent occupancy is the standard target capacity for beds, being the minimum level to increase the number of hospital beds (Green, 2006). Green and Nguyen (2001) address the issue with using occupancy levels for bed planning, pointing out a number of issues. Occupancy levels are based on the number of certified beds, which may differ from the number of staffed beds. Occupancy is typically measured as the midnight census, generally the lowest level of the day. Lastly, occupancy levels are yearly averages and do not incorporate weekly or seasonal variations. Green (2003) and de Bruin *et al.* (2007) both critique traditional occupancy levels, with Green (2003) also detailing the severity of bed delay and the need to plan properly. To further stress the importance of bed planning, we highlight that the lack of staffed critical care beds is the number one reason for ambulance diversion (American Hospital Association, 2007). The issue of *staffed* beds illustrates the need to properly allocate staffing and beds together.

Having established the necessity of accounting for external factors in bed planning, as well as the woes of current practice, we now summarize work done to combat the issue. Queueing Theory (Kao and Tung, 1981; Worthington, 1987; Huang, 1995) and simulation (Goldman *et al.*, 1968; Dumas, 1984) are established interventions in the realm of bed planning, and we will discuss them in more details in Section 3. Given the incorporation of variability present in these methods, we see them as the preferred approach of determining bed capacity and optimizing bed allocation. Green (2003) succeeds in this regard by determining bed capacity for an obstetrics and an ICU unit, where insufficient capacity exists 40% and 90% of the time, respectively. Modeling bed demand with a Poisson distribution for a pediatrics unit in Milne and Whitty (1995) provided more accurate results than formerly averaged data. Results from a simulation in Bagust *et al.* (1999) recommend to invest in a spare capacity of beds to minimize risk of bed shortage. Simulation is also used in Pines *et al.* (2011), where dynamic inpatient bed management reduces ED boarding times, providing significant financial benefit. Recent work has been done to identify bottlenecks and optimize bed allocation (de Bruin *et al.*, 2007; Elbeyli and Krishnan, 2000). For example, Elbeyli and Krishnan (2000) study the allocation of inpatient beds to different departments to maximize the effect on patient times. A simulation model in Harper and Shahani (2002) allow the user to perform a "what if" analysis, balancing hospital bed capacity with refused admissions. Cochran and Bharti (2006) create a queueing network and perform a discrete event simulation (DES) to optimize hospital bed allocation, resulting in 10% increase in throughput. Cochran and Roche (2008) generate a decision making tool using Queueing Theory to optimize inpatient bed planning.

Lovejoy and Desmond (2011) propose a simple solution to the issue of bed congestion by using a method more practical for physicians, requiring less hard-to-gather data and ultimately delivering speedy results: *Little's Law*. The analysis in Lovejoy and Desmond (2011) justifies the purchase of less costly dedicated beds for an observation unit to free up inpatient beds, thus relieving congestion upstream to the ED. A simple ratio method (Plati *et al.*, 1996) frequently used to calculate nursing or bed requirements based on patient census fails to incorporate variability of arrivals and service times in decision making. These methods are effective relative to their ease-of-use, but some caution should be taken, given that they do not incorporate variability. This is addressed in Nguyen *et al.* (2005), where a new method outperforms the past ratio method by minimizing the mean and standard deviation of occupancy-based parameters.

From a less technical operations perspective, active management of hospital beds, a task or dedicated full time job typically performed by physicians and RNs, has had a positive impact on patient flow, as well as benefiting safety and satisfaction (Howell *et al.*, 2008; Borders *et al.*, 2009). These decisions performed by "planners" or "czars" typically involve reallocating beds between the ED and internal wards based on demand. For example, multiple ICU and ED assessments performed by physicians acting as bed managers in Howell *et al.* (2010) saw a 28% reduction in LOS for ED patients admitted into ICU or CCU. The use of a dedicated bed planner is frequently coupled with other improvements as well. A dedicated RN bed planner, as well as the implementation of a bed management database and streamlined ED-ICU communication lowered one hospital's diversion hours by 63% (Hemphill and Nole, 2005). Deeper analysis into the use of a bed manager shows great potential, and the investment of resources into training staff for such a task may be a wise decision (Proudlove *et al.*, 2003, ,2007).

2.2.6. *Staffing and scheduling*

Personnel are typically the deciding factor in moving a patient through the ED effectively and efficiently. As personnel account for two thirds of a hospital's entire expenses (Warner, 2006), the importance of appropriate staffing and scheduling cannot be overstated. The overall wait to see a physician in the ED increased to 30 minutes in 2004, up from 22 minutes in 1997 (Green, 2008). It is common to see patient LWBS rates above 6% due to physician unavailability (Ding *et al.*, 2006). Also, patient discharge is often delayed because staff are tied up with more urgent patients (Kelly *et al.*, 2007), indicating that staffing and scheduling have a widespread effect on all areas of the ED. Green *et al.* (2007, p. 34) address the common issue with staffing in healthcare: "Hospital managers, while aware of the variability over the day, have not used queueing models, but instead allocate staff based on general perceptions and intuition." The lack of OR/OM driven decision making pointed out by Green *et al.* (2007) is also identified by Carter and Lapierre (2001) after interviewing physicians to determine their staffing methods. Similar to Beaulieu *et al.* (2000), their solution is to formulate a mathematical program that incorporates all the rules that should govern a physician's schedule. Below, we identify additional literature that focuses on OR/OM related techniques for ED staffing and scheduling to improve patient flow.

Considering physicians, nurses and exam rooms as variables in a simulation, Duguay and Chetouane (2007) test numerous variable settings to improve key performance indicators. Adding a doctor and nurse during regular business hours was found to have the best impact on patient waiting. Two heuristic algorithms are used to staff physicians, nurses and technicians in Sinreich *et al.* (2012). The two algorithms developed efficient work schedules which reduced patient waiting between 20% and 64% and LOS between 7% and 29%. A linear optimization model is used in Sinreich and Jabali (2007) to find a resource's contribution to ED operations, allowing for a reduction in LOS while also reducing staffing levels. Green *et al.* (2006) and Green *et al.* (2001) identify the variation of patient arrival through the day and use a Lag SIPP approach to create a weekday and weekend staffing model. The result in pursuing a demand-based staffing model was more than a 20% decrease in LWBS, despite an increase in patient volume. Yankovic and Green (2011) build a queueing model with nurses as servers to minimize delay probability through staffing, with the added benefit of a tool to identify overcrowding bottlenecks. Fullam (2002) presents an ED staffing success, where nurse staffing levels are determined by acuity data. In Patel and Vinson (2005), ED staff members are organized into teams consisting of one physician, two nurses and one technician in one suburban ED. The result of such a team assignment system was notable decreases in patient wait time and LWBS (Patel and Vinson, 2005). Traub *et al.* (2014b) study the rotational assignment of patients to physicians, finding a decrease in both LOS and LWBS. Increased satisfaction for patient waiting in DeBehnke and Decker (2002) further validates the use of patient care teams.

Finally, we note that several studies provide methods to forecast surges in ED volume, which can also be used to improve staffing and scheduling methods. For instance, Chase *et al.* (2012) consider the ratio of new patients requiring treatment over total physician capacity (a metric termed the care utilization ratio (CUR)), and finds it to be a robust and promising predictor.

2.3. *Patient flow out of the ED*

In this section, we focus on the final leg of ED patient flow: flow out of the ED. With a holistic systems approach, lean thinking has put focus on patient flow out of the ED so as to avoid congestion (Ben-Tovim *et al.*, 2008). Khare *et al.* (2009) uses a simulation model and finds that reducing admitted patients boarding time in the ED was the biggest influence of reduced congestion. Schneider *et al.* (2001) confirms these findings, where rapid removal of inpatients from the ED was the greatest relief of overcrowding in their study. With this, we recognize that improving patient flow out of the ED is every bit as essential to ED operations as is patient arrival or flow within the ED.

2.3.1. *Patient discharge*

Optimizing patient discharge is as crucial to improving patient flow as any other aspect of ED operations. It is shown that a lower admission-to-discharge ratio in the ED is crucial for a low LOS (Vermeulen *et al.*, 2009). Fatovich and Hirsch (2003) identify improving the discharge process as one of five major steps for addressing ED overcrowding. Despite a scarcity of data, Black and Pearson (2002) recognize that delayed discharge is a serious issue. The use of a discharge lounge can free up a needed inpatient bed while patients ready for discharge have their prescriptions filled, wait for transportation, receive care education or schedule follow-up appointments (Williams, 2006). Geer and Smith (2004) represent a success story in this regard; the implementation of a discharge room was part of several process improvements that resulted in a 79% reduction in diversion hours. This is also backed by Moskop *et al.* (2009b), who even endorse taking discharge further by having a "reverse triage" system for early discharge of hospital inpatients. This innovative solution is originally proposed for inpatients by Kelen *et al.* (2009b), where a five-category scale is used to identify patients for early discharge. Although "triage at discharge" has been proposed as a crisis measure (Hick *et al.*, 2009; Kelen *et al.*, 2009a), Kelen *et al.* (2009b) note that such a system could be used in daily ED operations.

Powell *et al.* (2012) illustrate the need to take a system wide approach in the hospital, particularly in discharge. Peck *et al.* (2012) achieve a creative success in this regard where a generalized linear regression model is one of a few ways to accurately predict inpatient admissions based on information gathered at ED triage. Several models of improvements in inpatient discharge time were shown to have positive impact towards ED boarding. Kravet *et al.* (2007) take a similar approach, whereby discharging inpatients earlier ultimately reduced crowding. Unfortunately, there has been limited research on improving patient flow through discharge. In the medical field, Samuels-Kalow *et al.* (2012) review literature relating to patient quality at discharge, particularly how discharge information is received.

2.3.2. *Patient routing*

Inspired by the success and simplicity of Erlang models in call centers, Armony *et al.* (2011) attempt to model routing from the ED to the inpatient ward with a series of time dependent processes. It is shown in Armony *et al.* (2011) that only 4.9% of patients were admitted to their internal ward within 30 minutes of being assigned. Lack of resource availability and poor routing decisions were identified as root causes, which may be optimized with queueing-based analysis.

Mandelbaum *et al.* (2012) introduce a RMI (randomized most-idle) policy to route from the ED to internal wards, achieving the same idle levels between server pools as LISF (longest-idle-server-first), without requiring idle times or pool capacity information. Given that patient flow out of the ED is greatly affected by staff unavailability (Kelly *et al.*, 2007), abandoning an LISF policy will not force staff to spend additional time collecting patient data. Ultimately, however, research on routing has lacked an emphasis on the ED. Armony and Mandelbaum (2011) provide a case of routing with homogeneous impatient customers and heterogeneous servers for large service systems, which can be applied to the ED.

2.3.3. *Bed-block*

"Bed-Block," or the patient boarding phenomenon, relates to ED patients admitted to the hospital who are unable to be transferred out of the ED due to unavailability of inpatient beds. In the medical literature, the discussion regarding ED flow problems does not begin without mentioning boarding, especially since boarded patients block ED beds and prevent from seeing new patients in a smooth and timely manner. Derlet and Richards (2008, p. 24) express the significance of this issue, stating "boarding of inpatients in the ED is unquestionably the leading cause of crowding." Decreasing boarding times has been found to be a major factor in reducing LOS (Khare *et al.*, 2009; Moskop *et al.*, 2009a). In fact, the number of patients who are boarded is such a significant problem that it is perhaps the most common trigger for ambulance diversion (Epstein and Tian, 2006; Ramirez *et al.*, 2009; Allon *et al.*, 2013). Pines *et al.* (2011) note that financial decisions further exacerbate the problem, enabling high boarding levels: hospital revenue is higher for non-ED admissions than for ED admissions, leading to non-ED admitted patients having a higher priority. ED success is largely gauged by the ability to manage boarding: the Institute of Health Improvement

(IHI) asks two questions as a measure of the success of patient flow, one of which relates directly to patient boarding: "Do you park more than 2% of your admitted patients at some time during the day for at least 50% of the time?" (Williams, 2006). Despite bed-block being one of the most well-known operational problems to afflict an ED, there is not a large body of OR/OM work present to directly combat these issues. The existing contributions are summarized below.

A few queueing networks that incorporate blocking into their model are outlined in Section 3.2. One of these blocking models identifies three major sources throughout the hospital from which blocking occurs (Osorio and Bierlaire, 2009). Wiler *et al.* (2013) also focus on minimizing patient boarding with a queueing perspective. Marshall *et al.* (2005) assess that simulation will be used in the future to carry out complexities of patient queueing systems like bed-block. Hoot *et al.* (2008) display the potential of simulating ED operations to forecast upcoming overcrowding, and discuss that such forecast could allow for better responsiveness to patient boarding. The systematic effect of bed-block in a geriatric department is modeled in El-Darzi *et al.* (1998); while their study is focused outside of the ED, their simulation indicates that a similar application in the ED is feasible. Saghafian *et al.* (2012) stands in a rare class of OR/OM oriented work where bed-block is a primary objective. The effect of boarding times on admitted patients is a significant reason "virtual streaming" is mentioned and discussed in Section 2.2.3. The complexity-augmented triage work of Saghafian *et al.* (2014) is also shown to benefit EDs in which bed-block is a significant issue. Shi *et al.* (2013) is another OR/OM related work that considers the ED bed-block effect; it provides a detailed study in addressing the effect of altering discharge times in inpatient units on ED bed-block durations.

