



Decision Analysis

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Optimizing the First Response to Sepsis: An Electronic Health Record-Based Markov Decision Process Model

Erik Rosenstrom, Sareh Meshkinfam, Julie Simmons Ivy, Shadi Hassani Goodarzi, Muge Capan, Jeanne Huddleston, Santiago Romero-Brufau

To cite this article:

Erik Rosenstrom, Sareh Meshkinfam, Julie Simmons Ivy, Shadi Hassani Goodarzi, Muge Capan, Jeanne Huddleston, Santiago Romero-Brufau (2022) Optimizing the First Response to Sepsis: An Electronic Health Record-Based Markov Decision Process Model. Decision Analysis

Published online in Articles in Advance 22 Jul 2022

<https://doi.org/10.1287/deca.2022.0455>

Full terms and conditions of use: <https://pubsonline.informs.org/Publications/Librarians-Portal/PubsOnLine-Terms-and-Conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval, unless otherwise noted. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2022, INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Optimizing the First Response to Sepsis: An Electronic Health Record-Based Markov Decision Process Model

Erik Rosenstrom,^{a,*} Sareh Meshkinfam,^{a,b} Julie Simmons Ivy,^a Shadi Hassani Goodarzi,^a Muge Capan,^c Jeanne Huddleston,^d Santiago Romero-Brufau^{e,f}

^aDepartment of Industrial and Systems Engineering, North Carolina State University, Raleigh, North Carolina 27606; ^bDynamic Ideas LLC, Waltham, Massachusetts 02452; ^cDepartment of Mechanical and Industrial Engineering, University of Massachusetts Amherst, Amherst, Massachusetts 01003; ^dRobert D. and Patricia E. Kern Center for the Science of Health Care Delivery, Mayo Clinic, Rochester, Minnesota 55902; ^eDepartment of Otolaryngology (ENT) / Head and Neck Surgery, Mayo Clinic, Rochester, Minnesota 55902; ^fDepartment of Biostatistics, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts 02115

*Corresponding author

Contact: erosens@ncsu.edu, <https://orcid.org/0000-0002-7755-0559> (ER); smeshki@ncsu.edu, <https://orcid.org/0000-0003-1895-5393> (SM); jsivy@ncsu.edu, <https://orcid.org/0000-0002-1138-4076> (JSI); shassan3@ncsu.edu (SHG); mcapan@umass.edu, <https://orcid.org/0000-0003-3194-6117> (MC); huddleston.jeanne@mayo.edu, <https://orcid.org/0000-0001-9101-4122> (JH); romerobrufau.santiago@mayo.edu, <https://orcid.org/0000-0002-9922-0083> (SRB)

Received: December 28, 2020

Revised: September 24, 2021; March 27, 2022

Accepted: April 19, 2022

Published Online in Articles in Advance: July 22, 2022

<https://doi.org/10.1287/deca.2022.0455>

Copyright: © 2022 INFORMS

Abstract. Sepsis is considered a medical emergency where delays in initial treatment are associated with increased morbidity and mortality, yet there is no gold standard for identifying sepsis onset and thus treatment timing. We leverage electronic health record (EHR) data with clinical expertise to develop a continuous-time Markov decision process (MDP) optimal stopping model that identifies the optimal first intervention action (anti-infective, fluid, or wait). To study the impact of initial treatment of patients at risk for developing sepsis, we define the delayed treatment population who received delayed treatment upon admission or during hospitalization and serves as an approximation of the natural history of sepsis. We apply the optimal first treatment policy to sample patient visits from the nondelayed treatment population. This analysis indicates the average risk of death could be reduced by approximately 2.2%, the average time until treatment could be reduced by 106 minutes, and the average severity of the treatment state could be reduced by 15.5% compared with the treatment they received in the hospital. We study the properties of the optimal policy to define an easily interpretable initial treatment heuristic that considers a patient’s organ dysfunction, location, and septic shock status. This generalizable framework can inform personalized treatment of patients at risk for sepsis.

History: This paper has been accepted for the *Decision Analysis* Special Issue on Emerging Topics in Health Decision Analysis.

Funding: This material is based upon work supported by the National Science Foundation [Grant 1522107 (North Carolina State University), 1522106 (Mayo Clinic), and 1833538 (Drexel University)].

Keywords: clinical decision support • sepsis treatment • electronic health records • continuous-time Markov decision process • optimal stopping problem

1. Introduction

Sepsis is a life-threatening condition associated with a dysregulated host response to infection, which can cascade into organ system dysfunction (OD) and septic shock (SS) (Singer et al. 2016). Septic shock is the most severe stage in the sepsis continuum in which profound circulatory, cellular, and metabolic abnormalities are associated with a greater risk of mortality than any other stage of sepsis (Hotchkiss et al. 2016). Sepsis

can be brought on by any infection and affects more than 1.5 million Americans each year (CDC 2019, NIH 2019). From 1999–2014, 6% of all deaths had sepsis listed among the causes of death (CDC 2019, NIH 2019). Many of those who survive, especially those who had pre-existing chronic diseases, may experience permanent organ damage (CDC 2019). The cost of sepsis was more than \$23 billion in 2013 for nearly 1.3 million hospitalizations in the United States (NIH 2019).

The incidence of sepsis is increasing as well. From 2012–2018, the annual number of Medicare beneficiaries that were hospitalized for sepsis increased from 811,644 to 1,136,889 (Buchman et al. 2020).

Although sepsis is considered a medical emergency, there is no gold standard for diagnosis. Delays in the initial treatment of sepsis are associated with increased morbidity and mortality (Kumar et al. 2006, SSC 2016). Guidelines for the management of sepsis and septic shock are designed by a committee of 55 international experts representing 25 international organizations every four years for the Surviving Sepsis Campaign. The committee uses an evidence-based approach to formulate and score statements on early management and resuscitation of patients with sepsis or septic shock. Assessing the state of the patient and responding as quickly as possible is a theme for the diagnosis of sepsis patients. It is recommended that a patient's lactate level is assessed and a blood culture is taken before administering antibiotics within one hour from the recognition of sepsis. The corresponding actions are administering broad-spectrum antibiotics (anti-infectives), beginning rapid administration of 30 mL/kg crystalloid (fluids) for hypotension or lactate ≥ 4 mmol/L, and applying vasopressors if hypotensive during or after fluid resuscitation to maintain a mean arterial pressure ≥ 65 mmHg (SSC 2016). Notice that recognition of sepsis is the trigger for treatment, for which there is no gold standard (Yealy et al. 2015).

Medically, there is a key tension between prioritizing fluids or anti-infectives first. This tension is that one ensures appropriate blood pressure and thus appropriate blood flow to organs, whereas the other attacks the infectious organism stopping its proliferation. Depending on how many organ dysfunctions the patient has (which are a marker for insufficient blood flow), it is an open question which of the two treatments is more important. Although ideally both would be administered simultaneously and as soon as possible, physicians are often forced to choose because of resource availability and time constraints of the treatment supply chain (limited nursing time and competing priorities with other patients). This leads to sequential administration of fluids or anti-infectives, requiring a decision of which to give first (SSC 2016).

