

DuBois, Eric; Schmidt, Adam; Albert, Laura A.

Article

Location of trauma care resources with inter-facility patient transfers

Operations Research Perspectives

Provided in Cooperation with:

Elsevier

Suggested Citation: DuBois, Eric; Schmidt, Adam; Albert, Laura A. (2021) : Location of trauma care resources with inter-facility patient transfers, Operations Research Perspectives, ISSN 2214-7160, Elsevier, Amsterdam, Vol. 8, pp. 1-17, <https://doi.org/10.1016/j.orp.2021.100206>

This Version is available at:

<https://hdl.handle.net/10419/325708>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

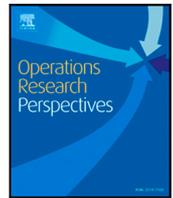
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Location of trauma care resources with inter-facility patient transfers

Eric DuBois, Adam Schmidt, Laura A. Albert*

University of Wisconsin-Madison, Industrial and Systems Engineering, 1513 University Avenue, WI 53706, United States of America

ARTICLE INFO

Keywords:

Location analysis
Emergency medicine
Trauma care system
Optimization

ABSTRACT

Traumatic injury is the third leading cause of death in the United States and is the leading cause of death among people aged 1–44. Regional trauma care systems can significantly reduce this mortality rate if resources are located and used appropriately. Trauma centers are categorized as able to either provide total care (Level I or II) or prompt care (Level III and IV). We study trauma system planning by introducing an integer programming model to optimally relocate existing air ambulances, locate new air ambulances, and upgrade Level III and IV trauma centers to Level I or II trauma centers within a mature trauma system. The goal of these models to improve timely access to care and ensure the appropriate level of care is provided to severely injured patients. A novel aspect of the modeling approach is the inclusion of patient transfers from Level III/IV trauma centers to Level I/II trauma centers. Patient transfers are of particular importance after mass casualty events when a surge of trauma patients can overwhelm local trauma center capacity. As a result, we introduce a mass casualty incident response stochastic programming model to optimally relocate air ambulances within the trauma system following a mass casualty incident but before patient demand is known with certainty. The two models inform planning decisions by considering the interrelated response decisions. We construct case studies for both models based on the Wisconsin, USA trauma care system. The results highlight the importance of relocating air ambulances for increasing the number of patients transported directly to a Level I/II trauma center. The results also indicate that considering patient transfers within the model can significantly alter the optimal locations of trauma care resources compared to when transfers are not considered.

1. Introduction

Traumatic injury is the third leading cause of death in the United States and is the leading cause of death among people aged 1–44 [1]. Regional trauma care systems have been shown to significantly reduce the mortality rate from traumatic injury [2]. However, access to trauma care is not equally distributed. Rural patients are typically far from a trauma center, and they have a 14% higher likelihood of death resulting from traumatic injury [3].

Within the United States, most states designate their trauma centers as one of four levels according to the American College of Surgeons (ACS) guidelines. Each level corresponds to the availability of specialized personnel and equipment to meet the needs of severely injured patients. The levels and a summary of their definitions are:

Level I facilities are regional trauma centers capable of providing advanced, specialized care for trauma patients. Level I facilities are expected to take a leadership role in their region and conduct research and education.

Level II facilities are “expected to be clinically equivalent [to Level I facilities] except for complex, specialized injuries such as

replantation” [4]. Level II facilities do not have a research and education requirement. They are expected to take a leadership role if no Level I facilities are located nearby.

Level III facilities are able to initially treat most patients and transfer more severely injured patients to Level I and II facilities.

Level IV facilities are only able to provide initial evaluation for most patients. Even more moderately injured patients are transferred to Level I and II facilities.

Trauma centers may be roughly divided into two categories: high-level trauma center facilities (Level I and II) that usually transfer patients in from other trauma centers, and low-level facilities (Level III and IV) that usually transfer patients out to other trauma centers.

Fig. 1 illustrates the potential pathways a patient may take within a trauma system to receive definitive care at an appropriate trauma center. After a patient is stabilized on-scene, either an air ambulance or a ground ambulance may *transport* the patient to a low- or high-level trauma center. Minimally injured patients may be treated at

* Corresponding author.

E-mail address: laura@engr.wisc.edu (L.A. Albert).

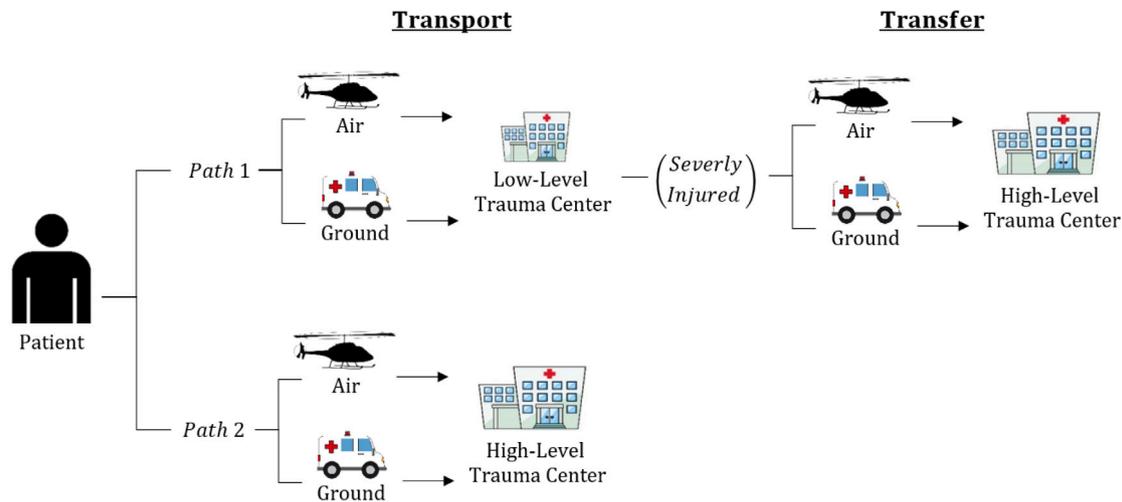


Fig. 1. Possible patient pathways through a trauma system.

either a low- or high-level trauma center. Severely injured patients should receive definitive trauma care at high-level trauma center [4]. Due to insufficient resources, severely injured patients are often first transported to low-level trauma centers to provide stabilization, and they must later be *transferred*, either by air or ground ambulance, to a high-level trauma center. While this may not represent every possible care pathway, it captures the typical patient pathways, and it highlights the importance of patient transfers in the treatment pathways. Note that when referring to ‘severely’ or ‘minimally’ injured patients, we consider them relative terms without direct correspondence to injury score [4].

The pathways through the trauma system are situationally dependent. Typically, it is preferable to transport severely injured patients directly to high-level trauma centers [4]. However, if a high-level trauma center is far from a patient, it may be preferable to transport the patient to a low-level trauma center so they receive timely care. Moreover, the optimal mode of transport and transfer depends on the trauma system. It is often faster to transport patients by an air ambulance, but ground ambulances can sometimes transport patients to a trauma center more quickly due to the location of the patient and the ambulance. Additionally, there are often few air ambulances. This limits the number of patients who can be transferred and makes the dispatch of an air ambulance to a particular patient a decision that impacts the care of other patients.

In this paper, we focus on ensuring that severely injured patients have access to care at high-level trauma centers. We evaluate access in two ways to reflect different patient pathways. First, access can be defined as being transported directly to and receiving all care at a high-level trauma center. This is preferred under the ACS guidelines [4] and helps to improve patients’ outcomes by limiting the impact of undertriaging [5]. However, direct transport to a high-level trauma center can result in long patient transport times, which delays patient care. Therefore, we consider a second definition of access pertaining to the ability of patients to be quickly transported to a low- or high-level trauma center and later be transferred to a high-level trauma center, when appropriate. This reduces the time until initial stabilizing care. Typically, the “golden hour” is used as the time threshold in trauma care, meaning that patients should arrive at a trauma center within 60 min of their injury to receive stabilizing care. By improving the timeliness of trauma care, the belief is that trauma patient morbidity and mortality can be decreased [6,7]. Since we only consider severely injured patients, any patient initially transported to a low-level facility is required to later transfer to a high-level facility by means of ground or air ambulance.

There is a trade-off between these two definitions of access when considering patient transfers that inform trauma center planning decisions. The first definition promotes locating resources in such a way

that allows patients to be initially transported to high-level trauma care centers to receive stabilizing and definitive care. The second promotes locating resources to facilitate the transport of patients to nearby trauma care centers and eventually the transfer of patients to high-level trauma care centers as necessary for definitive care. The importance of each definition of access is not arbitrary. We discuss results for multiple combinations of the objectives to provide insight into the trade-off between them. This provides insight into the implications of various policy decisions made by government and healthcare leaders.

In this paper, we introduce two mathematical models to study trauma center planning problems. We first introduce the *trauma system planning problem* as an integer programming model to locate air ambulances and high-level trauma centers. This model informs the long-term planning decisions of trauma system design for typical patient call volumes. Each of these changes may require significant investment, but we do not consider the costs directly as they can be compared a posteriori. The trauma system planning problem sheds light on the design of a trauma care system under typical patient volumes, but there is a need to understand how to improve patient access when there are patient surges. Therefore, we introduce a second problem, the *mass casualty incident response problem*, to study shorter-term planning decisions related to patient surges. We introduce a stochastic program to address this problem. The model studies how to temporarily relocate air ambulances in response to a localized mass casualty incident (MCI), but before details of the patient surge (when, where, and how many) is known, to improve patient access to care by transporting and transferring patients.

We introduce two case studies, one for each problem, based on the state of Wisconsin. The case studies illuminate the impact of the changes to the trauma system. We investigate the solutions to better understand why the decisions identified by the optimization models are made. For the trauma system planning problem, we consider trauma care patients spatially correlated to population levels and historical motor vehicle crashes. For the MCI response problem, we consider an added cohort of patients due to surges related to a tornado-related mass casualty incident. We find that relocating existing air ambulances provides a similar benefit to advanced trauma care access as locating new air ambulances and upgrading trauma centers. Thus, relocating air ambulances could be an effective policy to implement. We further use these models to quantify the impact of patient transfers by comparing patient pathways and access to care when patient transfers are and are not allowed. In one case, patients receive stabilizing and definitive care at a high-level trauma center to which they are initially transported. In the other, patients are able to receive more timely stabilizing care at a low-level trauma center, although they must be transferred to receive

definitive care. The results suggest that transfers are important, and the results highlight the need to consider the role of transfers in trauma system planning problems.

The remainder of this paper is organized into six sections. Section 2 reviews the facility location literature, with an emphasis on applications within the emergency response and trauma care context. Section 3 introduces the *trauma system planning problem* and integer programming formulation to improve access of trauma patients when experiencing typical patient demand volumes by making changes to the trauma care system. Section 4 introduces the *mass casualty incident response problem* and stochastic programming formulation to reposition air ambulances in response to a MCI. Section 5 presents the computational results for the trauma system planning model when applied to a real-world case study. In Section 6 we apply the mass casualty incident response model to a case-study of tornado MCIs. Finally, we conclude with a discussion of the models and their results in Section 7.

2. Literature review

The existing facility location literature is extensive and varied both in terms of context and modeling approach. Location optimization problems are often formulated as a covering problem in which a demand (e.g., patient) is either covered or not covered by a facility/server (e.g., trauma center, ambulance), where coverage of demand can be defined in many ways. Most location covering models can be traced to one of two ancestral problems. Toregas et al. [8] introduce the Location Set Covering Model (LSCM), which determines the minimum number of facilities necessary to cover all demands in a network. Church and ReVelle [9] present the Maximum Covering Location Problem (MCLP), which instead fixes the number of facilities to locate and determines their optimal placement to maximize the expected number of covered demands.

Many location covering models use a binary definition of coverage that imply facilities are either always or never able to cover demands. Several papers do not make this assumption and extend the MCLP to include backup coverage of demands. Hogan and ReVelle [10] introduce the Backup Covering Problem (BACOP), which provides additional rewards when demands are covered twice and is easily generalizable to n levels of backup coverage. A weakness of the backup coverage models is that they do not explicitly consider how many facilities or servers must cover a location to ensure adequate service. Other researchers address this gap. They consider a continuous coverage value that captures the probability that a demand is covered. Daskin [11] presents an MCLP variant that considers server availability with the Maximum Expected Covering Location Model (MEXCLP). MEXCLP uses the probabilistic availability of servers to weigh the coverage they provide. Daskin [12] introduces the MCLP with probabilistic response (MCLP+PR) model for emergency services as an extension of the MCLP model. In this formulation, coverage reflects the confidence that a server is able to provide service within a time threshold. With this, the objective becomes equivalent to maximizing the expected number of demands that are covered within the time threshold. Goldberg and Paz [13] introduce a model (MEXCLP+PR+SSBP) for locating emergency vehicle stations that merges these two models that considers probabilistic coverage and facility-specific busy probabilities. However, these models assume a prior knowledge of server availability.