Addressing a less technical perspective, Patel *et al.* (2006) found that boarding patients was a major hurdle to overcome. The first solution proposed for combating ED overcrowding and boarding is to expand the hospital capacity in Derlet and Richards (2008). We advise caution, however, considering one study where expansion of the ED actually increased the number of boarding patients (Han *et al.*, 2007). Further research into the proper expansion or allocation of resources to minimize bed-block is required, and we identify this as a research area which can have high-impacts for ED research. A similar conclusion has been made in a number of OR/OM papers throughout the last decade (Gallivan *et al.*, 2002; Broyles and Cochran, 2007; Deo and Gurvich, 2011).

3. OR/OM tools used for optimizing ED patient flow

In the previous section, we discussed challenges with respect to the three components of ED flow. We now review the main OR/OM tools that have been used to address such challenges: Mathematical Programming, Queueing Theory, Simulation, Markov models (and Markov Decision Processes, MDPs), and Game Theory.

3.1. *Mathematical programming and optimization*

We start by reviewing related work using Location Theory techniques such as the Maximal Covering Location Problem (MCLP) and the Maximum Availability Location Problem (MALP). For the purposes of this paper, we only highlight work that immediately affects patient flow into the ED – primarily ambulance decisions. The Maximum Covering Location Problem (MCLP), introduced by Church and ReVelle (1974) and having undergone many early extensions (see, e.g., Daskin, 1982, 1983; Batta *et al.*, 1989), seeks to maximize coverage within a certain distance by establishing fixed location points. Hogan and Revelle (1986) build on the MCLP by recognizing the need for secondary, or backup coverage. Murray *et al.* (2010) provide a recent enhancement of classic location problems. Genetic algorithms have been endorsed for their ability to arrive at near optimal solutions in a much more realistic time frame than integer programs or heuristics for MEXCLP (Maximum Expected Covering Location Problem) (Aytug and Saydam, 2002). MALP is structured as an integer program extending the MEXCLP (ReVelle and Hogan, 1989). It seeks to maximize the population that will be reached within a given time. As we will next briefly describe, these techniques have been instrumental in research related to transferring patients to EDs.

Multiple tabu search (TS) heuristics are frequently used to determine ambulance location (Adenso-Díaz and Rodríguez, 1997; Gendreau *et al.*, 1997, 1999, 2001). One of these is a reactive TS that iterates through an algorithm to improve ambulance coverage, stopping after additional iterations would mean negligible improvement (Rajagopalan *et al.*, 2011). Erdogan *et al.* (2010) use a TS not only for ambulance location, but also for crew scheduling, outperforming past literature on processing time and performance. The TS extends into the ED as well, where it is used by Gendreau *et al.* (2007) for physician scheduling. A comparison of genetic algorithms, simulated annealing and TS concluded that all three are robust, each with unique merits, though favor is expressed towards TS (Youssef *et al.*, 2001; Arostegui Jr. *et al.*, 2006). TS is utilized to address the patient flow within the ED as well, where it helps set favorable physician schedules (Carter and Lapierre, 2001). For instance, Gendreau *et al.* (2007) compares mathematical programming, column generation, TS and constraint programming approaches for physician scheduling in the ED.

Mathematical programming tools are also used for ED staffing in other forms. For a recent example, we refer interested readers to Wang (2013), where a model based on separated continuous linear programming (SCLP) is used to provide high-level staffing guidelines. Karnon *et al.* (2009)

summarize several case studies about types of mathematical models used to improve health care. Beaulieu *et al.* (2000) take a mathematical programming approach to ED physician scheduling, where a large number of scheduling constraints are incorporated into a program to staff 20 physicians, ultimately outperforming the "expert scheduler." Similar to Queueing Theory efforts, forecasting is often used along with mathematical programming to develop staffing and scheduling models that also incorporate variability of hour-by-hour patient arrivals. Mathematical programming tools have also been used to develop schedules for medical residents (see, e.g., Cohn *et al.*, 2009) and cyclic schedules for ED physicians (see, e.g., Ferrand *et al.*, 2011).

3.2. *Queueing theory*

Queueing Theory has had a prominent role in research related to patient flow optimization. Wang *et al.* (2013, p. 341) sum up the simplicity and efficiency of Queueing Theory: "Although analytical methods contain less details than simulation, and are based on simplified models, it could provide quick results and an opportunity to investigate system properties more efficiently under appropriate assumptions." Fomundam and Herrmann (2007) and Lakshmi and Iyer (2013) review Queueing Theory applications in healthcare, ranging from a single department to the regional healthcare level. Wiler *et al.* (2011) reviews modeling applications concerning patient flow and crowding in the ED, with a dedicated section on Queueing Theory. Green (2006) details some basic queueing models with healthcare application, such as $M/M/s$, $M/G/1$, and $G/G/s$.

We begin with some interesting extensions of the $M/M/s$ queue used in general (not necessarily ED) patient flow applications, where a queue with s servers follows a Poisson arrival distribution and exponential service distribution. Two interacting queueing networks are set up in Yankovic and Green (2011) with no blocking or balking. A closed queueing system is used in de Véricourt and Jennings (2011) for staffing purposes, where there is a finite population of n patients within the $M/M/s//n$ model. Singer and Donoso (2008) use an $M/G/s$ approximation to identify key performance indicators in ambulance services. Broyles and Cochran (2007) use an $M/M/1/K$ to measure the financial impact of patient reneging. Patient balking with an $M/M/1/K$ is also seen in Cochran and Broyles (2010). They point out that the accuracy coupled with the minimal data required for a queueing model makes it the preferred method for modeling over regression. An $M/M/s/K$ addresses patient flow in Roche and Cochran (2007), where s is calculated by the arrival rate, length of service and desired bed utilization level. The use of $M/M/s$ in bed planning throughout the hospital is detailed in Green and Nguyen (2001) and Green (2003), and within the ICU in Ridge *et al.* (1998) and Kim *et al.* (1999) (we point out later how bed management in the ICU affects ED operations). Extensions of $M/M/s$ are also seen in Green *et al.* (2001), Green *et al.* (2006), and Green *et al.* (2007), where arrivals are time dependent and are based on a Lag SIPP (Stationary Independent Period by Period) approach. Au *et al.* (2009) use a six-hour moving average to represent time-dependency in an $M/M/s$ model for predicting overflow. Au-Yeung *et al.* (2007) develop a queueing model with an approximate generating function analysis, designed to accommodate a larger state space than traditional models. This is modeled with a network of $M/M/s$ queues where patients are identified by arrival and acuity. Au-Yeung *et al.* (2006) turn a network of $M/M/s$ queues into a simulation model, although there is an expressed concern about the validity of some of the assumptions.

Infinite-server queues are also used to optimize the patient flow. For instance, an $M/G/\infty$ model is used to analyze hospital bed allocation and minimize overflows in Kao and Tung (1981). Ultimately, the appropriateness and suitability of the type of the queueing model used depends on (i) the goal of the study, and (ii) the underlying assumptions made. Generally speaking, we observe four main deficiencies in the typical queueing theory models used for optimizing ED patient flow: (i) they ignore blocking issues in the ED, (ii) they assume stationary arrival and service processes while these processes are indeed non-stationary in reality, (iii) they ignore abandonment issues in the ED, and (iv) they assume state-independent service processes. Furthermore, as we will discuss in Section 4, human behavioral elements of service delivery are also widely ignored in current queueing models. In what follows, we overview papers that try to address these deficiencies.

Some papers have recognized that the patient queue needs to be modeled as a finite queue with blocking to be robust (Cochran and Bharti, 2006; Osorio and Bierlaire, 2009). Koizumi *et al.* (2005) incorporate blocking into a queueing model and extend upon the work of a traditional single server model by modeling with a $M/M/s$ multi-server network. Bretthauer *et al.* (2011) point out flaws in past blocking models, however, criticizing the use of an infinite queue capacity. A heuristic is used to predict blocking probabilities from which optimal capacity can be derived.

As discussed earlier, an important part of flow out of the ED is the bed-block phenomenon, which refers to situations in which ED patients who need to be hospitalized cannot be transferred to their inpatient units due to lack of bed availability. It prevents EDs from serving new patients in a timely manner, and results in longer Length of Stay (LOS) as well as a percentage of patients who Left Without Being Seen (LWBS). Some OR/OM studies such as Saghafian *et al.* (2012) consider the effect of the bed-block phenomenon on various patient flow optimization techniques including "virtual streaming." Shi *et al.* (2013) provides a detailed study of issues related to altering discharge times in inpatient units, which directly affects the ED bed-block durations.

While the majority of queueing models assume time-stationarity, there is also some work that considers time dependency. Armony *et al.* (2011) study the ED as a queueing model representing a section of a bigger queueing network, i.e., the whole hospital. Three Markovian queueing models are analyzed to fit ED occupancy along with a simulation framework. Armony *et al.* (2011) discuss that the time dependent (i.e., non-stationary) model, $M_t/M_t/\infty$, is only accurate when the ED is occupied with less than 15 patients. The state dependent model, $M_i/M_i/\infty$, in Armony *et al.* (2011) is found to be very accurate for a queueing model. These results are intuitive as the state dependent model incorporates important factors beyond time, such as ambulance diversion. Armony *et al.* (2011) then suggest modeling ED occupancy as a black-box birth and death process with state dependent arrival and service distributions. An $M/M/\infty$ is also used in de Bruin *et al.* (2007) to model flow of emergency cardiac patients. Time dependency is also captured in the modeling of clinical wards, where it is recognized that understanding variability outside of the ED is essential for capacity planning (Bekker and de Bruin, 2009).

Wiler *et al.* (2013) incorporate patient abandonment (LWBS) by using a $M/GI/r/s + GI$ model introduced by Whitt (2005), where patient arrival follows a Poisson distribution, service times are a general distribution i.i.d, with r bed servers, s waiting area capacity and abandonment times are a general distribution i.i.d. Batt and Terwiesch (2013) provide innovative work in patient waiting, noting that factors beyond waiting time affect abandonment. Among these are the observed number of patients waiting, the flow of patients in and out, and the inferred severity of waiting patients—all visual information acquired by a patient in the waiting room.

Cochran and Roche (2009) address many complexities that are often ignored in high-level queueing models, taking into account patient acuity, variation of arrival, and different resource consumption across several EDs. This is also seen in Roche and Cochran (2007) to test an "extreme" fast track in the ED. Queueing Theory has also been applied as an extension to the Maximum Availability Location Problem (MALP), where it is used to relax the assumption that server availability is independent (Marianov and ReVelle, 1996; Ghani, 2012). Huang (2013) and Huang *et al.* (2013) split patients into two queueing networks, new patients and WIP patients, to optimize physician decisions of which patients to service. Gallivan *et al.* (2002) simplify the patient flow process through a heavy-traffic deterministic system, assuming that the number of patients per day, the probability of "success," and patient LOS is always the same. For some other studies regarding applications of queueing theory in the ED, we refer interested readers to Alavi-Moghaddam *et al.* (2012) and the references therein.

3.3. *Simulation analysis*

Simulation has provided strong decision making tools for ED operations even before global accessibility to computers and the development of widely available software. We see the use of an animated simulation that factors in random arrivals in Saunders *et al.* (1989), including individual service times and patient acuity, factors not incorporated into early ED queueing models. The ability to model processes in a great level of detail makes simulation a potential tool for virtually every aspect of the ED that impacts patient flow. There exists great depth in DES research in staffing and scheduling, fast track implementation, ambulance diversion, streaming, sequencing, performance tracking, and overall process improvement. It is the "what-if" analysis—the ability to test a high number of scenarios in a minimal amount of time — that have made simulation a widespread tool in the ED. Kolker (2008, p. 391) endorses the use of simulation above other methods: "Process model simulation approach seems to be much more flexible and versatile. It is free from assumptions of the particular type of the arrival process (Poisson or not), as well as the service time (exponential or not). The system structure (flow map) could be of any complexity, and custom action logic can be built in to mimic practically any features of the real system behavior." While simulation is widely used in various ED patient flow studies, we note that as Günal and Pidd (2010) discuss, there is a lack of generality in simulation studies: the objective, scope, level of details, and calibrations performed vary considerably among such studies, making simulation more of a "case-by-case" approach rather than a generically available tool.

With an emphasis on patient flow, we identify a number of reviews that survey the use of simulation in the ED (Jun *et al.*, 1999; Fone *et al.*, 2004; White, 2005; Brailsford, 2007; Günal and Pidd, 2010). Paul *et al.* (2010) review over 90 papers focused on the investigation of ED crowding with simulation. Sinreich and Marmor (2005) provide a walk-through of how simulation can be used in the ED. How to define research objectives, gather and classify data, and validate the simulation are explained in a general manner to be applied to any ED. High-fidelity simulation analysis of ED patient flow calibrated with hospital data can also be found in studies such as Saghafian *et al.* (2012), Saghafian *et al.* (2014), and the references therein.

Simulation is a strong tool for ambulance services as it allows for robust models in the absence of parameters (Goldberg, 2004). In one case, three ambulance diversion policies are compared with a simulation, with one policy being based off of a MDP (Ramirez-Nafarrate *et al.*, 2012). A number of alternatives are tested in order to reduce LOS in McGuire (1994) and patient waiting time in Komashie and Mousavi (2005). LOS is a common focal point of simulation, with the implementation of triage (Ruohonen *et al.*, 2006), proper expansion of resources and a fast track (Samaha *et al.*, 2003) or a combination of changes (Wang *et al.*, 2012) providing an operational benefit. Macdonald *et al.* (2005) use simulation to test alternatives for how changes within the ED can improve LWBS rates and other performance indicators. The experimental power of simulation was able to test 21 alternatives on nine performance criteria, finding that different alternatives are optimal depending on the desired improvement. Hoot *et al.*

(2008) are able to forecast ED crowding with DES. Kolker (2008) establishes a quantitative link between patient LOS and diversion. Predictably, it is found that low ambulance diversion levels correspond to low LOS times and number of patients in the waiting room.