In this paper, we address this clinical decision-making challenge associated with initial treatment conditions (e.g., type of organ system dysfunctions) and

initial treatment type, fluids or anti-infectives. We leverage electronic health record (EHR) data to develop a proxy for the natural history of sepsis, the delayed treatment population (DTP), and a continuous-time Markov chain (CTMC) model of this population's health during hospitalization. Utilizing this CTMC, we develop a continuous-time Markovian optimal stopping decision model to identify the optimal first treatment policy. This model addresses the conditions for and selection of the first treatment by quantifying the impact of this first treatment as a function of the patient's future health. Using this modeling framework, we develop a generalizable methodology to optimally treat patients at risk for sepsis when they enter a healthcare system or during hospitalization. In Section 2, we review relevant literature; in Section 3, we describe the data and define the study population; in Section 4, we formulate the Markov Decision Process (MDP) model; in Section 5, we examine the structural characteristics, performance of the optimal first treatment policy, and limitations; and we conclude with discussion and directions for future work in Section 6.

2. Literature Review

Markov decision process models have been used in the healthcare field for acute treatment of patients (Sonnenberg and Beck 1993, Shechter et al. 2008, Capan et al. 2017). Markovian and machine learning models have also been used to study sepsis. This research can be divided into two categories: (i) models to predict sepsis onset and associated mortality risk (Alberti et al. 2005, Marshall et al. 2005, Saka et al. 2007, Shapiro et al. 2007, Ribas et al. 2012, Paxton et al. 2013, Tsoukalas et al. 2015, Chen et al. 2017, Ghosh et al. 2017, van Wyk et al. 2019, Gupta et al. 2020, Parente et al. 2021) and (ii) models to identify optimal sepsis treatment strategies (Rangel-Frausto et al. 1998, Magni et al. 2000, Annane et al. 2004, Huang et al. 2007, Tsoukalas et al. 2015, Ghassemi et al. 2017, Raghu et al. 2017, Komorowski et al. 2018, Ayvaci et al. 2021).

i. Sepsis Progression Models: Parente et al. (2021) developed a hidden Markov model (HMM) for early detection and diagnosis of sepsis in patients that signals the beginning of treatment. They used a four-point sepsis score to model patient health to determine the unknown status of a patient's infection and the onset of sepsis. Gupta et al. (2020) utilized an HMM to

continuously model mortality risk for patients who are suspected of infection and at risk for developing sepsis. They found that the temporal nature of the HMM yields better performance compared with nontemporal machine learning techniques like random forest or support vector machine classification.

Other machine learning techniques have been used to study sepsis onset and mortality risk. van Wyk et al. (2019) used random forest classification on different sets of physiological data to classify patients who are likely to develop sepsis. Tsoukalas et al. (2015) developed support vector machine classifiers to determine sepsis patients' length-of-stay and mortality risk. Saka et al. (2007) designed an empirically based Monte Carlo model to simulate the progression of sepsis in 1,888 hospitalized patients enrolled in the Genetic and Inflammatory Markers of Sepsis (GenIMS) study over 30 days. Their model predicted a patient's change in health represented by the Sepsis-Related Organ Failure Assessment (SOFA) score as a function of their previous health state and hospital length of stay (LOS). Similarly, Shapiro et al. (2007), Ribas et al. (2012), and Chen et al. (2017) applied statistical approaches such as logistic regression to predict a patient's mortality from sepsis.

ii. Sepsis Treatment Models: Komorowski et al. (2018) and Raghu et al. (2017) used deep reinforcement learning to find optimal sepsis treatment strategies for intensive care unit (ICU) patients in the Multiparameter Intelligent Monitoring in Intensive Care (MIMIC-III v1.4) database (Johnson et al. 2016). They defined a five-by-five action space for interventions, including the fluid volume and vasopressor dosage for each four-hour window, and derived a sequential treatment policy that performs better than the corresponding clinical treatment. Tsoukalas et al. (2015) used a partially observable Markov decision process model with eight systemic inflammatory response syndrome (SIRS)-based health states to derive a sequential treatment policy for the combination of five different antibiotic treatments. Ayvaci et al. (2021) developed a personalized sepsis alert system using a discrete-time MDP model. They leveraged patient-specific attributes (e.g., race, age) with the SIRS score as the state space to construct patient-specific Markov chains and derive their optimal alert policy. Although these papers have developed Markovian models to find sequential treatment strategies, we specifically focus on the first treatment timing to provide a meaningful

treatment recommendation. We also create a more detailed state space to model patient health to tailor treatment recommendations more effectively.

Our approach combines these two categories of sepsis research in a single Markovian model. Our model of patient health explicitly captures multiple types of organ system dysfunction, inflammation, and the patient's location, which leads to a larger state space. This representation of patient health considers sepsis as an evolving condition and facilitates early intervention or prevention. We use this model to formulate a continuous-time MDP optimal stopping problem to identify the optimal initial treatment type as a function of the patient's health state. An optimal first treatment policy can provide clinical insights that reduce the uncertainty medical practitioners face when treating a patient with suspected infection that is at risk for developing sepsis.

Our approach differs from previous work as we define a unique patient population, the DTP, to approximate the natural history of sepsis and derive the optimal first treatment policy as a function of organ system dysfunction and patient location. Thematically, identifying the optimal first treatment is similar to identifying the onset of sepsis or the first stage of sepsis for which providers should intervene. Because we solve for the optimal first treatment, the action space of our Markovian model also differs from previous work as we identify the type of intervention, administration of fluids or anti-infectives. This action space considers the resource challenges faced by many treatment facilities that cannot provide all recommended treatments simultaneously.

3. Study Population

The study population includes adult patient visits (i.e., age 18 or older at the time of admission) for patients suspected of infection defined by a single dose of any anti-infective or a positive viral polymerase chain reaction (PCR) result. We exclude surgical patients and patients who left against medical advice. Time-stamped laboratory results (labs), vital signs, and care location are recorded for each visit. For each new observation (e.g., a newly measured temperature value), the updated value was stored with the corresponding data entry time. The interarrival time of updates is calculated from the difference in time stamps. During their visit, patients are

treated in various locations (i.e., the general floor, emergency department (ED), Step Down (STEPDN), and ICU) and could change locations during their visit. The discharge disposition is also recorded, which terminates the visit. The discharge dispositions are “Dead,” “Non-routine,” or “Routine.” The Dead disposition includes “Do Not Resuscitate” (DNR) and discharge to hospice. The Routine disposition indicates that patients are discharged to their home. The Nonroutine disposition refers to all other discharges, including transfers to another hospital or care facility or receiving a home health service.

3.1. Delayed and Nondelayed Subpopulations

This study focuses on the initial treatment of patients suspected of infection and at risk for sepsis. To study initial treatment, we must define a patient population that captures the disease dynamics of sepsis progression unimpeded by treatment, the natural history of the disease. However, there are only six visits with a positive PCR and no anti-infective or fluid that qualify as the natural history, as it is unethical not to treat patients. To address this issue, we identify a subpopulation, the DTP, that approximates the natural history of sepsis, as those visits with suspected infection during which a patient had delayed anti-infective treatment. All visits that do not belong to the DTP belong to the nondelayed treatment population (NDTP).

Clinically timely treatment is critical for sepsis management; for each hour treatment is delayed, mortality increases by 3%–8% (Kumar et al. 2006, Andersson et al. 2019). A significant delay in treatment can be as bad as receiving no treatment (Bone et al. 1989, Kumar et al. 2006). We explore the optimization of the first treatment for the DTP. We define two types of delayed treatment visits based on the time of infection suspicion: (i) infection suspicion at the time of admission and (ii) infection suspicion during their hospital stay. Based on clinical expert opinion, these two groups are defined as follows:

1. Patients who had a positive culture within 48 hours from admission:
 - a. Patients who received no anti-infective in the first 24 hours from admission and received anti-infective within the first 72 hours from admission.
 - b. Patients who received no anti-infective from culture order for more than 24 hours and received anti-infective within the first 72 hours from admission.

