Computers & Industrial Engineering xxx (xxxx) 109250

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Nested trauma network design considering equity and effectiveness in patient safety

Sagarkumar Hirpara [a], Pratik J. Parikh [a,*], Nan Kong [b]

[a] Department of Industrial Engineering, University of Louisville, Louisville, KY, USA
[b] Weldon School of Biomedical Engineering, Purdue University, West Lafayette, IN, USA

ARTICLE INFO

Keywords:
Operations Research in Health Services
Trauma Center Location
Patient safety
Bi-objective optimization

ABSTRACT

Trauma injuries continue to be the leading cause of mortality and morbidity among US citizens aged 44 years and under. Government agencies are often in charge of designing an effective trauma network in their region to provide prompt and definitive care to their citizens. This process is, however, largely manual, experience-based and often leads to a suboptimal network in terms of patient safety. To support effective decision making, we propose a Nested Trauma Network Design Problem (NTNDP), which can be characterized as a nested multi-level, multi-customer, multi-transportation, multi-criteria, capacitated model with the bi-objective of maximizing the weighted sum of equity and effectiveness in patient safety. We use mistriages (system-related under- and over-triages) as surrogates for patient safety. To add realism, we include intermediate trauma centers that are set up in many states in the US to serve as feeder centers to major trauma centers to improve patient safety and three criteria to mimic EMS's on-scene decisions. We propose a '3-phase' solution approach that first solves a relaxed version of the model, then solves a Constraint Satisfaction Problem, and a modified version of the original optimization problem (if needed), all using a commercial solver. Our findings suggest that solutions are sensitive to (i) the proportion of assignments attributed to various destination determination criteria, (ii) distribution of trauma patients, and (iii) relative emphasis on equity vs. effectiveness. We also illustrate the use of our approach using real data from a midwestern US state; results show over 30% performance improvement in the objective value.

1. Introduction

In the US, trauma is the leading cause of death for individuals aged 44 and under (#3 across all ages), resulting in almost 200,000 deaths and an economic burden of over $670 billion annually (ACS, 2016; CDC, 2022). Trauma is a serious public health problem with significant social and economic costs. A trauma care system in a state (or a region within a state) is often established in an attempt to provide prompt and definitive care to trauma patients. Timely access to a trauma center (TC) is one of the key determinants of patient outcomes (Branas et al., 2013; Jansen et al., 2015).

1.1. Types of trauma centers

The American College of Surgeons (ACS) verifies TCs as Levels I-V based on the presence of the type of trauma resources and their availability (American Trauma Society, 2022). ACS-verified Levels I and II are referred to as major trauma centers (MTCs) and capable of providing definitive care for patients suffering from major traumatic injuries (i.e., severely injured patients). MTCs are equipped with highly sophisticated surgical and diagnostic equipment, with 24/7 surgeon availability, to provide high-quality medical and nursing care. While timely access to a ~~H~~ (MTC) improves survival of severely injured patients by 25% relative to care delivered at a non-trauma center (MacKenzie et al., 2006).

According to the Centers for Disease Control and Prevention, "there is no access to an advanced trauma center for nearly 45 million Americans within the golden hour (60 min)" (ACS, 2016). The reason for this is the geographic maldistribution of MTCs in the U.S.; in 2010, 9 states had a clustered pattern, 22 had a dispersed pattern, and 10 had a random pattern (Brown et al., 2016). Further, there is a significant cost associated with building and operating MTCs, and it can be financially challenging to open an MTC in rural areas due to concerns of sufficient patient volume.

To circumvent this problem, Levels III-V TCs are set up to serve as feeder centers to MTCs for communities that do not have timely access

* Corresponding author at: Department of Industrial Engineering, 132 Eastern Parkway, University of Louisville, Louisville, KY 40292, USA.
E-mail address: pratik.parikh@louisville.edu (P.J. Parikh).

https://doi.org/10.1016/j.cie.2023.109250
Received 15 January 2023; Received in revised form 9 April 2023; Accepted 13 April 2023
0360-8352/© 20XX

to MTCs; we refer to such TCs as intermediate TCs (ITCs). ITCs provide a subset of services offered by ~~LI/LII~~ MTCs, but only during part of the day, and serve as centers for initial care, resuscitation, and subsequent transfer to major trauma centers (MTCs). It has been shown that an inclusion of ITCs in underserved counties decreases trauma-related mortality rates due to improved survival of transferred severely injured patients after stabilizing at those ITCs (Barringer et al., 2006; Tinkoff et al., 2007). After stabilizing, a patient is eventually transferred to an MTC as ITCs are not capable of providing definitive care to severely injured patients. All other hospitals are referred to as non-trauma centers (NTCs), which are the ideal destination for non-severely injured trauma patients.

1.2. On-field decisions and trauma triage

The majority of trauma deaths occur in the pre-hospital environment or within 4 h of the trauma event (ACEP, 1987). The pre-hospital trauma triage is designed to transport the right patient to the right hospital at the right time. The emergency medical service (EMS) is crucial in providing initial care to the injured patient and accurate pre-hospital triage. EMS providers' on-field decision making practice involves two components; (i) injury assessment (how severe the injuries are) and (ii) destination determination (which hospital to select and how to transport). An error in making any of these decisions can lead to pre-hospital mistriage.

Note that, besides mortality, mistriage has been used in the trauma literature as a surrogate for patient safety as it often increases the risk of short/long disability caused due to delay in provision of definitive care (Jansen et al., 2015; Hirpara et al., 2022, Parikh et al., 2022). Consequently, considering (i), an error in accurately assessing the injury type (severe or non-severe) can lead to 'clinical mistriage.' Similarly, for (ii), an error in determining the most suitable hospital type (trauma center or not) can lead to 'system-related mistriage.'

1.3. System-related mistriages

We define three types of system-related mistriages (as surrogates for patient safety). A situation when a severely injured patient is taken to an NTC because of a lack of access to an MTC experiences is referred to as 'system-related under-triage (srUT).' Further, in a trauma network with MTCs and ITCs, if a severely injured patient, who ideally should be transported to an MTC, is first transported to an ITC due to lack of access to MTC, then we refer to that as 'system-related under-triage stabilized ($srUT^{S}$).' We use the modifier 'stabilized' because an ITC has the ability to provide prompt assessment, resuscitation, limited surgery, intensive care and stabilization of injured patients and emergency operations, compared to an NTC (in which case we would have referred this patient as srUT). In contrast, an excess (or cluster per Brown et al., 2016) of MTCs and ITCs in the vicinity of an incidence location (also known as scene) could induce EMS to transport a less severely injured patient to such hospitals, which we refer to as 'system-related over triage (srOT).'

Generally, srUT (and $srUT^{S}$) and srOT have negative implications on patient safety. A srUT increases the likelihood of an adverse outcome such as disability, morbidity, and even mortality due to delay in receiving definitive care (Rotondo et al., 2014). In contrast, a srOT indirectly impacts patient safety by causing overcrowding at emergency departments (Lerner, 2006), unnecessary trauma activation resulting in additional charges to the patient, and loss of salvageable lives in mass casualty trauma (Frykberg, 2002; Armstrong et al., 2008).

1.4. Trauma network's influence on destination determination

It is during the destination determination phase when the network of MTCs and ITCs is critical. Table 1 shows three destination determination criteria used by EMS providers at the incidence location, the decision makers, and how the network of MTC/ITC impacts the corresponding decision. Clearly, the network of MTC/ITC influences the selection of an appropriate hospital for prompt and definitive care, eventually reducing mistriages and improving patient care.

Table 1
Influence of the trauma network on different destination determination criteria.

Influence of the MTC/ITC network
For severe injuries, take to an MTC (ideally) or ITC (if no MTC available)
Choices tend to favor MTC or ITC based on perception of different hospital types in the vicinity, past experience, and access time
Take to nearest hospital (even if NTC) during extreme weather condition or road closure

Although trauma literature alludes to the importance of network of MTC and ITC and implications on mistriages (a key patient safety metric), there key questions are yet to be addressed, which form the basis of our research.

1.5. Focus of this work

This paper ~~focus~~ on the strategic decision of jointly determining the number and location of MTCs and ITCs to improve patient safety. We address the following questions:

1. How do ITCs support patient safety?
2. What effect does destination determination criteria have on the MTC/ITC network?
3. How sensitive is the MTC/ITC network to the distribution of trauma patients?
4. What is the impact of focusing on equity of patient safety on the trauma network's performance?

The key contributions of our research are as follows. First, we propose a Nested Trauma Network Design Problem (NTNDP), which is a nested multi-level, multi-customer, multi-choice, multi-transportation capacitated model with a bi-objective of maximizing equity and effectiveness in patient safety. Multi-choice refers to the inclusion of all 3 dominant criteria for destination determination (see Table 1). While 'equity' quantifies the level of similarity in patient safety across regions in a geographical area (portion of a state or the state), 'effectiveness' quantifies overall patient safety (see Section 3.1 for details). Second, we propose a three-step approach to efficiently solve the proposed MIP model. This approach is able to find a near-optimal solution in a reasonable amount of time for instances of realistic problem sizes. Finally, to test our approach, we generate several datasets with different distributions of trauma patients using information available from the trauma system of Ohio, a midwestern US state. We also evaluate the sensitivity of the solution to variations in proportion attributed to the 3 destination determination criteria, weights associated with equity and effectiveness, and different distributions of patients. Finally, we illustrate the use of our approach for real data from a midwestern US state (i.e., the state of Ohio).

Our experiments suggest that destination determination criteria impact a trauma system's design and performance. While ACS and many state trauma agencies recommend using 'protocol' as the primary destination determination criteria, increased use of 'patient choice' criteria (often practiced in reality) results in more ITCs in suburban and rural zones; the corresponding mistriages are also high. Further, for the same number of patients, dispersed distribution of patients results in a 21.8% decrease in the trauma network performance (i.e., causes high mistriages) even with almost 3 times of ITCs in the network compared to cluster distribution. Further, if only equity among regions was emphasized (compared to effectiveness), the performance of the resulting network declines by over 8% given the limitations inherent in the equity objective. Using real data from OH for 2019, we demonstrate that the state could achieve a 31.2% and 33.1% reduction in mistriages by using our approach to redistribute and optimize their trauma network.

In the following sections, we first review the existing literature in Section 2. Our proposed optimization model for NTNDP and the solution approach are discussed in Sections 3 and 4, respectively. Next, we discuss our experimental study in Section 5 and illustrate the use of our approach on a real network in Section 6. Finally, in Section 7, we summarize our key findings and Discuss avenues for further research.

2. Literature review

Several approaches to address a variety of healthcare facility location problems have been proposed; e.g., primary health centers (Güneş et al., 2014), long-term care centers (Cardoso et al., 2015; Intrevado et al., 2019), preventive healthcare facilities (Zhang et al., 2009; Zhang et al., 2010), ambulance location and/or relocation (Reuter-Oppermann et al., 2017; Van Buuren et al., 2018), among others. For a comprehensive review, see Reuter-Oppermann et al. (2017), Ahmadi-Javid et al. (2017), and Güneş et al. (2019).

Because our work focuses on patient safety, our review suggests that two types of surrogate metrics for patient safety have been widely used in the literature; (i) minimizing total distance or travel time across all constituents (Cocking et al., 2012; Schmid, 2012; Beliën et al., 2013; Toro-Díaz et al., 2013; Chen et al., 2013; Bayram et al., 2015) and (ii) maximizing demand coverage within a fixed access time (Ingolfsson et al., 2008; Balcik & Beamon, 2008; Lim et al., 2011; Shariff et al., 2012; Kim & Kim, 2013; Salman & Yücel, 2015).

In terms of patient safety in trauma network design, Branas et al. (2000) proposed a model (known as TRAMAH) to simultaneously locate major trauma centers and air ambulances to maximize coverage of severely injured patients. Cho et al. (2014) also presented a model to simultaneously find major trauma centers and medical helicopters to maximize the expected number of patients transported to an MTC within 60 min. The authors incorporated busy fraction of medical helicopters in their model and developed the Shifting Quadratic Envelopes algorithm to optimize the problem. Lee and Jang (2018) extended this model to a multiperiod location model by introducing an additional decision on when to locate trauma centers and air ambulances over a planning horizon. Considering additional complexity, the authors proposed a solution approach that iteratively updates helicopters' availability using the previous step of optimization result. However, these approaches do not account for non-severely injured patients (who affect srOT) and intermediate trauma centers (which can improve access in rural areas).

Jansen et al. (2015) proposed a novel data-driven approach to locate MTCs and ITCs with the bi-objective of minimizing the total access time and the number of exceptions or srUT for Scotland. The same authors developed a multi-fidelity surrogate-management strategy to reduce the computation time for real-world data-driven optimization problems (Wang et al., 2016). They demonstrated the viability of their approach using real data from the state of Colorado's trauma system (Jansen et al., 2018). While this model considered ITCs, it failed to account for non-severely injured patients and various destination determination criteria.