Bagust *et al.* (1999) used Excel to create a DES model to measure bed availability versus demand. Duguay and Chetouane (2007) analyze the addition of constrained resources, staffing and exam rooms, to best reduce patient waiting time. Staffing is also addressed in Zeltyn *et al.* (2011), with a focus on three time horizons ranging from within a few hours or days to yearly. Overall, we see staffing and scheduling as one of the most popular uses of simulation in the ED, given the ability to rapidly test numerous alternatives (Draeger, 1992; Badri and Hollingsworth, 1993; Rosetti *et al.*, 1999; Sinreich *et al.*, 2012).

Wang *et al.* (2013) opt to use simulation to model patient flow, given its ability to capture ED complexities like re-entrant flow and resource limitations. Considering the patient flow through hospital (not merely ED), Harrison *et al.* (2005) model patient flow through the entire hospital to determine bed occupancy.

In addition to DES, system dynamics (SD) and agent-based simulation approaches are also used to study ED patient flow. Lane *et al.* (2000) use SD modeling and find that decreasing hospital beds does not significantly increase the waiting time for admitted patients through the ED; rather, it can cause a higher cancelation rate for elective surgeries in the hospital. Their study emphasizes a holistic approach of studying patient flow throughout the hospital, and highlights that looking only at one performance measure in the system can be misleading. Lane *et al.* (2003) discuss a case study of using SD in analyzing ED patient waiting times, with an emphasis on involving the client in the process of model building.

Agent-based simulations are also used to analyze ED patient flow. Laskowski and Mukhi (2009) develop an agent-based model (one that allows simulating a number of EDs) through which one can extract time patient data from EDs of a city to examine patient diversion policies. Jones and Evans (2008) use agent-based simulation for evaluating various physician schedules. Stainsby *et al.* (2009) emphasize the role of human elements, and presents another example of developing an agent-based simulation model for the ED.

Simulation using DES, SD, or agent-based techniques allow for considering various complex aspect of patient flow. It is this ability that has established simulation as a legitimate OR/OM tool in the ED. Undoubtedly, simulation will continue to see universal implementation and will continue to be a prominent tool for patient flow optimization.

3.4. *Markov models and Markov decision processes*

Markov models are briefly discussed in a review of LOS-based patient flow models by Marshall *et al.* (2005, p. 214) who state that "Markov models are based on well established statistical methodologies and provide a viable approach to measuring and modelling flow." The probabilistic nature of Markov models make them an ideal candidate for ED modeling. We identify unique uses of Markov models and speak to their breadth below.

Markov models and Markov Decision Processes (MDPs) are common in ambulatory research to support data-driven decisions for ambulance location and deployment (Maxwell *et al.*, 2010). For instance, the Hypercube model, used for decisions in the ambulance system, is based on a multidimensional Markov chain with multiple queues (Brandeau and Larson, 1986). In Ramirez-Nafarrate *et al.* (2012), an ambulance diversion policy is based on a MDP. The use of MDPs is simplified in Ramakrishnan *et al.* (2005), where continuous time Markov chains and discrete time Markov chains are used to model the ED and internal wards, respectively. MDPs are also extensively used in Saghafian *et al.* (2012) and Saghafian *et al.* (2014) (and some references therein) to gain insights into effective ED patient flow designs, triage and prioritization. Similar objectives are followed in Zayas-Caban *et al.* (2013), where MDP and sample path analyses are used to determine how ED patients should be dynamically prioritized. Overall, research involving Markov models present diverse application, with the ability to predict inpatient LOS (Kapadia *et al.*, 2000), make admission decisions (Nunes *et al.*, 2009), describe bed queues (Au *et al.*, 2009), represent reneging (Cochran and Broyles, 2010) and model patient flows (Davies and Davies, 1994; Wang *et al.*, 2013).

3.5. *Game theory*

Game Theory, a study of rational decision making that has been widely used in the field of economics, has been mostly obscure in ED operations. In rare, scattered cases, Game Theory offers a supplemental tool to a number of decision based applications. We identify those limited contributions below.

The review provided by Brandeau and Chiu (1989) recognizes that Game Theory models have been integrated with location models as an additional application. Hagtvedt *et al.* (2009) and Deo and Gurvich (2011) model a hospital's decision to go on ambulance diversion using Game Theory. Game Theory has also been proposed as an aid to measuring transport reliability (Bell, 2000). Mandelbaum *et al.* (2012) provide a unique perspective, suggesting that Game Theory could be used to model intricate decisions made by hospital staff such as decreasing or increasing service rates and/or the quality of care.

3.6. *Summary and level of use of OR/OM tools*

Table 1 summarizes the use of the OR/OM tools (discussed in the previous sections) in addressing the three components of ED flow. Areas indicated with "S" or "N/A" in this table also indicate opportunities for further research.

Table 1. The use of OR/OM tools in addressing challenges in ED patient flow optimization ("A": Ample; "S": Scarce; "N/A": not available)

Tool	Flow Into	Flow Within	Flow Out
Mathematical Programming	A	A	S
Queueing Theory	A	A	S
Simulation	A	A	A
Markov Models & MDPs	A	S	S
Game Theory	S	N/A	N/A

It is noteworthy that all the tools/techniques summarized in Table 1 have legitimate benefits and drawbacks. For instance, Simulation requires a much greater investment in time and resources to develop, as it can build a model with greater detail and hence complexity than Queueing Theory analysis. In itself, however, the complexity can also prove a hindrance. Coats and Michalis (2001) are forced to estimate some processing times in their simulation model, even given a large set of preexisting data in information systems. Modeling patient flow with Queueing Theory is also not without its complexities and shortcomings. Mayhew and Smith (2008) express the need to split patients into two distinct categories, and note that service time was not the same at each stage. Kolker (2008) argues that with all the transformations and adjustments required to accurately model the network, the main benefit of Queueing Theory – transparency and simplicity – is lost. Queueing Theory's reliance on closed-form mathematical solutions also presents an issue to Markov models: as Sinreich and Marmor (2005, p. 233) note, they both are "very sensitive to the size, complexity and level-of-detail required by the system under study."

Overall, simulation has been considered to be more conducive for modeling systems of patient care (Davies and Davies, 1994), and we see a high percentage of DES focused research in the literature. However, it is noteworthy that some researchers have preferences towards Queueing Theory (de Bruin *et al.*, 2007), others prefer simulation (Marshall *et al.*, 2005), and some incorporate both Queueing Theory and DES for robustness (Cochran and Bharti, 2006). In some cases, we see both approaches considered along with some additional approaches. This is true, for instance, in Saghafian *et al.* (2012) and Saghafian *et al.* (2014), where various queueing models, simulation analyses, Markov Decision Processes (MDPs), and hospital data are used to study issues related to patient flow, patient streaming, triage, and patient sequencing in EDs. Hagtvedt *et al.* (2009) use Game Theory as an extension to the analysis relating to an ambulance diversion model. A combination of stochastic and deterministic processes, such as a random walk, DES, and integer programming are used in Kokangul (2008) to optimize bed capacity for a hospital unit, further showing the integration of multiple approaches. Atallah and Lee (2013) utilize multiple OR/OM approaches, including simulation, mathematical programming and optimization for a complete performance overhaul, resulting in major LOS, waiting time, and LWBS improvements. Lee *et al.* (2014a) and Lee *et al.* (2014b) are further excellent examples of using a variety of approaches to significantly improve ED operations. These successes in using a combination of approaches endorse the integration of all these tools in the ED.

4. Concluding remarks and research prospects

In recent years, OR/OM tools have been widely used to optimize ED patient flow, and this paper provides a comprehensive survey of such contributions. In order to demonstrate the potential impact of OR/OM tools in the ED, we classified operational improvements into three categories: flow into the ED, flow within the ED, and flow out of the ED. The range of papers identified that model patient flow speak to the breadth of work in OR/OM. We identified significant problems facing the ED such as ambulance diversion, triage, sequencing, streaming, resource planning, scheduling, staffing, discharge, routing and bed-block, and identified successes in combating these issues.

While OR/OM studies in the ED are ample, we note that there still seems to be a lack of implementation (see also the related discussion in Brailsford *et al.*, 2009). This is mainly due to the low level of close collaboration between ED managers, hospital stakeholders, and OR/OM researchers. We believe involving the ED managers and other hospital stakeholders early on in the process of developing appropriate flow models can have a significant impact on the implementation of the results. Our view is also aligned with the recent President's Council of Advisors on Science and Technology (PCAST) report (PCAST, 2014) that identifies various ways to improve health care delivery and lower the costs. Specifically, aligned with the several suggestions provided by PCAST, we believe that sharing lessons learned from successful OR/OM implementations, increasing the access of OR/OM researchers to appropriate data sets, better use of information technology and health analytics by system modelers, and training health professionals in using and even developing OR/OM approaches for their practice are important changes that can result in higher levels of implementation in the near future.

Our review also suggests that the term "operations research" may have different meanings depending on whether it is used in the Emergency Medicine literature or in the OR/OM literature. In the Emergency Medicine literature, operations research may refer to simple before-and-after studies or limited logistic regressions that show an improvement or decline in operational variables (such as length of stay) after an intervention, whereas in the OR/OM literature the term is usually employed to describe robustly controlled interventions suggested by rigorous mathematical models. In the Emergency Medicine literature, there is seldom a discussion of mathematical models or engineering principles; in the OR/OM literature, there is a premium placed on them. In the day-to-day management of EDs – which is done almost exclusively by nurses and physicians without the input of engineers – anecdotal evidence suggests that the mathematical models developed in the OR/OM literature are often downplayed in favor of perceived experience or an administrative gestalt of "how things work." This disconnection, in our opinion, represents an extraordinary opportunity to improve ED operations through joint collaborations between OR/OM researchers and ED managers or other hospital stakeholders.

Successful OR/OM work in the ED is most prominent in the sections of ambulance management, fast track, patient streaming, bed planning, staffing and scheduling. Outside of these areas, depth of work is somewhat scarce. Acuity-based triage, a wide source of publication in the non-technical arena, has not seen abundant focus in OR/OM (with only a few exceptions such as Konrad *et al.*, 2013). We look for more prominent use upcoming in the area of triage, especially with the recent successful innovations reported in both OR/OM and Emergency Medicine literatures in areas such as complexity-based triage (Saghafian *et al.*, 2014), disposition-based triage and patient streaming (King *et al.*, 2006; Kelly *et al.*, 2007; Kinsman *et al.*, 2008; Saghafian *et al.*, 2012), and telemedicine-based triage (Traub *et al.*, 2013).

Towards the back end of the ED service process, new processes involving reverse triage, while unestablished, show promise for future work. We believe focusing on the final part of ED flow, flow out of the ED, is an important research direction for future studies. In particular, there has not been enough focus on issues of outflow such as "bed block/access block" (Proudlove *et al.*, 2003; Khare *et al.*, 2009; Saghafian *et al.*, 2012; Shi *et al.*, 2013), and we believe that this should change in the near future, particularly as flow out of the ED is often the bottleneck, causing poor overall ED flow. With a shockingly low percentage of OR/OM work focusing on effective ways of moving patients out of the ED (see, e.g., Table 1), a paradigm shift should be imminent.

Process improvement methodologies with origins in automotive and electronics industry seek to further immerse themselves in healthcare. The use of Lean in healthcare, which was virtually nonexistent a decade ago, is beginning to become prominent (Jimmerson *et al.*, 2005; King *et al.*, 2006; Ben-Tovim *et al.*, 2008; Decker and Stead, 2008; Dickson *et al.*, 2009; Eller, 2009; Ng *et al.*, 2010; Holden, 2011; Piggott *et al.*, 2011). This has recently been coupled with Six Sigma to simultaneously eliminate waste and improve quality (Langabeer *et al.*, 2009; Berwald *et al.*, 2010). Both of these methodologies also stress the importance of a system-wide optimization approach that would greatly benefit ED operations. Hence, combining them with OR/OM techniques can be another fruitful path for research.

Strong correlation between ED and inpatient LOS suggests that continuing work in improving ED patient flow can have further impact on downstream operations. We also see in Kelen *et al.* (2001) and McConnell *et al.* (2005) how the addition of resources or focus on process improvement outside of the ED still has a benefit on traditional ED metrics. On a broader scale, decisions made by neighboring hospitals may also have a major impact on patient flow within the ED (Deo and Gurvich, 2011). The effect of providing better access to primary care on improving ED metrics is another research direction that deserves more study in the future. In general, having a comprehensive knowledge of the system beyond a narrow scope of the ED will improve operations across the board, and we expect to see more contributions from OR/OM researchers in this vein in the future. Additionally, successes in patient flow improvement processes need to be validated across multiple institutions. An overwhelming majority of process optimization research presents results on a single hospital. Results spanning multiple institutions such as Borders *et al.* (2009) show promise but lack geographic diversity. There is a definite need for research on multiple institutions in disparate geographic locations to gain broader insights on effective ED interventions.

We also hope to see further advancements in OR/OM tools that can better represent ED patient flow. One fruitful direction is to incorporate behavioral aspects in care delivery in OR/OM models, which are important for an accurate representation of patient flow but are currently largely overlooked. Future research may also continue to develop more advanced queueing models that better represent such complex care delivery and patient flow.