2. Patients who had no culture, no anti-infective, and no met SIRS criteria in the first 48 hours from admission:

- a. Patients who did not have a culture or an anti-infective before meeting SIRS; the time between SIRS criteria and culture order was between 24 and 72 hours.
- b. Patients who did not have a culture or an anti-infective before meeting SIRS; the time between SIRS criteria and anti-infective was between 24 and 72 hours.

Clinical expertise suggests delayed treatment upon admission (Group 1) may arise in practice because of human error (possibly due to the load on the system or patient volume) as best treatment practices emphasize beginning treatment within one hour (SSC 2016). There are several possible delays in care, including failure to see the results of labs and vitals for timely treatment decisions to be made, errors in patient admission, patient results are switched, and false-positive culture results. EHR limitations do not allow for exact reconstruction of events in the hospital. It is possible that patients may be misclassified as belonging to the NDTP.

4. Markov Decision Process Optimal Stopping Problem Formulation

We use DTP EHR data to formulate an MDP stopping problem to identify the optimal first treatment policy for patients with suspected infection at risk for sepsis. The state space (Section 4.2) and transition probabilities (Section 4.3) of the MDP are constructed using the health trajectories from admission to discharge. The possible actions are to administer fluid, administer anti-infectives, or wait (Section 4.4). The optimal policy identified for each state minimizes a function that considers the cost of (i) treatment, (ii) the current health of the patient, (iii) the future expected health of the patient, and (iv) the expected discharge status (Section 4.5). The optimal first treatment policy is identified using policy iteration (Section 4.6). The concept of policy dominance is introduced in Section 4.7. Note that we use the postintervention portion of NDTP visits to inform action cost; thus, it receives similar treatment to the DTP in Sections 4.2 and 4.3. Table 1 introduces the required notation.

4.1. Modeling Intuition and Assumptions

The problem is modeled as an infinite horizon stopping problem to identify a time-independent or stationary

Table 1. Markov Decision Process Notation

Notation		Description
State Space	S^ω	Set of states in CTMC ω , $\omega \in \Omega = \{W \text{ (Delayed), F (Post Intervention Fluid), AI (Post-Intervention Anti-Infective)}\}$
	q^ω	Set of transient states, $q^\omega \subset S^\omega$,
Probability Transition Matrix	$p(j s, d(s))$	Probability of transitioning to state j given action $d(s)$ is taken in the current state s
	$\beta(s, d(s))$	Transition rate from state s under action $d(s)$
	$p(j t, s, d(s))$	Probability that the process occupies state j , t time units after a decision epoch, given that action $d(s)$ was chosen in state s at the current decision epoch has not occurred prior to time t . If the state of the process does not change until the next decision epoch, $p(j t, s, d(s)) = 1$
	$F(j t, s, d(s))$	Probability that a transition from state s to state j occurs within t time after action $d(s)$
	$M(s, \hat{s})$	Mean time to transition from state s to \hat{s} observed in EHRs
	T	Uniformized transition probability matrix (TPM)
	$T(s' s, \pi(s))$	Uniformized transition probability from s to s' under policy $\pi(s)$
	$T(s' s, d(s))$	Uniformized transition probability from s to s' under action $d(s)$
	P^ω	TPM of the CTMC ω , $\omega \in \Omega$
	Q^ω	Transient-to-transient TPM of P^ω , $\omega \in \{F, AI\}$
	R^ω	Transient-to-recurrent TPM of P^ω , $\omega \in \{F, AI\}$
	D^ω	$(I - Q^\omega)^{-1}$, $\omega \in \{F, AI\}$
	B^ω	$D^\omega R^\omega$, $\omega \in \{F, AI\}$
	Action Space	$d(s)$
Cost Function	$r(s, d(s))$	Total cost associated with taking action $d(s)$ in state s
	$k(s, d(s))$	Immediate cost of taking action $d(s)$ in state s
	$H(s, d(s))$	Expected remaining health cost associated with taking action $d(s)$ in state s
	$c_i(s)$	Cost of i organ dysfunctions in state s , $i \in I = \{1, 2, \dots, 7\}$
	$c_l(s)$	Cost of septic shock in state s
	$c_e(s)$	Location cost associated with state s
	$u(s)$	Probability fluid action has a positive effect in state s
	$q(s)$	Probability anti-infective action has a positive effect in state s
	$m_{d(s)}$	Measure which quantifies effect of action $d(s)$ in state s
	$c(s, d(s))$	Cost rate for taking action $d(s)$ in state s
	Γ	Terminal vector cost of the absorbing states
	$t_\omega(s, s')$	Time to absorption from s to s' , $\omega \in \{F, AI\}$
	$R(s, \pi(s))$	Uniformized cost for state s under policy $\pi(s)$
	$R(s, d(s))$	Uniformized cost for state s under action $d(s)$
Solution Method	α	Discount rate, $0 < \alpha < 1$
	c	Uniformization factor
	$\pi(s)$	Action in state s under deterministic policy π
	π^*	Optimal policy
	V_π	Expected infinite-horizon discounted reward vector under policy π , optimal under policy π^*

optimal treatment policy. We choose this approach because the patient's health condition is the key factor for identifying the optimal first treatment. The time scale in which treatment decisions are made in this

context is in the range of minutes or hours, as such our approach focuses on the patient's health condition as a trigger for action. Furthermore, it is generally accepted that for worse health states, the longer a patient stays in

a state, the worse the potential outcome. Therefore, the priority is to intervene as soon as a patient enters a health state rather than delay.

Clinical expertise suggests the current state is the most relevant factor for patients who have not received treatment supporting our Markovian framework. In practice, clinicians use the most recent information available about the patient at the current time to make first treatment decisions. In our model framework, we carry forward vitals and labs until new information is learned, so decisions are made based on the most recently available information about the patient's health. We assume that sojourn time between health states follows an exponential distribution. Additional analysis to support this assumption is in Appendix A.

4.2. State Space

The patient's health condition is defined by more than 40 vital signs and labs within the EHR. These vitals and labs are translated into the patient's health state, which

is tracked longitudinally. To reduce the complexity of the state space, the observed values for each laboratory and vital are aggregated to indicator values, indicating whether the values are in the normal or abnormal range. Further, sets of laboratory results and vitals are grouped to translate to inflammation, organ system function, and septic shock as defined in Table 2. The normal and abnormal value thresholds for each vital and laboratory are based on the SEPSIS Collaborative definition. They were developed through an iterative process considering clinicians' input and guidelines, such as the Surviving Sepsis Campaign (Singer et al. 2016, Rhodes et al. 2017), and are supported by Shankar-Hari et al. (2016).