To support decision making around trauma networks, the ACS Committee on Trauma (ACS COT) developed the Needs-Based Assessment of Trauma System (NBATS) tool (ACS-NBATS, 2015). NBATS uses six criteria to suggest the required number of MTCs in a given geographical area, also known as the trauma service area (TSA); population, median travel times, lead agency support, an existing number of major trauma centers, and where severely injured patients are transported (MTCs and NTCs). However, NBATS does not determine the location of the MTCs. To address this gap, Parikh et al. (2022) proposed a model for a Performance-based Assessment of Trauma System (PBATS) to find the minimum number and location of MTCs by keeping system-related under-triage (srUT) and over-triage (srOT) rates within a prespecified limit. Recently, Hirpara et al. (2022) proposed a bi-objective model for trauma center location problem (TCLP) to determine the number and location of MTCs and NTCs in order to minimize the weighted sum of srUT and srOT rates. They demonstrated their approach through a case study based on the existing network of a US state with focus on 'green-field' design and 'redistribution' of existing MTCs. While both these recent works consider both types of patients and associated mistriages, they do not explicitly consider ITCs (a critical trauma facility for a viable trauma system) and various destination determination criteria (that affect mistriages).

In terms of destination determination criterion, prior trauma location models have only considered a single criterion, often mimicking the ACS-suggested protocol. However, multiple criteria have been observed in practice besides this protocol, with patient choice and closest facility being dominant (Newgard et al., 2011, Newgard et al., 2013). Patient choice has been studied in many IE/OR journals to determine destination location in an optimization framework. Zhang et al. (2012) studied the impact of client choice behavior on the preventive care facility network configuration. The authors presented two alternative models; (i) probabilistic-choice model based on the multinomial logit (MNL) model, where a client may patronize each facility with a certain probability based on the attractiveness of the facilities, and (ii) optimal-choice model, where each client will go to the most attractive facility. Zhang and Atkins (2019) presented several models for designing a network of walk-in medical facilities. For a choice model, they considered travel time, attractiveness, and waiting time at the facility to calculate the utility of receiving care at a given facility. Further, they also considered deterministic patient choice, where a patient chooses the facility with the highest utility to receive care. Closest facility criteria have been considered in Cardoso et al. (2015), Mestre et al. (2015), and Nasrabadi et al. (2020).

Our review of the literature suggests the following gaps:

- All prior trauma system design approaches failed to explicitly consider multiple destination determination criteria alluded in medical literature and followed in practice.
- None of the prior research considered both types of patients (severe and non-severe), along with consideration of intermediate trauma centers.
- Further, equity in safety among regions, along with effectiveness, have not been considered jointly in the trauma literature (further elaborated in Section 3).

To fill the above gaps, we propose a nested multi-level, multi-customer, multi-destination determination criteria and multi-transportation bi-objective (equity and effectiveness) capacitated model. Our proposed NTNDP model not only accounts for both types of patients (severely and non-severely injured) and associated mistriages, but also explicitly considers several other factors that affect system performance; i.e., ITCs, three criteria for destination determination, and equity and effectiveness in patient safety. We now present our proposed model.

3. A bi-objective model for NTNDP

Our generic model is developed for a Trauma Service Area (TSA); a geographical area comprising a collection of counties in a state, the state itself, or even collection of states, similar to the definition in NBATS tool. Further, this TSA is divided into subareas known as regions or districts, which have the oversight to providing trauma care within that region. Because of the existence of such regions within a TSA, it becomes critical to consider the equity of patient safety among regions when designing a trauma network.

A variety of equity measures have been proposed in the literature when allocating public resources; e.g., minimax, variance, range, sum of absolute deviations, sum of absolute deviation from desire standard, squared coefficient of variation, and Gini index (Burkey et al., 2012; Lejeune and Prasad, 2013; Smith et al., 2013; Chanta et al., 2014; Wang et al., 2015; Ares et al., 2016; Enayati et al., 2019). However, little consensus exists concerning which equity measure researchers should employ (Stone, 1997; McLay & Mayorga, 2013). Based on our interactions with trauma collaborators, their general focus is to improve patient safety in the worst-performing region (among all regions) of the TSA. Therefore, we use the minimax equity measure as it intrinsically focuses on improving the performance of the worst one.

However, any equity measure as a standalone objective often results in undesirable, sometimes meaningless, solutions (Burkey et al., 2012; Smith et al., 2013; Enayati et al., 2019). For instance, minimax cannot distinguish between two networks with identical worst performing regions; however, one solution could have better performance in other regions than the other solution. In some situations, if a higher aggregated network performance can be achieved with a slightly less equity among individual regions, then it may be a preferred network for the decision makers. Considering both these factors, recent literature has proposed 'effectiveness' as a supporting metric, alongside equity (Burkey et al., 2012; Smith et al., 2013; Enayati et al., 2019). We, therefore, use both equity and effectiveness as objective terms in the proposed model. That is, the NTNDP is to determine the optimal number and location of MTCs and ITCs to maximize the weighted sum of equity in patient safety (among regions) and effectiveness (across the TSA).

Recall that patients with traumatic injuries can be classified into two categories; (i) severely injured patients with life-threatening injuries and (ii) non-severely injured with other trauma injuries. Due to limited patient-level on-scene vitals data, we estimate the severity of injury at the incidence location using the Injury Severity Score (ISS) as a surrogate, in line with the existing trauma literature. We also define two thresholds: 'access' threshold as a clinically-driven time (specified in trauma literature) to reach a hospital (MTC, ideally) and 'bypass' threshold as a resource-driven value that specifies the maximum additional minutes (compared to a nearby MTC/ITC) that EMS can dedicate to transport them to an NTC (ideal hospital).

Table 2
Classification of triage type based on injury severity and destination hospital type.

		Injury Severity Score (ISS)	
		ISS > 15 (severely injured)	*ISS ≤ 15 (non-severely injured)*
Destination hospital type	*MTC*	System-related appropriate-triage (srAT^{P})	System-related over-triage (srOT)
	ITC	System-related under-~~trigae~~ stabilized (srUT^{S})	
	NTC	System-related under-triage (srUT)	System-related appropriate-triage (srAT^{N})

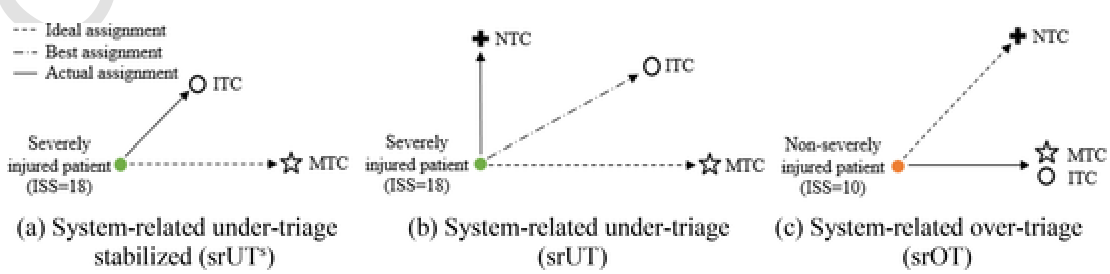


Fig. 1. Mistriages based on severity of injury and destination hospital type.

Before we present the model, we first present some preliminaries around destination determination and triage classification.

3.1. Triage classification

Table 2 classifies the triage types based on injury severity and destination hospital type. Irrespective of the destination determination criteria, if a patient is transported to the ideal hospital type based on their injury severity, then it is deemed as appropriate triage; i.e., severely injured transported to MTC is classified as srAT^{P} and non-severely injured transported to NTC is classified as srAT^{N}. Mismatch in injury severity and destination hospital type results in mistriage (see Fig. 1); recall that ISS > 15 is considered a severely injured patient.

As mentioned earlier, delay in definitive care for severely injured patients (i.e., srUT^{S} or srUT) increases the likelihood of an adverse outcome due to the life-threatening nature of those injuries. We combine both mistriages associated with severely injured patients and refer to it as 'system-related aggregated under-triage (srAU).' It defines as a weighted sum of srUT and srUT^{S}. In contrast, mistriage of the non-severely injured patients indirectly impacts patient safety and is relatively non-serious. Therefore, we consider mistriages of severely injured patients (i.e., srUT and srUT^{S}, aggregated as srAU in the model) as the primary patient safety metric, while mistriage of non-severely injured patients (i.e., srOT) as a secondary patient safety metric.

3.2. Destination determination

Recall that in Section 2, we mentioned that we incorporate three dominant destination determination criteria that EMS use at the incidence location; protocol, patient choice, and closest facility.

3.2.1. Protocol

The protocol criterion is essentially the Notional Tasking Algorithm (NTA) that attempts to mimic the EMS decision making process at the incidence location, as proposed by the American College of Surgeons (ACS). It considers clinical and resource factors for destination determination. In this paper, we extend the NTA used in Hirpara et al. (2022) to consider ITCs for severely injured patients (see Fig. 2). The NTA follows an ordered priority list based on patient's injury severity, the thresholds, and the vicinity of MTCs, ITCs, and NTCs.

3.2.1.1. Severely injured patient. The top priority is to assign a severely injured patient to any MTC (ideal hospital) within the 'access' threshold via ground; if this occurs, we refer to it as system-related appropriate triage positive (srAT^{P}). If no MTC is accessible via ground, then the second priority is assigning them to an MTC that is accessible via air ambulance (if available); this is also considered as srAT^{P}. For air ambulance transport, the NTA considers inbound-to- incidence location, loading, and transport-to-MTC times and compares it against the 'access' threshold; if below, then such transport is feasible.

If one of the first two priorities satisfy, then the third and fourth priorities are to assign them to an accessible ITC (not ideal, but better equipped than NTC) via ground and air. Because the ideal trauma hos-

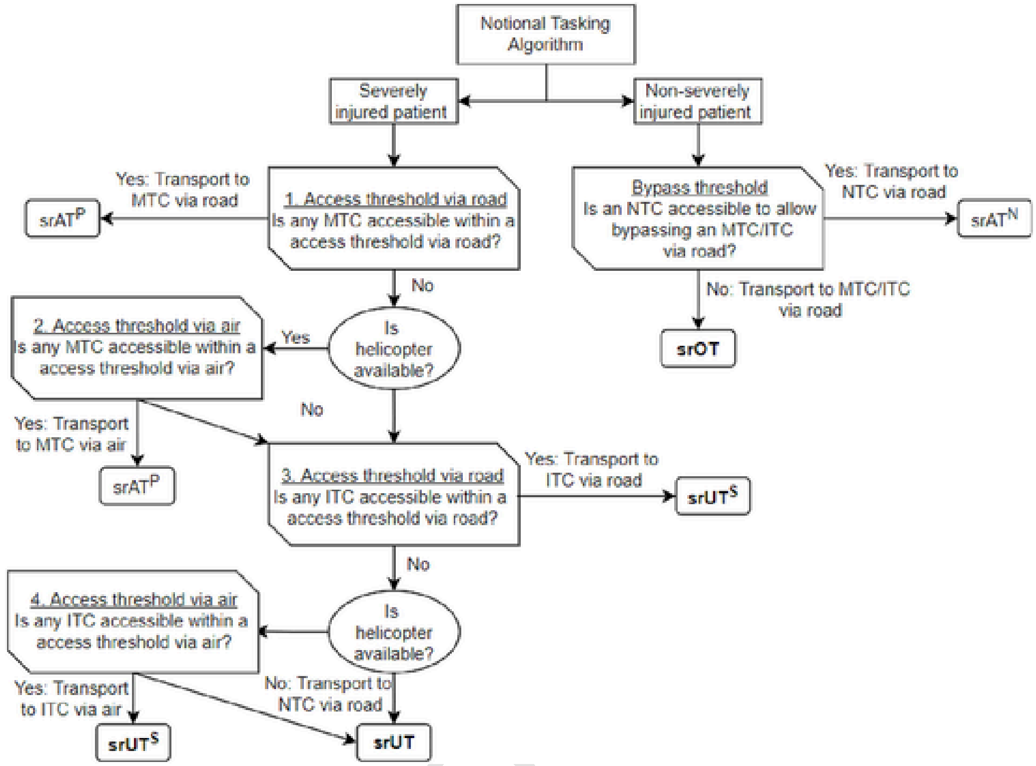


Fig. 2. Notional Tasking algorithm.

pital (i.e., MTC) is not chosen, we consider such a patient as system-related under-triage stabilized (srUT^{S}). The modifier ‘stabilized’ is used because ITCs are often capable to stabilize a severely injured patient. If all of the above are infeasible, then, as the last option for EMS, the patient is assumed to be transported to a nearby NTC; this results in system-related under-triage (srUT).