A significant amount of research has been devoted to location covering models within the context of trauma care systems. Some of the research focuses on locating healthcare facilities while others focus on locating resources, such as ambulances. Some research locates both simultaneously. Branas et al. [14] introduce the TRAMAH model, which locates trauma centers and air ambulances to maximize patient-weighted coverage of locations. Patient demand nodes are considered covered if the patients at that location can be reached and then transported by ground or air ambulance to a trauma center within the time threshold. Røislien et al. [15] use a similar model to locate

air ambulances to maximize patient-weighted coverage. In contrast to Branas et al. [14], Røislien et al. [15] do not include trauma centers or ground ambulances when determining coverage. Instead, they only consider whether air ambulances are able to reach the patients within a given time threshold. Both of these models assume that travel times are deterministic and that air ambulances are always available to cover patient demands. More recently, Erdemir et al. [16] extend the idea of backup coverage to find the optimal mix of ground and air ambulances in a trauma system. In their model, they consider the inability of air ambulances to land at certain locations and require that patients be transferred from ground ambulances to air ambulances at designated transfer points. Like BACOP, their model provides an additional reward if a patient is covered twice, in this case by both a ground and an air ambulance, rather than by either just a ground or an air ambulance. Cho et al. [17] introduce a trauma model to locate both trauma centers and air ambulances that includes air ambulance busy proportions as an endogenous variable. However, their model is a mixed-integer nonlinear program requiring a specialized algorithm to solve and does not necessarily lend itself to significant extension. For more information on research into health care facility location optimization, refer to the survey by Ahmadi-Javid et al. [18].

Throughout the trauma care literature, trauma center location models consider trauma centers as indistinguishable facilities. Location covering models that consider different level facilities are notably absent from the trauma literature despite being prevalent for other applications [19]. For example, Erdemir et al. [16] group all trauma centers together when considering coverage. This may overestimate coverage by ignoring the varying quality of care provided by different level trauma centers [20]. Similarly, Røislien et al. [15] seek to optimally locate air ambulances to cover as many patients as possible. By not considering the locations of trauma centers, they effectively assume equal care is provided to patients regardless of what level of trauma center is within range of the air ambulances. In contrast, Branas et al. [14] group level I and II trauma centers together and ignore lower-level trauma centers. This may underestimate coverage by dismissing the ability of lower-level facilities to stabilize and transfer patients to high-level facilities. While we focus on severely injured patients in this paper, we consider the ability of low-level trauma centers to provide stabilization before the patients are transferred to high-level trauma centers.

Finally, to our knowledge, only one other paper has included inter-hospital patient transfers when optimizing resource placement although over half of all air ambulance flights are for transferring patients between hospital facilities [21]. Nezamoddini and Khasawneh [22] use a capacitated network model to optimize the placement of resources at hospitals to minimize penalties associated with wait times and unmet patient demand. Patients may be transferred between hospitals, allowing hospitals to better balance resource utilization by offloading patients to less busy hospitals. They show a significant decrease in both total resource requirements and patient wait times. Nezamoddini and Khasawneh [22] are primarily concerned with balancing patient demand between hospitals using inter-facility transfers. However, inter-facility patient transfers have not been included within existing location covering models [18]. The models introduced in this paper address this gap. In contrast to Nezamoddini and Khasawneh [22], we consider the pre-hospital transport of patients, and how locating resources geographically within the system affects its ability to transport and transfer patients. Table 1 summarizes the contribution of this paper by comparing it to the closest papers in the literature.

In this paper, we take a probabilistic response approach to coverage. We seek to locate two types of resources that provide coverage. The first is air ambulances that can transport patients to trauma centers more quickly than ground ambulances, and therefore, provide a higher probabilistic coverage to patients. Second, we seek to locate new high-level trauma centers that increase advanced trauma care coverage for patients. For literature related to the patient demand coverage

Table 1
Comparison of existing literature.

Paper	Multiple trauma center care levels	Locate air ambulances	Locate trauma centers	Inter-facility transfers
Erdemir et al. [16]	✗	✓	✗	✗
Røislien et al. [15]	✗	✓	✗	✗
Branas et al. [14]	✗	✓	✓	✗
Nezamoddini and Khasawneh [22]	✗	✗	✗	✓
This paper	✓	✓	✓	✓

methodology applied in this paper, see the survey by Farahani et al. [19]. The stochastic model we introduce to solve the MCI response problem in this paper continues the stream of literature begun with MEXCLP by directly considering the availability of servers to cover patient demand.

3. Trauma system planning problem

In this section, we introduce the trauma system planning problem and bi-objective coverage model formulation to locate resources to maximize the access to real-time care when experiencing typical patient arrival rates of severely injured patients. The decision to locate resources is considered a one-time decision that impacts patient care for years following the decision.

3.1. Planning case description

Let there be a set of patient demand locations, I . Let each $i \in I$ represent a location (e.g., a census block group) with an average of p_i severely injured patients entering the trauma care system per day who must receive definitive care at a high-level trauma center. The patients may initially be transported to either a high- or low-level trauma center. We let J represent the collection of trauma centers, J_h the set of current high-level trauma centers, and J_ℓ the set of current low-level trauma centers, with $J = J_h \cup J_\ell$ and $J_h \cap J_\ell = \emptyset$. Trauma centers in the same level – within either J_ℓ or J_h – are assumed to provide equivalent care. While differences may exist in practice, such granular data is unlikely to be readily available or reliable. Table 2 summarizes the notation used in the model formulation and discussion.

A patient may be transported and transferred to a trauma center by one of two classes of ambulance: ground or air. Ambulances are located at depots within the set K . We let K_g represent the depots that house ground ambulances and K_a represent the depots that can house air ambulances, with $K = K_g \cup K_a$, $K_g \cap K_a = \emptyset$. We assume all depots in K_g currently house ground ambulances and depots in the subset $K'_a \subseteq K_a$ currently house air ambulances, while $K_a \setminus K'_a$ are locations at which an air ambulance may be located in the future. We assume a depot location houses at most one ambulance. This is usually, but not always, true in practice [23]. When this is not true, a depot can be added representing the same location but with a unique identifier.

Patients initially transported to low-level trauma centers are eventually transferred to high-level facilities to receive the appropriate level of definitive care. Patients can be transferred from low-level trauma center $j \in J_\ell$ to trauma centers in the set $F_{jk} \subseteq J$ by an ambulance at depot $k \in K$. This allows for trauma system restrictions to be taken into an account. One such restriction may be that air ambulances cannot travel excessively far to transfer a patient to limit flight time. Since low-level trauma centers may be upgraded to high-level trauma centers, F_{jk} may include both high- and low-level trauma centers. However, a patient may only be transferred to a currently high-level trauma center or a low-level trauma center that is upgraded to a high-level one.

We are concerned primarily with mature trauma systems. We assume that the existing number of ground ambulances is sufficient to transport and transfer any number of patients assigned to them when experiencing typical patient volumes. Thus, ground ambulances are not constrained by capacity. However, air ambulances are likely to be limited by a capacity constraint even when experiencing typical patient volumes. This assumption is reasonable, since of the 797 licensed EMS providers in Wisconsin, only 11 are air medical providers [24] operating a total of 14 air ambulance depots [23]. We assume an air ambulance at depot $k \in K_a$ cannot perform more than c_k transports or transfers on average per day. This limits the patient load for a given air ambulance, reducing the likelihood that a patient will be turned away in real-time care due to the ambulance being busy.

Within this framework, we consider the planning problem to upgrade up to u low-level trauma centers, add up to n new air ambulances, or relocate up to m existing air ambulances to new depots. Since ground ambulances are often under local control and we assume ground ambulances are uncapacitated, adding or relocating ground ambulances is not considered. We acknowledge that it is not necessarily realistic to be able to upgrade low-level trauma centers; however, different combinations of u , n , and m can be used to explore different planning choices and identify bottlenecks within patient care pathways.

We consider two goals when locating trauma care resources. First, patients should experience low transport times between injury and arriving at a trauma center, whether it be a high- or low-level facility. We employ probabilistic coverage [12] to maximize the expected number of patients that can be transported to a trauma center within the pre-hospital time threshold. The probability distribution of transportation times arises from differences in ambulance travel times. This goal necessitates the transfer of patients from low-level trauma centers to high-level trauma centers, where patients receive definitive care.

Second, patients should receive the appropriate level of care at the first trauma center to which they are transported. Since we only consider severely injured patients, the appropriate level of care is provided at a high-level trauma center. As a result, the second objective is to maximize the proportion of patients transported directly to a high-level trauma center.

The two goals are somewhat aligned, but their difference highlights the role of patient transfers. Note that if transfers were not considered, the two objectives are aligned and the overall goal is to transport patients to a high-level trauma center (the second objective).

Together, each objective informs how and when transfers occur. When it is important to maximize the number of patients transported to a trauma center of either level within the time threshold (the first goal), often the closest trauma center will be a low-level one. Patients must later be transferred to a high-level trauma center. In many cases, the transfers are completed by an air ambulance since the distance is too far for a ground ambulance. Conversely, if it is important to transport patients directly to a high-level trauma center (the second goal), then very few patients are transferred within the system.

Note that in the planning problem we assume ground ambulances and trauma centers are uncapacitated in the sense that they can serve typical levels of patient demand. Thus, minimally injured patient demand could be considered within this problem definition without changing any optimal decision, under the assumption that minimally injured patients would never be transported by an air ambulance and can receive definitive care at either low- or high-level trauma centers. Therefore, we do not consider minimally injured patients within the formulation.

3.2. Model formulation

We introduce an integer programming formulation of the trauma system planning problem. We introduce a binary variable v_j that equals 1 if the low-level trauma center $j \in J_\ell$ is upgraded to a high-level trauma center. Otherwise, $v_j = 0$. Similarly, we introduce a binary

Table 2
Planning model notation.

Sets		
I	=	Set of patient locations.
$J_h (J_\ell)$	=	Set of trauma centers starting at high- (low-) level.
$J = J_h \cup J_\ell$	=	Set of all trauma centers.
$K_g (K_a)$	=	Set of depots that may host ground (air) ambulances.
$K = K_g \cup K_a$	=	Set of all ambulance depots.
$K'_a \subseteq K_a$	=	Subset of ambulance depots starting with air ambulances.
F_{jk}	=	Subset of trauma centers to which patients may be transferred from $j \in J_\ell$ by an ambulance located at depot $k \in K$.
Variables		
v_j	=	1 if the low-level facility $j \in J_\ell$ is upgraded; 0 otherwise.
w_k	=	1 if an air ambulance is located at depot $k \in K_a$; 0 otherwise.
x_{ijk}	=	Number of patients from location $i \in I$ initially transported to facility $j \in J$ using an ambulance at depot $k \in K$.
$y_{jj'k}$	=	Number of patients transferred from un-upgraded facilities $j \in J_\ell$ to high-level or upgraded facility $j' \in J$ by an ambulance located at depot k .
h_{ij}	=	Number of patients at location $i \in I$ transported to high-level or upgraded facility $j \in J$.
t_{ij}	=	Number of patients at location $i \in I$ transported to low-level facility j then transferred to a high-level facility.
Parameters		
m	=	Number of current air ambulances that may be relocated.
n	=	Number of new air ambulances that may be located.
u	=	Number of low-level facilities that may be upgraded.
p_i	=	Mean daily number of patients at location $i \in I$.
c_k	=	Number of patients each air ambulance $k \in K$ may transport or transfer daily.
ϕ_{ijk}	=	Probability that a patient at location $i \in I$ can be transported to facility $j \in J$ by an ambulance from depot $k \in K$ within the time threshold.

variable w_k that equals 1 if an air ambulance is located at depot $k \in K_a$ in the new trauma system. Otherwise, $w_k = 0$. The value of these variables describe the location of trauma system resources.

Depending on the location of trauma system resources, transport and transfer decisions for patients may change. We introduce the x_{ijk} variables to represent the number of patients transported from patient demand node $i \in I$ to facility $j \in J$ using an ambulance at depot $k \in K$. Since we use average daily patient demands, ambulances may transport fractions of a patient, and the variable x_{ijk} may take a non-negative real value. The proportion of patients at location $i \in I$ that are expected to arrive at the trauma center $j \in J$ within the pre-hospital time threshold using ambulance $k \in K$ is captured by the parameter ϕ_{ijk} , which reflects probabilistic travel times.

If a patient is transported to a low-level trauma center, the patient must then be transferred to a high-level trauma center. We introduce the variable $y_{jj'k}$ to represent the number of patients transferred from low-level trauma center j (non-upgraded) to high-level trauma center j' (currently or upgraded) using ambulance $k \in K$. The high-level trauma center may either be an existing high-level trauma center or an upgraded low-level trauma center. Finally, the continuous variable h_{ij} reflects the number of patients from each location $i \in I$ that is transported to a high-level trauma center $j \in J_h$ or low-level trauma center $j \in J_\ell$ that has been upgraded to a high-level trauma center. Similarly, t_{ij} captures the number of patients from each patient demand node $i \in I$ that are initially transported to a low-level trauma center $j \in J_\ell$ and later transferred to a high-level trauma center.