The state is defined as an 11-element vector of 0 (normal), 1 (abnormal), and NA (i.e., not measured) flags for the patient's vitals and laboratory results corresponding to two types of inflammation, seven organ system dysfunctions, septic shock, and the location within the hospital. When a patient enters the hospital,

Table 2. Health State Element Definition

Health state element	Abbreviation	Criteria for abnormal indicator
Cellular Inflammation	IC	White blood cell count (>12,000 cells/mL) OR Bandemia (>10%) OR Sedimentation rate (>20 mm/h) OR C reactive protein (>8 mg/L) OR Procalcitonin (>15 ng/mL)
Physiological Inflammation	IP	[Heart rate (≥90 beats/min) AND Respiration rate (≥20 breaths/min)] OR Temperature (≥100.4 F or <96.8 F)
Cardiovascular Organ System Dysfunction	OC	Systolic blood pressure (<90 mmHg) OR Mean arterial pressure (<65 mmHg) OR Decrease in SBP >40 mmHg in eight-hour period
Gastrointestinal Organ System Dysfunction	OG	Bilirubin (>2 mg/dL)
Hemopoietic Organ System Dysfunction	OH	White blood cell count (<4,000 cells/mL) OR Platelets (<100,000 cells/mL)
Renal Organ System Dysfunction	OK	Creatinine (>1.2 mg/dL) OR Urine output (<500 mL) OR Blood urea nitrogen (>20 mg/dL)
Metabolic Organ System Dysfunction	OM	Lactate (>2.0 mmol/L)
Neurological Organ System Dysfunction	ON	Glasgow coma score (<14)
Respiratory Organ System Dysfunction	OR	Pulse oximetry (<90%) OR Fraction of inspired oxygen (>21%)
Septic Shock	SS	Persistent hypotension shown through two consecutive readings (≥ 30 min apart) of (i) SBP < 90 mmHg, (ii) MAP < 65mmHg, (iii) decrease in SBP > 40 mmHg in eight-hour period OR Vasopressor administration

all state vector elements are assigned NA except the location. The elements of the state vector change as new information from vitals and labs is accrued.

4.3. Transition Matrix

The health evolution of the DTP and NDTP are formulated as CTMCs. We define the “Time Between Transitions” as the interarrival time between recorded values of the 11-element health state vector components. For each possible transition between two health states, s and \hat{s} , in a CTMC, we calculate the average time to transition, $M(s, \hat{s})$. Because the CTMCs assume exponentially distributed interarrival times, the rate of each transition between any state s and state \hat{s} is defined as

$$Rate(s, \hat{s}) = \frac{1}{M(s, \hat{s})}. \quad (1)$$

The transition probability matrix (TPM), P^ω , that characterizes the transitions of the CTMC ω , $\omega \in \Omega = \{D$ (Delayed), F (Postintervention Fluid), AI (Postintervention Anti-Infective)}. The (s, \hat{s}) element of P^ω describes the probability of transitioning from state s to state \hat{s} and is defined as follows:

$$P^\omega(s, \hat{s}) = \frac{Rate(s, \hat{s})}{\sum_{s'' \in S^\omega} Rate(s, s'')}. \quad (2)$$

4.4. Action Space

Treatment is decided each time the health state is updated. The treatment actions for the decision model are “wait,” “administer fluid,” or “administer anti-infective,” which are based on clinical guidance associated with sepsis management (SSC 2016). Treatment actions are modeled as absorbing states in the CTMC.

4.5. Cost Structure

The costs in the MDP are not monetary, but rather they model decrements associated with the patient’s health state. The expected discounted cost, $r(s, d(s))$, is composed of three elements that consider different aspects of a patient’s health trajectory and treatment. The expected discounted cost associated with choosing action $d(s)$ in state s is the summation of (i) the immediate cost of action $d(s)$ in state s , $k(s, d(s))$; (ii) the cost for the time the patient spends in state s using the cost rate, $c(s, d(s))$; and (iii) the expected cost for the remaining patient health trajectory from state s , $H(s, d(s))$. Relative costs are assigned based on the severity of illness (e.g.,

number of organ system dysfunctions), the complexity of care based on location (e.g., ICU), and the severity of the discharge location (e.g., home or death). These costs are developed with clinical expertise. Each element of the cost function is described in the sections that follow.

$$r(s, d(s)) = k(s, d(s)) + H(s, d(s)) + \int_0^\infty \sum_{j \in S} c(s, d(s)) \int_0^u e^{-at} p(j|t, s, d(s)) dt F(du|s, d(s)) \quad (3)$$

4.5.1. Immediate Cost of Taking Action. To incorporate the negative effect of administering anti-infective or fluids when it is not clinically recommended, we assign a cost penalty to these actions.

$$k(s, d(s)) = \begin{cases} 0 & \text{for wait} \\ \kappa & \text{for Fluid (0\% to 20\% of the cost of death)} \\ \delta & \text{for Anti-infective (0\% to 20\% of the cost of death)} \end{cases} \quad (4)$$

Based on clinical input, we assume these costs to be constant values ranging between 0% and 20% of the cost of death. Sensitivity analysis is conducted on κ and δ in Section 5.4.

4.5.2. Cost of Time the Patient Spends in State s . The cost rate, $c(s, d(s))$, for state s and action $d(s)$ is the penalty for each unit of time a patient spends in a health state. It quantifies the severity of the health state and is defined as a function of the number of organ dysfunctions, the presence of septic shock, and the patient’s location. Specifically,

$$c(s, d(s)) = (c_i(s) + c_l(s))m_{d(s)} + c_e(s). \quad (5)$$

The cost associated with i organ system dysfunctions in health state s , $c_i(s)$, captures the cumulative effect of organ dysfunctions. Experiencing multiple organ dysfunctions is worse than any single organ dysfunction (Levy et al. 2005, Rubulotta et al. 2009). We assign a “unit” cost for each organ dysfunction in state s . This follows SOFA score criteria, which does not value one organ dysfunction more than another (Singer et al. 2016). The cost of septic shock, $c_l(s)$, captures that septic shock is the most severe health state (Schoenberg et al.

1998, Bauer et al. 2020).

$$c_i(s) = i \quad \forall i \in I = \{1, 2, \dots, 7\} \quad (6)$$

$$c_l(s) = \begin{cases} 9 & \text{if septic shock is present} \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

We incorporate the positive or negative effective of action $d(s)$ in state s through the multiplier $m_{d(s)}$. We define the $m_{d(s)}$ as

$$m_{d(s)} = \begin{cases} 1 & \text{for } d(s) = \text{Wait} \\ 0.9u(s) + 1.5(1 - u(s)) & \text{for } d(s) = \text{Fluid} \\ 0.8q(s) + 2(1 - q(s)) & \text{for } d(s) = \text{Anti-infective.} \end{cases} \quad (8)$$

Clinically, the first treatment impacts the discharge disposition as a patient’s risk of mortality increases by 3%–8% for every hour they are untreated (Kumar et al. 2006, Andersson et al. 2019), so using the discharge disposition provides a reasonable estimate for capturing the effectiveness of a treatment given the treatment state. We define survival (Routine and Nonroutine discharge) as the positive effect of treatment and death as the negative effect of treatment. The probability of survival, $u(s)$ and $q(s)$, given anti-infective or fluid treatment, respectively, in state s is defined by the absorption probabilities of state s from the NDTP postintervention TPMs. The probability of death is the complement, $1 - u(s)$ and $1 - q(s)$, for anti-infective and fluid, respectively. The coefficient for the negative effect of anti-infective is greater than fluid as misuse or overuse of anti-infective should be avoided to prevent the development of antibiotic resistance (Abraham 2015, CDC 2021).

Given the NDTP postintervention TPMs for the anti-infective and fluid treatments, we calculate the absorption probabilities using $B^\omega = (I - Q^\omega)^{-1}R^\omega$, where I is the identity matrix, R^ω is the transient-to-absorbing state submatrix of P^ω , and Q^ω is the transient-to-transient state submatrix of P^ω for $\omega \in \{F, AI\}$ (Puterman 2014). The probabilities $u(s)$ and $q(s)$ are the sum of the probabilities of Routine and Nonroutine discharge found in B^ω . Under the wait action, a multiplier of one is used.