3.2.1.2. Non-severely injured patient. For such ISS $\leq$ 15 patients, the ‘bypass’ threshold is a resource-driven value that specifies the maximum additional minutes (compared to a nearby MTC/ITC) EMS can dedicate to transport the patient to an NTC (ideal hospital). For example, suppose the additional time to reach an NTC beyond the time to nearest MTC/ITC (say, 10 min) is within the ‘bypass’ threshold (say 15 min), then, in practice, the EMS is often likely to take the patient to that NTC. We refer to this type of a situation as system-related appropriate triage negative (srAT^{N}). Otherwise, the EMS would likely take the patient to the nearby MTC/ITC (due to longer drive or other operational criteria) resulting in system-related under-triage (srOT).

3.2.2. Patient choice

Anecdotal evidence and discussions with EMS suggest that patients often choose bigger hospitals over nearby hospitals due to their perception that the bigger the hospital, the better the care. However, travel time to the hospital also impacts their decision as they want to reach the hospital soon to avoid delay in receiving the care. In line with literature in the healthcare domain, we model patients’ choices through a linear utility model (Zhang et al., 2012; Haase & Müller, 2015; Zhang & Atkins, 2019). Accordingly, linear function comprises two dominant components that impact patients’ decision making; (i) the attractiveness of hospitals and (ii) ground travel time to those hospitals. The below equation calculates the utility of patient i receiving care at hospital j as a linear function of attractiveness of facility j (A_j) and ground travel time from the location of patient i to hospital j (TG_{ij}):

$$u_{ij} = \beta_1 A_j - \beta_2 TG_{ij}$$

here, we let the attractiveness of a hospital for a patient (A_j) depend upon the hospital type (MTC, ITC, or NTC) and that it is identical for all patients. The coefficients β_1 and β_2 denote the sensitivity to the two components, respectively, and can be estimated empirically based on available data or existing literature. Each patient is assigned to a hospital that has the highest utility across all hospitals.

3.2.3. Closest facility

In case of extreme weather considerations, road closures, or other unforeseen circumstances, EMS providers tend to prioritize closest facility over protocol or patient choice, irrespective of patient’s injury severity and closest hospital’s type. We model this by assigning such a patient to the closest hospital from the incidence location.

3.3. Optimization model

With this background, we now present the model under the following assumptions:

- The candidate locations for the MTCs, ITCs, and NTCs are known and finite.
- The number of patients, their locations, and severity are deterministic and known.
- The destination determination criteria for each patient is preassigned based on the given %-allocation among the three criteria.
- All severely injured patients, if initially transported to an ITC or an NTC, will eventually be transferred via ground to the nearest MTC from the incidence location (to allow them access to definitive care); patients are categorized as srUT and srUT^{S} accordingly because of delays in reaching MTC.
- The attractiveness of the facility to patients is given and depends only on the hospital’s type.
- A severely injured patient can be assigned to any MTC/ITC accessible within the access threshold in protocol criteria.

Further, in keeping up with the existing literature and what was observed in the data we had access to, we make the following assumptions about transportation modes:

- Ground and air transport times are known and deterministic.
- Air ambulance is only allowed to transport severely injured patients to MTCs and ITCs in the protocol criteria.
- While ground ambulance services are available without constraints, the availability of air ambulances was restricted to 15% of total severely injured patients based on data from state trauma agencies reports.

Tables 3 and 4 summarize the parameters and decision variables, respectively, used in our model.

$$\text{minimize}: \omega_1\, au^{\max} + \omega_2 \frac{\sum_k au_k}{|K|}$$

subject to:

Calculation of region-wise and maximum aggregated under-triage

$$au_k = \gamma \sum_{i:S_i=1} R_{ik} \sum_j (y^3_{ij} - y^4_{ij} - y^5_{ij}) + \delta \sum_{i:S_i=1} R_{ik} \sum_j (y^4_{ij} + y^5_{ij});\ \forall k \in K \quad (1)$$

$$au^{\max} \geq au_k;\ \forall k \in K \quad (2)$$

Limit on number of MTCs, and their minimum and maximum volume

Table 3
Parameters in the model.

Notation	*Definition*
I	Set of incidences for trauma patients, divided into the subsets I^O; subset of patients assigned via protocol criteria; $i \in I^O \subseteq I$; I^P subset of patients assigned via patient choice criteria; $i \in I^P \subseteq I$; I^C subset of patients assigned via closest facility criteria; $i \in I^C \subseteq I$
J	Set of candidate locations (for MTC, ITC, and NTC); $j \in J$
K	Set of regions in the TSA; $k \in K$
L	Set of hospital type; $l \in L$; $l = 1, 2, 3$ represent MTC, ITC, and NTC, respectively
ω_1, ω_2	Weights for equity and effectiveness in the objective function; $\omega_1 + \omega_2 = 1$
γ, δ	Weight for srUT and srUT^S patient
S_i	Injury severity of patient i; 1, if severely injured (ISS $>$ 15); 0, otherwise
R_{ik}	Region indicator, 1 if patient i is from a region k; 0, otherwise
TG_{ij}, TA_{ij}	Travel time from patient i to any candidate location j via ground and air
SG_{ij}	Subset of set J corresponding to each i-j pair and includes all other locations $t \in J$ such that ground travel time from patient i to t is greater than from i to j (i.e., $t \in SG_{ij}$, if $TG_{ij} < TG_{it}$; $j, t \in J$)
α	'Access' time threshold to determine srUT (for protocol criteria only)
β	'Bypass' time threshold to determine srOT (for protocol criteria only)
T_{in}, T_{load}	Inbound time from base-to-incidence location and loading time of patient at the incidence location for an air ambulance
Z	Maximum allowable patients via air ambulance
A^l	Attractiveness of hospital level l
β_1, β_2	Coefficient for attractiveness and travel time in the utility function
$V^{\min}_{MTC}, V^{\max}_{MTC}$	Minimum and maximum allowable volume of a severely injured patient at MTC
$V^{\min}_{ITC}, V^{\max}_{ITC}$	Minimum and maximum allowable volume of a severely injured patient at ITC
Ψ	Minimum allowable ratio of number of ITCs to MTCs
OT^{max}	Maximum allowable overall over-triage patients
A^G_{ij}, A^A_{ij}	Accessibility of candidate location j from patient i within α via ground and air; 1, if candidate location j is accessible from patient i; 0, otherwise
ρ	Equivalent fraction of an MTC corresponding to an ITC
C	Maximum equivalent MTCs allowed in the network
M	Big number

Table 4
Decision variables in the model.

Notation	*Definition*
x^l_j	1, if a candidate location j is designated to be level l; 0, otherwise
au_k, $au^{\max}$	System-related aggregated under-triage in region k; $au^{\max} = \max_k \{au_k\}$
y^1_{ij}	1, if patient i is transported via ground to location j (i.e., if j is an MTC, then patient i is srAT^P and if j is an NTC, then patient i is srAT^N); 0, otherwise;
y^2_{ij}	1, if severely injured patient i ($i \in I^O \subseteq I$) is transported via air to location j that is marked as MTC (i.e., srAT^P via air); 0, otherwise
y^3_{ij}	1, if severely injured patient i is transferred (from ITC or NTC) to location j that is marked as MTC (i.e., transferred srUT or srUT^S patient); 0, otherwise
y^4_{ij}	1, if severely injured patient i is transported via ground to location j that is marked as ITC (i.e., srUT^S via ground); 0, otherwise
y^5_{ij}	1, if severely injured patient i ($i \in I^O \subseteq I$) is transported via air to location j that marked as ITC (i.e., srUT^S via air); 0, otherwise
$ne^{MTC\_ITC}_{ij}$	1, if location j is marked as MTC or ITC and is the nearest non-NTC via ground for patient i ($i \in I^O \subseteq I$); 0, otherwise
ne^{NTC}_{ij}	1, if location j is marked as NTC and is the nearest NTC via ground for patient i ($i \in I^O \subseteq I$); 0, otherwise
u_{ij}, $u^{\max}_i$	Utility of patient i receiving care at hospital j; $u^{\max}_i = \max_j \{u_{ij}\}$
n_{ij}	1, if candidate location j is the nearest hospital for patient i ($i \in I^C \subseteq I$) or if the highest utility for patient i ($i \in I^P \subseteq I$) occurs for a hospital j; 0, otherwise

$$\sum_l x^l_j = 1;\ \forall j \in J \quad (3)$$

$$\sum_j x^1_j + \rho \sum_j x^2_j \leq C \quad (4)$$

$$x^1_j V^{\min}_{MTC} \leq \sum_{i:S_i=1} (y^1_{ij} + y^2_{ij} + y^3_{ij}) \leq x^1_j V^{\max}_{MTC};\ \forall j \in J \quad (5)$$

Allowable number of ITCs, and their minimum and maximum volume

$$\sum_j x^2_j \geq \psi \sum_j x^1_j \quad (6)$$

$$x^2_j V^{\min}_{ITC} \leq \sum_{i:S_i=1} (y^4_{ij} + y^5_{ij}) \leq x^2_j V^{\max}_{ITC};\ \forall j \in J \quad (7)$$

Limit on state-wide srOT

$$\sum_i (1 - S_i) - \sum_{i:S_i=0} \sum_j y^1_{ij} \leq OT^{\max} \quad (8)$$

Assignment and triage classification using protocol criteria

$$\sum_j (y^1_{ij} + y^2_{ij} + y^3_{ij}) = 1;\ \forall i \in I^O : S_i = 1 \quad (9)$$

$$y^1_{ij} = 0;\ \forall i \in I^O : S_i = 1, \forall j \in J, TG_{ij} > \alpha \quad (10)$$

$$y^2_{ij} = 0;\ \forall i \in I^O : S_i = 1, \forall j \in J, TA_{ij} + T_{in} + T_{load} > \alpha \quad (11)$$

$$\sum_j A^G_{ij} x^1_j \leq M(1 - \sum_j y^2_{ij});\ \forall i \in I^O : S_i = 1 \quad (12)$$

$$x^1_j + \sum_{t \in SG_{ij}} y^3_{it} \leq 1;\ \forall i \in I^O : S_i = 1, \forall j \in J \quad (13)$$

$$\sum_j (y^4_{ij} + y^5_{ij}) \leq \sum_j y^3_{ij};\ \forall i \in I^O : S_i = 1 \quad (14)$$

$$y^4_{ij} = 0;\ \forall i \in I^O : S_i = 1, \forall j \in J, TG_{ij} > \alpha \quad (15)$$

$$\sum_j A^G_{ij} x^1_j \leq M(1 - \sum_j y^4_{ij});\ \forall i \in I^O : S_i = 1 \quad (16)$$

$$y^5_{ij} = 0;\ \forall i \in I^O : S_i = 1, \forall j \in J, TA_{ij} + T_{in} + T_{load} > \alpha \quad (17)$$

$$\sum_j A^G_{ij} x^2_j + \sum_j A^G_{ij} x^1_j + \sum_j A^A_{ij} x^1_j \le M\Big(1 - \sum_j y^5_{ij}\Big); \forall i \in I^O : S_i = 1 \quad (18)$$

$$\sum_i \sum_j (y^2_{ij} + y^5_{ij}) \le Z \quad (19)$$

$$ne^{NTC}_{ij} \le x^3_j; \forall i \in I^O : S_i = 0, \forall j \in J \quad (20)$$

$$\sum_j ne^{NTC}_{ij} = 1; \forall i \in I^O : S_i = 0 \quad (21)$$

$$x^3_j + \sum_{t \in SG_{ij}} ne^{NTC}_{it} \le 1; \forall i \in I^O : S_i = 0, \forall j \in J \quad (22)$$

$$ne^{MTC\_ITC}_{ij} \le x^1_j + x^2_j; \forall i \in I^O : S_i = 0, \forall j \in J \quad (23)$$

$$\sum_j ne^{MTC\_ITC}_{ij} = 1; \forall i \in I^O : S_i = 0 \quad (24)$$

$$x^1_j + x^2_j + \sum_{t \in SG_{ij}} ne^{MTC\_ITC}_{it} \le 1; \forall i \in I^O : S_i = 0, \forall j \in J \quad (25)$$

$$\sum_j \left(ne^{NTC}_{ij} TG_{ij}\right) - \sum_j \left(ne^{MTC\_ITC}_{ij} TG_{ij}\right) - \beta \le M\left(1 - \sum_j y^1_{ij}\right); \forall i \in I^O : S_i = 0 \quad (26)$$

$$\sum_j y^1_{ij} \le 1; \forall i \in I^O : S_i = 0 \quad (27)$$

Utility calculation for an assignment using patient choice criteria

$$u_{ij} = \beta_1 \sum_l A^l x^l_j - \beta_2 TG_{ij}; \forall i \in I^P, \forall j \in J \quad (28)$$

$$u^{\max}_i \ge u_{ij}; \forall i \in I^P, \forall j \in J \quad (29)$$

$$(u^{\max}_i - u_{ij}) - M(1 - n_{ij}) \le 0; \forall i \in I^P, \forall j \in J \quad (30)$$