Using these variables, we introduce the objective functions. The first objective function captures the expected number of patients with a transport time less than the pre-hospital time threshold, $f_1(\mathbf{x}) = \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} \phi_{ijk} x_{ijk}$, based on initial patient transport decisions, x_{ijk} . The second objective function, $f_2(\mathbf{h}) = \sum_{i \in I} \sum_{j \in J} h_{ij}$, captures the expected number of patients transported directly to a high-level trauma center. We now introduce the model formulation.

$$\begin{aligned} \max \quad & \left[f_1(\mathbf{x}) = \sum_{i \in I} \sum_{j \in J} \sum_{k \in K} \phi_{ijk} x_{ijk}, f_2(\mathbf{h}) = \sum_{i \in I} \sum_{j \in J} h_{ij} \right] \\ \text{s.t.} \quad & \sum_{j \in J_\ell} v_j \leq u \end{aligned} \tag{1}$$

$$\sum_{k \in K'_a} w_k \geq |K'_a| - m \tag{2}$$

$$\sum_{k \in K_a} w_k \leq |K'_a| + n \tag{3}$$

$$\begin{aligned} \sum_{j \in J} \sum_{k \in K} x_{ijk} &= p_i, \\ \forall i \in I \end{aligned} \tag{4}$$

$$\begin{aligned} h_{ij} &= \sum_{k \in K} x_{ijk}, \\ \forall i \in I, j \in J_h \end{aligned} \tag{5}$$

$$\begin{aligned} h_{ij} + t_{ij} &= \sum_{k \in K} x_{ijk}, \\ \forall i \in I, j \in J_\ell \end{aligned} \tag{6}$$

$$\begin{aligned} h_{ij} &\leq v_j p_i, \\ \forall i \in I, j \in J_\ell \end{aligned} \tag{7}$$

$$\begin{aligned} t_{ij} &\leq p_i(1 - v_j), \\ \forall i \in I, j \in J_\ell \end{aligned} \tag{8}$$

$$\begin{aligned} \sum_{i \in I} t_{ij} &= \sum_{k \in K} \sum_{j' \in F_{jk}} y_{jj'k}, \\ \forall j \in J_\ell \end{aligned} \tag{9}$$

$$\begin{aligned} y_{jj'k} &\leq v_{j'} \sum_{i \in I} p_i, \\ \forall j \in J_\ell, k \in K, j' \in F_{jk} \cap J_\ell \end{aligned} \tag{10}$$

$$\begin{aligned} \sum_{i \in I} \sum_{j \in J} x_{ijk} + \sum_{j \in J_\ell} \sum_{j' \in F_{jk}} y_{jj'k} &\leq c_k w_k, \\ \forall k \in K_a \end{aligned} \tag{11}$$

$$\begin{aligned} v_j, w_k &\in \{0, 1\}, \\ \forall j \in J_\ell, k \in K_a \end{aligned} \tag{12}$$

$$\begin{aligned} x_{ijk}, y_{jj'k}, h_{ij}, t_{ij} &\geq 0, \\ \forall i \in I, j \in J, k \in K, j' \in F_{jk} \end{aligned} \tag{13}$$

The bi-objective function seeks to maximize the expected number of patients transported to a trauma center within the time threshold ($f_1(x)$) and maximize the expected number of patients transported directly to a high-level trauma center ($f_2(h)$). Constraints (1)–(3) describe the trauma system design decisions. Constraint (1) allows at most u low-level facilities to be upgraded, constraint (2) allows at most m air ambulances to be relocated, and constraint (3) permits up to n new air ambulances to be added to the system. Constraint set (4) requires all patients to be transported to a trauma center. Constraint sets (5) and (6) require all patients to be either transported to a high-level trauma center or later transferred to a high-level trauma center. Constraint set (7) defines if a low-level trauma center is upgraded and, if not, ensures no patient receives definitive care there. Constraint set (8) requires that if a low-level trauma center is upgraded, then patients cannot be transferred to another high-level trauma center. Constraint set (9) requires that all transfers are assigned a receiving high-level trauma center and ambulance to perform the transfer. Constraint set (10) ensures that each low-level facility that receives a transfer is upgraded to a high-level facility. Finally, constraint set (11) limits the number of patients that an air ambulance transports or transfers.

The integer programming formulation may have many variables. However, it is possible to significantly limit the number of variables that need to be instantiated. First note that it is feasible to send all patients from location $i \in I$ by ground ambulance to the nearest existing high-level trauma center, since neither the ground ambulances $k \in K_g$ nor the high-level trauma centers $j \in J_h$ are capacitated. The highest proportion of patients that can be transported from location i to an existing high-level trauma center by ground ambulance is $\max_{j \in J_h, k \in K_g} \phi_{ijk}$. Therefore, we only need to include a variable x_{ijk} for patient location $i \in I$, trauma center $j \in J$, and ambulance depot $k \in K$ in the formulation if $\phi_{ijk} \geq \max_{j' \in J_h, k' \in K_g} \phi_{ij'k'}$. Similarly, if it is possible to transfer patients from a low-level trauma center $j \in J_\ell$ to an existing high-level trauma center $j' \in J_h$ by ground ambulance, no other receiving trauma center or ambulance types must be considered for patients transported to j .

Furthermore, if $m = 0$, $n = 0$, or $u = 0$, the integer programming formulation can be simplified and variables removed. For example, if $u = 0$ then no low-level facilities can be upgraded, and we can replace the associated variables v_j for $j \in J_\ell$ with zero in the constraint sets (1), (7), (8), and (10). By utilizing these strategies, the number of variables needed to solve the model to optimality can be significantly reduced and allow for the solution of large instances.

4. Trauma system mass casualty incident response problem

A limitation of the trauma system planning problem is that it does not directly consider the capacity of trauma centers or the real-time availability of air ambulances. These capacities are relevant during mass casualties incidents (MCIs) and inform shorter-term planning decisions. Air ambulances are typically limited to transporting one patient at a time [21], and there may not be sufficient capacity to respond to all patients caused by a MCI. Air ambulances and trauma centers located near the MCI are likely to be overwhelmed, and patients must be transferred to trauma centers elsewhere.

Within this section, we introduce a two stage stochastic program. In the first stage, it must be decided where to locate and relocate air ambulances in preparation of patients caused by the MCI, for who the location and time of need is a random variable. In the second stage, the patient pathways must be decided to provide the best trauma care. The decisions in the second stage depend on the decisions in the first stage and inform the first stage decisions. We present the deterministic equivalent formulation [25] of this stochastic program to replace random variables with parameters, which is achieved by enumerating scenarios that capture the stochasticity of the problem. We first discuss the notation used in the formulation while introducing the problem in detail and then introduce the formulation. We finish by discussing an algorithm to improve the solution time of the model.

4.1. Incident response case description

The trauma system MCI response problem is motivated by shorter time-frame response decisions where hospital capacity and air ambulance availability are a concern due to a surge of patients in a small geographic region. We consider the problem of locating or relocating new air ambulances in response to a MCI from which the patient demand is unknown at the time decisions must be made. In reality, it would be infeasible to procure and locate new air ambulances quickly enough following an MCI to be impactful, while relocating is more feasible. Instead, locating new air ambulances serves as a proxy for if an air ambulance and team from a neighboring area is lent in response to the MCI (e.g., an out-of-state air ambulance). We do not consider upgrading trauma centers, since the time-frame to complete such a project is too long. We use the same notation as in the trauma system planning (see Table 2). The changes in notation from those described in Section 3 is presented here.

Patient arrivals (e.g., location, time) into the trauma system after an MCI occurs (e.g., four hours following a tornado) are modeled with a random variable. The distribution of the random variable depends on the MCI, and it captures both patients caused by the MCI as well as patients not caused by the MCI who would typically arrive anyway. We introduce scenario-based deterministic equivalent formulation, where ξ captures a scenario in a set of Ξ of total scenarios and occurs with probability $Pr(\xi)$. Each scenario ξ represents a realization of a set of random patient arrivals and locations that follow a distribution describing all potential spatial and temporal patient scenarios following the MCI. The time of the patient arrival is the time at which the patient requires an ambulance transport. We define $I(\xi)$ as the set of patients in scenario ξ . We let each element in $I(\xi)$ represent a single patient, rather than a location representing p_i patients as with the trauma system planning problem.

Similar to the trauma system planning problem, we limit the number of transports and transfers that an air ambulance at depot $k \in K_a$ can perform to be at most c_k in each scenario. It is possible that trauma centers may reach their patient capacity limits following an MCI, and therefore, we introduce a trauma center capacity. We let e_j be the maximum number of patients trauma center $j \in J$ can treat in each scenario.

Since a MCI will likely cause an influx of patients that is higher than typical, it is no longer reasonable to assume that at the time of a patient's arrival, an air ambulance is available to transport the patient to a trauma center, since it may be transporting a different patient. Therefore, we require that if an air ambulance is conducting another patient transport when patient $i \in I(\xi)$ must be transported to a trauma center, then the ambulance cannot transport patient i to a trauma center. We let $P_{ik}(\xi)$ denote the set of patient transport restrictions, and each $(i', j') \in P_{ik}(\xi)$ of patient $i' \in I(\xi)$ represents a transport to trauma center $j' \in J$ that would make an air ambulance stationed at depot $k \in K_a$ busy when patient $i \in I(\xi)$ ($i \neq i'$) must be transported to a trauma center in scenario ξ . This set depends on the estimated time to transport patient i' to trauma center j' using ambulance k . If an air ambulance at depot $k \in K_a$ is assigned to any of the patient transports in $P_{ik}(\xi)$ in scenario ξ then it cannot transport patient i to any trauma center. If no air ambulance is available to transport patient i to a trauma center, then a ground ambulance transports the patient, since we still assume ground ambulances are uncapacitated. We do not consider the specific time that patient transfers occur, since there is significant variance and unpredictability in how long patients will be delayed at a low-level trauma center before being transferred to a high-level trauma center [26,27]. We assume that patient transfers occur when the transfer does not interfere with the ability of an air ambulance to transport patients to a trauma center, since these patients require the most immediate emergency care.

We let $F_j \subseteq J$ represent the set of trauma centers either to which trauma center j can transfer patients or from which patients can be

Table 3
Response model notation.

Sets		
Ξ	=	Set of scenarios.
$Pr(\xi)$	=	Probability of scenario $\xi \in \Xi$.
$I(\xi)$	=	Set of patients in scenario $\xi \in \Xi$.
J_h (J_ℓ)	=	Set of high-level (low-level) trauma centers.
$J = J_h \cup J_\ell$	=	Set of all trauma centers.
K_g (K_a)	=	Set of depots that may host ground (air) ambulances.
$K = K_g \cup K_a$	=	Set of all depots.
$K'_a \subseteq K_a$	=	Subset of depots starting with air ambulances.
F_j (F'_j)	=	Set of trauma centers from/to which patients can be transferred to/from $j \in J$ by a ground (air) ambulance.
$D'_j \subseteq K_a$	=	Subset of depots that host air ambulances able to transfer patients to/from facility $j \in J$ from/to one in F'_j .
$P_{ik}(\xi)$	=	Set of patient-depot pairs, (i', j') , such that an ambulance located at depot $k \in K$ would be busy when demand $i \in I(\xi)$ must be transported to a trauma center given the patient arrivals in scenario $\xi \in \Xi$.
Variables		
r_k	=	1 if the air ambulance at depot $k \in K'_a$ relocated; 0 otherwise.
w_k	=	Number of air ambulances located at depot $k \in K_a$.
x_{ijk}^ξ	=	1 if patient $i \in I$ is transported to trauma center $j \in J$ by an ambulance at depot $k \in K$ in scenario $\xi \in \Xi$; 0 otherwise.
$y_{jj'}^\xi$	=	Number of transfers between trauma center $j \in J$ and $j' \in J$ in scenario $\xi \in \Xi$.
z_{jk}^ξ	=	Number of transfers to/from trauma center $j \in J$ performed by air ambulances at depot $k \in D'_j$ in scenario $\xi \in \Xi$.
Parameters		
m	=	Number of current air ambulances that may be relocated.
n	=	Number of new air ambulances that may be located.
c_k	=	Number of transports or transfers an air ambulance located at depot $k \in K_a$ may perform.
e_j	=	Number of patients a trauma center $j \in J$ may treat.
ϕ_{ijk}^ξ	=	Probability that patient $i \in I(\xi)$ can be transported to trauma center $j \in J$ by an ambulance at depot k within the time threshold in scenario $\xi \in \Xi$.

transferred to j using a ground ambulance, depending on if j is a low- or high-level trauma center. Some transfers between may only be feasible when using an air ambulance. We introduce the set F'_j to define the trauma centers to/from which transfers can be facilitated using an air ambulance. If $j \in J_\ell$, then F_j and F'_j include high-level trauma centers ($F_j, F'_j \subseteq J_h$) to which j can transfer patients. If $j \in J_h$, then F_j and F'_j include low-level trauma centers ($F_j, F'_j \subseteq J_\ell$) from which patients can be transferred to j . In contrast to the trauma system planning problem, we make a simplifying assumption that if an ambulance depot can facilitate a patient transfer to/from a trauma center $j \in J$ and a trauma center in F_j or F'_j , then the ambulance depot can also facilitate a transfer to/from every other trauma center in F_j or F'_j , respectively. Note that if the set F_j or F'_j is non-empty, there must be at least one ground ambulance or air ambulance that can facilitate these transfers. We denote $D'_j \subseteq K_a$ as the set of air ambulance depots that can facilitate a patient transfer between patients at j and a trauma center in F'_j .

There are two objectives when locating or relocating air ambulances. First, the expected pre-hospital time should be low for each patient in each potential realization of patient arrivals. We again use probabilistic coverage to maximize the expected proportion of patients transported to a trauma center within a defined pre-hospital time threshold. Second, we maximize the proportion of patients transported directly to a high-level trauma center. The patients not transported to a high-level trauma center must be transferred to one to receive definitive care.

This model and the subsequent analysis does not consider the much more catastrophic case of a hospital or trauma center struck by a tornado, which has happened in the past [28]. This would greatly increase the number of patients localized in a single area as well as reduce the number of high- or low-level facilities within the trauma system.

4.2. Model formulation

We introduce a stochastic programming model formulation of the mass casualty incident response problem based on, for example, weather

alerts. The planning decisions are to locate or relocate air ambulances in preparation of patient arrivals. The recourse decisions are to determine the optimal movement of patients in each scenario through the trauma care system given the location of air ambulances. The decision of where to locate or relocate air ambulances must consider the real-time use of the air ambulances and trauma centers during each patient arrival scenario. We introduce a scenario-based deterministic equivalent formulation of the stochastic programming model with a finite set of scenarios. See Table 3 for a summary of the notation, decision variables, and parameters used in the model formulation.