In the near term, we believe that information technology will help to move OR/OM forward in the ED. If "Big Data" in the ED refers to, *inter alia*, precise time stamps to record significant events (such as patient admission/discharge/transfer, patient movement through the ED, and order entry and completion), then Big Data will provide the raw materials needed to feed the theoretical constructs of OR/OM in order to develop appropriate solutions to extraordinarily complex problems. It is, therefore, essential to develop data-driven OR/OM techniques that can take advantage of this new opportunity.

In conclusion, we expect to see more contributions in the future from (i) close collaborations between OR/OM researchers and hospital stakeholders, (ii) innovative work for moving patients out of the ED, (iii) combining the process improvement methodologies with origins in automotive and electronics industry (e.g., Lean, Six Sigma, etc.) with OR/OM techniques, (iv) comprehensive views of the system (e.g., the effect of hospital inpatient units, neighboring hospitals, access to primary care, etc.), and (v) improved OR/OM patient flow models that benefit from data-driven and/or behavioral-driven approaches. If carefully developed and implemented, these can have significant impacts on ED operations.

Acknowledgments

The authors are grateful for invaluable comments and discussions provided by Wallace Hopp. The authors are also thankful for the helpful comments provided by the DE, anonymous AE, and referees.

Funding

This work was partially supported by Mayo Clinic through grant XSS0133.

References

Abelson, R. (2013) E.R.'s account for half of hospital admissions, study says. *The New York Times*, http://www.nytimes.com/2013/05/21/business/half-of-hospital-admissions-from-emergency-rooms.html?_r=0.

Adenso-Díaz, B., and Rodríguez, F. (1997) A simple search heuristic for the MCLP: Application to the location of ambulance bases in a rural region. *Omega - Int. J. Manage. S.*, **25**(2), 181–187.

Alavi-Moghaddam, M., Forouzanfar, R., Alamdari, S., Shahrami, A., Kariman, H., Amini, A., Pourbaba, S., and Shirvani, A. (2012) Application of queuing analytic theory to decrease waiting times in emergency department: Does it make sense? *Archives of Trauma Res.*, **1**(3), 101–107.

Allon, G., Deo, S., and Lin, W. (2013) The impact of size and occupancy of hospital on the extent of ambulance diversion: Theory and evidence. *Oper. Res.*, **61**(3), 544–562.

American Hospital Association. (2007) *Survey of Hospital Leaders*. American Hospital Association.

Andersson, T., and Värbrand, P. (2007) Decision support tools for ambulance dispatch and relocation. *J. Oper. Res. Soc.*, **58**, 195–201.

Armony, M., Israelit, S., Mandelbaum, A., Marmor, Y. N., Tseytlin, Y., and Yom-Tov, G. B. (2011) *Patient Flow in Hospitals: A Data-Based Queueing-Science Perspective*. Working Paper, New York University, New York.

Armony, M., and Mandelbaum, A. (2011) Routing and staffing in large-scale service systems: The case of homogeneous impatient customers and heterogeneous servers. *Oper. Res.*, **59**(1), 50–65.

Arostegui Jr, M. A., Kadipasaoglu, S. N., and Khumawala, B. M. (2006) An empirical comparison of tabu search, simulated annealing, and genetic algorithms for facilities location problems. *Int. J. Prod. Econ.*, **103**(2), 742–754.

Arrow, K. J., and Hurwicz, L. (1977) *Studies in Resource Allocation Processes*. Cambridge University Press, Cambridge, UK.

Asplin, B. R. (2003) Does ambulance diversion matter? *Can. Med. Assoc. J.*, **41**, 477–480.

Atallah, H. Y., and Lee, E. K. (2013) *Modeling and Optimizing Emergency Department Workflow*. Working Paper, Georgia Institute of Technology, Atlanta, GA.

Au, L., Byrnes, G. B., Bain, C. A., Fackrell, M., Brand, C., Campbell, D. A., and Taylor, P. G. (2009) Predicting overflow in an emergency department. *IMA J. Manage. Math.*, **20**, 39–49.

Au-Yeung, S. W. M., Harrison, P. G., and Knottenbelt, W. J. (2006) A queueing network model of patient flow in an accident and emergency department. *Proceedings of the 20th Annual European and Simulation Modelling Conference*. 60–67.

Au-Yeung, S. W. M., Harrison, P. G., and Knottenbelt, W. J. (2007) Approximate queueing network analysis of patient treatment times. *Proceedings of the 2nd International Conference on Performance Evaluation Methodologies and Tools*. 1–12.

Aytug, H., and Saydam, C. (2002) Solving large-scale maximum expected covering location problems by genetic algorithms: A comparative study. *Eur. J. Oper. Res.*, **141**, 480–494.

Badri, M. A., and Hollingsworth, J. (1993) A simulation model for scheduling in the emergency room. *Int. J. Oper. Prod. Manage.*, **13**, 13–24.

Baer, R. B., Pasternack, J. S., and Zwemmer Jr, F. L. (2001) Recently discharged inpatients as a source of emergency department overcrowding. *Acad. Emerg. Med.*, **8**(11), 1091–1094.

Bagust, A., Place, M., and Posnett, J. W. (1999) Dynamics of bed use in accommodating emergency admissions: stochastic simulation model. *Brit. Med. J.*, **319**, 155–158.

Ball, M. O., and Lin, F. L. (1993) A reliability model applied to emergency service vehicle location. *Oper. Res.*, **41**(1), 18–36.

Batt, R. J., and Terwiesch, C. (2012) *Doctors under load: An empirical study of state-dependent service times*. Working Paper, University of Pennsylvania, Philadelphia, PA.

Batt, R. J., and Terwiesch, C. (2013) *Waiting patiently: An empirical study of queue abandonment in an emergency department*. Working Paper, University of Pennsylvania, Philadelphia, PA.

Batta, R., Dolan, J. M., and Krishnamurthy, N. N. (1989) The maximal expected covering location problem: Revisited. *Transport. Sci.*, **23**(4), 277–287.

BBC News., (2002) Police improve heart attack survival. http://news.bbc.co.uk/2/hi/health/2188852.stm.

Beaulieu, H., Ferland, J. A., Gendron, B., and Michelon, P. (2000) A mathematical programming approach for scheduling physicians in the emergency room. *Health Care Manage. Sci.*, **3**, 193–200.

Bekker, R., and de Bruin, A. M. (2009) Time-dependent analysis for refused admissions in clinical wards. *Ann. Oper. Res.*, **178**(1), 45–65.

Bell, M. G. H. (2000) A game theory approach to measuring the performance reliability of transport networks. *Transport. Res. B-Meth.*, **34**, 533–545.

Ben-Tovim, D. I., Bassham, J. E., Bennett, D. M., Dougherty, M. L., Martin, M. A., O'Neill, S. J., Sincock, J. L., and Szwarcbord, M. G. (2008) Redesigning care at the Flinders Medical Centre: clinical process redesign using "lean thinking". *MJA*, **188**(6), S27–S31.

Berwald, N., Morisano, F., Ardolic, B., Silich, S., and Coleman, C. (2010) Reducing patient turnaround time in the emergency department using six sigma methodology. *Ann. Emerg. Med.*, **56**(3), S86–S87.

Beveridge, R. (1998) The Canadian Triage and Acuity Scale: A new and critical element in health care reform. *J. Emerg. Med.*, **16**(3), 507–511.

Black, D., and Pearson, M. (2002) Average length of stay, delayed discharge, and hospital congestion. *Brit. Med. J.*, **325**, 610–611.

Borders, M., Boyle, J., Turner, P., and Williams, M. (2009) *Reducing ambulance diversion in California: Strategies and best practices*. California Healthcare Foundation. http://www.chcf.org/∼/media/MEDIA%20LIBRARY%20Files/PDF/R/PDF%20ReducingAmbulanceDiversionInCA.pdf

Boudreaux, E. D., and O'Hea, E. L. (2004) Patient satisfaction in the emergency department: A review of the literature and implications for practice. *J. Emerg. Med.*, **26**(1), 13–26.

Brailsford, S. C. (2007) Advances and challenges in healthcare simulation modeling: Tutorial. *Proceedings of the 39th Winter Simulation Conference*, 1436–1448.

Brailsford, S. C., Harper, P. R., Patel, B., and Pitt, M. (2009) An analysis of the academic literature on simulation and modelling in health care. *J. of Simulation*, **3**, 130–140.

Brandeau, M., and Larson, R. C. (1986) *Extending and applying the hypercube queueing model to deploy ambulances in Boston: Delivery of urban services*. National Emergency Training Center, New York: North Holland.

Brandeau, M. L., and Chiu, S. S. (1989) An overview of representative problems in location research. *Manage. Sci.*, **35**(6), 645–674.

Breen, N., Woods, J., Bury, G., Murphy, A. W., and Brazier, H. (2000) A national census of ambulance response times to emergency calls in Ireland. *J. Accid. Emerg. Med.*, **17**, 392–395.

Brennan, J. A., Allin, D. M., Calkins, A. M., Enguidanos, E. R., Heimbach, Sr, L. J., Pruden, J. N., and Stilley, D. G. (2000) Guidelines for ambulance diversion. *Ann. Emerg. Med.*, **36**(4), 376–377.

Bretthauer, K. M., Heese, S., Pun, H., and Coe, E. (2011) Blocking in healthcare operations: A new heuristic and an application. *Prod. Oper. Manage.*, **20**(3), 375–391.

Brotcorne, L., Laporte, G., and Semet, F. (2003) Ambulance location and relocation models. *Eur. J. Oper. Res.*, **147**, 451–463.

Broyles, J. R., and Cochran, J. K. (2007) Estimating business loss to a hospital emergency department from patient reneging by queuing-based regression. *Proceedings of the 2007 Industrial Engineering Research Conference*, 613–618.

Burnett, M. G., and Grover, S. A. (1996) Use of the emergency department for nonurgent care during regular business hours. *Can. Med. Assoc. J.*, **154**(9), 1345–1351.

Burt, C. W., and McCaig, L. F. (2005) National Hospital Ambulatory Medical Care Survey: 2003 emergency department summary. *CDC Advance Data from Vital and Health Statistics*.

Burt, C. W., and McCaig, L. F. (2006) Staffing, capacity, and ambulance diversion in emergency departments: United States, 2003–04. *CDC Advance Data from Vital and Health Statistics*.

Burwell, T. H., Jarvis, J. P., and McKnew, M. A. (1993) Modeling co-located servers and dispatch ties in the hypercube model. *Comput. Oper. Res.*, **20**(2), 113–119.

Carret, M. L. V., Fassa, A. G., and Domingues, M. R. (2009) Inappropriate use of emergency services: a systematic review of prevalence and associated factors. *Cad. Saúde Pública*, **25**(1), 7–28.

Carter, G. M., Chaiken, J. M., and Ignall, E. (1972) Response areas for two emergency units. *Oper. Res.*, **20**(3), 571–594.

Carter, M. W., and Lapierre, S. D. (2001) Scheduling emergency room physicians. *Health Care Manage. Sci.*, **4**(4), 347–360.

Chan, T. C., Killeen, J. P., Kelly, D., and Guss, D. A. (2005) Impact of rapid entry and accelerated care at triage on reducing emergency department patient wait times, lengths of stay, and rate of left without being seen. *ACEP*, **46**(6), 491–497.

Chase, V. J., Cohn, A. E., Peterson, T. A., Lavieri, M. S. (2012) Predicting emergency department volume using forecasting methods to create a "surge response" for noncrisis events. *Acad. Emerg. Med.*, **19**(5), 569–576.

Chonde, S. J., Ashour, O. M., Nembhard, D. A., Okudan Kremer, G. E. (2013) Model comparison in emergency severity index level prediction. *Expert Syst. Appl.*, **40**, 6901–6909.

Church, R. L., and ReVelle, C. (1974) The maximal covering location problem. *Pap. Reg. Sci.*, **32**, 101–118.

CMS., (2013) Additional information regarding proposed eligible hospital and CAH clinical quality measures for 2014 EHR incentive programs. http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityMeasures/Downloads/Hospitals-and-CAH-2014-Proposed-EHR-Incentive-Program-CQM.pdf.

Coats, T. J., and Michalis, S. (2001) Mathematical modelling of patient flow through an accident and emergency department. *Emerg. Med. J.*, **18**, 190–192.

Cochran, J. K., and Bharti, A. (2006) Stochastic bed balancing of an obstetrics hospital. *Health Care Manage. Sci.*, **9**, 31–45.

Cochran, J. K., and Broyles, J. R. (2010) Developing nonlinear queuing regressions to increase emergency department patient safety: Approximating reneging with balking. *Comput. Ind. Eng.*, **59**, 378–386.

Cochran, J. K., and Roche, K. T. (2008) A queuing-based decision support methodology to estimate hospital inpatient bed demand. *J. Oper. Res. Soc.*, **59**(11), 1471–1482.

Cochran, J. K., and Roche, K. T. (2009) A multi-class queuing network analysis methodology for improving hospital emergency department performance. *Comput. Oper. Res.*, **36**(5), 1497–1512.

Cohn, A., Root, S., Kymissis, C., Esses, J., and Westmoreland, N. (2009) Scheduling medical residents at Boston University School of Medicine. *Interfaces*, **39**(3), 186–195.

Considine, J., Kropman, M., and Kelly, E. et al. (2008) Effect of emergency department fast track on emergency department length of stay: a case-control study. *Emerg. Med. J.*, **25**, 815–819.

Cooke, M. W., Wilson, S., and Pearson, S. (2002) The effect of a separate stream for minor injuries on accident and emergency department waiting times. *Emerg. Med. J.*, **19**, 28–30.