Closest hospital for patients assigning through closest facility criteria

$$1 + \sum_{t \in SG_{ij}} n_{it} \le 1; \forall i \in I^C, \forall j \in J \quad (31)$$

$$\sum_j n_{ij} = 1; \forall i \in I^C \cup I^P \quad (32)$$

Assignment of patients through patient choice and closest facility criteria

$$\sum_j (y^1_{ij} + y^3_{ij}) = 1; \forall i \in I^C \cup I^P : S_i = 1 \quad (33)$$

$$n_{ij} + x^1_j \ge 2y^1_{ij}; \forall i \in I^C \cup I^P : S_i = 1, \forall j \in J \quad (34)$$

$$x^1_j + \sum_{t \in SG_{ij}} y^3_{it} \le 1; \forall i \in I^C \cup I^P : S_i = 1, \forall j \in J \quad (35)$$

$$\sum_j y^4_{ij} \le \sum_j y^3_{ij}; \forall i \in I^C \cup I^P : S_i = 1 \quad (36)$$

$$n_{ij} + x^2_j \ge 2y^4_{ij}; \forall i \in I^C \cup I^P : S_i = 1, \forall j \in J \quad (37)$$

$$n_{ij} + x^3_j \ge 2y^1_{ij}; \forall i \in I^C \cup I^P : S_i = 0, \forall j \in J \quad (38)$$

Bounds on decision variables

$$x^l_j \in \{0, 1\}; \forall j \in J, \forall l \in L \quad (39)$$

$$au_k, au^{\max} \ge 0; \forall k \in K \quad (40)$$

$$y^1_{ij} \in \{0, 1\}; \forall i \in I, \forall j \in J \quad (41)$$

$$y^3_{ij}, y^4_{ij} \in \{0, 1\}; \forall i \in I, \forall j \in J \quad (42)$$

$$y^2_{ij}, y^5_{ij} \in \{0, 1\}; \forall i \in I^O, \forall j \in J \quad (43)$$

$$ne^{NTC}_{ij}, ne^{MTC\_ITC}_{ij} \in \{0, 1\}; \forall i \in I^O, \forall j \in J \quad (44)$$

$$u_{ij}, u^{\max}_i \in \mathbb{R}; \forall i \in I^P, \forall j \in J \quad (45)$$

$$n_{ij} \in \{0, 1\}; \forall i \in I^C \cup I^P, \forall j \in J \quad (46)$$

The model minimizes a weighted sum of maximum srAU patients among regions (equity measure) and average srAU patients across regions (effectiveness measure). We use y_{ij} variables in the model to classify triage types and record destination hospitals for further volume calculation.

For each region, Constraints (1) calculate the total aggregated under-triage patients (srAU), which is a weighted sum of overall srUT (first term) and $srUT^S$ (second term) in a region. Constraints (2) calculate maximum srAU patients among all regions. Constraints (3) ensure that each candidate location is designated as either MTC, ITC, or NTC. Constraint (4) ensures that the total number of MTCs and ITCs must be less than or equal to the maximum allowable equivalent MTCs (which allows the model to find the best combination of MTCs and ITCs considering budgetary constraints). Constraints (5) bound volume of severely injured patients (directly transported to MTC or transferred from ITC or NTC) at candidate location *j* if it is designated as MTC. Constraint (6) ensures that the ratio of ITCs to MTCs is within a prespecified value. Constraints (7) limit the volume of severely injured patients (transported to ITC via ground or air) at candidate location *j* if it is designated as ITC. Constraint (8) ensures that total TSA-wide srOT patients (difference of total non-severely injured patients and $srAT^N$) is within an allowable limit.

For patients assigned via protocol criteria, Constraints (9)-(27) assign them to hospitals and classify their triage types. Constraints (9) ensure that each severely injured patient is either initially transported to an MTC via ground or air, or eventually transferred to MTC from an ITC/NTC. Constraints (10) and (11) rule out an assignment of severely injured patient *i* to every inaccessible MTC via ground and air, respectively. Note that an MTC is considered not accessible via ground if ground travel time is higher than the 'access' threshold; it is not accessible via air if the total airtime (sum of inbound, loading, and air travel) is higher than the 'access' threshold. Constraints (12) rule out an assignment of severely injured patient *i* to all MTCs via air if any MTC is accessible via ground. That is, in an effort to preserve the limited air ambulance trips, a patient is only airlifted if no MTC is accessible via ground.

Constraints (13) capture the transfer of severely injured patients from an ITC or NTC to the nearest MTC to receive definitive care. Constraints (14) ensure that severely injured patient *i* is assigned to an ITC (via ground or air) if initially not assigned to any MTC ($\sum_j y^3_{ij} = 1$). Constraints (15) rule out an assignment of severely injured patient *i* to ITCs not accessible via ground. However, if any MTC is accessible via ground, then Constraints (16) rule out assignment of severely injured patient *i* to all ITCs. Constraints (17) rule out an assignment of severely injured patient *i* to ITCs not accessible via air, while Constraints (18) rule out assignments to all ITCs if any ITC is accessible via ground or any MTC is accessible via ground or air. Constraints (16) and (18) ensure priority-based assignment of severely injured patients discussed in the section 3.2.1.1. Constraints (19) ensure that the air transport usage does not exceed their availability.

For each non-severely injured patient *i*, Constraints (20)-(22) determine the nearest NTC. Constraints (20) ensure that a candidate location *j* must be an NTC to be considered as the nearest NTC, Constraints (21) make sure that for a non-severely injured patient *i*, only one NTC should be considered as the nearest NTC. For any pair of patient *i* and candidate location *j*, if a candidate location *j* is marked as NTC, then Constraints (22) rule out the assignment of patient *i* to candidate location(s) *t* that are located further (in terms of time) than *j*. Constraints (23)-(25) serve the same purpose as (20)-(22), respectively, for the nearest non-NTC (MTC or ITC) via ground. For non-severely injured patient *i*, Constraints (26) rule out the assignment to all NTCs if the 'bypass' threshold criterion is not met; this patient is marked as srOT. Note that srOT occurs when MTC or ITC is closer than the nearest NTC. We do not need to explicitly assign srOT patients to an MTC or ITC as they are not counted towards trauma volume; these patients are often discharged from the ED of an MTC or ITC without admission to the inpatient trauma unit.

For patients assigned via patient choice criteria, Constraints (28)-(30) capture patients' choices using the utility model. For each patient *i*, Constraints (28) calculate utility of receiving care at each hospital (candidate location), Constraints (29) find maximum utility among all hospitals, while Constraints (30) record the hospital with the maximum utility. Constraints (31) and (32) find the closest facility for each patient assigned via closest facility criteria. For patients assigned via patient choice and closest facility criteria, variable n_{ij} capture patient choice and the closest facility, respectively. Constraints (32) ensure that each patient has only one closest facility and select one hospital with maximum utility for the closest facility and patient choice criteria, respectively.

For patients assigned through patient choice and closest facility criteria, Constraints (33)-(38) assign them to hospitals and classify their triage types. Constraints (33) ensure that each severely injured patient is initially assigned to MTC or eventually transferred to MTC from ITC/NTC. Constraints (34) assign and classify severely injured patient *i* as $\mathrm{srAT^P}$ if the nearest hospital or patient's choice is MTC. Constraints (35) ensure that each severely injured patient is transferred to the nearest MTC after being initially transported from the incidence location to an ITC ($\mathrm{srUT^S}$ patient) or an NTC (srUT patient). Constraints (36) ensure that severely injured patient is assigned to an ITC if initially not assigned to any MTC; i.e., $\sum_j y_{ij}^3$ Constraints (37) classify a severely injured patient *i* as $\mathrm{srUT^S}$ if the criterion of nearest hospital or patient choice results in ITC. Constraints (38) classify a non-severely injured patient *i* as $\mathrm{srUT^N}$ if the criterion of nearest hospital or patient choice results in NTC. Constraints (39)-(46) define bound on decision variables.

Note that the NTNDP can be characterized as a hierarchical, discrete, multi-facility location problem. Such problems are combinatorial in nature and have been shown to be NP-hard (Daskin, 2011). For even 50 candidate hospital locations, there are $3^{50} = 7.18 \times 10^{23}$ solutions. Our preliminary experiments suggested that commercial software such as CPLEX and Gurobi encountered out-of-memory issues for realistic problem instances that normally have 100 + locations and 1000+ patients. We, therefore, explored a tailored '3-phase' approach to avoid such issues and find a near-optimal solution. We now discuss our proposed approach.

4. A 3-phase solution approach

A primary goal of any trauma system is to provide prompt care to severely injured patients. Data from the state of OH indicated that severely injured patients made up about 15% of the total patients. The problem complexity can thus be reduced if relaxed the model to first focus on severely injured patients, and then the non-severely injured patients. Considering this, we propose a '3-phase' approach that systematically reduces the problem complexity into different phases to decrease the number of decision variables and constraints (see Fig. 3).

In Phase 1, we only consider severely injured patients ($S_i = 1$) and determine the optimal location of MTCs and ITCs based on the NTNDP model presented earlier. Essentially, we remove all decision variables

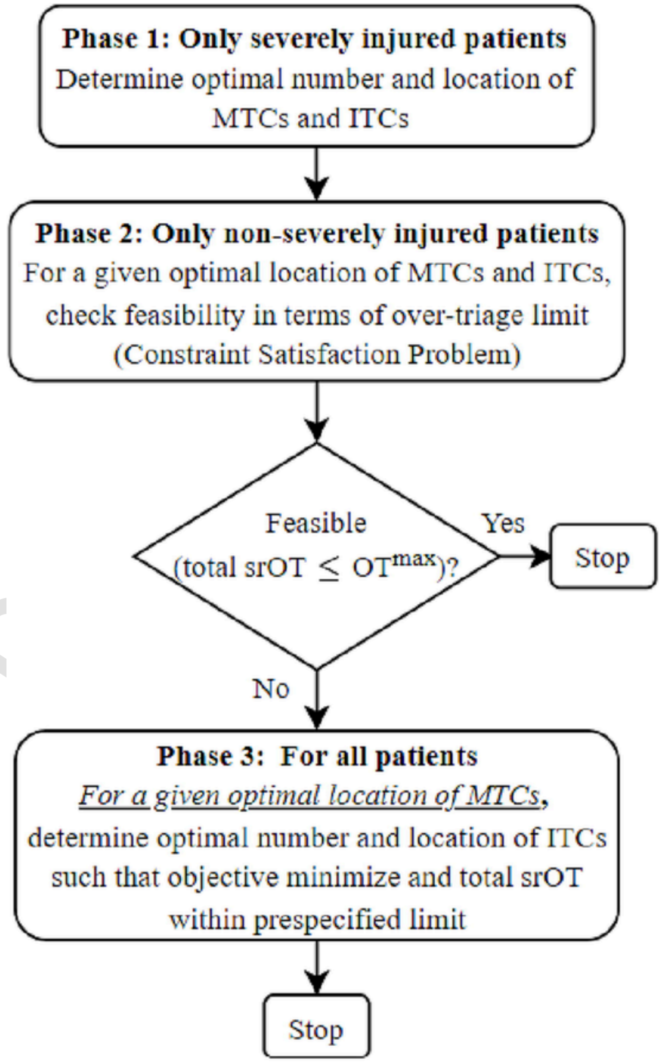


Fig. 3. Flowchart of the '3-phase' approach.

and constraints related to non-severely patients. The Phase 1 problem is as follows:

Phase 1 problem

minimize: $\omega_1 au^{\max} + \omega_2 \frac{\sum_k au_k}{|K|}$

s.t.

Constraints (1)-(7), (9)-(19), (28)-(37), (39)-(43), (45)-(46)

In Phase 2, we use the solution from Phase 1 and solve a Constraint Satisfaction Problem for non-severely injured patients. Essentially, we check the feasibility of the Phase 1 solution in terms of total over-triage patients (which are triggered by non-severely injured patients). Feasibility in Phase 2 means overall srOT patients are within the prespecified limit, and that the solution found in Phase 1 is optimal to the entire problem. However, infeasibility in Phase 2 indicates the need for changes in the solution from Phase 1 to keep total srOT patients within the limit, which then invokes Phase 3. As non-severely injured patients do not directly impact the objective of the NTNDP, the model for the Phase 2 can be defined by Constraints (8), (20)-(32), (38), (41), and (44)-(46).