We introduce the variables associated with the formulation. The integer variable w_k equals the number of air ambulances located at depot $k \in K_a$ in the trauma system. Due to the concentration of patients in a single location due to the MCI and temporary relocation of resources, we lift the assumption that only one air ambulance may be assigned to any particular depot and allow w_k to be a non-negative integer variable. We let a binary variable r_k equal one if the ambulance currently located at depot $k \in K'_a$ is relocated and is zero otherwise. For each scenario ξ , we determine the optimal movement of patients through the trauma care system given the location of air ambulances. We let the binary variable x_{ijk}^ξ equal one if patient $i \in I(\xi)$ is transported to trauma center $j \in J$ by ambulance $k \in K$ in scenario ξ and zero otherwise. We let the integer variable $y_{jj'}^\xi$ equal the number of transfers between trauma center $j \in J$ and trauma center $j' \in J$ conducted by an ambulance in scenario ξ . Note that $y_{jj'}^\xi + y_{j'j}^\xi$ represents the total number of patients transported between j and j' in scenario ξ . The direction of the flow depends on which trauma center is high-level and which is low-level. We let the integer variable z_{jk}^ξ capture the number of patients that air ambulances at depot $k \in D'_j$ transfer to/from facility $j \in J$ in scenario ξ .

Using these variables, we introduce the objective functions. We let $Q_1(\mathbf{w}, \mathbf{r}, \xi)$ capture the expected number of patients transported to a trauma center within the time threshold given the relocation decisions \mathbf{r} , the location of the air ambulances \mathbf{w} , and the patient arrivals in

scenario ξ . The first objective is to maximize the expected value of $Q_1(\mathbf{w}, \mathbf{r}, \xi)$ across all scenarios ξ , with $Q_1(\mathbf{w}, \mathbf{r}, \xi) = \sum_{i \in I(\xi)} \sum_{j \in J} \sum_{k \in K} \phi_{ijk}^\xi x_{ijk}^\xi$. Note that the value of ϕ_{ijk}^ξ is dependent on the patient arrivals, since the location of patient i may change with ξ . Similarly, $Q_2(\mathbf{w}, \mathbf{r}, \xi)$ captures the expected number of patients transported directly to a high-level trauma center given the relocation decisions \mathbf{r} , the location of the air ambulances \mathbf{w} , and the patient arrivals in scenario ξ . The second objective is to maximize the expected value of $Q_2(\mathbf{w}, \mathbf{r}, \xi)$ across all scenarios ξ , with $Q_2(\mathbf{w}, \mathbf{r}, \xi) = \sum_{i \in I(\xi)} \sum_{j \in J_h} \sum_{k \in K} x_{ijk}^\xi$. Later in the case study we interpret both objectives as proportions by rescaling by the total number of patients.

$$\max \left[\mathbb{E}_\xi [Q_1(\mathbf{w}, \mathbf{r}, \xi)] = \sum_{\xi \in \Xi} \sum_{i \in I(\xi)} \sum_{j \in J} \sum_{k \in K} Pr(\xi) \phi_{ijk}^\xi x_{ijk}^\xi, \right. \\ \left. \mathbb{E}_\xi [Q_2(\mathbf{w}, \mathbf{r}, \xi)] = \sum_{\xi \in \Xi} \sum_{i \in I(\xi)} \sum_{j \in J_h} \sum_{k \in K} Pr(\xi) x_{ijk}^\xi \right] \quad (14)$$

$$\text{s.t. } \sum_{k \in K_a} w_k \leq |K'_a| + n \quad (15)$$

$$\sum_{k \in K'_a} r_k \leq m \quad (16)$$

$$w_k \geq 1 - r_k, \quad \forall k \in K'_a \quad (17)$$

$$w_k \in \mathbb{Z}^+, \quad \forall k \in K_a \quad (18)$$

$$r_k \in \{0, 1\}, \quad \forall k \in K'_a \quad (19)$$

$$\sum_{j \in J} \sum_{k \in K} x_{ijk}^\xi = 1, \quad \forall i \in I(\xi), \xi \in \Xi \quad (20)$$

$$\sum_{i \in I(\xi)} \sum_{k \in K} x_{ijk}^\xi + \sum_{j' \in F_j \cup F'_j} y_{jj'}^\xi + \sum_{j' : j \in F_j \cup F'_j} y_{j'j}^\xi \leq e_j, \quad \forall j \in J_h, \xi \in \Xi \quad (21)$$

$$\sum_{i \in I(\xi)} \sum_{k \in K} x_{ijk}^\xi \leq e_j, \quad \forall j \in J_\ell, \xi \in \Xi \quad (22)$$

$$\sum_{i \in I(\xi)} \sum_{j \in J} x_{ijk}^\xi + \sum_{j \in J : k \in D'_j} z_{jk}^\xi \leq c_k w_k, \quad \forall k \in K_a, \xi \in \Xi \quad (23)$$

$$\sum_{j \in J} x_{ijk}^\xi + \sum_{(i', j') \in P_k(\xi)} x_{i'j'k}^\xi \leq w_k, \quad \forall i \in I(\xi), k \in K_a, \xi \in \Xi \quad (24)$$

$$\sum_{i \in I(\xi)} \sum_{k \in K} x_{ijk}^\xi = \sum_{j' \in F_j \cup F'_j} y_{jj'}^\xi + \sum_{j' : j \in F_j \cup F'_j} y_{j'j}^\xi, \quad \forall j \in J_\ell, \xi \in \Xi \quad (25)$$

$$\sum_{j' \in F'_j \setminus F_j} y_{jj'}^\xi \leq \sum_{k \in D'_j} z_{jk}^\xi, \quad \forall j \in J, \xi \in \Xi \quad (26)$$

$$x_{ijk}^\xi \in \{0, 1\}, \quad \forall i \in I(\xi), j \in J, k \in K, \xi \in \Xi \quad (27)$$

$$y_{jj'}^\xi \in \mathbb{Z}^+, \quad \forall j \in J, j' \in F_j, \xi \in \Xi \quad (28)$$

$$z_{jk}^\xi \in \mathbb{Z}^+, \quad \forall j \in J, k \in D'_j, \xi \in \Xi \quad (29)$$

The bi-objective function (14) is to simultaneously maximize the expected number of patients transported to a trauma center within the defined pre-hospital time threshold and to maximize the expected number of patients transported directly to a high-level trauma center. Constraint (15) limits the number of air ambulances that can be added to the system to n . Constraint (16) limits the number of air ambulances that can be relocated to m . Constraint set (17) evaluates if a current air ambulance is relocated. Constraint sets (15)–(19) capture the planning decisions. Constraint sets (18) and (19) require the planning variables to be positive integer and binary values. The remaining constraints capture the response to a MCI by determining the patient movement through the system in each scenario. Constraint set (20) requires that

each patient in scenario ξ is transported to exactly one trauma center. Constraint sets (21) and (22) limit the number of patients in scenario ξ that high-level and low-level trauma centers may treat, respectively. Constraint set (23) limits the number of transports and transfers an air ambulance may perform in scenario ξ . Constraint set (24) only permits an air ambulance to transport one patient at a time. Constraint set (25) requires that all patients in scenario ξ who are transported to low-level trauma centers are later transferred to high-level trauma centers. Constraint set (26) ensures that there are sufficient air ambulances available in scenario ξ to complete the transfers requiring them when coupled with constraint set (23). The final two sets of constraints require that the variables are binary and positive integers, respectively.

4.3. Solution algorithm

With a large number of scenarios, the number of model variables is large, especially with regards to x_{ijk}^ξ . We take advantage of a similar methodology to that discussed in Section 3 to reduce the number of variables. We also employ the ϵ -constraint method to generate weakly Pareto optimal solutions by enforcing a minimum value for the second objective $\mathbb{E}_\xi [Q_2(\mathbf{w}, \mathbf{r}, \xi)]$. That is, we maximize the expected number of patients transported within the time threshold ($\mathbb{E}_\xi [Q_1(\mathbf{w}, \mathbf{r}, \xi)]$) while requiring that an ϵ proportion of them are transported directly to high-level facilities, in expectation ($\mathbb{E}_\xi [Q_2(\mathbf{w}, \mathbf{r}, \xi)] \geq \epsilon \times \mathbb{E}_\xi [I(\xi)]$). For any given ϵ , often either the closest low-level or closest high-level trauma center is the optimal transport destination for patients. We exploit this pattern as follows. Similar to the planning problem, we exclude the variable x_{ijk}^ξ from the model (i.e., $x_{ijk}^\xi = 0$), if for ξ and patient i :

$$\phi_{ijk}^\xi \leq \max_{j' \in J_h, k' \in K_g} \phi_{ij'k'}^\xi.$$

Our solution procedure removes constraint sets (21), (22), and (24) when solving the model so that not all capacity constraints on the trauma centers and air ambulances are enforced. We solve this model relaxation and determine which of the capacity constraints were violated (if any). Capacity constraints that were violated are added back to the model. Additionally, in the case a high-level trauma center capacity constraint was violated (i.e., a constraint in constraint set (21) was violated), we add specific variables back into the model. Let \bar{J}_1 be the set of high-level trauma centers whose capacities were violated. Then, we include a previously excluded variable x_{ijk}^ξ in the formulation if for patient i :

$$\phi_{ijk}^\xi \geq \max_{j' \in J_h \setminus \bar{J}_1, k' \in K_g} \phi_{ij'k'}^\xi$$

This adds back the variables relating to patient i such that transporting patient i to trauma center j using ambulance at depot $k \in K$ has a higher likelihood of being transported to j within the time threshold than when transported by ground ambulance to a high-level trauma center not at capacity. We re-solve the formulation and repeat this procedure until no constraints are violated.

The benefit of this procedure is a result of the localized nature of the MCI. The vast majority of trauma centers outside of the immediate area of the MCI do not reach capacity. Therefore, we do not need to capitate them or consider transport/transfer options for patients outside of the MCI area other than the closest trauma centers. Similarly, patients outside of the MCI tend to be dispersed enough both spatially and temporally that the air ambulances serving them are unlikely to violate constraint set (24) when the constraints are not explicitly enforced.

This solution algorithm finds the optimal solution if a feasible solution exists. A feasible solution exists as long as the total capacity of the high-level facilities is greater than the number of patients for each possible ξ , which can be easily checked.

Further analysis may highlight additional solution methods that reduce computational time to solve the optimization problem. For example, solutions algorithms that address symmetry in the formulation may reduce solution times [29].

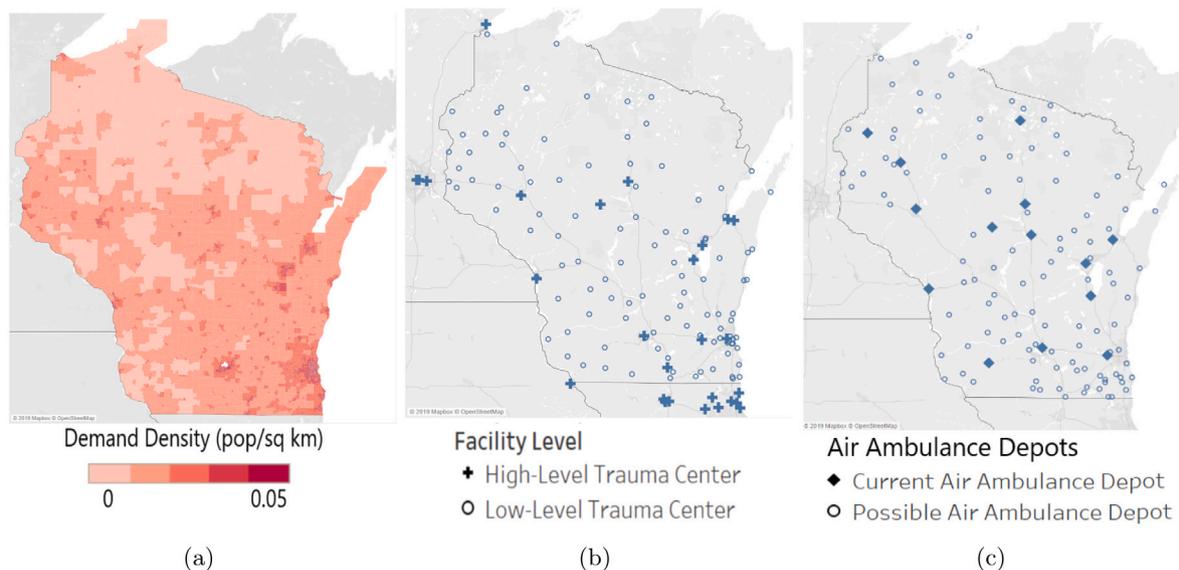


Fig. 2. Locations of (a) Patients with average daily arrival rate, (b) Trauma centers, and (c) Air ambulance depots.

5. Trauma care system planning problem: A case study of Wisconsin

We investigate the trauma system planning problem introduced in Section 3 using a case study based on the state of Wisconsin. The real-world case study is based on average patient demand in the Wisconsin trauma care system to demonstrate the benefit of the model and investigate the optimal resource location decisions. Multiple changes have been made to the air ambulance system considered in our case study during this research [30,31], which highlights the relevance of this research.

We begin by describing the data used to generate this case study in detail. We then describe and analyze the computational results, including changes to air ambulances and facilities and the flow of patients within the trauma care system. Finally, we discuss the results of a sensitivity analysis that provides insight into key properties of the solutions.

5.1. Case study design

Table 4 provides a complete list of parameter values considered in this case study. We first describe average patient demand within the trauma system. Historically, the average number of daily trauma patients in Wisconsin is approximately 54 [32]. However, not all of these patients require care at a high-level facility. We estimate the number of severely injured patients to be 30, based in part on the historical rate at which patients are transferred from low-level facilities to high-level facilities (54%) [32].