The location of the patient is incorporated via $c_e(s)$, which has values based on the severity of the disease

typically seen within these locations. The location provides additional information about the severity of patient health that may be otherwise unobservable in EHRs. Discharge dispositions are modeled as absorbing states. For these states, $c_e(s)$ is the discharge cost.

$$c_e(s) = \begin{cases} 0 & \text{for GENERALFLOOR} \\ 1 & \text{for STEPDN} \\ 3 & \text{for ED} \\ 5 & \text{for ICU} \\ 0 & \text{for Routine discharge} \\ \Delta & \text{for Nonroutine discharge} \\ \Delta & \text{for Death} \end{cases} \quad (9)$$

This location cost of the discharge disposition is leveraged when calculating the future expected health cost. The relative magnitude of these costs is decided with clinical expertise. Sensitivity analysis is conducted on Δ in Section 5.4.

4.5.3. Expected Remaining Health Cost. To capture the post-anti-infective and fluid treatment health trajectories, we incorporate the future expected health cost for the patient postintervention. This cost is calculated using posttreatment NDTP health trajectories that capture the intervention’s effect on a patient’s future health and discharge disposition. The postintervention CTMCs for each treatment action are transient with absorbing states representing patient discharge. The transient state TPM for each action, Q^ω defined for the set states q^ω , $\omega \in \{F, AI\}$, can be inverted, and the expected time spent in any state s' starting from state s , $t_\omega(s, s')$ is an element of the $D^\omega = (I - Q^\omega)^{-1}$ matrix (Puterman 2014). We calculate the cost rate of the health state s' using this expected time in state s' and the state severity costs defined in Equations (6)–(9). The sum over all transient states given the state s in which the action $d(s)$ is taken defines the future expected health cost until discharge. If the action is to wait, then the patient’s health evolution is defined by the CTMC for the delayed population. We incorporate the discharge disposition in the future expected health cost. Using the absorption probability matrix, B^ω , defined by $B^\omega = D^\omega R^\omega$, this total expected remaining health cost is calculated for each state

and action combination as follows:

$$\begin{aligned}
 & H(\mathbf{s}, d(\mathbf{s})) \\
 & = \begin{cases} \sum_{s' \in q^F} (c_i(\mathbf{s}') + c_l(\mathbf{s}'))t_F(\mathbf{s}, \mathbf{s}') + B^F \Gamma & \text{for } d(\mathbf{s}) = \text{Administer Fluid} \\ \sum_{s' \in q^{AI}} (c_i(\mathbf{s}') + c_l(\mathbf{s}'))t_{AI}(\mathbf{s}, \mathbf{s}') + B^{AI} \Gamma & \text{for } d(\mathbf{s}) = \text{Administer Anti-Infective} \\ 0 & \text{for } d(\mathbf{s}) = \text{Wait,} \end{cases}
 \end{aligned} \tag{10}$$

where Γ is the cost matrix associated with the absorbing states defined by $c_e(\mathbf{s})$.

4.6. Uniformized Value Function

The uniformization of the continuous time Markov decision process applies to the TPM and the rewards. Here $p(j|\mathbf{s}, d(\mathbf{s}))$ is the probability j is the next state given the current state is \mathbf{s} and action $d(\mathbf{s})$ is taken, $\beta(\mathbf{s}, d(\mathbf{s}))$ is the transition rate from state \mathbf{s} under action $d(\mathbf{s})$, and $\beta(\mathbf{s}, \pi(\mathbf{s}))$ is the transition rate from state \mathbf{s} under action $\pi(\mathbf{s})$ defined by policy π . Using discount factor, α , the uniformization constant c where

$$[1 - p(j|\mathbf{s}, d(\mathbf{s}))] \beta(\mathbf{s}, d(\mathbf{s})) \leq c < \infty \quad \forall \mathbf{s} \in S, \tag{11}$$

the uniformized transition probability associated with state \mathbf{s} and action $d(\mathbf{s})$, $T(j|\mathbf{s}, d(\mathbf{s}))$, and uniformized cost associated with state \mathbf{s} and action $d(\mathbf{s})$, $R(\mathbf{s}, d(\mathbf{s}))$, are defined as follows:

$$T(j|\mathbf{s}, d(\mathbf{s})) \approx \begin{cases} 1 - \frac{[1 - p(j|\mathbf{s}, d(\mathbf{s}))] \beta(\mathbf{s}, d(\mathbf{s}))}{c}, & j = s \\ \frac{p(j|\mathbf{s}, d(\mathbf{s})) \beta(\mathbf{s}, d(\mathbf{s}))}{c}, & j \neq s, \end{cases} \tag{12}$$

and

$$R(\mathbf{s}, d(\mathbf{s})) \approx r(\mathbf{s}, d(\mathbf{s})) \frac{\alpha + \beta(\mathbf{s}, d(\mathbf{s}))}{\alpha + c}. \tag{13}$$

The value function associated with policy π for the discounted continuous time MDP is defined as

$$\begin{aligned}
 & V_{\pi(\mathbf{s})} \\
 & = R(\mathbf{s}, \pi(\mathbf{s})) + \frac{\beta(\mathbf{s}, \pi(\mathbf{s}))}{\beta(\mathbf{s}, \pi(\mathbf{s})) + \alpha} \sum_{j \in S} T(j|\mathbf{s}, \pi(\mathbf{s})) V_{\pi(j)} \quad \forall \mathbf{s} \in S.
 \end{aligned} \tag{14}$$

For which the optimality equation is

$$\begin{aligned}
 V_{\pi^*(\mathbf{s})} = \min_{d(\mathbf{s})} & \left\{ R(\mathbf{s}, d(\mathbf{s})) \right. \\
 & \left. + \frac{\beta(\mathbf{s}, \pi(\mathbf{s}))}{\beta(\mathbf{s}, \pi(\mathbf{s})) + \alpha} \sum_{j \in S} T(j|\mathbf{s}, d(\mathbf{s})) V_{\pi^*(j)} \right\} \quad \forall \mathbf{s} \in S.
 \end{aligned} \tag{15}$$

We use the policy iteration algorithm to solve the continuous-time MDP (Puterman 2014). We initiate the algorithm with a randomly generated policy, π , which prescribes administering fluid, administering anti-infective, or waiting with equal probability. To apply the policy iteration, we use the method introduced by Puterman (2014). We use R software to extract the population from the EHR, create the states, and calculate the matrix of transition rates and probabilities. We use Python to implement the policy iteration algorithm.

4.7. Treatment Policy Dominance

By modeling this as a stopping problem, the optimal first treatment policy behaves as a treatment alert system and a treatment recommendation for patients at risk for developing sepsis. We examine the structure of the optimal policy from this perspective. Patients enter the hospital with little known about their condition; as more is discovered, our policy provides guidance for when to intervene and what treatment should be used. Because of potential noise in the optimal policy, we aggregate states by the number of organ dysfunctions, septic shock status, and location to identify the dominant action for each group. An action (anti-infective or fluid) is defined as the dominant action if it is more prevalent among states than the other treatment option and waiting. Waiting is the dominant action if it is more prevalent among states than the sum of anti-infective and fluid. Otherwise, the sum of anti-infective and fluid is more prevalent among states than waiting and an alert for treatment is dominant. The aggregation of states does not differentiate by the inflammation status because there is no cost associated with inflammation.