In Phase 3, we fix the location of MTCs (obtained from Phase 1) while considering both types of patients and solve the original model for NTNDP. Basically, for a given location of MTCs, we find the optimal location of ITCs to minimize the objective while keeping total srOT within a limit. Fixing MTCs

Phase-3 problem

minimize: $\omega_1 au^{\max} + \omega_2 \frac{\sum_k au_k}{|K|}$

s.t.

Constraints (1)-(46)

$$x_j^1 = 1; \forall j \in J' \quad (47)$$

is reasonable as the impact of MTCs on the objective is relatively higher than ITCs due to their capability of providing definitive care to severely injured patients. In the formulation for Phase 3, we add Constraints (47) to fix the location of MTCs obtained from Phase 1 (represented by set J' where $J' \in J$).

We used Gurobi solver on Dell I7-10700 CPU @2.90 GHz Desktop with 32 GB RAM to find an optimal solution in each phase.

5. Computational study

We now detail our experimental study starting with the test area generation (referred to as a TSA), sources of data collection, evaluation of the solution approach, sensitivity analysis, and insights.

5.1. TSA determination

We consider the collection of counties in an existing midwestern US state as TSA. In so doing, we can use the underlying transportation network to estimate actual ground transportation times from the incidence locations to the candidate hospitals. Fig. 4 illustrates the TSA with 34 counties and 64 hospitals. In this TSA, 21 counties are rural (61.7%), in line with the % of rural counties in the US (i.e., 62%). All 64 hospitals in the TSA were considered as candidate locations for an MTC, ITC, or

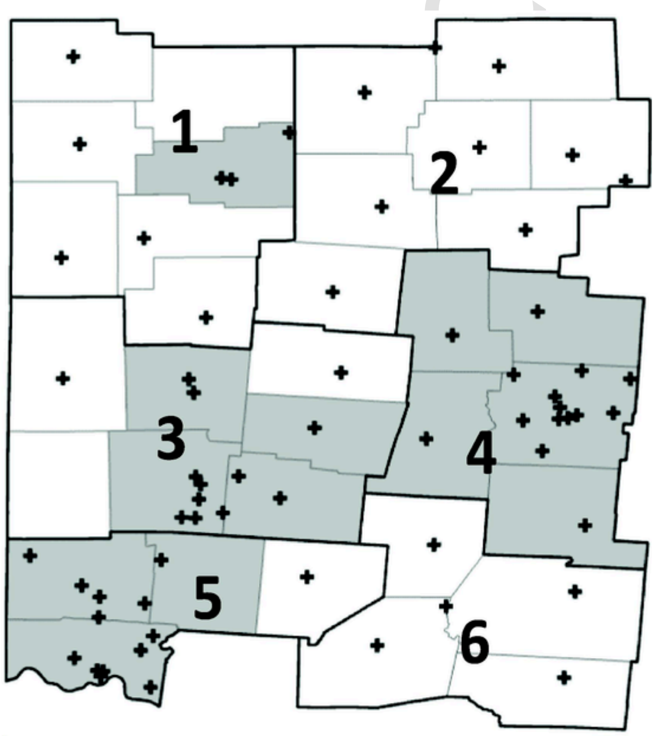


Fig. 4. TSA with counties, and region; grey filled areas are urban counties; ‘ + ’represents candidate locations.

NTC. We also grouped counties to represent regions similar to several state trauma agencies (MDHHS, 2022; TDSHS, 2022). In our chosen TSA, Regions 2 and 6 are entirely rural regions, while Region 4 is an entirely urban. Further, Region 1 is dominantly rural (with higher % of rural counties compared to urban counties), while Regions 3 and 5 are dominantly urban.

For all analyses, we used ArcGIS Pro 2.9.1 to calculate actual drive times to generate the ground time matrix (TG_{ij}) and the Haversine formula (assuming the helicopter speed of 120 mph) to generate the airtime matrix (TA_{ij}). Note that both these time matrices are pre-generated and serve as ~~a~~ look-up table during the solution process.

5.2. Performance of the 3-Phase approach

For performance evaluation of the ‘3-phase’ approach, we considered 10 problem instances using this TSA, each with 5000 patients (15.63% of severely injured patients), and between 15 and 60 candidate locations. We used 200 and 50 as a lower bound for the volume of severely injured patients at MTCs and ITCs, respectively. The limit for srOT patients is set as 50% of total non-severely injured patients. The attractiveness for MTCs, ITCs, and NTCs is set as 5, 4, and 1, respectively, and coefficients for attractiveness and ground travel time were set as 0.825 and 0.566. All other parameters are the same as the base case mentioned in section 5.4. We set the CPU-time limit as 12 h for solving the Gurobi MIP solver.

Table 5 presents our computational experiments that compare the solution quality and runtime of the ‘3-phase’ approach and ‘Original model’ (per Section 3.3) for several problem instances. The ‘% Difference’ column represents the difference between the objective of the ‘3-phase’ and ‘Original model,’ where positive value represents a ‘3-phase’ approach outperformed the ‘Original model.’ The number in a bracket of the ‘Original model’ column of ‘Solution Quality’ represents the gap between best solution and lower bound when the solver reached the time limit.

These computational experiments verify that our ‘3-phase’ approach can achieve high-quality solutions in a short amount of time; therefore, we used this approach for further experiments to generate insights.

5.3. Patient volume and sampling

We collected state-wide trauma data across various states from published annual reports (available on state trauma websites) and observed a substantial variation across these states. The patient volume varied between 11,000 and 72,000 per year, with 3.2 to 8.2 variation in number of trauma patients per thousand citizens. Additionally, patient vol-

Table 5
Performance evaluation of the ‘3-phase’ approach.

Problem instance	*Candidate Locations*	*Solution Quality*			*Runtime in Hours*	
		Original model	*3-phase*	*% Difference*	*Original model*	*3-phase*
1	15	158.3	159.8	−0.93%	1.65	0.07
2	20	169.3	169.6	−0.22%	1.94	0.14
3	25	141.5	142.3	−0.53%	2.16	0.19
4	30	120.9	120.9	0%	3.29	0.28
5	35	122.5	122.5	0%	3.77	0.32
6	40	100.9	100.9	0%	6.70	0.4
7	45	113.1 (4.95)	112.9	0.17%	12	1.17
8	50	100.3 (6.17)	99.9	0.35%	12	1.53
9	55	Out of Memory	–	–	Out of Memory	1.68
10	60	Out of Memory	–	–	Out of Memory	5.2

ume at a county level was observed to be highly correlated with the population of that county. For the experimental study, our TSA attempts to mimic the trauma patient volume of a median US state; i.e., we used 5.2 as the average trauma patients per thousand citizens and median population of a US state as 4.5 million to arrive at 23,680 trauma patients. In line with the literature, we considered 15.63% of patients as severely injured and the rest as non-severely injured.

Through preliminary experiments, we also noticed that the computational time to reach a solution was prohibitively high when considering all 23,680 patients (65 h). Instead of aggregating patients at the county or zip level (which would lose the granularity required for our problem), we adopted a sampling approach. We selected a representative sample among these 23,680 patients such that the underlying distribution of patients (Gini index) was highly correlated with the distribution of these 23,680 patients. All other parameters were appropriately scaled.

Fig. 5 illustrates the solution quality and runtime comparison at various sampling rates. To balance quality and computational time, we selected 15% as the sample size; it reduced the time by 95% with about 1.6% difference in the solution compared to the problem being solved with complete data. Essentially, this sampling allowed us to solve problem instances (per Table 6) on average in 3.5 h.

5.4. Experimental setting

During the preliminary experiments, we also noticed that the solutions appeared to be sensitive to three key factors. Table 6 summarizes these factors and their levels with bold entries in the last column indicating the base case. While ACS or state trauma agencies typically propose a protocol for destination determination based on the severity of injuries, only 40% of patients were assigned using protocol criteria according to literature and data from our collaborators. Patient choice (PC) is the second dominant criteria for destination determination, followed by assignment to the closest facility, which is inevitable. Therefore, we considered four scenarios to quantify the impact of assignment criteria on the performance and design of the network. We used the (40, 40, 20) combination as a base case based on our interactions with our

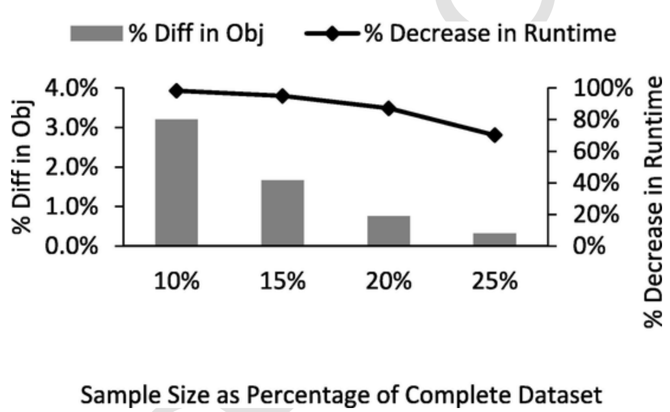


Fig. 5. Solution quality and runtime for different sample size compared to complete data set.

Table 6
Summary of the parameters, levels, and values in the sensitivity analysis.

Parameter	Level	Values
Percentage of assignment using protocol, patient choice and closest facility criteria	4	**(40, 40, 20)**, (60, 20,20), (80, 0,20), (100,0,0)
Distribution of trauma patients	3	Disperse (0.25), **Regular (0.5)**,Cluster (0.75)
Weights combination for equity and effectiveness	3	(0.1, 0.9), **(0.5, 0.5)**, (0.9, 0.1)

trauma collaborators and the trauma literature; i.e., 40% of patients were assigned using protocol and patient choice (PC) criteria, while the remaining 20% used the closest facility criteria in all counties. Assignment criteria and injury severity are preassigned to each patient as part of the data preprocessing step.

Distribution of patients in the TSA was distributed using 3 levels quantified through the Gini index, where 0 and 1 represent fully-dispersed and fully-clustered distributions. Accordingly, dispersed corresponded to Gini = 0.25 (patients are less clustered and more homogeneously distributed), clustered corresponded to Gini = 0.75 (patients are highly clustered around urban zones), and regular corresponded to Gini = 0.5 (patients are moderately clustered around urban zones). While the dispersed distribution attempted to mimic states such as New Jersey, Delaware, and Vermont, the clustered distribution mimicked states such as Nevada, Texas, and Arizona.

Three weight combinations are used to evaluate the impact of emphasis on equity vs. effectiveness. We also use 1 and 0.5 as the γ and δ, respectively. Following ACS recommendation, we used 240 severely injured patients as a lower bound for MTCs and 60 for ITCs considering their limited resources (specialist surgeons, equipment and capacity). Additionally, per trauma literature, we used the upper bound on volume as 1,000 at both MTCs and ITCs, access time threshold as 30 min, bypass time threshold as 0 min, and C = 64 (total candidate locations). For air transport (via helicopter), we set the inbound time (time from the helicopter depot to the incidence location) as 10 min and the loading time as 5 min. Based on the range calculated from state trauma reports, we used 15% of severely injured patients as upper bound for helicopter use. We set 70% of total non-severely injured as the maximum allowable number of over-triage patients in the TSA.

For the utility model representing the patient choice, the attractiveness for MTC, ITC, and NTC is set as 5, 3, and 1, respectively, as a way to differentiate the relative perception of trauma centers among citizens. The coefficients for attractiveness and ground travel time were estimated as 0.1 and 0.05 using the optimization framework and data from the state of Ohio (see Appendix for details).

5.5. Insights from the experiments: Below, we summarize key insights from our experimental study

Insight 1: Destination determination criteria impacts patient safety; while 100% protocol usage improves it, increased use of patient choice lowers it.

As alluded earlier, the ACS and/or state trauma agencies prefer that EMS paramedics determine the destination of patients based on a protocol. Our results suggest that if all such determinations were done using this protocol (i.e., 100, 0, 0), we observed a 92.4% reduction in objective compared to the base case of (40, 40, 20) (i.e., 1.71 vs. 22.54) with fewer ITCs (1 vs. 9) and MTCs (14 vs. 15). That is, if patient choice and closest facility considerations were not part of the destination determination, the trauma network could be optimized and substantial performance benefits could be achieved.

In terms of the distribution of MTCs, we noticed a disperse distribution that accessed by most of the TSA by at least one MTC or ITC such that most severely injured patients had access to one of them within the access time, thus, reducing under-triage patients (see Fig. 6). Further, a dispersed network of MTCs and ITCs also means a higher chance of having NTCs within the bypass threshold for non-severely injured patients resulting in lower over-triage patients.

However, as shown in Fig. 7, higher assignments through patient choice (PC) criterion require more ITCs (3 in *a* vs. 9 in *c*) and result in a 73.1% increase in objective (12.96 vs. 22.54).