Cooper, L. (1963) Location-allocation problems. *Oper. Res.*, **11**(3), 331–343.

Côté, M. J. (2000) Understanding patient flow. *Decision Line*, **31**, 8–10.

Côté, M. J., and Stein, W. E. (2000) An erlang-based stochastic model for patient flow. *Omega*, **28**, 347–359.

Daskin, M. S. (1982) Application of an expected covering location model to emergency medical service system design. *Decision Sci.*, **13**(3), 416–439.

Daskin, M. S. (1983) A maximum expected covering location model: Formulation, properties and heuristic solution. *Transport. Sci.*, **17**(1), 48–70.

Davies, R., and Davies, H. T. O. (1994) Modelling patient flows and resource provision in health systems. *Omega - Int. J. Manage. S.*, **22**(2), 123–131.

Davis, E. W., and Patterson, J. H. (1975) A comparison of heuristic and optimum solutions in resource-constrained project scheduling. *Manage. Sci.*, **21**(8), 944–955.

de Bruin, A. M., van Rossum, A. C., Visser, M. C., and Koole, G. M. (2007) Modeling the emergency cardiac in-patient flow: an application of queuing theory. *Health Care Manage. Sci.*, **10**, 125–137.

de Véricourt, F., and Jennings, O. B. (2011) Nurse staffing in medical units: A queueing perspective. *Oper. Res.*, **59**(6), 1320–1331.

DeBehnke, D., and Decker, M. C. (2002) The effects of a physician-nurse patient care team on patient satisfaction in an academic ED. *Am. J. Emerg. Med.*, **20**(4), 267–270.

Decker, W. W., and Stead, L. G. (2008) Application of lean thinking in health care: a role in emergency departments globally. *Int. J. Emerg. Med.*, **1**, 161–162.

Deo, S., and Gurvich, I. (2011) Centralized vs. decentralized ambulance diversion: A network perspective. *Manage. Sci.*, **57**(7), 1300–1319.

Derlet, R. W., Kinser, D., Ray, L., Hamilton, B., and McKenzie, J. (1995) Prospective identification and triage of nonemergency patients out of an emergency department: A 5-year study. *Ann. Emerg. Med.*, **25**(2), 215–223.

Derlet, R. W., and Richards, J. R. (2008) Ten solutions for emergency department crowding. *Western J. Emerg. Med.*, **IX**(1), 2.

Dickson, E. W., Singh, S., Cheung, D. S., Wyatt, C. C., and Nugent, A. S. (2009) Application of lean manufacturing techniques in the emergency department. *J. Emerg. Med.*, **37**(2), 177–182.

Ding, R., McCarthy, M. L., Li, G., Kirsch, T. D., Jung, J. J., and Kelen, G. D. (2006) Patients who leave without being seen: Their characteristics and history of emergency department use. *Ann. Emerg. Med.*, **48**(6), 686–693.

Draeger, M. A. (1992) An emergency department simulation model used to evaluate alternative nurse staffing and patient population scenarios. *Proceedings of the 1992 Winter Simulation Conference*. 1057–1064.

Duguay, C., and Chetouane, F. (2007) Modeling and improving emergency department systems using discrete event simulation. *Simulation*, **83**(4), 311–320.

Dumas, M. B. (1984) Simulation modelling for hospital bed planning. *Simulation*, **43**, 69–78.

El-Darzi, E., Vasilakis, C., Chaussalet, T., and Millard, P. H. (1998) A simulation modelling approach to evaluating length of stay, occupancy, emptiness and bed blocking in a hospital geriatric department. *Health Care Manage. Sci.*, **1**, 143–149.

Elbeyli, S., and Krishnan, P. (2000) In-patient flow analysis using ProModel simulation package. University of Delaware, Newark, DE. FREC Staff Paper.

Eller, A. (2009) Rapid assessment and disposition: applying LEAN in the emergency department. *J. Health Care Qual.*, **31**(3), 17–22.

Epstein, S. K., and Tian, L. (2006) Development of an emergency department work score to predict ambulance diversion. *Acad. Emerg. Med.*, **13**, 421–426.

Erdogan, G., Erkut, E., Ingolfsson, A., and Laporte, G. (2010) Scheduling ambulance crews for maximum coverage. *J. Oper. Res. Soc.*, **61**, 543–550.

Falvo, T., Grove, L., and Stachura, R., et al. (2007) The financial impact of ambulance diversions and patient elopements. *Acad. Emerg. Med.*, **14**, 58–62.

Farrokhnia, N., and Goransson, K. E. (2011) Swedish emergency department triage and interventions for improved patient flows: a national update. *Scand. J. Trauma Resusc. Emerg. Med.*, **19**(72), 1–5.

Fatovich, D. M., and Hirsch, R. L. (2003) Entry overload, emergency department overcrowding, and ambulance bypass. *Emerg. Med. J.*, **20**, 406–409.

Fernandes, C. M. B., Price, A., and Christenson, J. M. (1997) Does reduced length of stay decrease the number of emergency department patients who leave without seeing a physician? *J. Emerg. Med.*, **15**(3), 397–399.

Fernandes, C. M. B., Tanabe, P., Gilboy, N., Johnson, L. A., McNair, R. S., Rosenau, A. M., Sawchuk, P., Thompson, D. A., Travers, D. A., Bonalumi, N., and Suter, R. E. (2005) Five-level triage: A report from the ACEP/ENA Five-Level Triage Task Force. *J. Emerg. Nurs.*, **31**(1), 39–50.

Fernandes, C. M. B., Wuerz, R., Clark, S., and Djurdjev, O. (1999) How reliable is emergency department triage? *Ann. Emerg. Med.*, **34**(2), 141–147.

Ferrand, Y., Magazine, M., Rao, U. S., and Glass, T. F. (2011) Building cyclic schedules for emergency department physicians. *Interfaces*, **41**(6), 521–533.

FitzGerald, G., Jelinek, G. A., Scott, D., Gerdtz, M. F. (2010) Emergency department triage revisited. *Emerg. Med. J.*, **27**, 86–92.

Fitzsimmons, J. A. (1973) A methodology for emergency ambulance deployment. *Manage. Sci.*, **19**(6), 627–636.

Fomundam, S., and Herrmann, J. (2007) *A Survey of Queuing Theory Applications in Healthcare*. Tech. rep., The Institute for Systems Research, A. James Clark School of Engineering, University of Maryland, College Park, MD.

Fone, D., Hollinghurst, S., Temple, M., Round, A., Lester, N., Weightman, A., Roberts, R., Coyle, E., Bevan, G., and Palmer, S. (2004) Systematic review of the use and value of computer simulation modelling in population health and health care delivery. *J. Public Health Med.*, **25**(4), 325–335.

Fullam, C. (2002) Acuity-based ED nurse staffing: A successful 5-year experience. *J. Emerg. Nurs.*, **28**(2), 138–140.

Gallivan, S., Utley, M., Treasure, T., and Valencia, O. (2002) Booked inpatient admissions and hospital capacity: mathematical modelling study. *Brit. Med. J.*, **324**, 280–282.

Galvão, R., Chiyoshi, F. Y., and Morabito, R. (2005) Towards unified formulations and extensions of two classical probabilistic location models. *Comput. Oper. Res.*, **32**, 15–33.

GAO. (2003) Hospital emergency departments: Crowded conditions vary among hospitals and communities. *GAO Report (GAO-03-460)*.

GAO. (2009) Hospital emergency departments: Crowding continues to occur, and some patients wait longer than recommended time frames. *GAO Report (GAO-09-347)*.

Gárcia, M. L., Rivera, C., Centeno, M. A., and DeCario, N. (1995) Reducing time in an emergency room via a fast-track. *Proceedings of the 1995 Winter Simulation Conference*, 1048–1053.

Geer, R., and Smith, J. (2004) Strategies to take hospitals off (revenue) diversion. *Health Care Financ. Manage.*, **58**(3), 70–74.

Gendrau, M., Laporte, G., and Semet, F. (2006) The maximal expected coverage relocation problem for emergency vehicles. *J. Oper. Res. Soc.*, **57**, 22–28.

Gendreau, M., Ferland, J., Gendron, B., Hail, N., Jaumard, B., Lapierre, S., Pesant, G., and Soriano, P. (2007) *Physician Scheduling in Emergency Rooms in Practice and Theory of Automated Timetabling VI*. Springer, Springer-Verlag Berlin Heidelberg.

Gendreau, M., Guertin, F., Potvin, J. Y., and Taillard, E. (1999) Parallel tabu search for real-time vehicle routing and dispatching. *Transport. Sci.*, **33**(4), 381–390.

Gendreau, M., Laporte, G., and Semet, F. (1997) Solving an ambulance location model by tabu search. *Loc. Sci.*, **5**(2), 75–87.

Gendreau, M., Laporte, G., and Semet, F. (2001) A dynamic model and parallel tabu search heuristic for real-time ambulance relocation. *Parallel Comput.*, **27**, 1641–1653.

Ghani, N. A. (2012) Multi-server queuing maximum availability location problem with stochastic travel times. *Proceedings of the World Congress on Engineering*, **1**, 137–143.

Goldberg, J. B. (2004) Operations research models for the deployment of emergency services vehicles. *EMS Manage. J.*, **1**(1), 20–39.

Goldman, J., Knappenberger, H. A., and Eller, J. C. (1968) Evaluating bed allocation policy with computer simulation. *Health Serv. Res.*, **3**(2), 119–129.

Görannson, K. E., Ehrenberg, A., and Ehnfors, M. (2005) Triage in emergency departments: National survey. *J. Clin. Nurs.*, **14**(9), 1067–1074.

Gottschalk, S. B., Wood, D., DeVries, S., Wallis, L. A., and Bruijns, S. (2006) The cape triage score: A new triage system South Africa. Proposal from the cape triage group. *Emerg. Med. J.*, **23**(2), 149–153.

Green, L. V. (2003) How many hospital beds? *Inquiry (Winter) - Blue Cross and Blue Shield Association*, 400–412.

Green, L. V. (2006) *Queueing Analysis in Healthcare, in Patient Flow: Reducing Delay in Healthcare Delivery*. Springer, New York.

Green, L. V. (2008) Using operations research to reduce delays for healthcare. *Tutorials in Operations Research*.

Green, L. V. (2012) OM Forum The vital role of operations analysis in improving healthcare delivery. *Manuf. Serv. Oper. Mang.*, **14**(4), 488–494.

Green, L. V., and Kolesar, P. J. (2004) Improving emergency responsiveness with management science. *Manage. Sci.*, **50**(8), 1001–1014.

Green, L. V., Kolesar, P. J., and Soares, J. (2001) Improving the SIPP approach for staffing service systems that have cyclic demands. *Oper. Res.*, **49**(4), 549–564.

Green, L. V., Kolesar, P. J., and Whitt, W. (2007) Coping with time-varying demand when setting staffing requirements for a service system. *Prod. Oper. Manage.*, **16**(1), 13–39.

Green, L. V., and Nguyen, V. (2001) Strategies for cutting hospital beds: The impact on patient service. *Health Serv. Res.*, **36**(2), 421–442.

Green, L. V., Soares, J., Giglio, J. F., and Green, R. A. (2006) Using queueing theory to increase the effectiveness of emergency department provider staffing. *Acad. Emerg. Med.*, **13**, 61–68.

Günal, M. M., and Pidd, M. (2010) Discrete event simulation for performance modelling in health care: A review of the literature. *J. of Simulation*, **4**, 42–51.

Hagtvedt, R., Ferguson, M., and Griffin, P., et al. (2009) Cooperative strategies to reduce ambulance diversion. *Proceedings of the 2009 Winter Simulation Conference*, 1861–1874.

Hall, R., Belson, D., Murali, P., and Dessouky, M. (2006) Modeling patient flows through the healthcare system, in *Patient Flow: Reducing Delay in Healthcare Delivery*. Springer, New York.

Hampers, L. C., Cha, S., Gutglass, D. J., Binns, H. J., Krug, S. E. (1999) Fast track and the pediatric emergency department: Resource utilization and patient outcomes. *Acad. Emerg. Med.*, **6**(11), 1153–1159.

Han, J. H., France, D. J., Levin, S. R., Jones, I. D., Storrow, A. B., and Aronsky, D. (2010) The effect of physician triage on emergency department length of stay. *J. Emerg. Med.*, **39**(2), 227–233.

Han, J. H., Zhou, C., and France, D. J., et al. (2007) The effect of emergency department expansion on emergency department overcrowding. *Acad. Emerg. Med.*, **14**(4), 338–343.

Handel, D. (2011) The case against diversion. *EP Monthly*, http://www.epmonthly.com/features/current-features/the-case-against-diversion/.

Harding, K. E., Taylor, N. F., and Leggat, S. G. (2011) Do triage systems in healthcare improve patient flow? A systematic review of the literature. *Aust. Health Rev.*, **35**, 371–383.

Harper, P. R., and Shahani, A. K. (2002) Modelling for the planning and management of bed capacities in hospitals. *J. Oper. Res. Soc.*, **53**, 11–18.

Harrison, G. W., Shafer, A., and MacKay, M. (2005) Modelling variability in hospital bed occupancy. *Health Care Manage. Sci.*, **8**, 325–334.

Hauswald, M. (2002) Can paramedics safely decide which patients do not need ambulance transport or emergency department care? *Prehosp. Emerg. Care*, **6**(4), 383–386.

Heckerling, P. S. (1985) Time study of an emergency room. Identification of sources of patient delay. *Illinois Med. J.*, **166**(6), 437–440.