We use the Wisconsin census block group regions as the patient demand locations. Each block group region may have multiple patients. Fig. 2(a) illustrates the distribution of estimated patient demands within the state. These estimates are calculated from historical data assuming that patients are spatially distributed according to population density and historical motor vehicle crash (MVC) locations. It has been shown that trauma patient locations are well-correlated with population density, except for those resulting from industrial accidents and MVCs [33]. Since industrial accidents create less than 2% of trauma patients [34], we do not model them separately. Historically, 25% of all Wisconsin trauma patients are due to MVCs [35]. We use the spatial distribution of 2011–2016 MVCs to determine the spatial distribution of patients, assuming that an MVC in any location has the same likelihood of creating a severely injured patient.

At the time this research was conducted, the Wisconsin trauma care system has 14 air ambulance depots [23] and 124 trauma centers, 12 of which are adult high-level facilities [36]. The set J of trauma centers is comprised of these 124 trauma centers plus all trauma centers in neighboring states within 20 miles of the Wisconsin border. In total this is 138 trauma centers (27 high). We include trauma centers in neighboring states, since patients from Wisconsin may be closer to these trauma centers than those in Wisconsin. Fig. 2(b) maps the location of the $|J_h| = 27$ high-level trauma centers and $|J_\ell| = 111$ low-level trauma centers. We do not include pediatric trauma centers, since not all patients can receive care at these locations.

Patients are transported and transferred to trauma centers by either ground or air ambulances. We let K_g be the set of EMS provider locations in Wisconsin. We let potential air ambulance depots be the $|K_a| = 124$ publicly accessible airfields in the state of Wisconsin. Fig. 2(c) illustrates the location of all possible and current air ambulance depots. For simplicity, we do not include hospitals as air ambulances locations. If an air ambulance is currently located at a hospital, we use the nearest airfield as its current location. We do so to ensure there is not an air ambulance located at a hospital and also one at a nearby airfield. We let each air ambulance depot have the same capacity $c_k = 3$ and in the sensitivity analysis consider values $c_k = 2$ and $c_k = 4$.

We let the probabilistic coverage [12] of a patient equal the probability that the patient can be transported to the initial trauma center within the pre-hospital time threshold of 60 min. Pre-hospital time is the interval between when the call arrives and when the patient arrives at the first hospital. While it is common convention to use 60 min, Alarhayem et al. [6] have found that a time threshold shorter than 60 min may be necessary to prevent mortality in torso injury victims. As a result, we also consider a pre-hospital time of 45 min in the sensitivity analysis.

The pre-hospital time interval can be divided into several sub-intervals (dispatch, chute, response, on-scene, and transport) as illustrated in Fig. 3. We use these times to compute the pre-hospital times. We calculate the pre-hospital time distribution as follows based on data from the East North Central Region in the National EMS Information System (NEMSIS) [37]. The first component of pre-hospital time is dispatch time. In some cases, air ambulances may not be dispatched until a ground ambulance has determined that the patient's condition requires air ambulance transport, leading to a long dispatch time. Similarly, air ambulances often have highly trained personnel and

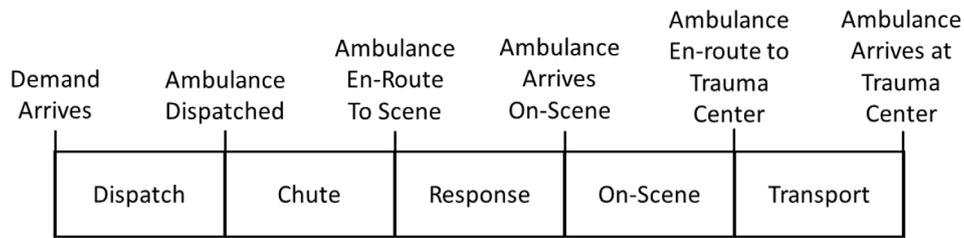


Fig. 3. Time intervals prior to a patient arriving at a trauma center.

more equipment allowing crews to provide more effective, but time-consuming, treatment on-scene. As a result, air-transported patients have similar or better outcomes on average than ground-transported patients with the same level of injury despite having longer average pre-hospital times [38,39]. To avoid potential bias within the NEMSIS database, we let there be a standard total dispatch plus on-scene time equaling 15 min for both ground and air ambulances.

Distributions for chute times for both ground and air ambulances are created from historical data in NEMSIS. For ground ambulances, a distribution of response times is estimated from NEMSIS while the transport times are calculated directly. We employ the strategy introduced in Budge et al. [40] to calculate transport duration between the patient location and the trauma center using travel distances retrieved from Bing Maps. Air ambulance response and transport times are calculated directly assuming a fixed airspeed of 220 km/h [15] and using the Vincenty’s formula to calculate air distances. Based on discussions with the staff at UW Health’s Med Flight, it is evident that flight times are far more predictable than driving times (F. Erdman, personal communication, September 27, 2018), and therefore it is reasonable to use a point estimate. Using these distributions, we calculate the probability ϕ_{ijk} that a patient at location $i \in I$ can reach trauma center $j \in J$ within the pre-hospital time threshold using an ambulance at depot $k \in K$.

Lastly, we define between which trauma centers an ambulance can perform a patient transfer. We assume that patient transfers by ground ambulance can occur if the road distance between the two trauma centers is no more than 30 km. When this is the case, we assume there is a ground ambulance near enough to one of the trauma centers to complete the transfer. Favoring air ambulances over ground ambulances for transfers at distances greater than 30 km has been shown to reduce patient mortality [41]. In the sensitivity analysis, we also consider a 50 km maximum transfer distance for ground ambulances. We assume that an air ambulance stationed at a depot k that is within 15 min of travel time (55 km) to trauma center j may perform transfers in/out of that trauma center. This enforces that air ambulances are located reasonably close to trauma centers to maintain efficiency when conducting transfers. The air ambulance can then transfer patients to any trauma center within 60 min (220 km) from trauma center j . This limits unreasonably long transfers. These two policies define F_{jk} . In practice, few transfers occur over a distance greater than 100 km [42]. We vary these time thresholds in the sensitivity analysis.

5.2. Case study results

In this section, we describe the results of the case study using the values described in Table 4. Solutions are generated using the ϵ -constraint method, where we seek to maximize the proportion of patients transported to a trauma center within the pre-hospital time threshold, $f_1(\mathbf{x})$, while ensuring that an ϵ proportion of patients are transported directly to a high-level facility, $f_2(\mathbf{h}) \geq \epsilon \times \sum_{i \in I} p_i$. We first discuss the trade-off between the two objectives using base-parameter values. We then analyze patient flows for the extreme values of ϵ using base-parameter values. We follow by evaluating the optimal placement of resources within the trauma system as ϵ changes and discuss how

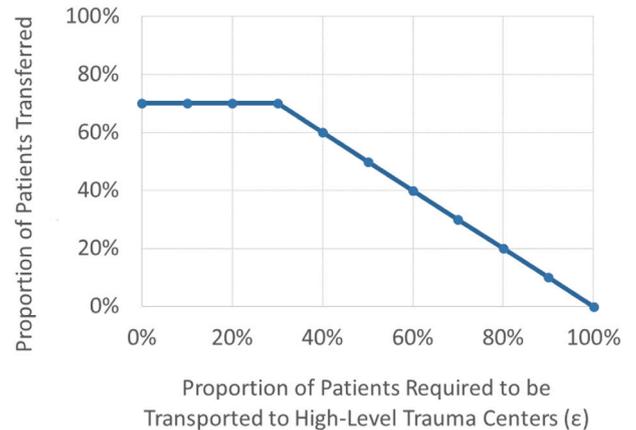


Fig. 4. Proportion of patients transferred from low-level facilities to high-level facilities as a function of the proportions of patients required to be transported to high-level facilities ($f_2(\cdot)$) when $u = 0$, $m = 0$, and $n = 0$.

the objective values are impacted using base-parameter values. Lastly, we highlight the parameters to which the solutions are most sensitive.

We first discuss the trade-off between the two objectives. At present, most patients are closest to a low-level trauma center. This can be seen in Fig. 4, which shows the proportion of patients transferred from the initial trauma center to which they are transferred (related to objective $f_1(\cdot)$) as a function of the proportion of patients transported directly to a high-level trauma center ($f_2(\cdot)$). The latter patients do not undergo transfer. When all patients are allowed to be transferred ($\epsilon = 0\%$), approximately 70% of patients are transported to a low-level trauma center before being transferred to a high-level trauma center. As we increase ϵ , a greater proportion of patients are required to be transported directly to high-level facilities. When $\epsilon = 100\%$, no patient transfers are performed, since all patients are transported directly to a high-level facility.

We find that for most values of $f_2(\cdot)$ (until about 70%–80% of patients), a marginal increase in $f_2(\cdot)$ results in a relatively small decrease in $f_1(\cdot)$ for any value of u , m , or n . This is because most patients and high-level trauma centers are located near urban centers, and there is little change in the expected proportion of individuals reaching a high-level trauma center within 60 min instead of a low-level trauma center. Increasing the proportion of patients transported directly to high-level trauma centers within the range of 70%–80% to 100% substantially decreases the proportion transported to a trauma center within 60 min due to the increasing rurality of the population being served. Instead, these patients can reach nearby low-level trauma centers within the 60 min pre-hospital time threshold where they may be stabilized. Therefore, it may be desirable to transport these patients to low-level trauma centers, since they are located too far from high-level trauma centers to permit quick transport even with changes to the trauma system. These patients can later be transferred to regional high-level trauma centers. This is already common in practice for rural patients [43].

Table 4
Base and alternative parameter values for the case study planning model.

Model parameter or input	Related symbol	Base value	Sensitivity analysis
Upgraded trauma centers	u	0	{1, 2}
Relocated air ambulances	m	0	{1, 2, 3, 13 ^a }
New air ambulances	n	0	{1, 2, 3}
Transport time threshold	ϕ_{ijk}	60 min	45 min
Patient demand distribution	I	Historical data	–
Trauma centers	J	$ J = 138$	–
Air ambulance depots	K_a	$ K_a = 124$	–
Current air ambulances	K'_a	$ K'_a = 14$	–
Air ambulance capacity	$c_k, k \in K_a$	3 routes	{2 routes, 4 routes}
Max air transfer distance	$F_{jk}, k \in K_a$	220 km	{165 km, 330 km}
Air ambulance distance to trauma center for transfers	$F_{jk}, k \in K_a$	55 km	{18.3 km, 110 km}
Max ground transfer distance	$F_{jk}, k \in K_g$	30 km by road	50 km by road

^aRelocate all air ambulances currently in use.

The different values of ϵ lead to different patient pathways through the trauma care system, illustrated in Figs. 5 and 6. We find dramatically different patterns in patient pathways based on the settings selected. We compare pathways for the extreme values of ϵ to highlight the difference in a system in which transfers are not allowed ($\epsilon = 100\%$, shown in Fig. 5) and a system in which transfers are allowed ($\epsilon = 0\%$, shown in Fig. 6). Figs. 5 and 6 show the results when no changes to the current trauma system are made ($u = 0, m = 0, n = 0$). When all patients are transported directly to high-level trauma centers ($\epsilon = 100\%$), shown in (Fig. 5), the flow is fairly simple. Patients are transported to the nearest high-level trauma center. Within a given time-based radius, this transport is performed by ground ambulances (blue lines), and beyond that radius by air ambulances (orange lines). These radii are not the same for all high-level trauma centers due to differences in the locally located air ambulances and their capacities. If a patient cannot be transported to a trauma center within 60 min with high probability by either a ground or air ambulance, a ground ambulance is likely to be chosen so the patient does not impact the capacity of the air ambulance. This is why some patients far from a high-level trauma center receive a ground transport (blue line) even though they are far away. In real time, emergency response teams must make non-trivial decisions regarding which patients would benefit the most from an air ambulance transport and when air ambulances should not be sent to distant patients to ensure available capacity for other patients who may benefit more from a quicker transport.

Patient pathways when transfers to high-level trauma centers are allowed ($\epsilon = 0\%$) are presented in Figs. 6(a) and 6(b). Patients are far more likely to be transported to their initial trauma center by ground due to the much larger number of trauma centers to which patients can be transported. These patients are then transferred from the local trauma center, often by air ambulance, to regional high-level trauma centers. This allows a substantially higher proportion of patients (approx. 94%) to be transported to a trauma center within 60 min compared to when transfers are not allowed (no more than 86%).