To understand this further, imagine a network without ITC. Under the PC criterion, the absence of ITCs in the vicinity of the incidence location in a suburban or rural zone would leave a severely injured patient (or their family) to choose between a nearby NTC (say, at 5 min)

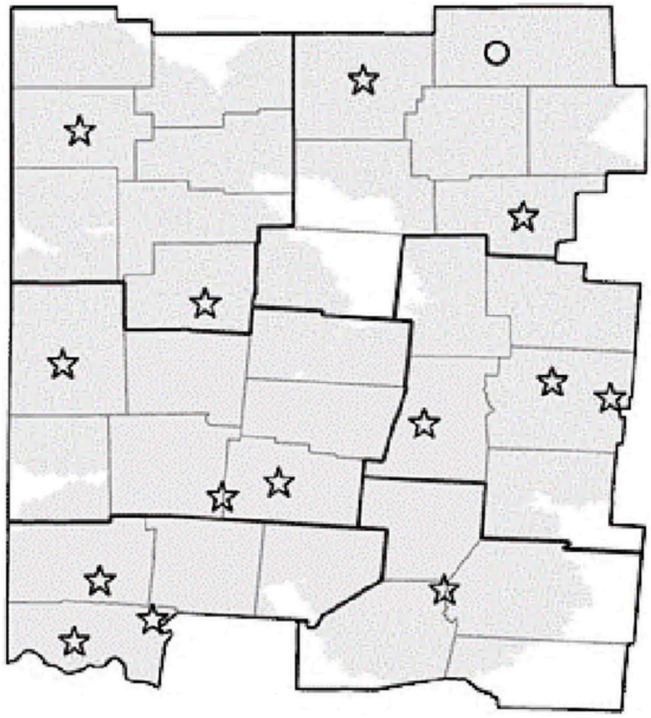


Fig. 6. Dark areas indicate 30-minute access from the incidence location to at least one MTC or ITC; stars indicate MTCs and circle indicates ITC.

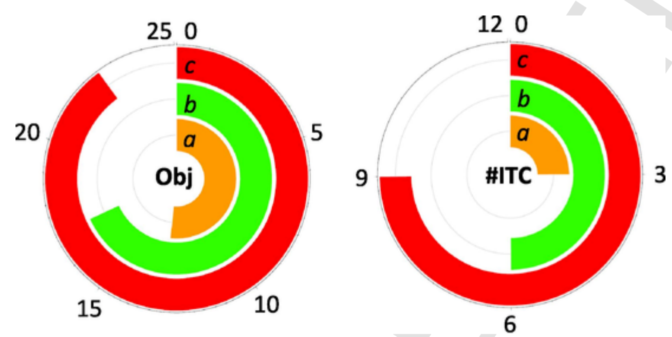


Fig. 7. Objective and #ITCs for different destination assignment criteria scenarios; $a=$ (80, 0, 20), $b=$(60, 20, 20), and $c=$(40, 40, 20), where each element represents protocol, patient choice, and closest facility, respectively.

and far located MTC (say, at 25 min). Considering a lower travel time, that patient will likely choose the NTC over MTC. This would result in that patient experience under-triage. While at least one MTC in that zone would mitigate such an under-triage, it may not be feasible due to MTC's minimum volume requirement. This is where an ITC could play a compromising role as it would likely induce the patient (or their family) to choose this ITC over an NTC, and eventually getting better care (see Fig. 8-c).

Insight 2: TSA with clustered distribution of patients appear to improve patient safety compared to other distributions.

To delineate different distributions of severely injured patients, we use D, R, and C to represent disperse, regular, and cluster with Gini indexes of 0.25, 0.5, and 0.75, respectively (see Fig. 9). Our results indicate that as the patient distribution changes from disperse (D) to cluster (C), the overall objective decreases by 18.8% (22.42 vs. 18.5). The number of ITCs is almost three times (14 vs. 5) in the dispersed situation compared to cluster situation.

This is reasonable as clustered distribution increases opportunities for treating more patients at the same MTCs and ITCs located around the clustered zones. However, the performance of cluster distribution highlights that better performance can be achieved with even fewer resources which is counterintuitive.

Moreover, we observed distinct location of MTCs and ITCs across three patient distribution scenarios (see Fig. 10). In the cluster distribution, due to higher patients from dominantly urban regions (#3-#5), 13 out of 14 MTCs are located in those regions; however, in the dispersed situation, the MTCs are spread across the TSA. In addition, higher patients from suburban and rural zones in disperse distribution trigger the opening of ITCs in those zones as patients are still not enough to make an MTC feasible from a minimum volume perspective. As a result, 11 out of 14 ITCs are located in dominantly rural regions in dispersed distribution compared to zero in the case of cluster distribution. That is, the distribution of the patients tends to drive the number and location of MTCs and ITCs across the TSA.

Insight 3: An emphasis on equity in a network may lead to a decline in overall patient safety.

As expected, in an equitable network, most regions performed equally. This is evident from the histogram (depicting equity per region under $\omega_1 = 0.9$) in Fig. 11. Despite this, the performance of many regions is worse than the performance observed with lower values of ω_1 (a less equitable network). To quantify this, we used skewness of the distributions for each of the three ω_1 values. For the most equitable network ($\omega_1 = 0.9$), the skewness was −2.4, while it reduced to −0.5 for the least equitable network (alternately, network with higher effectiveness, ($\omega_1 = 0.1$). The corresponding TSA-wide AU (A_{avg}) increased by 8% (20.75 vs. 22.42) indicating an overall decline in the system performance. Note that an 8% increase is equivalent to an annual increase of

(a) (80, 0, 20) (b) (60, 20, 20) (c) (40, 40, 20)

Fig. 8. Locations of MTCs and ITCs for different percentage of assignments; dense areas represent higher number of patients, stars represent MTCs, and circles represent ITCs.

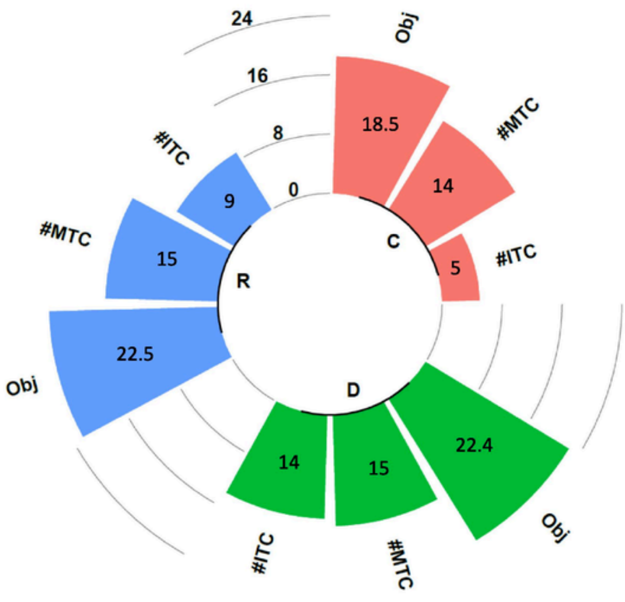


Fig. 9. Objective, #MTC and #ITCs for disperse (D), regular (R) and cluster (C) distribution of patients.

67 severely injured patients who will suffer aggregated under-triage (considering 23,680 data); they all could experience disabilities or mortality. The reason for this increase is due to the relocation of a few MTCs and ITCs to improve the performance of worse-performing regions. However, those relocations decrease AU of worse regions at the cost of a relatively higher increase of AU patients from the well-performing regions; the net effect is an overall increase of AU at the TSA level.

In a nutshell, results emphasize that the trauma decision maker should choose weights wisely as a higher focus on equity of patient safety can lead to higher under-triage patients, eventually increasing the likelihood of disability and mortality.

6. Case study

To illustrate the practical benefit of the proposed approach, we considered the state of Ohio as a TSA and used actual data from the state for 2019. Among 71,971 trauma patients recorded in 2019, we received 17,757 de-identified patient records resulting after data linkage performed by the ODPS (Ohio Department of Public Safety). This data was further cleansed to remove missing data and unresolved addresses using ArcGIS. The resulting 11,313 patients in the cleansed dataset had a correlation of 0.99 with the 17,757 patients based on county-level case comparison, which indicated a similar spatial distribution of incidences between the original and cleansed datasets. This TSA consisted of a network of 163 hospitals in 2019, which included 21 MTCs, 21 ITCs, and the remaining 121 NTCs. Fig. 12 illustrates the heat map of 11,313 incidences and the location of hospitals.

In this data, destination determination through protocol, patient choice, and closest facility criteria were around 20%, 50%, and 30%, respectively. We empirically derived ‘access’(α) as 25 min and ‘bypass’ (β) as -12 min such that the estimated srAU and srOT closely matched the observed values in the 20% of patients assigned through protocol criteria in the existing data. Similarly, patients assigned through patient choice criteria were used to estimate $\beta_1 = 0.1$ and $\beta_2 = 0.05$ through an optimization model shown in Appendix. In line with the discussion in Section 5.3, we further sampled 3,552 patients (correlation of 0.999 with 11,313 data) to limit the computational burden; we scaled the MTC and ITC volume requirements accordingly. We set maximum allowable number of over-triage patients in the state as 67.31% of total non-severely injured patients (similar to observed in 11,313 data) and remaining parameters values are as used in Section 5.

Using $\omega_1 = 0.5$, we derived two optimal networks, one for the case when the number of effective MTCs is the same (Redistributed) and the other where this number is also optimally determined by the model (Greenfield). In addition, we also derived an optimal trauma network with all assignments through protocol criteria as recommended by ACS and/or state trauma agencies.

6.1. Existing vs. Redistributed vs. Greenfield trauma network

The Existing network had 31.5 effective MTCs (21 MTCs + 0.5*21 ITCs). Hence, for the Redistributed network, we set C = 31.5. However, for the Greenfield network, we let C = 163 (all candidate locations) allowing the model to select as many MTCs and ITCs it needs to minimize the objective function. Table 7 summarizes the performance of the three networks, while Fig. 13 illustrates the trauma network superimposed on heat maps of incidences.

The Gini index for the 2019 data is 0.751, representing a clustered distribution of patients. In the Existing network, many ITCs were observed in the clustered areas (darker areas in Fig. 13); however, the Redistributed and Greenfield networks, we observed several MTCs in those clustered areas in line with Insight 2. Due to these additional MTCs, redistribution reduces the average srAU patients by 31.2% compared to the Existing network (10 vs. 6.88). The Greenfield network only did slightly better than the Redistributed network; the average srAU patients decreased by 33.1% (10 vs. 6.69) at the cost of an additional 3.5 effective MTCs. Further, in both the Redistributed and Greenfield networks, ITCs were located in moderately dense areas (light grey areas in Fig. 13) as highlighted in Insight 1.

(a) Disperse distribution (D) (b) Regular distribution (R) (c) Cluster distribution (C)

Fig. 10. Locations of MTCs and ITCs superimposed over heatmap of patient distribution.

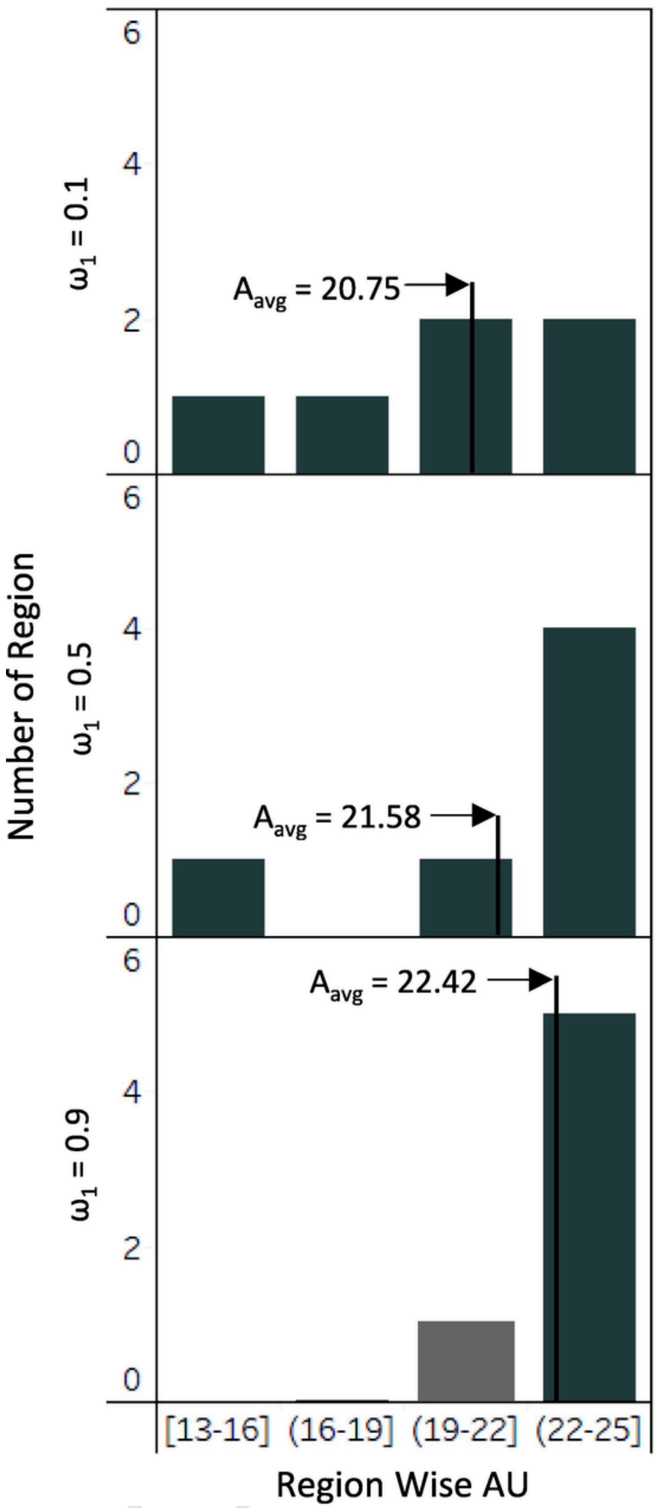


Fig. 11. Region-wise AU (A_{avg}) for different values of ω_1; black line represents A_{avg} of the TSA.