Hemphill, R., and Nole, B. (2005) Relieving an overcrowded ED and increasing capacity for regional transfers: One hospital's bed management strategies. *J. Emerg. Nurs.*, **31**(3), 243–246.

Hick, J. L., Barbera, J. A., and Kelen, G. D. (2009) Refining surge capacity: Conventional, contingency, and crisis capacity. *Disaster Med. Public Health Prep.*, **3**(1), 59–67.

Hogan, K., and Revelle, C. (1986) Concepts and applications of backup coverage. *Manage. Sci.*, **32**(11), 1434–1444.

Holden, R. J. (2011) Lean thinking in emergency departments: A critical review. *Ann. Emerg. Med.*, **57**(3), 265–278.

Holdgate, A., Morris, J., Fry, M., and Zecevic, M. (2007) Accuracy of triage nurses in predicting patient disposition. *Emerg. Med. Australas.*, **19**, 341–345.

Hoot, N. R., LeBlanc, L. J., Jones, I., Levin, S. R., Zhou, C., Gadd, C. S., and Aronsky, D. (2008) Forecasting emergency department crowding: A discrete event simulation. *Ann. Emerg. Med.*, **52**(2), 116–125.

Hopp, W. J., and Lovejoy, W. S. (2013) *Hospital Operations, Principles of High Efficiency Health Care*. FT Press, Upper Saddle River, NJ.

Howell, E., Bessman, E., Kravet, S., Kolodner, K., Marshall, R., and Wright, S. (2008) Active bed management by hospitalists and emergency department throughput. *Ann. Intern Med.*, **149**, 804–810.

Howell, E., Bessman, E., Marshall, R., Wright, S. (2010) Hospitalist bed management effecting throughput from the emergency department to the intensive care unit. *Crit. Care*, **25**, 184–189.

Hu, B., and Benjaafar, S. (2009) Partitioning of servers in queueing systems during rush hour. *Manuf. Serv. Oper. Manage.*, **11**(3), 416–428.

Huang, J. (2013) Patient flow management in emergency departments. Ph.D. Dissertation, National University of Singapore.

Huang, J., Carmeli, B., and Mandelbaum, A. (2013) *Control of Patient Flow in Emergency Departments, or Multiclass Queues with Deadlines and Feedback*. Working Paper, National University of Singapore.

Huang, J. A., Tsai, W. C., Chen, Y. C., Hu, W. H., and Yang, D. Y. (2003) Factors associated with frequent use of emergency services in a medical center. *J. Formos. Med. Assoc.*, **102**(4), 345–353.

Huang, X. M. (1995) A planning model for requirement of emergency beds. *IMA J. Math. Appl. Med.*, **12**, 345–353.

Ieraci, S. (2003) Streaming by case complexity: Evaluation of a model for emergency department Fast Track. *Emergen. Med.*, **15**, 528–529.

Ieraci, S., Digiusto, E., Sonntag, P., Dann, L., and Fox, D. (2008) Streaming by case complexity: Evaluation of a model for emergency department Fast Track. *Emerg. Med. Australas.*, **20**, 241–249.

Ingolfsson, A., Budge, S., and Erkut, E. (2008) Optimal ambulance location with random delays and travel times. *Health Care Manage. Sci.*, **11**, 262–274.

Iserson, K. V., and Moskop, J. C. (2007) Triage in medicine, Part I: Concept, history, and types. *Ann. Emerg. Med.*, **49**(3), 275–287.

Jaslow, D., Barbera, J. A., Johnson, E., and Moore, W. (1998) Application of discrete-event simulation in health care clinics: A survey. *Prehosp. Emerg. Care*, **2**(1), 18–22.

Jimmerson, C., Weber, D., and Sobek II, D. K. (2005) Reducing waste and errors: piloting lean principles at Intermountain Healthcare. *Jt. Comm. J. Qual. Patient Saf.*, **31**(5), 249–257.

Jones, S. S., and Evans, R. S. (2008) An agent based simulation tool for scheduling emergency department physicians. *AMIA Annu. Symp. Proc.*, 338–342.

Jun, J. B., Jacobson, S. H., and Swisher, J. R. (1999) Application of discrete-event simulation in health care clinics: A survey. *J. Oper. Res. Soc.*, **50**, 109–123.

Kao, E. P. C., and Tung, G. G. (1981) Bed allocation in a public health care delivery system. *Manage. Sci.*, **27**(5), 507–520.

Kapadia, A. S., Chan, W., Sachdeva, R., Moye, L. A., and Jefferson, L. S. (2000) Predicting duration of stay in a pediatric intensive care unit: A Markovian approach. *Eur. J. Oper. Res.*, **124**, 353–359.

Karchere, A., and Hoeber, F. P. (1953) Combat problems, weapon systems, and the theory of allocation. *J. Oper. Res. Soc. Am.*, **1**(5), 286–302.

Karnon, J., Mackay, M., and Mills, T. M. (2009) Mathematical modelling in health care. *18th World IMACS / MODSIM Congress*. 44–56.

Keen, W. W. (1917) *The Treatment of War Wounds*. W.B. Saunders, PhiladelphiaPA.

Kelen, G. D., Kraus, C. K., McCarthy, M. L., Bass, E., Hsu, E. B., Li, G., Scheulen, J. J., Shahan, J. B., Brill, J. D., and Green, G. B. (2009a) Creation of surge capacity by early discharge of hospitalized patients at low risk for untoward events. *Disaster Med. Public Health Prep.*, **3**(1), S1–S7.

Kelen, G. D., McCarthy, M. L., Kraus, C. K., Ding, R., Hsu, E. B., Li, G., Shahan, J. B., Scheulen, J. J., and Green, G. B. (2009b) Inpatient disposition classification for the creation of hospital surge capacity: a multiphase study. *Lancet*, **368**, 1984–1990.

Kelen, G. D., Scheulen, J. J., and Hill, P. M. (2001) Effect of an emergency department (ED) managed acute care unit on ED overcrowding and emergency medical services diversion. *AEM*, **8**(11), 1095–1100.

Kelly, A-M, Bryant, M., Cox, L., and Jolley, D. (2007) Improving emergency department efficiency by patient streaming to outcomes-based teams. *Aust. Health Rev.*, **31**(1), 16–21.

Kelly, A-M, Kerr, D., Patrick, I., Walker, T. (2002) Benchmarking ambulance call-to-needle times for thrombolysis after acute myocardial infarction in Australia: a pilot study. *Internal Med.*, **32**, 138–142.

Kennedy, J., Rhodes, K., and Walls, C. A., et al. (2004) Access to emergency care: Restricted by long waiting times and cost and coverage concerns. *Ann. Emerg. Med.*, **43**(5), 567–573.

Khare, R. K., Powell, E. S., Reinhardt, G., and Lucenti, M. (2009) Adding more beds to the emergency department or reducing admitted patient boarding times: Which has a more significant influence on emergency department congestion? *Ann. Emerg. Med.*, **53**(5), 575–585.

Kim, S., Horowitz, I., Young, K. K., and Buckley, T. A. (1999) Analysis of capacity management of the intensive care unit in a hospital. *Eur. J. Oper. Res.*, **115**, 36–46.

King, D. L., Ben-Tovim, D. I., and Bassham, J. (2006) Redesigning emergency department patient flows: Application of lean thinking to health care. *Emerg. Med. Australas.*, **18**, 391–397.

Kinsman, L., Champion, R., Lee, G., Martin, M., Masman, K., May, E., Mills, T., Taylor, M. D., Thomas, P., Williams, R. J., and Zalstein, S. (2008) Assessing the impact of streaming in a regional emergency department. *Emerg. Med. Australas.*, **20**, 221–227.

Knapp, B. J., Kerns, B. L., Riley, I., and Powers, J. (2009) EMS-initiated refusal of transport: The current state of affairs. *J. Emerg. Med.*, **36**(2), 157–161.

Koizumi, N., Kuno, E., and Smith, T. E. (2005) Modeling patient flows using a queuing network with blocking. *Health Care Manage. Sci.*, **8**, 49–60.

Kokangul, A. (2008) A combination of deterministic and stochastic approaches to optimize bed capacity in a hospital unit. *Comput. Meth. Prog. Bio.*, **90**, 56–65.

Kolker, A. (2008) Process modeling of emergency department patient flow: Effect of patient length of stay on ED diversion. *J. Med. Syst.*, **32**, 389–401.

Komashie, A., and Mousavi, A. (2005) Modeling emergency departments using discrete event simulation techniques. *Proceedings of the 2005 Winter Simulation Conference*. 2681–2685.

Konrad, R., Desotto, K., Grocela, A., McAuley, P., Lyons, J., Bruins, M., and Wang, J. (2013) Evaluation of the split-flow concept in emergency departments using discrete-event simulation. *Oper. Res. for Health Care*, **2**(4), 66–74.

Kosowsky, J. M., Shindel, S., Liu, T., Hamilton, C., and Pancioli, A. M. (2001) Can emergency department triage nurses predict patients' dispositions? *Am. J. Emerg. Med.*, **19**(1), 10–14.

Kowalczyk, L. (2008) State orders ERs to halt diversions: Bid to ease overcrowding seen to sometimes delay care. *Boston Globe*.

Koziol-McLain, J., Price, D. W., Weiss, B., Quinn, A. A., and Honigman, B. (2000) Seeking care for nonurgent medical conditions in the emergency department: Through the eyes of the patient. *J. Emerg. Nurs.*, **26**(6), 554–563.

Kravet, S. J., Levine, R. B., Rubin, H. R., and Wright, S. M. (2007) Discharging patients earlier in the day: A concept worth evaluating. *Health Care Manage.*, **26**(2), 142–146.

Kwa, P., and Blake, D. (2008) Fast track: Has it changed patient care in the emergency department? *Emerg. Med. Australas.*, **20**, 10–15.

Lakshmi, C., and Appa Iyer, S. (2013) Application of queueing theory in health care: A literature review. *Oper. Res. Health Care*, **2**, 25–39.

Lane, D. C., Monefeldt, C., and Husemann, E. (2003) Client involvement in simulation model building: Hints and insights from a case study in a london hospital. *Health Care Mang. Sci.*, **6**(2), 105–116.

Lane, D. C., Monefeldt, C., and Rosenhead, J. V. (2000) Looking in the wrong place for healthcare improvements: A system dynamics study of an accident and emergency department. *J. of the Oper. Res. Soc.*, **51**(5), 518–531.

Langabeer, J. R., DelliFraine, J. L., Heineke, J., and Abbass, I. (2009) Implementation of Lean and Six Sigma quality initiatives in hospitals: A goal theoretic perspective. *Oper. Manage. Res.*, **2**, 13–27.

Larson, R. C. (1974) A hypercube queuing model for facility location and redistricting in urban emergency services. *Comput. Oper. Res.*, **1**(1), 67–95.

Larson, R. C. (1975) Approximating the performance of urban emergency service system. *Oper. Res.*, **23**(5), 845–868.

Laskowski, M., and Mukhi, S. (2009) Agent-based simulation of emergency departments with patient diversion. D. Weerasinghe, ed., *Electronic Healthcare, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering*, vol. 0001. Springer Berlin Heidelberg, 25–37.

Lawler, E. L., and Moore, J. M. (1969) A functional equation and its application to resource allocation and sequencing problems. *Manage. Sci.*, **16**(1), 77–84.

Lee, E. K., Atallah, H. Y., Wright, M. D., Post, E. T., Thomas IV, C., Wu, D. T., and Haley, L. L. (2014a) Transforming emergency department workflow and patient care. *Interfaces (To Appear)*.

Lee, E. K., Yuan, F., Zhou, R. L., Lahlou, Ss., Post, E., Wright, M., Atallah, H., and Haley, L. L. (2014b) Modeling and optimizing emergency department workflow of large urban public hospital. *Interfaces* (to appear).

Lee, M. H., Schuur, J. D., and Zink, B. J. (2013) Owning the cost of emergency medicine: Beyond 2%. *Annals of Emergency Medicine*, **62**(5), 498–505.

Lee, S. (2011) The role of preparedness in ambulance dispatching. *J. Oper. Res. Soc.*, **62**, 1888–1897.

Levine, S. D., Colwell, C. B., Pons, P. T., Gravitz, C., Haukoos, J. S., and McVaney, K. E. (2006) How well do paramedics predict admission to the hospital? a prospective study. *J. Emerg. Med.*, **31**(1), 1–5.

Lovejoy, W. S., and Desmond, J. S. (2011) Little's law flow analysis of observation unit impact and sizing. *Acad. Emerg. Med.*, **18**(2), 183–189.

Lowe, R. A., and Bindman, A. B. (1997) Judging who needs emergency department care: A prerequisite for policy-making. *Am. J. Emerg. Med.*, **15**(2), 133–136.

Lowthian, J. A., Cameron, P. A., Stoelwinder, J. U., Curtis, A., Currell, A., Cooke, M. W., and McNeil, J. J. (2011) Increasing utilisation of emergency ambulances. *AHR*, **35**, 63–69.

Macdonald, S. J., Karkam, I., Al-Shirrawi, N., Chowdhary, R. K., Escalante, E. M., and Afandi, A. (2005) Emergency department process improvement. *Proceedings of the 2005 Systems and Information Engineering Design Symposium*. 253–262.

Mandelbaum, A., Momcilovic, P., and Tseytlin, Y. (2012) On fair routing from emergency departments to hospital wards: QED queues with heterogeneous servers. *Manage. Sci.*, **58**(7), 1273–1291.

Marianov, V., and ReVelle, C. (1996) The queueing maximal availability location problem: A model for the siting of emergency vehicles. *Eur. J. Oper. Res.*, **93**, 110–120.