We examine two types of aggregation with notation Table 3: (i) an unweighted count of states in the group and (ii) a weighted sum that captures the expected proportion of time spent in the states of the DTP TPM. For aggregation method (ii), the starting states are weighted

Table 3. Treatment Policy Dominance Notation

Notation	Description
S^W	Set of states in the DTP TPM
N	Set of the first health states observed for each visit in the DTP, $N \subset S^W$
Q^W	Transient-to-transient TPM of P^W
D^W	$(I - Q^W)^{-1}$
A	Meantime spent in state j
E_i	Proportion of DTP visits that start in state i
W_j	Expected proportion of time that will be spent in state j given the set of starting states N

by the number of times each state is observed in the DTP. The calculation of weight for each state W_j is

$$W_j = \frac{\sum_{i \in N} D_{ij}^W A_j E_i}{\sum_{k \in S^W} \sum_{i \in N} D_{ik}^W A_k E_i}, \quad (16)$$

where D_{ij}^W is the (i, j) element of the $D^W = (I - Q^W)^{-1}$ matrix, A_j corresponds to the mean time spent in state j , and E_i is the observed proportion of patient trajectories that begin with health state i . By weighting the states to study the structure of the optimal policy, we aim to capture the optimal policy structure of states that appear more often in patient visits.

5. Results

First, we summarize our analysis and characterize properties of the study population, emphasizing characterization of the DTP. Second, we discuss the structure of the optimal intervention policy under the baseline cost scenario and develop a heuristic for the initial treatment. Lastly, we compare the performance of the optimal policy to the care patients in NDTP received.

5.1. Analysis of the Study Population

Retrospective data from adult inpatient visits in a health system in the Northeast from July 2013 to April 2016 are used to create DTP and NDTP. There are 88,849 patient visits (corresponding to 53,411 unique patients). After exclusion, the resulting study population corresponds to 58,382 visits (related to 41,540 unique patients). In total, 672 visits meet the criteria for delayed treatment. The remaining patient population meets the study inclusion criteria and did not have delayed treatment, therefore, belonging to the NDTP. The NDTP is divided into groups by the first treatment received: anti-infective or

fluid. There are 49,710 (33,547) and 8,000 (7,321) visits (patients) where the first treatment was anti-infective and fluids, respectively.

Table 4 summarizes the age, demographic, and length of stay distribution of all visits by population type. The DTP is significantly older ($p = 2E-8$) and has a significantly larger proportion of males ($p = 0.002$) than the NDTP. There are no significant differences in race or ethnicity with $\alpha = 0.05$. The literature is conflicted on the impact of gender on sepsis outcomes (Papathanasoglou et al. 2017). However, the difference in age is important from a clinical perspective. The age difference between populations may occur because older patients are more likely to experience atypical manifestations of sepsis, which make them harder to diagnose (Girard et al. 2005). In addition, older patients are at a higher risk of severe outcome than younger patients and older patients' health can also deteriorate faster than young patients (Martin et al. 2006). These factors may influence the data from the NDTP.

Table 5 summarizes the number of health states and the number of unique transitions in the TPMs for the DTP and NDTP. There are 2,719 states observed in all three populations. These states will be referred to as the filtered states of the DTP.

5.1.1. Analysis of the DTP. To better understand the evolution of sepsis in the DTP, we study the state transitions by location. Table 6 describes the number of unique state transitions and the total number of state transitions that occur between locations in the DTP Markov chain and the distribution of the patients' discharge dispositions. Table 6 shows that the ICU has the most unique and total transitions and the largest number of unique health states. The majority of transitions occur between states that are in the same location; only 2.9% of the transitions occur between locations. However, this accounts for 11% of the unique transitions. With respect to transitions to absorbing states, the majority (76%) of the health state transitions to death are from the ICU; the majority of transitions to Nonroutine and Routine Discharge come from the general floor, 67% and 81.5%, respectively. Of the 672 DTP patients, 38, 402, and 232 patients had a Routine Discharge, Nonroutine Discharge, or died in the hospital, respectively.

Table 4. Demographic Distribution of the Delayed and Non-Delayed Treatment Populations

Attribute	Response	Delayed treatment population	Nondelayed treatment population	P-value
Age	Minimum	20	18	
	1st Quartile	55	48	
	Median	65	63	3E-6
	Mean	64.3	61.7	2E-08
	3rd Quartile	77	76	
	Maximum	90	90	
Gender	Male	330 (49.1%)	17,614 (43.1%)	0.002
	Female	342 (50.9%)	23,254 (56.9%)	0.002
Race/ethnicity	Non-Hispanic White	460 (68.5%)	29,098 (71.2%)	0.175
	Non-Hispanic Black	169 (25.2%)	8,991 (22.0%)	0.075
	Non-Hispanic Asian	14 (2.1%)	572 (1.4%)	0.119
	Non-Hispanic American Indian	2 (0.3%)	82 (0.2%)	0.371
	Non-Hispanic Hawaiian/Pacific Islander	0 (0.0%)	16 (0.04%)	0.611
	Hispanic	22 (3.2%)	1,798 (4.4%)	0.185
	Other	3 (0.4%)	159 (0.4%)	0.861
	Unknown	2 (0.3%)	152 (0.4%)	0.393
Length of stay in days	Minimum	2.75	0.1	
	1st Quartile	12.6	2.3	
	Median	23.3	4.1	1E-17
	Mean	29.9	6.3	2E-16
	3rd Quartile	37.6	7.2	
	Maximum	336.8	292.4	

Table 7 shows the average proportion of time a patient spent in each health state by the element and status. The median length of stay for the DTP is approximately 23 days. For each patient, the proportion of time spent in each health state is calculated by dividing the interarrival times between state changes by the patient’s LOS. These portions are summed if the corresponding condition is met. The average of the proportions across all patients is then calculated. Patients spend the largest proportion of their visits with abnormal physiological inflammation (IP), respiratory organ system dysfunction (OR), and renal organ system dysfunction (OK) health state elements and the smallest proportion of their visit with abnormal hemopoietic organ system dysfunction (OH), metabolic organ

system dysfunction (OM), and gastrointestinal organ system dysfunction (OG). SS occurs in the DTP on average for 41% of their LOS.

In Table 8, the average proportion of time patients spent in health states by the number of organ system dysfunctions and presence of septic shock is shown by the patient’s discharge disposition. This is calculated in a similar manner to Table 7. Patients that died spent more time in health states with three or more organ system dysfunctions and in states with septic shock than patients who had a Routine or Nonroutine discharge. Patients who were discharged, routinely or nonroutinely, spent a majority of their time in states with two or fewer organ system dysfunctions.

Table 5. Size Attributes of Delayed and Non-Delayed Treatment Population Markov Chains

	Delayed treatment population	Nondelayed treatment populations: post anti-infective	Nondelayed treatment population: post fluid
Number of health states	3,637	8,945	5,768
Number of unique transitions	12,162	49,083	22,576
Total number of visits	672	49,710	8,000

Table 6. Unique and Total Health State Transitions of the Delayed Treatment Population

Current Location	Next location				Discharge disposition		
	ED	General floor	STEPDN	ICU	Routine	Nonroutine	Death
ED	1,805	133	102	212	0	0	2
(966 states)	3,381	182	118	249	0	0	2
General floor	0	2,615	130	267	24	137	16
(822 states)	0	22,217	259	335	31	270	17
STEPDN	0	207	2,556	171	5	74	27
(789 states)	0	319	24,400	224	5	120	36
ICU	0	185	206	3,190	2	11	83
(1,051 states)	0	260	411	26,578	2	12	177

Note. In each cell, the top value indicates the number of unique transitions and the bottom number indicates the total number of transitions.