Overall, our results indicate that for the same number of effective MTCs, the potential to improve patient safety is considerably high in the Redistributed approach. Further, the law of diminishing returns applies in the Greenfield network, where an increase from 31.5 to 35 effective MTCs does not yield a significant benefit in the performance. However, the Greenfield solution can enable benchmarking of existing, redistributed, or any other network that the state trauma decision makers may be considering.

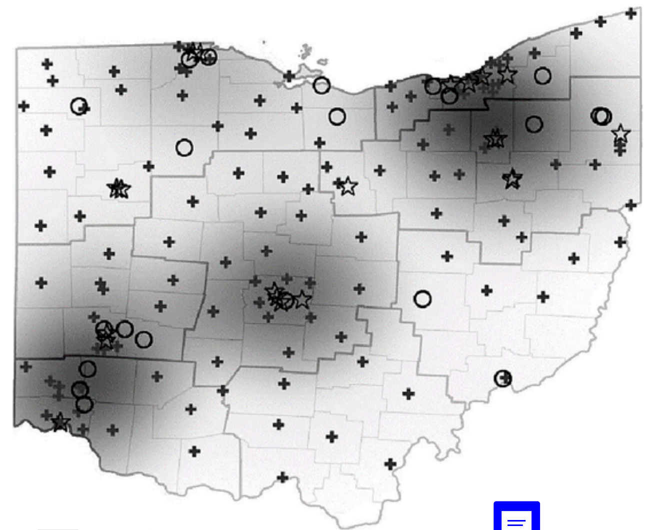


Fig. 12. Trauma network in OH for 8 regions; star indicates MTCs, circle indicates ITCs and cross represents NTCs. Darker shades of grey indicate higher volume of incidences.

Table 7
Performance of Existing, Redistributed, and Greenfield network.

Network	Obj	Average srAU	Max srAU	# MTC	# ITC	Effect. MTC
Existing	16.50	10	23	21	21	31.5
Redistributed	9.69	6.88	12.5	25	13	31.5
Greenfield	9.59	6.69	12.5	29	12	35

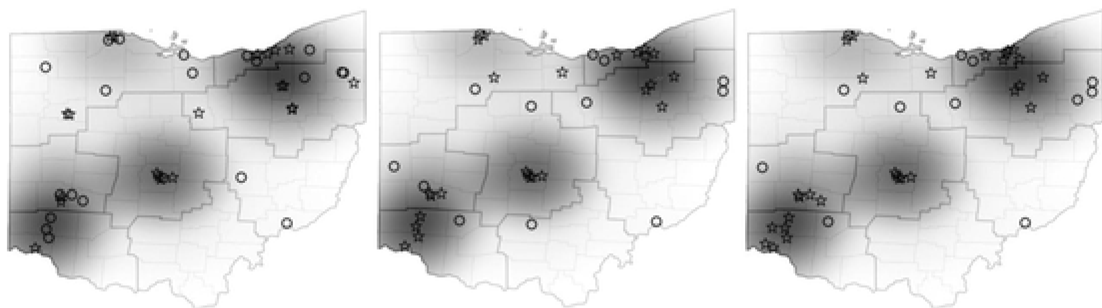
6.2. *Greenfield network with 100% protocol criteria*

Fig. 14 compares the Greenfield network with two destination determination allocations over 11,313 patients. We observed that a network with 100% protocol criteria results in a reduction in the objective by more than 50% (9.59 vs. 4.75) compared to a network with 20% protocol, 50% patient choice and 30% closest facility (similar to the Existing network). This observation is similar to Insight 1. We did observe that 100% protocol led to fewer number of effective MTCs (24 vs. 31.5; MTCs increase to 29 from 22, but ITCs decreased to 4 from 12).

Further, maximum srAU among all regions decreased by 52% (a 47.7% reduction in average srAU patients), which corresponds to 27 ({6.89–3.5}*8) fewer severely injured patients who would suffer a mistriage (srUT or srUT^s) in sample (3,552) data. Considering 71,971 patients reported in 2019 in OH, this would correspond to 547 fewer patients annually, which is substantial. Clearly, following ACS recommendation of using protocol as the primary criteria can lead to substantial benefits in patient safety; however, this will require the state to introduce new EMS training initiatives to promote protocol, while attempting to mitigate other reasons for destination determination.

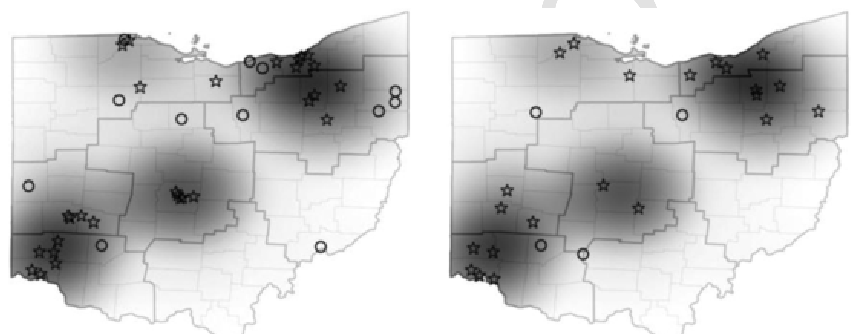
7. Summary and future research

Our research introduced the nested trauma network design problem (NTNDP) that determines the number and location of major, intermediate, and non-trauma centers for a trauma service area. The NTNDP minimizes a weighted sum of equity among regions and effectiveness across the TSA. Several practical considerations, compared to existing trauma literature, were incorporated in the NTNDP, such as multiple patient types, multiple choices for transportation, multiple destination determination criteria and multiple level of hospitals. Specific to the trauma literature, NTNDP generalizes TCLP by including intermediate trauma centers, a vital element of a trauma network that improves access to trauma care for regions that do not have access to a major trauma cen-



(a) Existing 2019 network (b) Redistributed network (c) Greenfield network

Fig. 13. Comparison of trauma networks (darker shades indicate higher volume of incidences; stars indicate MTCs; circles indicate ITCs).



a) 20% protocol, 50% patient choice, 30% closest facility (similar to Existing network) b) 100% protocol

Fig. 14. Optimal trauma network superimposed on heat maps of incidence.

ter. The inclusion of three dominant destination determination criteria ensures faithful representation of the on-scene EMS decision making process. In addition, consideration of both equity and effectiveness in the objective allow trauma decision makers to trade off these two contradicting measures in network design.

We modeled the NTNDP as a mixed-integer linear program and proposed a ‘3-Phase’ solution approach to find near-optimal solutions in a reasonable amount of time. We generated a TSA using data from a state trauma system in the US and quantified the impact of system parameters on the performance of the trauma system. We illustrated the use of the proposed approach on 2019 data for the state of Ohio. The key findings from our study are as follows:

- If EMS providers could exclusively use ‘protocol’ as the criterion for destination determination at the incidence location, then a trauma network with low levels of under-triages can be realized with fewer MTCs and ITCs. Increased use of the ‘patient choice’ criterion could result in higher number of ITCs required in suburban and rural zones and increased under-triages.
- A clustered distribution of severely injured patients in the TSA appears to improve trauma system performance with fewer MTCs and ITCs.
- Solely focusing on equity of patient safety among regions as an objective function appears myopic; balancing it with effectiveness across the TSA appears to result in a better performing network.
- Illustration of our approach on real data from a midwestern US state indicated an over 30% improvement in patient safety; Greenfield network can enable benchmarking of existing, redistributed, and other networks. Importance of 100% use of protocol for destination determination was also verified.

These findings have several practical implications. Trauma decision makers can use our approach to comprehend the compromised role offered by ITCs on patient safety (via provision of intermediate care), especially in suburban and rural zones where MTCs are financially not viable (due to a low number of patients). Further, they can quantify the impact of various destination determination criteria used in practice on patient safety. They can, subsequently, design training programs for EMS providers that help them employ ‘protocol’ (which is the preferred approach suggested by ACS) during on-scene decision making.

Future research in this domain could include incorporating on-scene patient vitals available in the EMS registries to improve injury determination. Further, it would be interesting to consider the variability in the EMS providers’ on-scene injury assessment and its impact on subsequent decisions. Considering migration patterns among regions in a state and other uncertainties in patient distribution to design a robust network could also be considered.

CRediT authorship contribution statement

Sagarkumar Hirpara : Conceptualization, Methodology, Software, Data curation, Formal analysis, Investigation, Visualization, Writing – original draft. **Pratik J. Parikh :** Conceptualization, Methodology, Formal analysis, ~~Writing – original draft,~~ Writing – original draft, Supervision. **Nan Kong :** Conceptualization, Methodology.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgements

This work was partially supported by the US National Science Foundation (CMMI##2100979 and #1761022).

Appendix A:. Optimization model to estimate coefficients of patient choice utility model

Our proposed model determines the coefficients of the utility model in order to minimize the misclassification of patients. A patient is misclassified if the destination hospital type estimated through the utility model differs from the actual destination type; e.g., a misclassification would be when a patient was taken to MTC (according to actual data), while the utility model's estimated choice is NTC. The optimization model is presented below with parameter and decision variables in Table A1 and A2, respectively.

minimize: $\sum_i m_i$
subject to:

$$u_{ij} = \beta_1 \sum_l A^l X_j^l - \beta_2 TG_{ij}; \forall i \in I, \forall j \in J \tag{1}$$

$$u_i^{\max} \geq u_{ij}; \forall i \in I, \forall j \in J \tag{2}$$

$$(u_i^{\max} - u_{ij}) - M(1 - n_{ij}) \leq 0; \forall i \in I, \forall j \in J \tag{3}$$

$$\sum_j n_{ij} = 1; \forall i \in I \tag{4}$$

$$m_i \geq n_{ij} X_j^l - C_i^l; \forall i \in I, \forall j \in J, \forall l \in L \tag{5}$$

$$0 \leq \beta_1, \beta_2 \leq 1 \tag{6}$$

$$u_{ij}, u_i^{\max} \in \mathbb{R}; \forall i \in I, \forall j \in J \tag{7}$$

$$n_{ij} \in \{0, 1\}; \forall i \in I, \forall j \in J \tag{8}$$

$$m_i \in \{0, 1\}; \forall i \in I \tag{9}$$

The objective of the model is to minimize the total misclassification of patients. Constraints (1)–(4) are similar to the NTNDP model to capture patients' choices using the utility model. For each patient i, Constraints (1) calculate the utility of receiving care at each hospital, Constraints (2) find maximum utility among all hospitals, while Constraints (3) and (4) record the hospital with the maximum utility. For each patient i, Constraints (5) record misclassification by comparing the estimated and actual hospital type. Constraints (6)-(9) define bound on decision variables.

The attractiveness for MTC, ITC, and NTC is set as 5, 3, and 1, respectively. We used 5627 cases assigned through patient choice criteria in the cleaned 2019 data from the state of Ohio, along with corresponding 2019 network of hospitals and their types. Further, we used ArcGIS to generate the ground travel time matrix and the Gurobi solver to find an optimal solution.

Table A1
Parameters in the model.

Notation	*Definition*
I	Set of trauma patients assigned via protocol criteria; $i \in I$
J	Set of candidate hospital locations (for MTC, ITC, and NTC); $j \in J$
L	Set of hospital type; $l \in L$; l = 1, 2, 3 represent MTC, ITC, and NTC, respectively
A^l	Attractiveness of hospital level l
C_i^l	1, if patient i chose hospital type l; 0, otherwise
TG_{ij}	Travel time from patient i to any candidate location j via ground
X_j^l	1, if a candidate location j is designated to be level l; 0, otherwise
M	Big number

Table A2
Decision variables in the model.