Marshall, A., Vasilakis, C., and El-Darzi, E. (2005) Length of stay-based patient flow models: Recent developments and future directions. *Health Care Manage. Sci.*, **8**, 213–220.

Martins, H. M., Cuña, L. M., and Freitas, P. (2009) Is Manchester (MTS) more than a triage system? A study of its association with mortality and admission to a large Portuguese hospital. *Emerg. Med. J.*, **26**(3), 183–186.

Maxwell, M. S., Restrepo, M., Henderson, S. G., and Topaloglu, H. (2010) Approximate dynamic programming for ambulance redeployment. *INFORMS J. Comput.*, **22**(2), 266–281.

Mayer, J. D. (1979) Emergency medical service: Delays, response time and survival. *Med. Care*, **17**(8), 818–827.

Mayhew, L., and Smith, D. (2008) Using queuing theory to analyse the governments 4-h completion time target in accident and emergency departments. *Health Care Manage. Sci.*, **11**, 11–21.

McConnell, K. J., Richards, C. F., and Daya M., et al. (2005) Effect of increased ICU capacity on emergency department length of stay and ambulance diversion. *Ann. Emerg. Med.*, **45**(5), 702–710.

McConnell, K. J., Richards, C. F., and Daya, M., et al. (2006) Ambulance diversion and lost hospital revenues. *Ann. Emerg. Med.*, **48**(6), 702–710.

McGuire, F. (1994) Using simulation to reduce length of stays in emergency department. *Proceedings of the 1994 Winter Simulation Conference*, 861–867.

Meislin, H. W., Coates, S. A., Cyr, J., and Valenzuela, T. (1988) Fast track: Urgent care within a teaching hospital emergency department: Can it work?*Ann. Emerg. Med.*, **17**(5), 453–456.

Mihal, N., and Moilanen, R. (2005) *When Emergency Rooms Close: Ambulance Diversion in the West San Fernando Valley*. Tech. rep., The Ralph and Goldy Lewis Center for Regional Policy Studies, University of California, Los Angeles.

Milne, E., and Whitty, P. (1995) Calculation of the need for paediatric intensive care beds. *Arch. Dis. Child.*, **73**, 505–507.

Miró, Ó., Sánchez, M., Espinosa, G., Coll-Vinent, B., Bragulat, E., and Millá, J. (2003) Analysis of patient flow in the emergency department and the effect of an extensive reorganisation. *Emerg. Med. J.*, **20**, 143–148.

Moskop, J. C., Sklar, D. P., Geiderman, J. M., Schears, R. M., and Bookman, K. J. (2009a) Emergency department crowding, Part 1 - Concept, causes, and moral consequences. *Ann. Emerg. Med.*, **53**(5), 605–611.

Moskop, J. C., Sklar, D. P., Geiderman, J. M., Schears, R. M., and Bookman, K. J. (2009b) Emergency department crowding, Part 2 - Barriers to reform and strategies to overcome them. *Ann. Emerg. Med.*, **53**(5), 612–617.

Murray, A. T., Tong, D., Kim, K. (2010) Enhancing classic coverage location models. *Int. Regional Sci. Rev.*, **33**(2), 115–133.

Nash, K., Zachariah, B., Nitschmann, J., and Psencik, B. (2007) Evaluation of the fast track unit of a university emergency department. *J. Emerg. Nurs.*, **33**(1), 14–20.

Ng, D., Vail, G., Thomas, S., and Schmidt, N. (2010) Applying the lean principles of the Toyota production system to reduce wait times in the emergency department. *Can. J. Emerg. Med.*, **12**(1), 50–57.

Nguyen, J. M., Six, P., Antonioli, D., Glemain, P., Potel, G., Lombrail, P., and Le Beux, P. (2005) A simple method to optimize hospital beds capacity. *Int. J. Med. Inform.*, **74**, 39–49.

Northington, W. E., Brice, J. H., Zou, B. (2005) Use of an emergency department by nonurgent patients. *Am. J. Emerg. Med.*, **23**, 131–137.

Nunes, L.G. Nadal, Venancio de Carvalho, S., de Cassia Menses Rodrigues, R. (2009) Markov decision process applied to the control of hospital elective admissions. *Artif. Intell. Med.*, **47**, 159–171.

O'Brien, D., Williams, A., Blondell, K., and Jelinek, G. A. (2006) Impact of streaming "fast track" emergency department patients. *AHR*, **30**(4), 525–532.

Olshaker, J. S., and Rathlev, N. K. (2006) Emergency department overcrowding and ambulance diversion: The impact and potential solutions of extended boarding of admitted patients in the emergency department. *J. Emerg. Med.*, **30**(3), 351–356.

Oredsson, S., Jonsson, H., Rognes, J., Lind, L., Goransson, K. E., Ehrenberg, A., Aspluind, K., Carstrén, M., and Farrohknia, N. (2011) A systematic review of triage-related interventions to improve patient flow in emergency departments. *Scand. J. Trauma Resusc. Emerg. Med.*, **19**(43), 996–1007.

Osorio, C., and Bierlaire, M. (2009) An analytic finite capacity queueing network model capturing the propagation of congestion and blocking. *Eur. J. Oper. Res.*, **9**, 996–1007.

Ozcan, Y. (2009) *Quantitative Methods in Health Care Management, Techniques and Applications*. Wiley (Jossey-Bass), San Francisco, CA.

Partovi, S. N., Nelson, B. K., Bryan, E. D., and Walsh, M. J. (2001) Faculty triage shortens emergency department length of stay. *Acad. Emerg. Med.*, **8**(10), 990–995.

Patel, P. B., Derlet, R. W., and Vinson, D. R., et al. (2006) Ambulance diversion reduction: the Sacramento solution. *Am. J. Emerg. Med.*, **24**, 206–213.

Patel, P. B., and Vinson, D. R. (2005) Team assignment system: Expediting emergency department care. *Am. J. Emerg. Med.*, **46**(6), 499–506.

Paul, S. A., Reddy, M. C., and DeFlitch, C. J. (2010) A systematic review of simulation studies investigating emergency department overcrowding. *Simulation*, **86**(8-9), 559–571.

PCAST. (2014) Report to the President: Better health care and lower costs: Accecelerating improvement through systems engineering. *Executive Office of the President Presidents Council of Advisors on Science and Technology Report*.

Peck, J., and Kim, S.-G. (2009) Improving patient flow through axiomatic design of hospital emergency departments. *Proceedings of the 19th CIRP Design Conference - Competitive Design*, 451–455.

Peck, J. S., Benneyan, J. C., Nightingale, D. J., and Gaehde, S. A. (2012) Predicting emergency department inpatient admissions to improve same-day patient flow. *Acad. Emerg. Med.*, **19**, 1045–1054.

Peleg, K., and Pliskin, J. S. (2004) A geographic information system simulation model of EMS: Reducing ambulance response time. *Am. J. Emerg. Med.*, **22**(3), 164–170.

Peters, J., and Hall, G. B. (1999) Assessment of ambulance response performance using a geographic information system. *Social Science and Medicine*, **49**, 1551–1566.

Pham, J. C., Patel, R., and Millin, M. G., et al. (2006) The effects of ambulance diversion: A comprehensive review. *Acad. Emerg. Med.*, **13**, 1220–1227.

Piggott, Z., Weldon, E., Strome, T., and Chochinov, A. (2011) Application of Lean principles to improve early cardiac care in the emergency department. *Can. J. Emerg. Med.*, **13**(5), 325–332.

Pines, J. M., Batt, R. J., Hilton, J. A., and Terwiesch, C. (2011) The financial consequences of lost demand and reducing boarding in hospital emergency departments. *Ann. Emerg. Med.*, **58**(4), 331–340.

Pitts, S. R., Carrier, E. R., and Rich, E. C., et al. (2010) Where Americans get acute care: Increasingly, it's not at their doctor's office. *Health Aff (Millwood)*, **29**, 91620–91629.

Pitts, S. R., Niska, R. W., Xu, J., and Burt, C. W. (2008) National Hospital Ambulatory Medical Care Survey: 2006 emergency department summary. *National Health Statistics Report*, **7**, 1–39.

Plati, C., Lemonidou, C., Priami, M., Baltopoulos, G., and Mantas, J. (1996) The intensive care units in greater Athens: needs and resources. *Intensive and Critical Care Nursing*, **12**, 340–345.

Pointer, J. E., Levitt, M. A., Young, J. C., Promes, S. B., Messana, B. J., and Adér, M. E. J. (2001) Can paramedics using guidelines accurately triage patients?*Ann. Emerg. Med.*, **38**(3), 268–277.

Powell, E. S., Khare, R. K., Venkatesh, A. K., Van Roo, B. D., Adams, J. G., and Reinhardt, G. (2012) The relationship between inpatient discharge timing and emergency department boarding. *J. Emerg. Med.*, **42**(2), 186–196.

Price, T. G., Hooker, E. A., and Neubauer, J. (2005) Prehospital provider prediction of emergency department disposition: Implications for selective diversion. *Prehosp. Emerg. Care*, **9**(3), 322–325.

Proudlove, N., Boaden, R., and Jorgensen, J. (2007) Developing bed managers: the why and the how. *J. Nurs. Manage.*, **15**, 34–42.

Proudlove, N. C., Gordon, K., and Boaden, R. (2003) Can good bed management solve the overcrowding in accident and emergency departments?*Emerg. Med. J.*, **20**, 149–155.

Purnell, L. D. (1991) A survey of emergency department triage in 185 hospitals: physical facilities, fast-track systems, patient-classification systems, waiting times, and qualification, training, and skills of triage personnel. *J. Emerg. Nurs.*, **17**(6), 402–407.

Rajagopalan, H. K., Saydam, C., Sharer, E., and Setzler, H. (2011) Ambulance deployment and shift scheduling: An integrated approach. *J. Serv. Sci. Manage.*, **4**, 66–78.

Rajagopalan, H. K., Saydam, C., Xiao, J. (2008) A multiperiod set covering location model for dynamic redeployment of ambulances. *Comput. Oper. Res.*, **35**, 814–826.

Ramakrishnan, M., Sier, D., and Taylor, P. G. (2005) A two-time-scale model for hospital patient flow. *IMA J. Manage. Math.*, **16**(3), 197–215.

Ramirez, A., Fowler, J. W., and Wu, T. (2009) Analysis of ambulance diversion policies for a large-size hospital. *Proceedings of the 2009 Winter Simulation Conference*, 1875–1886.

Ramirez-Nafarrate, A., Fowler, J. W., Wu, T. (2011) Design of centralized ambulance diversion policies using simulation-optimization. *Proceedings of the 2011 Winter Simulation Conference*. 1251–1262.

Ramirez-Nafarrate, A., Hafizoglu, A. B., Gel, E. S., and Fowler, J. W. (2012) Comparison of ambulance diversion policies via simulation. *Proceedings of the 2012 Winter Simulation Conference*, 967–978.

Repede, J. F., and Bernardo, J. J. (1994) Developing and validating a decision support system for locating emergency medical vehicles in Louisville, Kentucky. *Eur. J. Oper. Res.*, **75**, 567–581.

ReVelle, C., and Hogan, K. (1989) The maximum availability location problem. *Transport. Sci.*, **23**(3), 192–200.

Richards, J. R., Ferrall, S. J. (1999) Inappropriate use of emergency medical services transport: Comparison of provider and patient perspectives. *Acad. Emerg. Med.*, **6**(1), 14–20.

Richardson, D. (1998) No relationship between emergency department activity and triage categorization. *Acad. Emerg. Med.*, **5**, 141–145.

Richardson, D. B. (2003) Reducing patient times in emergency department. *Med. J. Aust.*, **179**(17), 516–517.

Ridge, J. C., Jones, S. K., Nielsen, M. S., and Shahani, A. K. (1998) Capacity planning for intensive care units. *Eur. J. Oper. Res.*, **105**, 346–355.

Roberts, R. R., Frutos, P. W., Ciavarella, G. C., Gussow, L. M., Mensah, E. K., Kampe, L. M., Straus, H. E., Joseph, G., and Rydman, R. J. (1999) Distribution of Variable vs Fixed Costs of Hospital Care. *J. Amer. Med. Assoc.*, **281**(7), 644–649.

Roche, K. T., and Cochran, J. K. (2007) Improving patient safety by maximizing fast-track benefits in the emergency department: A queuing network approach. *Proceedings of the 2007 Industrial Engineering Research Conference*, 619–624.

Rodi, S. W., Grau, M. V., and Orsini, C. M. (2006) Evaluation of a fast track unit: Alignment of resources and demand results in improved satisfaction and decreased length of stay for emergency department patients. *Qual. Manage. Health Care*, **15**(3), 163–170.

Rosetti, M. D., Trzcinski, G. F., and Syverud, S. A. (1999) Emergency department simulation and determination of optimal attending physician staffing schedules. *Proceedings of the 1999 Winter Simulation Conference*, 1532–1540.

Ruohonen, T., Neittaanmaki, P., and Teittinen, J. (2006) Simulation model for improving the operation of the emergency department of special health care. *Proceedings of the 2006 Winter Simulation Conference*, 453–458.

Russ, S., Jones, I., Aronsky, D., Dittus, R. S., and Slovis, C. M. (2010) Placing physician orders at triage: The effect on length of stay. *Ann. Emerg. Med.*, **56**(1), 27–33.

Saghafian, S., Hopp, W. J., Van Oyen, M. P., Desmond, J. S., and Kronick, S. L. (2012) Patient streaming as a mechanism for improving responsiveness in emergency departments. *Oper. Res.*, **60**(5), 1080–1097.