5.1.2. Filtered States of the Delayed Treatment Population. The optimal treatment policy is limited to health states observed in the DTP. The number of states for which an optimal treatment policy can be found is limited further by the states that occur in all three groups (i.e., the states observed in the wait-delayed population, anti-infective, and fluid CTMCs). If a state does not appear in the AI and F TPMs, then the health state progression postintervention cannot be determined.

If a patient is in a health state that was not observed in either postintervention CTMC state space, no treatment is taken and the patient must wait. Because it is not possible to assess the impact of taking an action, no optimal action can be identified. An analysis, shown in Table 9, compares the filtered states of the DTP to all states in the DTP that are not in the set of filtered states (i.e., complement of filtered states). We conduct hypothesis tests to determine if the difference of each statistic in the table is statistically significant. Table 9 indicates that there are no significant differences in the severity of the health states based on the average number of organ dysfunctions and proportion of states with septic shock. There are significant differences based on the location of the state. The filtered states of the DTP have a larger proportion of states in the General Floor and STEPDN. The complement of

the filtered states of the DTP has a larger proportion of states in the ED and ICU.

5.2. Structure of Optimal Policy

The optimal policy defines the cost-minimizing treatment action (wait, administer fluid, or administer anti-infective) for each health state in the set of filtered states of the DTP. For the states that are not in the set of filtered states of the DTP states, the optimal action is assumed to be wait by setting the future expected cost of each action to infinity. Because waiting may not be the optimal action for these states, they are excluded from the optimal policy. An optimal policy is only identified for health states in the set of filtered states of the DTP.

We use a baseline set of cost parameters when finding the optimal initial treatment policy. General Floor, STEPDN, ED, and ICU have location costs 0, 1, 3, and 5, respectively; Routine, Nonroutine, and Death have discharge costs 0, 1,000, and 2,000, respectively; administering fluid and anti-infective both have a treatment cost of 200. The corresponding optimal policy assigns 307 states to wait and 375 and 2,036 states to treat with fluid and anti-infective, respectively. To understand the attributes of this baseline optimal policy, the

Table 7. Average Proportion of Time Delayed Treatment Population Patients Spent in Health States During Their Stay by Element

Status	Health state element									
	IC	IP	OC	OG	OH	OK	OM	ON	OR	SS
Normal	0.500	0.406	0.554	0.698	0.841	0.355	0.615	0.000	0.196	0.580
Abnormal	0.493	0.593	0.438	0.095	0.152	0.637	0.145	0.266	0.801	0.410
Not measured	0.007	0.001	0.008	0.207	0.007	0.008	0.240	0.734	0.003	0.010

Table 8. Average Proportion of Time Delayed Treatment Population Patients Spent in Health States by Severity

Discharge disposition	Health state element							Septic shock		
	0	1	2	3	4	5	6	7	Yes	No
Dead	0.033	0.119	0.242	0.294	0.183	0.082	0.037	0.010	0.520	0.480
Nonroutine	0.054	0.198	0.343	0.255	0.111	0.028	0.008	0.003	0.381	0.619
Routine	0.037	0.252	0.362	0.246	0.081	0.022	0.000	0.000	0.379	0.621
All	0.044	0.171	0.304	0.272	0.138	0.048	0.018	0.005	0.410	0.590

severity of the health states (i.e., the number of organ system dysfunctions and the proportion of states with septic shock) for each optimal action group are studied. Intuitively, the wait group had the lowest average number of organ system dysfunctions and the lowest proportion of states with septic shock, 1.85 and 0.195, respectively. In comparison with anti-infective, fluid treatment did not have a statistically different average number of organ system dysfunctions but did have a statistically higher proportion of states with septic shock, 2.61 compared with 2.65 ($p = 0.55$) and 0.38 compared with 0.30 ($p = 0.002$), respectively.

We examine aggregation of less severe health states to establish treatment guidance that is location and organ system specific using the notion of dominance described in Section 4.7. We begin by examining states with no organ dysfunction or septic shock and progress incrementally by adding organ dysfunction and septic shock combinations. Appendix A summarizes the analysis of more severe health states and different cost scenarios.

5.2.1. States with No Organ System Dysfunction and No Septic Shock. Figure 1 shows the unweighted and weighted distribution of the optimal policy action under the baseline cost scenario for states with no organ dysfunction or SS by location. In the ED, waiting

is the dominant action under the weighted aggregation; but under the unweighted aggregation, the number of states that take an action and wait are equivalent. The weighted aggregation considers the importance of the state within the Markovian network. Cases where no treatment dominates under the aggregations are important, as they highlight the potential importance of other patient attributes, for example, comorbidities.

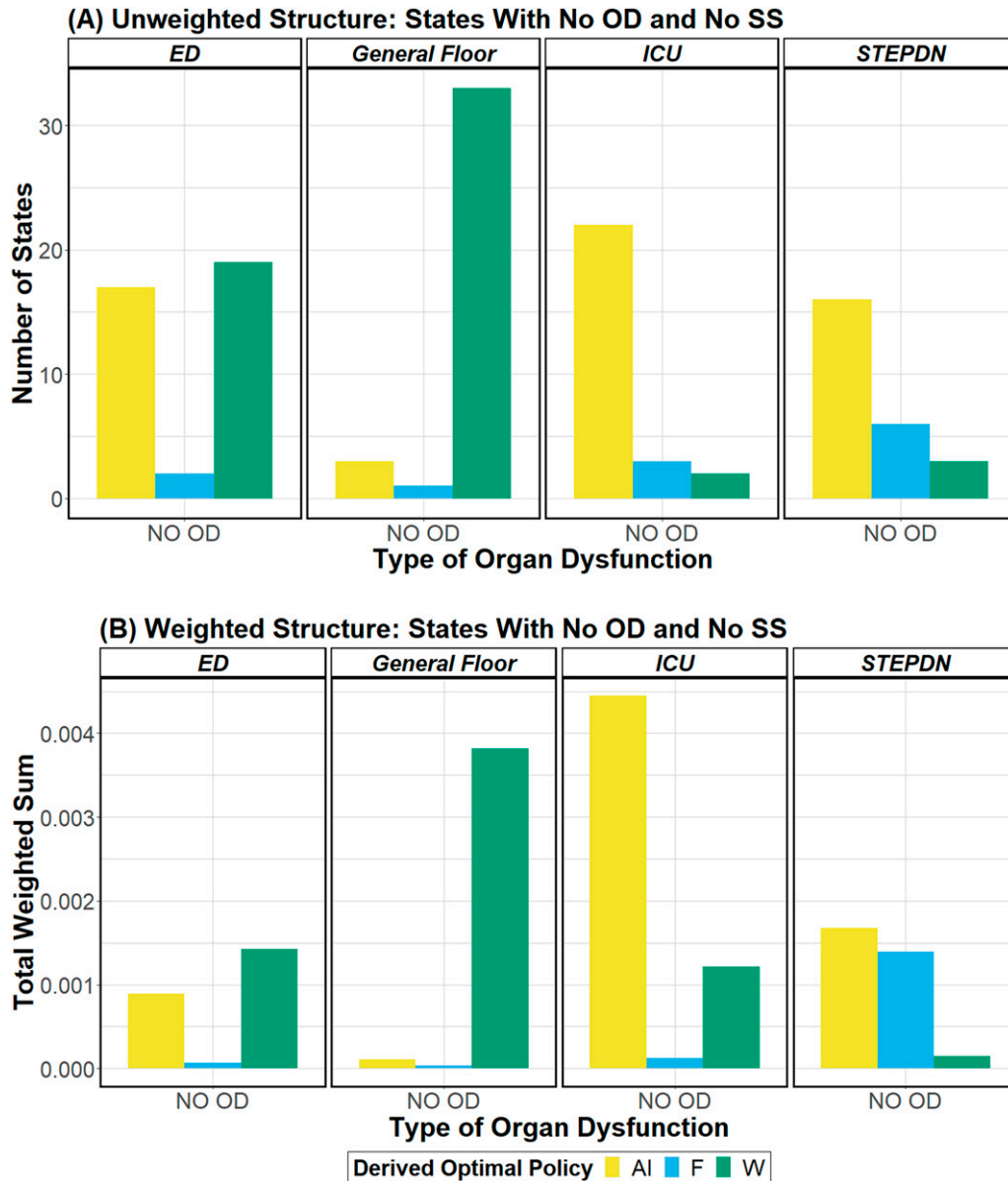
Under both aggregation schemes, waiting is dominant in the General Floor. For the ICU and STEP DN locations, anti-infective treatment is dominant. If patients do not have organ dysfunction and septic shock, their location can inform treatment need. The optimal policy suggests that when patients with no organ dysfunction or SS are in an intensive treatment location, they should be considered for treatment with anti-infective.