Notation	*Definition*
β_1	Coefficient for the attractiveness of hospital
β_2	Coefficient for travel time between incidence location and hospital
u_{ij}, $u_i^{\max}$	Utility of patient i receiving care at hospital j; $u_i^{\max} = \max_j \{u_{ij}\}$
n_{ij}	1, if the highest utility for patient i occurs for a hospital j; 0, otherwise
m_i	1, if estimated choice through utility model is different than chosen hospital type; 0, otherwise

References

Ahmadi-Javid, A., Seyedi, P., & Syam, S.S. (2017). A survey of healthcare facility location. Computers & Operations Research, 79, 223–263.

Ares, J.N., De Vries, H., & Huisman, D. (2016). A column generation approach for locating roadside clinics in Africa based on effectiveness and equity. European Journal of Operational Research, 254(3), 1002–1016.

Armstrong, J.H., Hammond, J., Hirshberg, A., & Frykberg, E.R. (2008). Is overtriage associated with increased mortality? The evidence says "yes". Disaster Medicine and Public Health Preparedness, 2(1), 4–5.

Balcik, B., & Beamon, B.M. (2008). Facility location in humanitarian relief. International Journal of Logistics, 11(2), 101–121.

Barringer, M.L., Thomason, M.H., Kilgo, P., & Spallone, L. (2006). Improving outcomes in a regional trauma system: Impact of a level III trauma center. The American Journal of Surgery, 192(5), 685–689.

Bayram, V., Tansel, B.Ç., & Yaman, H. (2015). Compromising system and user interests in shelter location and evacuation planning. Transportation Research Part B: Methodological, 72, 146–163.

Beliën, J., De Boeck, L., Colpaert, J., Devesse, S., & Van den Bossche, F. (2013). Optimizing the facility location design of organ transplant centers. Decision Support Systems, 54(4), 1568–1579.

Branas, C.C., MacKenzie, E.J., & ReVelle, C.S. (2000). A trauma resource allocation model for ambulances and hospitals. Health Services Research, 35(2), 489.

Branas, C.C., Wolff, C.S., Williams, J., Margolis, G., & Carr, B.G. (2013). Simulating changes to emergency care resources to compare system effectiveness. Journal of Clinical Epidemiology, 66(8), S57–S64.

Brown, J.B., Rosengart, M.R., Billiar, T.R., Peitzman, A.B., & Sperry, J.L. (2016). Geographic distribution of trauma centers and injury related mortality in the United States. The Journal of Trauma and Acute Care Surgery, 80(1), 42.

Burkey, M.L., Bhadury, J., & Eiselt, H.A. (2012). A location-based comparison of health care services in four US states with efficiency and equity. Socio-Economic Planning Sciences, 46(2), 157–163.

Cardoso, T., Oliveira, M.D., Barbosa-Póvoa, A., & Nickel, S. (2015). An integrated approach for planning a long-term care network with uncertainty, strategic policy and equity considerations. European Journal of Operational Research, 247(1), 321–334.

Centers for Disease Control and Prevention (2022). Injuries and Violence Are Leading Causes of Death. Retrieved from https://www.cdc.gov/injury/wisqars/animated-leading-causes.html.

Chanta, S., Mayorga, M.E., & McLay, L.A. (2014). Improving emergency service in rural areas: A bi-objective covering location model for EMS systems. Ann. Oper. Res., 221(1), 133–159.

Chen, Z., Chen, X., Li, Q., & Chen, J. (2013). The temporal hierarchy of shelters: A hierarchical location model for earthquake-shelter planning. International Journal of Geographical Information Science, 27(8), 1612–1630.

Cho, S.-H., Jang, H., Lee, T., & Turner, J. (2014). Simultaneous location of trauma centers and helicopters for emergency medical service planning. Operations Research, 62(4), 751–771.

Cocking, C., Flessa, S., & Reinelt, G. (2012). Improving access to health facilities in Nouna district, Burkina Faso. Socio-Economic Planning Sciences, 46(2), 164–172.

Daskin, M.S. (2011). Network and discrete location: Models, algorithms, and applications. John Wiley & Sons.

Enayati, S., Mayorga, M.E., Toro-Díaz, H., & Albert, L.A. (2019). Identifying trade-offs in equity and efficiency for simultaneously optimizing location and multipriority dispatch of ambulances. International Transactions in Operational Research, 26(2), 415–438.

Rotondo, M. F., Cribari, C., & Smith, R. S. (2014). Resources for optimal care of the injured patient. Retrieved from https://www.facs.org/media/yu0laoqz/resources-for-optimal-care.pdf.

Frykberg, E.R. (2002). Medical management of disasters and mass casualties from terrorist bombings: How can we cope? Journal of Trauma and Acute Care Surgery, 53(2), 201–212.

Güneş, E., Yaman, H., Cekyay, B., & Verter, V. (2014). Matching patient and physician preferences in designing a primary care facility network. Journal of the Operational Research Society, 65(4), 483–496.

Güneş, E.D., Melo, T., & Nickel, S. (2019). Location problems in healthcare. Location Science (pp. 657–686). Springer.

Haase, K., & Müller, S. (2015). Insights into clients' choice in preventive health care facility location planning. OR Spectrum, 37(1), 273–291.

Hirpara, S., Vaishnav, M., Parikh, P.J., Kong, N., & Parikh, P. (2022). Locating trauma centers considering patient safety. Health Care Management Science, 1–20.

Ingolfsson, A., Budge, S., & Erkut, E. (2008). Optimal ambulance location with random delays and travel times. Health Care Management Science, 11(3), 262–274.

Intrevado, P., Verter, V., & Tremblay, L. (2019). Patient-centric design of long-term care networks. Health Care Management Science, 22(2), 376–390.

Jansen, J.O., Moore, E.E., Wang, H., Morrison, J.J., Hutchison, J.D., Campbell, M.K., & Sauaia, A. (2018). Maximizing geographical efficiency: An analysis of the configuration of Colorado's trauma system. Journal of Trauma and Acute Care Surgery, 84(5), 762–770.

Jansen, J.O., Morrison, J.J., Wang, H., He, S., Lawrenson, R., Hutchison, J.D., & Campbell, M.K. (2015). Access to specialist care: Optimizing the geographic configuration of trauma systems. The Journal of Trauma and Acute Care Surgery, 79(5), 756.

Kim, D.-G., & Kim, Y.-D. (2013). A Lagrangian heuristic algorithm for a public healthcare facility location problem. Annals of Operations Research, 206(1), 221–240.

Lee, T., & Jang, H. (2018). An iterative method for simultaneously locating trauma centers and helicopters through the planning horizon. Operations Research for Health Care, 19, 185–196.

Lejeune, M.A., & Prasad, S.Y. (2013). Effectiveness–equity models for facility location problems on tree networks. Networks, 62(4), 243–254.

Lerner, E.B. (2006). Studies Evaluating current field triage: 1966–2005. Prehospital Emergency Care, 10(3), 303–306.

Lim, C.S., Mamat, R., & Braunl, T. (2011). Impact of ambulance dispatch policies on performance of emergency medical services. IEEE Transactions on Intelligent Transportation Systems, 12(2), 624–632.

MacKenzie, E.J., Rivara, F.P., Jurkovich, G.J., Nathens, A.B., Frey, K.P., Egleston, B.L., … Scharfstein, D.O. (2006). A national evaluation of the effect of trauma-center care on mortality. New England Journal of Medicine, 354(4), 366–378.

McLay, L.A., & Mayorga, M.E. (2013). A dispatching model for server-to-customer systems that balances efficiency and equity. Manufacturing & Service Operations Management, 15 (2), 205–220.

Mestre, A.M., Oliveira, M.D., & Barbosa-Póvoa, A.P. (2015). Location–allocation approaches for hospital network planning under uncertainty. European Journal of Operational Research, 240(3), 791–806.

Nasrabadi, A.M., Najafi, M., & Zolfagharinia, H. (2020). Considering short-term and long-term uncertainties in location and capacity planning of public healthcare facilities. European Journal of Operational Research, 281(1), 152–173.

Newgard, C.D., Mann, N.C., Hsia, R.Y., Bulger, E.M., Ma, O.J., Staudenmayer, K., … Western Emergency Services Translational Research Network (WESTRN) Investigators (2013). Patient choice in the selection of hospitals by 9-1-1 emergency medical services providers in trauma systems. Academic Emergency Medicine, 20(9), 911–919.

Newgard, C.D., Nelson, M.J., Kampp, M., Saha, S., Zive, D., Schmidt, T., … Warden, C. (2011). Out-of-hospital decision-making and factors influencing the regional distribution of injured patients in a trauma system. The Journal of Trauma, 70(6), 1345.

Parikh, P.P., Parikh, P., Hirpara, S., Vaishnav, M., Sebastian, S., McCarthy, M.C., … Winchell, R.J. (2022). Performance-based assessment of trauma systems: Estimates for the State of Ohio. The American Surgeon. 00031348211065095.

Physicians, A. C. o. E. (1987). Guidelines for trauma care systems. Annals of Emergency Medicine, 16(4), 459-463.

Reuter-Oppermann, M., van den Berg, P.L., & Vile, J.L. (2017). Logistics for emergency medical service systems. Health Systems, 6(3), 187–208.

Salman, F.S., & Yücel, E. (2015). Emergency facility location under random network damage: Insights from the Istanbul case. Computers & Operations Research, 62, 266–281.

Schmid, V. (2012). Solving the dynamic ambulance relocation and dispatching problem using approximate dynamic programming. European Journal of Operational Research, 219(3), 611–621.

Michigan Department of Health and Human Services (2022). Michigan Statewide Trauma System. Retrieved from https://www.michigan.gov/mdhhs/doing-business/trauma.

Texas Department of State Health Services (2022). Trauma System Development. Retrieved from https://dshs.texas.gov/emstraumasystems/etrauma.shtm.

Shariff, S.R., Moin, N.H., & Omar, M. (2012). Location allocation modeling for healthcare facility planning in Malaysia. Computers & Industrial Engineering, 62(4), 1000–1010.

Smith, H.K., Harper, P.R., & Potts, C.N. (2013). Bicriteria efficiency/equity hierarchical location models for public service application. Journal of the Operational Research Society, 64(4), 500–512.

American Trauma Society (2022). Trauma Center Levels Explained. Retrieved from https://www.amtrauma.org/page/traumalevels.

Stone, D. A. (1997). Policy paradox: The art of political decision making (Vol. 13). Ww Norton New York.

American College of Surgeons Committee on Trauma (2016). Strengthening Our National Trauma System. Retrieved from https://docs.house.gov/meetings/IF/IF14/20160712/105121/HHRG-114-IF14-20160712-SD005.pdf.

Tinkoff, G.H., O'connor, R.E., Alexander, E.L., III, & Jones, M.S. (2007). The Delaware trauma system: Impact of Level III trauma centers. Journal of Trauma and Acute Care Surgery, 63(1), 121–127.

Toro-Díaz, H., Mayorga, M.E., Chanta, S., & McLay, L.A. (2013). Joint location and dispatching decisions for emergency medical services. Computers & Industrial Engineering, 64(4), 917–928.

American College of Surgeons Committee on Trauma (2015). Needs-Based Assessment of Trauma Systems (NBATS) Tool. Retrieved from https://lern.la.gov/wp-content/uploads/Appendix-I-acs-nbats-tool.pdf.

Van Buuren, M., Jagtenberg, C., Van Barneveld, T., Van Der Mei, R., & Bhulai, S. (2018). Ambulance dispatch center pilots proactive relocation policies to enhance effectiveness. Interfaces, 48(3), 235–246.

Wang, H., Jin, Y., & Jansen, J.O. (2016). Data-driven surrogate-assisted multiobjective evolutionary optimization of a trauma system. IEEE Transactions on Evolutionary Computation, 20(6), 939–952.

Wang, X.J., Yang, M., & Fry, M.J. (2015). Efficiency and equity tradeoffs in voting machine allocation problems. Journal of the Operational Research Society, 66(8), 1363–1369.

Zhang, Y., & Atkins, D. (2019). Medical facility network design: User-choice and system-optimal models. European Journal of Operational Research, 273(1), 305–319.

Zhang, Y., Berman, O., Marcotte, P., & Verter, V. (2010). A bilevel model for preventive healthcare facility network design with congestion. IIE Transactions, 42(12), 865–880.

Zhang, Y., Berman, O., & Verter, V. (2009). Incorporating congestion in preventive healthcare facility network design. European Journal of Operational Research, 198(3), 922–935.

Zhang, Y., Berman, O., & Verter, V. (2012). The impact of client choice on preventive healthcare facility network design. OR Spectrum, 34(2), 349–370.