Moreover, the benefit of non-zero values of $u, m,$ and n on the system depends on if transfers are allowed. When the value of $f_2(\cdot)$ is low, changes to the trauma system have little impact of the maximal value of $f_1(\cdot)$ ($< 1\%$). This because the trauma system is fairly mature and adding resources are unlikely to substantially increase the expected number of individuals within 60 min of a trauma center due to stochasticity in transport times. In our case study, there are areas within the state with no close high- or low-level trauma center, and new trauma center locations must be considered to substantially improve access. However, this is likely very costly and is outside the scope of this study. If local health officials aim to transport more patients directly to high-level trauma system (high values of $f_2(\cdot)$), changing the trauma system has a larger impact. By relocating air ambulances or adding air ambulances, up to an additional 3.5% of the patients are expected to reach a trauma center within 60 min of pre-hospital

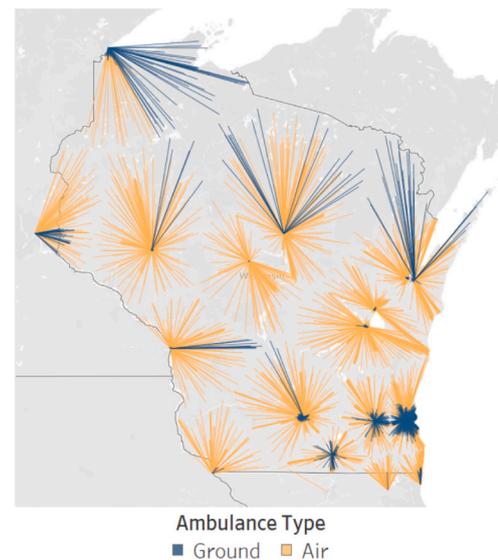


Fig. 5. Flow of patients through the trauma care system when no changes are made to the current air ambulances or trauma centers ($u = 0, m = 0, n = 0$) and all patients are transported directly to a high-level trauma center ($\epsilon = 100\%$). Flow is differentiated by ambulance type.

time, depending on the number of ambulances added or relocated. By upgrading up to two trauma centers, up to an additional $\sim 3\%$ of the patients are able to reach a trauma center within 60 min of pre-hospital time. The largest improvements are seen when one air ambulance is added, one air ambulance is relocated, or one trauma center is upgraded. Moreover, a direct comparison of relocating air ambulances to other changes with the same model inputs indicates that relocating air ambulances is roughly as valuable to system performance as adding new air ambulances or upgrading trauma centers, which most likely have a higher cost. Although no cost-benefit analysis was done, these results suggest that relocating air ambulances is likely a cost-effective approach to improving access to high-level trauma care. This is an important finding from this analysis

The importance of each of the objectives impacts the optimal location of trauma systems resources. When we add, relocate, or upgrade resources within the system, the differences in patient pathways for different values of ϵ lead to fundamentally different optimal resource locations. When patients must be directly transported to a high-level trauma center ($\epsilon = 100\%$), the implicit goal is to spread air ambulances and high-level trauma centers across the state so that every patient has a relatively high likelihood of being transported to one of the facilities within 60 min. As seen in Figs. 7(a) and 7(b), the optimal location of added or relocated air ambulances are in locations where they will

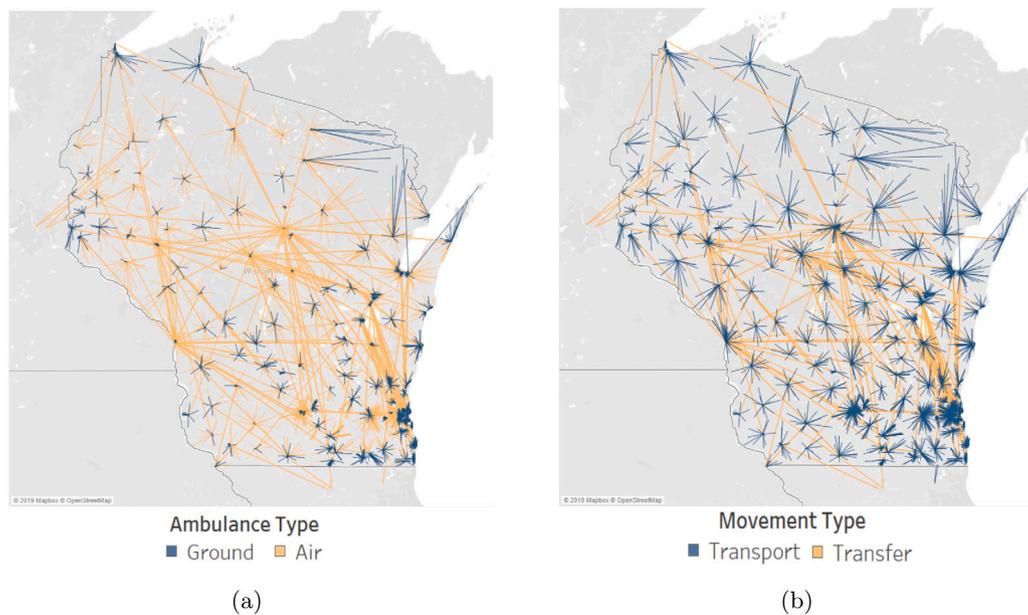


Fig. 6. Flow of patients through the trauma care system when no changes are made to the current air ambulances or trauma centers ($u = 0$, $m = 0$, $n = 0$) and all patients may be transferred ($\epsilon = 0\%$). The flow is differentiated by (a) The type of ambulance used to move the patient; and (b) Whether the movement was the result of an initial transport or a inter-facility transfer.

cover the most densely populated areas without current air ambulance coverage, instead of covering areas that would be served by a ground ambulance or another air ambulance. Trauma center upgrades, seen in Fig. 7(c), are applied to low-level trauma centers in areas with relatively poor access to a high-level trauma center and in areas that can increase the capacity of air ambulances to transport more rural patients to a high-level trauma center.

When patients can be transferred to a high-level trauma center ($\epsilon = 0\%$), the implicit goal when locating air ambulances is to ensure air ambulances are close enough to high-level trauma centers so that transfers can be completed, while still being available to transport patients to a trauma center if ground ambulance coverage is insufficient. Air ambulances more often transfer patients than respond to the scene of patient incidents [21]. Figs. 8(a) and 8(b) present the optimal location of air ambulances when three are relocated or added to the system, respectively. Similarly, trauma center upgrades, illustrated in Fig. 8(c), are located in areas at which there exist air ambulances that can transfer patients to them and in areas where air ambulances had previously conducted patient transports so that capacity of those air ambulances is freed.

Through the sensitivity analysis, we also find that the quantitative benefit of changes to the trauma systems are relatively robust to input parameter changes. While comprehensive results are not reported within this paper, we provide a brief overview of sensitivity analysis results. Additional details of the sensitivity analysis can be found in [44]. Compared to the other measures, a pre-hospital time threshold of 45 min significantly reduces the proportion of patients transported to a trauma center within the time threshold ($f_1(\mathbf{x})$) for any value of $f_2(\mathbf{x})$. This is unsurprising, since it is a more restrictive definition of “coverage”. We find the difference in the proportion of patients with pre-hospital times within 60 min and 45 min in the current trauma system vary between 14.8–30.4% of patients, and the true value depends on ϵ . We also find that relocated air ambulances, new air ambulance, and upgraded trauma centers lead to relatively larger improvements in $f_1(\mathbf{x})$ (between 0.5–7.0%, 1.1–7.0%, and 0.3–4.1%, across $\epsilon = 0\%$ and $\epsilon = 100\%$) when a pre-hospital time threshold of 45 min is used. Different values for the other parameters considered in the sensitivity analysis result in no more than a 1% impact on the objective function values for the values considered for any considered change to the

system. A change to the maximum distance between an air ambulance depot and a facility for which they conduct transfers lead to different locations of resources. As the maximum allowable distance between an ambulance depot to trauma center for air transfers increases, the air ambulances are located further from trauma centers. This improves the expected number of patients that can be transported to a trauma center within the pre-hospital time threshold while still being able to conduct transfers. Similarly, the maximum ground transfer distance impacts the optimal location of resources. A larger maximum ground transfer distance allows some air ambulances to be located further away from trauma centers since there is less reliance on air ambulances to complete patient transfers. However, in both cases, the patient density in areas that can be served with a larger distance is very low, so the improvement to the objective function values is relatively low.

6. Mass casualty incident response problem: A case study of Wisconsin tornadoes

We investigate the trauma system mass casualty incident response problem introduced in Section 4 using a case study based on a tornado-based MCI. Specifically, we consider how to add or reposition air ambulance resources in preparation of patient surges following an F5 tornado in Wisconsin. Patient demands are based on historical tornado locations [45] and previous work on patients volumes [46], locations, and arrival times [47] following a tornado.

We begin by discussing the data used to instantiate the model. We then discuss the results for a specific tornado event to gain a better understanding of how the model performs and how it makes decisions in preparation of a specific weather event. We investigate the sensitivity of model results to various parameters used within the model instances. Finally, we present and broadly analyze the results of fifty model instances, each corresponding to a different tornado MCI.

6.1. Case study design

We consider a F5 tornado in Wisconsin as the MCI to which the trauma system must respond. We assume the tornado has a starting location, bearing, width, and length that follows the historical distribution of tornadoes in Wisconsin from 1950–2017 [45].

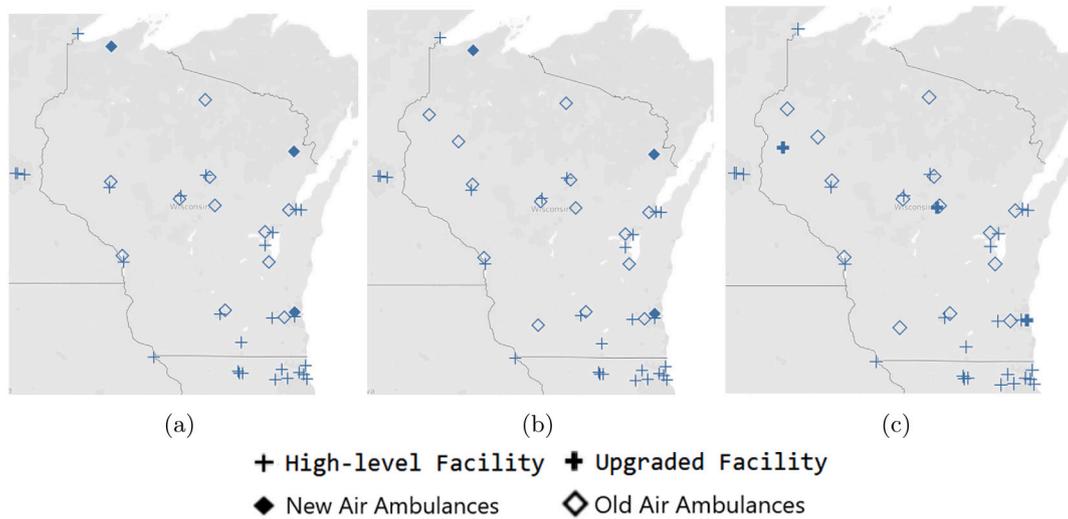


Fig. 7. When patients must be transported directly to high-level facilities ($\epsilon = 100\%$), the optimal locations to (a) relocate three current air ambulances ($m = 3$), (b) locate three new air ambulances ($n = 3$), and (c) upgrade three low-level trauma centers ($u = 3$).

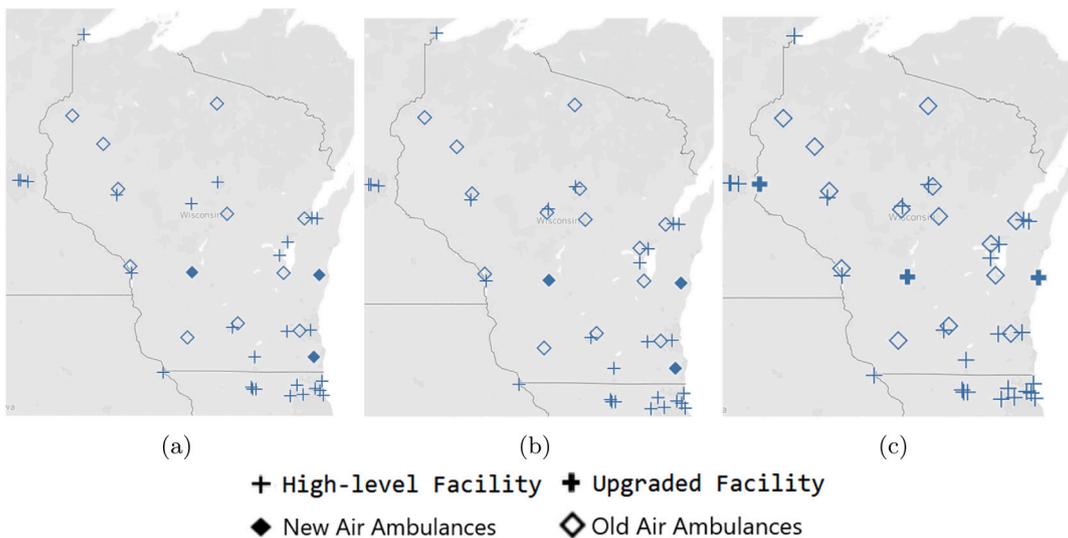


Fig. 8. When patients may be transferred ($\epsilon = 0\%$), the optimal locations to (a) relocate three current air ambulances ($m = 3$), (b) locate three new air ambulances ($n = 3$), and (c) upgrade three low-level trauma centers ($u = 3$).

Much of the data, such as travel time calculations, trauma center locations, and ambulance depot locations, used in this case study are identical to those described previously in Section 5.1. Therefore, in this section, we describe only the changes and additions to the data presented in Section 5.1. Table 5 provides the base parameter values of the case study and parameter values used in the sensitivity analysis.

Patients entering the trauma system following the tornado are from one of two groups: patients whose injuries are caused by the MCI and patients whose injuries are not caused by the MCI and would enter the trauma system regardless. Patients not caused by the MCI are, like before, considered to be related to population density and historical motor vehicle crashes. We create scenarios of patient locations and arrival times by sampling from the following distributions. MVC victim arrivals are sampled from historical crash data. For the remaining patients, we assume they enter the trauma system according to a Poisson distribution with arrival rates equal to those in the NEMSIS database proportionally split by census block group population.