Saghafian, S., Hopp, W. J., Van Oyen, M. P., Desmond, J. S., and Kronick, S. L. (2014) Complexity-augmented triage: A tool for improving patient safety and operational efficiency. *Manuf. Serv. Oper. Mang.*, **16**(3), 329–345.

Samaha, S., Armel, W. S., and Starks, D. W. (2003) The use of simulation to reduce the length of stay in an emergency department. *Proceedings of the 2003 Winter Simulation Conference*, 1907–1911.

Samuels-Kalow, M. E., Stack, A. M., and Porter, S. C. (2012) Effective discharge communication in the emergency department. *Ann. Emerg. Med.*, **60**(2), 152–159.

Sanchez, M., Smally, A. J., Grant, R. J., and Jacobs, L. M. (2006) Effects of a fast-track area on emergency department performance. *J. Emerg. Med.*, **31**(1), 117–120.

Saunders, C. E. (1987) Time study of patient movement through the emergency department: Sources of delay in relation to patient acuity. *Ann. Emerg. Med.*, **16**(11), 85–89.

Saunders, C. E., Makens, P. K., Leblanc, L. J. (1989) Modeling emergency department operations using advanced computer simulation systems. *Ann. Emerg. Med.*, **18**(2), 134–140.

Savas, E. S. (1969) Simulation and cost-effectiveness analysis of New York's emergency ambulance service. *Manage. Sci.*, **15**(12), B608–B627.

Schaefer, R. A., Rea, T. D., Plorde, M., Peiguss, K., Goldberg, P., Murray, J. A. (2000) Evaluation of protocols allowing emergency medical technicians to determine need for treatment and transport. *Acad. Emerg. Med.*, **7**(6), 663–669.

Schaefer, R. A., Rea, T. D., Plorde, M., Peiguss, K., Goldberg, P., and Murray, J. A. (2002) An emergency medical services program of alternate destination of patient care. *Prehosp. Emerg. Care*, **6**(3), 309–314.

Schneider, S., Zwemer, F., Doniger, A., Dick, R., Czapranski, T., and Davis, E. (2001) Rochester, New York: A decade of emergency department overcrowding. *Acad. Emerg. Med.*, **8**(11), 1044–1050.

Schneider, S. M., Galley, M. E., Schafermeyer, R., Zwemer, F. L. (2003) Emergency department crowding: A point in time. *Ann. Emerg. Med.*, **42**(2), 167–172.

Schull, M. J., Lazier, K., and Vermeulen, M., et al. (2003) Emergency department contributors to ambulance diversion: A quantitative analysis. *Ann. Emerg. Med.*, **41**(4), 467–476.

Schuur, J. D., and Venkatesh, A. K. (2012) The growing role of emergency departments in hospital admissions. *New England J. of Medicine*, **367**, 391–393.

Sempere-Selva, T., Peiró, S., Sendra-Pina, P., Martínez-Espín, C., and López-Aguilera, I. (2001) Inappropriate use of an accident and emergency department: Magnitude, associated factors, and reasons an approach with explicit criteria. *Ann. Emerg. Med.*, **37**(6), 568–579.

Shi, P., Chou, M. C., Dai, J. G., Ding, D., and Sim, J. (2013) *Hospital inpatient operations: Mathematical models and managerial insights*. Working Paper, Georgia Institute of Technology.

Simon, H. K., McLario, D., Daily, R., Lanese, C., Castillo, J., and Wright, J. (1996) "Fast tracking" patients in an urban pediatric emergency department. *Am. J. Emerg. Med.*, **14**(3), 242–244.

Singer, M., and Donoso, P. (2008) Assessing an ambulance service with queuing theory. *Comput. Oper. Res.*, **35**, 2549–2560.

Sinreich, D., and Jabali, O. (2007) Staggered work shifts: a way to downsize and restructure an emergency department workforce yet maintain current operational performance. *Health Care Manage. Sci.*, **10**, 293–308.

Sinreich, D., Jabali, O., and Dellaert, N. P. (2012) Reducing emergency department waiting times by adjusting work shifts considering patient visits to multiple care providers. *IIE Trans.*, **44**, 163–180.

Sinreich, D., and Marmor, Y. (2005) Emergency department operations: The basis for developing a simulation tool. *IIE Trans.*, **37**(3), 233–245.

Snooks, H. A., Dale, J., Hartley-Sharpe, C., and Halter, M. (2004) On-scene alternatives for emergency ambulance crews attending patients who do not need to travel to the accident and emergency department: a review of the literature. *Emerg. Med. J.*, **21**, 212–215.

Sprivulis, P. (2004) Pilot study of metropolitan emergency department workload complexity. *Emerg. Med. Australas.*, **16**, 59–64.

Stainsby, H., Taboada, M., Luque, E. (2009) Towards an agent-based simulation of hospital emergency departments. *Proc. of IEEE International Conference on Services Computing*. 536–539.

Stoykova, B., Dowie, R., Bastow, P., Roswell, K. V., and Gregory, R. P. F. (2004) Ambulance emergency services for patients with coronary heart disease in Lancashire: achieving standards and improving performance. *Emerg. Med. J.*, **21**, 99–104.

Sun, B. C., Mohanty, S. A., and Weiss, R., et al. (2006) Effects of hospital closures and hospital characteristics on emergency department ambulance diversion, Los Angeles County, 1998 to 2004. *Ann. Emerg. Med.*, **47**(4), 309–316.

Swoveland, C., Uyeno, D., Vertinsky, I., and Vickson, R. (1973) Ambulance location: A probabilistic enumeration approach. *Manage. Sci.*, **20**(4), 686–698.

Takeda, R. A., Widmer, J. A., and Morabito, R. (2007) Analysis of ambulance decentralization in an urban emergency medical service using the hypercube queueing model. *Comput. Oper. Res.*, **34**, 727–741.

Thompson, D. A., Yarnold, P. R., Williams, D. R., and Adams, S. L. (1996) Effects of actual waiting time, perceived waiting time, information delivery, and expressive quality on patient satisfaction in the emergency department. *Ann. Emerg. Med.*, **29**(6), 657–665.

Traub, S. J., Butler, R., Chang, Y. H., Lipinski, C. (2013) Emergency department physician telemedical triage. *Telemed J. E. Health.*, **19**(11), 841–845.

Traub, S. J., Stewart, C., Didehban, R., Nestler, D., Chang, Y. H., Saghafian, S., and Lipinski, C. A. (2014a) *Effects of Rotational Patient Assignment on Emergency Department Operational Metrics and Equity of Physician Workload*. Working Paper, Mayo Clinic.

Traub, S. J., Stewart, C., Didehban, R., Nestler, D., Chang, Y. H., Saghafian, S., and Lipinski, C. A. (2014b) Emergency department rapid medical assessment: Overall effect and mechanistic consideration. *J. of Emergency Medicine* (to appear).

Trzeciak, S., and Rivers, E. P. (2003) Emergency department overcrowding in the United States: an emerging threat to patient safety and public health. *Emerg. Med. J.*, **20**, 402–405.

van der Wulp, I., van Baar, M. E., and Schrijvers, A. J. (2008) Reliability and validity of the Manchester triage system in a general emergency department patient population in the Netherlands: results of a simulation study. *Emerg. Med. J.*, **25**(7), 431–434.

Vance, J., and Sprivulis, P. (2005) Triage nurses validly and reliably estimate emergency department patient complexity. *Emerg. Med. Australas.*, **17**, 382–386.

Vermeulen, M. J., Ray, J. G., Bell, C., Cayen, B., Stukel, T. A., and Schull, M. J. (2009) Disequilibrium between admitted and discharged hospitalized patients affects emergency department length of stay. *Ann. Emerg. Med.*, **54**(6), 794–804.

Vertesi, L. (2004) Does the Canadian Emergency Department Triage and Acuity Scale identify non-urgent patients who can be triaged away from the emergency department?*Can. J. Emerg. Med.*, **6**(5), 337–342.

Vilke, G. M., Brown, L., and Skogland, P., et al. (2004) Approach to decreasing emergency department ambulance diversion hours. *J. Emerg. Med.*, **26**(2), 189–192.

Wang, J., Li, J., and Howard, P. K. (2013) A system model of work flow in the patient room of hospital emergency department. *Health Care Manage. Sci.*, **16**(4), 341–351.

Wang, J., Li, J., Tussey, K., Ross, K. (2012) Reducing length of stay in emergency department: A simulation study at a community hospital. *IEEE T. Syst. Man. Cy. A*, **42**(6), 1314–1322.

Wang, Q. (2004) Modeling and analysis of high risk patient queues. *Eur. J. Oper. Res.*, **155**, 502–515.

Wang, X. (2013) Emergency department staffing: A separated continuous linear programming approach. *Math. Problems in Eng,* 680152.

Warden, C. R., Bangs, C., Norton, R., and Huie, J. (2003) Temporal trends in ambulance diversion in a mid-sized metropolitan area. *Prehosp. Emerg. Care*, **7**(1), 109–113.

Warner, M. (2006) Personnel Staffing and Scheduling, in *Patient Flow: Reducing Delay in Healthcare Delivery*. Springer, New York.

Washington, D. L., Stevens, C. D., Shekelle, P. G., Baker, D. W., Fink, A., and Brook, R. H. (2000) Safely directing patients to appropriate levels of care: Guideline-driven triage in the emergency service. *Ann. Emerg. Med.*, **36**(1), 15–22.

Welch, S. (2009) Patient segmentation: Redesigning flow. *Emerg. Med. News*, **31**(8).

Welch, S., Augustine, J., Camargo Jr, C. A., and Reese, C. (2006) Emergency department performance measures and benchmarking summit. *Acad. Emerg. Med.*, **13**(10), 1074–1080.

Welch, S. J., Asplin, B. R., Stone-Griffith, S., Davidson, S. J., Augustine, J., and Schuur, J. (2011) Emergency department operational metrics, measures and definitions: Results of the second Performance measures and benchmarking summit. *Ann. Emerg. Med.*, **58**(1), 33–40.

White, K. P., Jr, (2005) A survey of data resources for simulating patient flows in healthcare delivery systems. *Proceedings of the 2005 Winter Simulation Conference*, 926–935.

Whitt, W. (1999) Partitioning customers into service groups. *Manage. Sci.*, **45**(11), 1579–1592.

Whitt, W. (2005) Engineering solution of a basic call-center model. *Manage. Sci.*, **51**, 221–235.

WHO. (2010) *Health Expenditure, Total (% of GDP)*. http://data.worldbank.org/indicator/SH.XPD.TOTL.ZS.

Wiler, J. L., Bolandifar, E., Griffey, R. T., Poirier, R. F., and Olsen, T. (2013) An emergency department patient flow model based on queueing theory principles. *Acad. Emerg. Med.*, **20**(9), 939–946.

Wiler, J. L., Gentle, C., Halfpenny, J. M., Heins, A., Mehrotra, A., Mikhail, M. G., and Fite, D. (2010) Optimizing emergency department front-end operations. *Ann. Emerg. Med.*, **55**(2), 142–160.

Wiler, J. L., Griffey, R. T., and Olsen, T. (2011) Review of modeling approaches for emergency department patient flow and crowding research. *Acad. Emerg. Med.*, **18**(12), 1371–1379.

Williams, M. (2006) *Hospitals and clinical facilities, processes and design for patient flow, in patient flow: Reducing delay in healthcare delivery*. Springer, New York.

Worthington, D. J. (1987) Queueing models for hospital waiting lists. *J. Oper. Res. Soc.*, **38**(5), 413–422.

Wright, S. W., Erwin, T. L., Blanton, D. M., and Covington, C. M. (1992) Fast track in the emergency department: A one-year experience with nurse practitioners. *J. Emerg. Med.*, **10**, 367–373.

Wuerz, R., Fernandes, C. M. B., and Alarcon, J. (1998) Inconsistency of emergency department triage. *Ann. Emerg. Med.*, **32**(4), 431–435.

Wuerz, R., Milne, L. W., Eitel, D. R., Travers, D., and Gilboy, N. (2000) Reliability and validity of a new five-level triage instrument. *Acad. Emerg. Med.*, **7**(3), 236–242.

Yankovic, N., and Green, L. V. (2011) Identifying good nursing levels: A queuing approach. *Oper. Res.*, **59**(4), 942–955.

Young, G. P., Wagner, M. B., Keilermann, A. L., Ellis, J., and Bouley, D. (1996) Ambulatory visits to hospital emergency departments patterns and reasons for use. *J. Amer. Med. Assoc.*, **276**(6), 460–465.

Youssef, H., Sait, S. M., and Adiche, H. (2001) Evolutionary algorithms, simulated annealing and tabu search: a comparative study. *Eng. Appl. Artif. Intel.*, **14**, 167–181.

Zayas-Caban, G., Xie, J., Green, L. V., and Lewis, M. (2013) *Optimal control of an emergency room triage and treatment process*. Working Paper.

Zeltyn, S., Marmor, Y. N., Mandelbaum, A., Carmeli, B., Greenshpan, O., Mesika, Y., Wasserkrug, S., Vortman, P., Shtub, A., Lauterman, A., Schwartz, D., Moskovitch, K., Tzafrir, S., and Basis, F. (2011) Simulation-based models of emergency departments: Operational, tactical, and strategic staffing. *ACM T. Model. Comput. S.*, **21**(4).

Zilm, F., Crane, J., and Roche, K. T. (2010) New directions in emergency service operations and planning. *J Ambulatory Care Mang.*, **33**(4), 296–306.