We construct scenarios representing the number of patients and location of injury due to the tornado by sampling from the distribution of the random variables as follows. As a result, $Pr(\xi) = \frac{1}{|\Xi|}$. Injuries

occur in few tornadoes, but when injuries do occur there are often many [46]. In this case study, we consider tornadoes that result in an average of at least 50 injuries as impacted areas are likely to require the additional air ambulances only with a large patient surge. We assume the rate of injury is distributed according to a Negative Binomial distribution based on the work of Elsner et al. [46], who find a relationship between the mean number of injuries, the energy of the tornado, the population density of the area affected using Negative Binomial regression. Some injuries are severe enough that the injured individual becomes a trauma patient. Historical data suggests that up to 25% of injured individuals result in hospitalizations [47]. Therefore, we assume the number of patients entering the trauma system follows a Binomial distribution where the probability an injured individual becomes a trauma patient is $20\%(54\%) \approx 11\%$, where 54% arises from the historical proportion of patients transported to a low-level trauma center are transferred to a high-level trauma center [32]. Injuries are usually spread over the entire area of destruction [47]. We assume patients caused by the MCI are randomly distributed by population density across the affected area (determined by the path and width of the tornado). Historical data suggests that most of these severely

Table 5

Values used to instantiate the response case study and sensitivity analysis. The ‘Symbol’ column indicates the symbol the input impacts, but not necessarily the direct value the symbol takes.

Model parameter or input	Symbol	Base value	Sensitivity analysis
Relocated air ambulances	m	0	{1, 2, 3}
New air ambulances	n	0	{1, 2, 3}
Transport time threshold	ϕ_{ijk}^{ξ}	60 min	45 min
Patient demand distribution	I	Historical data	–
Trauma centers	J	$ J = 138$	–
Air ambulance depots	K_a	$ K_a = 124$	–
Current air ambulances	K'_a	$ K'_a = 14$	–
Air ambulance capacity	$c_k, k \in K_a$	3 routes	{2 routes, 4 routes}
Trauma center capacity	$e_j, j \in J_e$ and $j \in J_h$	30 patients	{20 patients, 40 patients}
Time until air travel permitted	–	2 h	{1 h, 3 h}
Max air transfer distance	$F'_j, j \in J$	220 km	–
Air ambulance distance to trauma center for transfers	$D'_j, j \in J$	55 km	–
Max ground transfer distance	$F_j, j \in J$	30 km by road	–
Air ambulance busy time	$P_{ik}(\xi), k \in K, i \in I(\xi), \xi \in \Xi$	2.5 h	–

injured patients arrives between 1 and 4 h after the tornado [47]. Thus, we assume all trauma patients arrive uniformly between one and four hours following the tornado.

As before, we assume that 3 patient transports or transfers may be completed by each air ambulance and investigate capacities of 2 and 4 within the sensitivity analysis. Moreover, during a tornado MCI, it is reasonable to assume that an air ambulance cannot be used to transport the initial trauma patients due to high winds and limited visibility caused by the storm [48]. In our case study, we assume that within the first two hours following the tornado, only ground ambulances may transport patients to a trauma center. Within the sensitivity analysis, we consider one and three hours following the tornado.

We assume the capacity of each trauma center is 30 patients in the base case and investigate a capacity of 20 and 40 patients within the sensitivity analysis. These numbers may be lower than official capacities of trauma centers. This is due to the difference between surge capacity and surge capability at various trauma centers. Surge capacity measures how many patients a hospital can treat within a fixed period of time (e.g., one hour) [48]. By comparison, surge capability refers to a hospital’s ability to treat patients requiring complex or specialized care [48]. Therefore, even if a hospital has significant surge capacity, it may have low surge capability and patients suffer worse outcomes if too many patients are treated at a single hospital. As a result, surge capacities may be overly optimistic. Thus, we use capacity values that are actually closer to surge capabilities.

6.2. Case study results — single tornado

In this section, we present the results of the stochastic model applied to a single, randomly generated tornado. The path of this tornado is illustrated in 9(a). Rural MCI’s like this are not a rarity. Elsner et al. [46] found that a tornado’s energy is far more predictive of the number of resulting injuries compared to the population density of the area that the tornado impacts. For this tornado, we consider 50, 100, and 200 possible scenarios of patient arrivals. We approximate the Pareto frontier (potentially including weakly optimal solutions) using the ϵ -constraint method. For each number of scenarios, the planning decisions (i.e., air ambulance locations) were identical. Due to the homogeneity of the results, we only discuss results from using the 100 scenarios. Unless otherwise noted, the results presented in this section refer to solutions using the base parameter values in Table 5.

The approximate Pareto frontier strongly indicates that, in this rural instance, it is better to first triage patients at lower-level hospitals before transferring them (lower ϵ) if pre-hospital times are a factor in decision-making. For any level of new or relocated air ambulances, we find a substantial decrease in the number of patients that can be transported to a trauma center within 60 min as a larger proportion of patients must be transported directly to a high-level trauma center

(i.e., as ϵ increases). This is due to the restricted capacity of high-level trauma centers near the MCI causing patients to be transported to trauma centers further away, rather than low-level trauma centers near the MCI. Allowing patients to be transferred to high-level trauma centers after stabilization at low-level trauma centers results in approximately 50% more of the patients expected to be transported to a trauma center within 60 min and receive at least stabilizing care. Beyond shorter pre-hospital times to low-level trauma centers, there are also logistical considerations. For example, tornadoes create significant debris, which impedes the use of air ambulances to transport patients from the scene [48], especially for crews not trained to navigate this debris. It is also easier to restock medical supplies when air ambulances are located at depots near trauma centers, which is a result of a system with many patient transfers. While this does not imply that air ambulance transports should always be avoided, it suggests that there are substantial trade-offs between both characteristics of high quality care.

Similar to our observation within the planning model solutions in the previous section, the improvements from relocating and locating air ambulances are more significant when patients must be transported directly to a high-level trauma center ($\epsilon = 100\%$) compared to when patients can be transferred to high-level facilities later ($\epsilon = 0\%$). As seen in Figs. 9(a) and 9(b), a majority of the tornado path crosses low population areas. As a result, most MCI patients are far from high-level trauma centers. This causes additional air ambulances to have a significant improvement on the likelihood a patient in an impacted area may be transported to a high-level trauma center within the 60 min pre-hospital time.

Figs. 9(a) and 9(b) illustrate the locations of air ambulances when patients are transported directly to high-level trauma centers or can be transferred to them, respectively. When air ambulance are added to the trauma system, rather than relocated, they are placed at the same depots as relocated air ambulances. We see a similar pattern in results of the incident response model to that of the planning model. When patients are transported directly to high-level trauma centers ($\epsilon = 100\%$), shown in Fig. 9(a), air ambulances are relocated closer to the locations expected to have patients that are not currently within the range of a high-level facility in order to facilitate those patient transports. When transfers to high-level trauma centers are allowed ($\epsilon = 0\%$), shown in Fig. 9(b), patients are almost always transported to local low-level trauma centers by ground ambulance. To facilitate patient transfers, the air ambulances are relocated to a nearby high-level trauma centers without sufficient air ambulance resources.

No parameter, other than the transport time threshold, had a large impact on the planning decisions regarding air ambulance locations or relocations. When patient transfers are allowed ($\epsilon = 0\%$), lowering the transport time threshold to 45 min decreases the proportion of patients transported to a trauma center within the time threshold by about

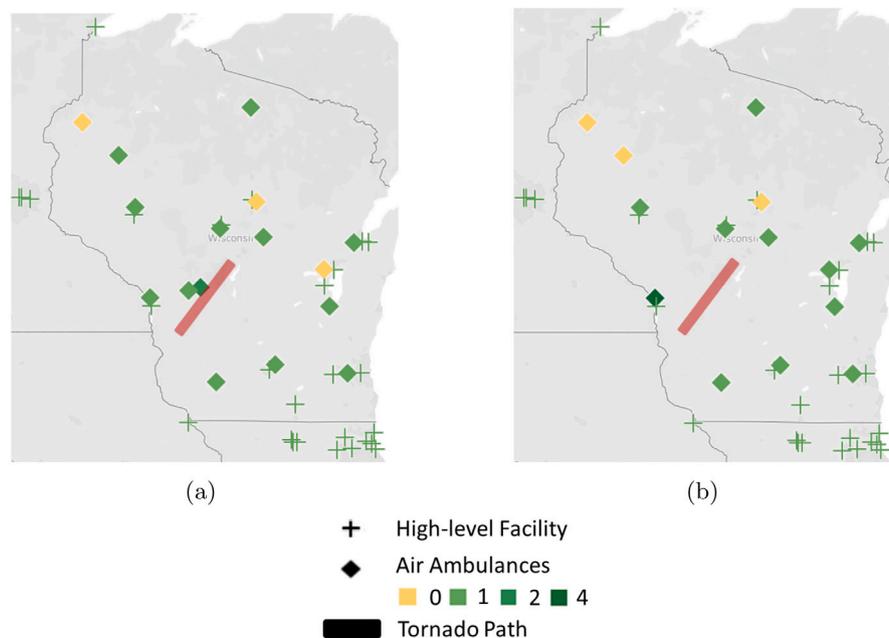


Fig. 9. The locations and number of air ambulances when three air ambulances are relocated. Depots shown with 0 air ambulances host an air ambulance that was relocated. Other possible depots are not illustrated. In Figure (a) patients are transported directly to high-level facilities ($\epsilon = 100\%$); and in Figure (b) patients may be transferred ($\epsilon = 0\%$).

30%, but does not change patient pathways through the trauma system. By comparison, when direct transport to a high-level trauma center is required ($\epsilon = 100\%$), a 45 min threshold significantly disincentives using air ambulance resources to respond to and transport MCI-caused patients to a trauma center, since the probability of reaching a high-level trauma center within 45 min is low even with an air ambulance. As a result, using a 45-minute time threshold when optimizing a rural MCI response may not provide desired solutions.

6.3. Case study results — 50 tornadoes

The results discussed in Section 6.2 are for a single tornado path. In this section, we briefly discuss the average results of fifty model solutions, each corresponding to a unique tornado path, using the base parameter values and solution times for these problem instances. Each tornado is generated as described in Section 6.1 and is given a starting location, bearing, width, and length according to historical data. For each tornado, we generate 100 scenarios of patient arrivals to the trauma system.

Table 6 presents the mean proportion of patients transported to a trauma center within 60 min after the 50 tornadoes. We see similar trends in the results to those discussed in Section 6.2. The proportion of patients that are transported to a trauma center within 60 min begins to decrease significantly, even when only 50% of patients must be transported directly to high-level trauma center ($\epsilon = 50\%$). For each tornado path and scenario, the optimal location of air ambulances and patient pathways are heterogeneous, reflecting the various tornado paths.

We find a noticeable difference in the results when the tornado impacts rural and urban areas. Urban areas, despite the increased number of patients, have far greater trauma care and air ambulance capacity. Thus, urban areas often have sufficient capacity to care for patients without transporting or transferring patients out of the region. Rural tornadoes see a far greater decrease in the proportion of patients transported within 60 min as we increase the proportion of patients being transported directly to high-level trauma centers (larger values of ϵ). Likewise, we find a much larger improvement in the objective function values when air ambulances are relocated in response to tornadoes in rural areas compared to urban areas.

Table 7 presents the mean computational solution times for 50 model solutions. Unsurprisingly, using $\epsilon = 50\%$ results in significantly longer average computational times versus $\epsilon = 0\%$ or $\epsilon = 100\%$. This is due to the additional complexity of determining which patients to transport versus transfer to a high-level facility. This does not occur when $\epsilon = 100\%$ as all patient must be transported directly to a high-level facility. When $\epsilon = 0\%$ it is usually optimal to transport all patients to the nearest facility and thus the computational requirement is relatively low.

Additionally, due to the aforementioned heterogeneity of the tornadoes, we see significant variation in the computational solution times. This may lead to unacceptably long practical solution times using sample sizes that are at least 100 scenarios, especially when $0\% < \epsilon < 100\%$. However, for a given tornado, if 50 scenarios are considered instead of 100, almost all first-stage solutions are the same, but computational times are much smaller. Only in one of the 50 model instances was a relocated ambulance location different, and this occurred when only one air ambulance was allowed to relocate. Therefore, while it is likely better to use 100 scenarios per tornado, using 50 scenarios often provides the same first-stage solutions and can be found in far less time, as shown in Table 7.

7. Conclusion

We introduce models for locating resources within a trauma system. The trauma system planning problem is to determine the location of new air ambulances, where to relocate air ambulances, and which low-level trauma centers to upgrade and transform into high-level trauma centers. The incident response problem focuses on a shorter time-frame decision of determining the optimal location to temporarily reposition air ambulances shortly after a mass casualty incident, but before the patient volume is known with certainty.

Our results suggest that the optimal locations of trauma resources change significantly depending if patient transfers are considered or if patients must be transported directly to a high-level trauma center. For example, when all patients must be transported directly to a high-level trauma center, air ambulances tend to be spread further away from trauma centers to increase the geographic area of patients that may be transported to a high-level trauma center within the time

Table 6
Mean proportion of patients transported within 60 min, $\mathbb{E}[Q_1(\mathbf{w}, \mathbf{r}, \xi)]$, for 50 tornadoes with 100 scenarios per tornado.

Relocated air Ambulances	New air Ambulances	Proportion transported to a high-level facility Within 60 min, $\mathbb{E}[Q_1(\mathbf{w}, \mathbf{r}, \xi)]$, if		
m	n	$\mathbb{E}[Q_2(\mathbf{w}, \mathbf{r}, \xi)] \geq 0\%$	$\mathbb{E}[Q_2(\mathbf{w}, \mathbf{r}, \xi)] \geq 50\%$	$\mathbb{E}[Q_2(\mathbf{w}, \mathbf{r}, \xi)] = 100\%$
0	0	88.2%	79.0%	46.3%
1	0	88.6%	81.0%	48.5%
2	0	89.9%	81.4%	50.4%
3	0	89.1%	82.8%	51.3%
0	1	88.7%	81.3%	49.3%
0	2	89.0%	82.2%	50.4%
0	3	89.3%	83.3%	52.6%

Table 7
Mean computational solution times and standard errors for 50 tornadoes.

$\mathbb{E}[Q_2(\mathbf{w}, \mathbf{r}, \xi)]$	Number of scenarios Per Tornado	Mean time (s) \pm Standard Error
$\geq 0\%$	50	59 \pm 10
$\geq 50\%$	50	150 \pm 54
$= 100\%$	50	58 \pm 9
$\geq 0\%$	100	636 \pm 515
$\geq 50\%$	100	2400 \pm 2633
$= 100\%$	100	523 \pm 748

threshold. When patients may be transferred to a high-level trauma center, some air ambulances are located close to trauma centers to be able to facilitate the transfer of patients between facilities. The case studies also suggest that increasing the proportion of patients who must be transported directly to a high-level trauma center decreases the expected number of patients arriving at a trauma center within a desired pre-hospital time. We also find that changes to the trauma system have a larger impact on patient pre-hospital time when patients must be transported directly to a high-level trauma center. When patients may be transferred between facilities, the benefit of changes to the trauma system are relatively small due to the maturity of the existing system.

Both problems highlight the importance of air ambulances in transferring patients between trauma centers. Air ambulances should be located in areas to perform both patient transports from the scene of the injury to a trauma center and transfers between centers. Some patients, such as those in rural areas, are unlikely to be transported directly to a high-level trauma center and are instead far more likely to reach a low-level trauma center to receive stabilizing care if a later transfer to a high-level trauma center is allowed. The decisions on where to locate air ambulances and high level trauma centers change significantly depending on the proportion of patients that can be transferred after receiving initial care at a low-level trauma center. Therefore, we present multiple solutions each corresponding to a unique importance of each objective. The trade-off in these objectives highlights a larger question: should resources be allocated to improve the ability of the trauma care system to transfer patients between facilities or to improve the number of people with direct access to the highest level of trauma care. Determining which is best using cost-analysis and patient outcomes is an area for future research.

In both problems, the need to locate at least some air ambulances near trauma centers to efficiently facilitate transfers differs from what has been previously seen in the literature. The optimal resource locations we find in this research have more in common with the current location of air ambulances than previous research, since many air ambulances are located at and share personnel with high-level trauma centers.

In this paper, we have only considered patient data at a highly aggregated level. In practice, dispatchers and EMS personnel have significantly more information on the health of each patient. In real-time operations, this information informs the decision to either transport or transfer patients and by what means. A direction for future research

is to incorporate this information into model formulations to more accurately capture the movement of patients in the trauma system and model the patients' impact on resource usage with more granularity. For example, some patients could be directly discharged from low-level trauma centers without being transferred. Moreover, future research can consider investment to increase trauma center capacity or to stand-up new trauma centers altogether.

CRedit authorship contribution statement

Eric DuBois: Methodology, Software, Validation, Data curation, Writing – original draft, Visualization. **Adam Schmidt:** Writing – original draft, Writing – review & editing. **Laura A. Albert:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing.

Declaration of competing interest

Laura Albert reports financial support was provided by National Science Foundation.

Acknowledgments

This work was in part funded by the National Science Foundation, USA Award 1444219. The views and conclusions contained in this document are those of the author and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the National Science Foundation. The authors would also like to thank the anonymous reviewers, whose suggestions for improvement led to a substantially improved manuscript.

References

- [1] Centers for Disease Control and Prevention, National Centers for Injury Prevention and Control. Web-based injury statistics query and reporting system (WISQARS). 2005, URL www.cdc.gov/injury/wisqars.
- [2] Celso Brian, Tepas Joseph, Langland-Orban Barbara, Pracht Etienne, Papa Linda, Lottenberg Lawrence, et al. A systematic review and meta-analysis comparing outcome of severely injured patients treated in trauma centers following the establishment of trauma systems. *J Trauma Inj Infect Crit Care* 2006;60(2):371–8, arXiv:16508498.
- [3] Jarman Molly P, Castillo Renan C, Carlini Anthony R, Kodadek Lisa M, Haider Adil H. Rural risk: geographic disparities in trauma mortality. *Surgery* 2016;160(6):1551–9.
- [4] American College of Surgeons. Resources for the optimal care of the injured patient. 2014, URL <https://www.facs.org/-/media/files/quality-programs/trauma/vrc-resources/resources-for-optimal-care.ashx>.
- [5] Haas Barbara, Gomez David, Zagorski Brandon, Stukel Therese A, Rubinfeld Gordon D, Nathens Avery B. Survival of the fittest: the hidden cost of undertriage of major trauma. *J Am Coll Surg* 2010;211(6):804–11.
- [6] Alarhayem AQ, Myers JG, Dent D, Liao L, Muir M, Mueller D, et al. Time is the enemy: mortality in trauma patients with hemorrhage from torso injury occurs long before the “golden hour”. *Am J Surg* 2016;212(6):1101–5.
- [7] Gonzalez Richard P, Cummings Glenn R, Phelan Herbert A, Mulekar Madhuri S, Rodning Charles B. Does increased emergency medical services prehospital time affect patient mortality in rural motor vehicle crashes? A statewide analysis. *Am J Surg* 2009;197(1):30–4.

- [8] Toregas Constantine, Swain Ralph, ReVelle Charles, Bergman Lawrence. The location of emergency service facilities. *Oper Res* 1971;19(6):1363–73.
- [9] Church Richard, ReVelle Charles. The maximal covering location problem. 1974, 32(1):101–18 . <https://doi.org/10.1007/BF01942293>.
- [10] Hogan Kathleen, ReVelle Charles. Concepts and applications of backup coverage. *Manage Sci* 1986;32(11):1434–44.
- [11] Daskin Mark S. A maximum expected covering location model: formulation, properties and heuristic solution. *Transp Sci* 1983;17(1):48–70.
- [12] Daskin Mark S. Location, dispatching and routing models for emergency services with stochastic travel times. 1987, p. 224–65.
- [13] Goldberg Jeffrey, Paz Luis. Locating emergency vehicle bases when service time depends on call location. *Transp Sci* 1991;25(4):264–80.
- [14] Branas Charles C, MacKenzie Ellen J, ReVelle Charles S. A trauma resource allocation model for ambulances and hospitals. *Health Serv Res* 2000;35(2):489–507.
- [15] Røislien Jo, van den Berg Pieter L, Lindner Thomas, Zakariassen Erik, Aardal Karen, van Essen J Theresia. Exploring optimal air ambulance base locations in Norway using advanced mathematical modelling. *Inj Prev* 2017;23(1):10–5.
- [16] Erdemir Elif Tokar, Batta Rajan, Rogerson Peter A, Blatt Alan, Flanigan Marie. Joint ground and air emergency medical services coverage models: A greedy heuristic solution approach. *European J Oper Res* 2010;207(2):736–49.
- [17] Cho Soo-Haeng, Jang Hoon, Lee Taesik, Turner John. Simultaneous location of trauma centers and helicopters for emergency medical service planning. *Oper Res* 2014;62(4):751–71.
- [18] Ahmadi-Javid Amir, Seyedi Pardis, Syam Siddhartha S. A survey of healthcare facility location. *Comput Oper Res* 2017;79:223–63.
- [19] Farahani Reza Zanjirani, Asgari Nasrin, Heidari Nooshin, Hosseininia Mahtab, Goh Mark. Covering problems in facility location: A review. *Comput Ind Eng* 2012;62(1):368–407.
- [20] Demetriades Demetrios, Martin Matthew, Salim Ali, Rhee Peter, Brown Carlos, Doucet Jay, et al. Relationship between American College of Surgeons Trauma Center designation and mortality in patients with severe trauma (injury severity score > 15). *J Am Coll Surg* 2006;202(2):212–5.
- [21] US Government Accountability Office. GAO-10-907: air ambulance: effects of industry changes on services are unclear. 2010, GAO-10-907. URL <https://www.gao.gov/assets/320/310532.html>.
- [22] Nezamoddini Nasim, Khasawneh Mohammad T. Modeling and optimization of resources in multi-emergency department settings with patient transfer. *Oper Res Health Care* 2019;10:23–34.
- [23] Marjka Kevin, Miller Maile, editors. Atlas & Database of Air Medical Services. 16th ed.. CentTIR/AAMS/NHTSA; 2018, URL <http://www.adamsairmed.org>.
- [24] Wisconsin Department of Health Services. Wisconsin EMS system statistics. 2014, URL <https://www.dhs.wisconsin.gov/ems/data.htm>.
- [25] neos Guide. Stochastic Linear Programming | NEOS. URL <https://neos-guide.org/content/stochastic-linear-programming>.
- [26] Harrington David T, Connolly Michael, Biffel Walter L, Majercik Sarah D, Cioffi William G. Transfer times to definitive care facilities are too long: A consequence of an immature trauma system. *Ann Surg* 2005;241(6):961–6.
- [27] Whedon James M, von Recklinghausen Friedrich M. An exploratory analysis of transfer times in a rural trauma system. *J Emerg Trauma Shock* 2013;6(4):259–63.
- [28] MSNBC. Hospital patients evacuated after destruction. 2011, URL http://www.nbcnews.com/id/43137115/ns/health-health_care/t/hospital-patients-evacuated-after-destruction/.
- [29] Pfetsch Marc E, Rehn Thomas. A computational comparison of symmetry handling methods for mixed integer programs. *Math Program Comput* 2019;11(1):37–93. <http://dx.doi.org/10.1007/s12532-018-0140-y>.
- [30] Novak Bill. Med flight adding helicopter at portage; third location for the UW health service. 2018, URL https://madison.com/news/local/health-med-fit-med-flight-adding-helicopter-at-portage-third-location-for-the/article_6ea83bd9-16bb-5c12-a18f-5669ad8134b2.html.
- [31] Roznik Sharon. Flight for life moving to new location to consolidate Fond Du Lac/Waukesha operations. 2018, URL <https://www.fdlreporter.com/story/news/2018/11/30/flight-life-leaving-fond-du-lac-waukesha/2165954002/>.
- [32] Wisconsin Department of Health Services. Wisconsin trauma registry. 2019, URL <https://www.dhs.wisconsin.gov/trauma/registry.htm>.
- [33] Haas Barbara, Doumouras Aristithes G, Gomez David, de Mestral Charles, Boyes Donald M, Morrison Laurie, et al. Close to home: an analysis of the relationship between location of residence and location of injury. *J Trauma Acute Care Surg* 2015;78(4):860–5.
- [34] Stewart Kenneth, Garwe Tabitha, Bhandari Naresh, Danford Brandon, Albrecht Roxie. Factors associated with the use of helicopter inter-facility transport of trauma patients to tertiary trauma centers within an organized rural trauma system. 2016, 20(5):601–8. <https://doi.org/10.3109/10903127.2016.1149650>, URL <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5012941/>.
- [35] Office of Preparedness and Emergency Health Care. Wisconsin trauma at a glance 2015–2016. Technical report. P-01564, Wisconsin Department of Health Services; 2016, p. 2, URL <https://www.dhs.wisconsin.gov/publications/p01564.pdf>.
- [36] Wisconsin Department of Health Services. Hospital trauma classification process. 2014, URL <https://www.dhs.wisconsin.gov/trauma/classification.htm>.
- [37] National Highway Traffic Safety Administration. National EMS information system. 2018, URL <https://nemsis.org>, [Data Set].
- [38] Andruszkow Hagen, Schweigkofler Uwe, Lefering Rolf, Frey Magnus, Horst Klemens, Pfeifer Roman, et al. Impact of helicopter emergency medical service in traumatized patients: which patient benefits most? *PLoS One* 2016;11(1):e0146897.
- [39] Froehler Michael, Fusco Matthew, Chitale Rohan, Liebeskind David, Zaidat Osama, Jahan Reza, et al. Air versus ground transport for interhospital transfer prior to endovascular stroke treatment. *Stroke* 2019;50.
- [40] Budge Susan, Ingolfsson Armann, Zerom Dawit. Empirical analysis of ambulance travel times: The case of calgary emergency medical services. *Manage Sci* 2010;56(4):716–23.
- [41] Kim Oh Hyun, Roh Young-Il, Kim Hyung-Il, Cha Yong Sung, Cha Kyoung-Chul, Kim Hyun, et al. Reduced mortality in severely injured patients using hospital-based helicopter emergency medical services in interhospital transport. *J Korean Med Sci* 2017;32(7):1187–94.
- [42] Di Rocco Damien, Pasquier Mathieu, Albrecht Eric, Carron Pierre-Nicolas, Dami Fabrice. HEMS inter-facility transfer: A case-mix analysis. *BMC Emerg Med* 2018;18(1):13.
- [43] Holst Jenelle A, Perman Sarah M, Capp Roberta, Haukoos Jason S, Ginde Adit A. Undertriage of trauma-related deaths in U.S. emergency departments. *West J Emerg Med* 2016;17(3):315–23, arXiv:27330664.
- [44] DuBois Eric. Optimization models with applications to homeland security systems [Ph.D. thesis], Madison, WI: University of Wisconsin-Madison; 2019.
- [45] National Weather Service. Severe weather database. 2018, URL <https://www.spc.noaa.gov/wcm/#data>.
- [46] Elsnor James B, Fricker Tyler, Berry William D. A model for U.S. Tornado casualties involving interaction between damage path estimates of population density and energy dissipation. *J Appl Meteorol Climatol* 2018;57(9):2035–46.
- [47] Bohonos Jay J, Hogan David E. The medical impact of tornadoes in North America. *J Emerg Med* 1999;17(1):67–73.
- [48] Smith Jack E, Gebhart Mark E. Ciottoni's disaster medicine. 2nd ed.. Elsevier; 2016.