5.2.2. States with Septic Shock and No Additional Organ System Dysfunctions. Figure 2 shows the unweighted and weighted distribution of the optimal policy action under the baseline cost scenario for states with one organ dysfunction and SS by location. As defined in Table 2, any state with SS also has cardiovascular organ system dysfunction (OC); however, not all OC includes SS. SS is a severe form of OC. This is studied to understand the treatment guidance

Table 9. Comparison of Filtered States and Complement Filtered States of the Delayed Treatment Population

Statistic	Filtered states	Complement of filtered states	Significance level of difference
Count of states	2,719	918	
Average number of organ dysfunction	2.55	2.56	$P = 0.82$
Proportion of septic shock	0.30	0.32	$P = 0.37$
Proportion of states in ED	0.21	0.44	$P = 2E-16$
Proportion of states in general floor	0.27	0.09	$P = 2E-16$
Proportion of states in STEP DN	0.24	0.14	$P = 1E-10$
Proportion of states in ICU	0.28	0.31	$P = 0.08$

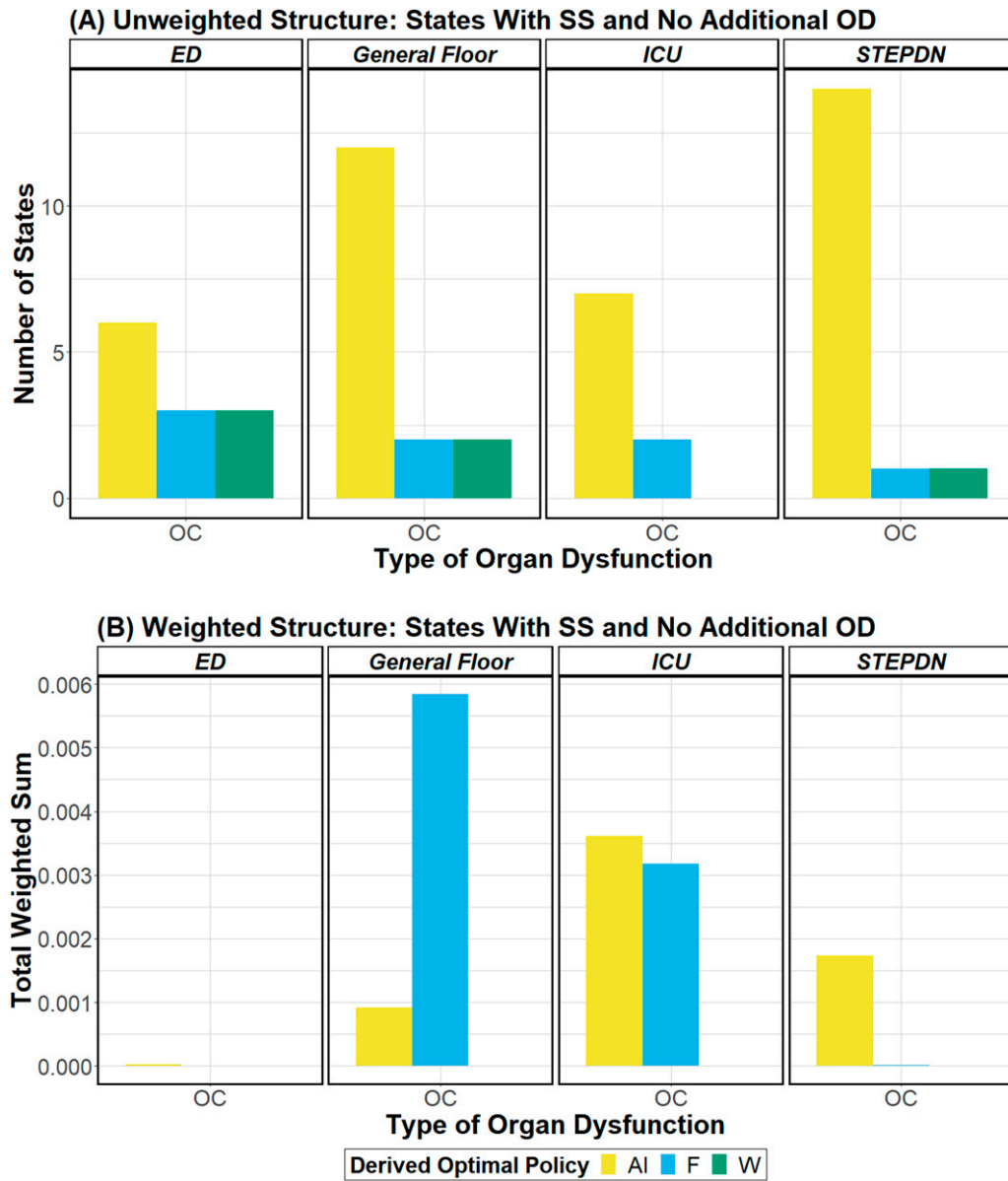
Figure 1. (Color online) Unweighted (a) and Weighted (b) Aggregations of States with No Organ Dysfunction and No SS



if a patient develops SS without developing other organ system dysfunctions. Compared with Figure 1, the introduction of SS warrants treatment in all locations. In the ED and STEPDN, anti-infective is dominant under both aggregations. For the ICU, it is never optimal to wait. Although anti-infective is the more common optimal treatment, fluid is a close second under the weighted aggregation. On the General

Floor, the optimal policy structure differs based on the weighting of the states. Anti-infective is the dominant action in the unweighted aggregation compared with fluid in the weighted. This implies that a larger number of states in the General Floor are optimally treated with anti-infective, but they are collectively less likely to be visited than the states treated with fluid.

Figure 2. (Color online) Unweighted (a) and Weighted (b) Aggregations of States with Septic Shock and No Additional Organ System Dysfunctions

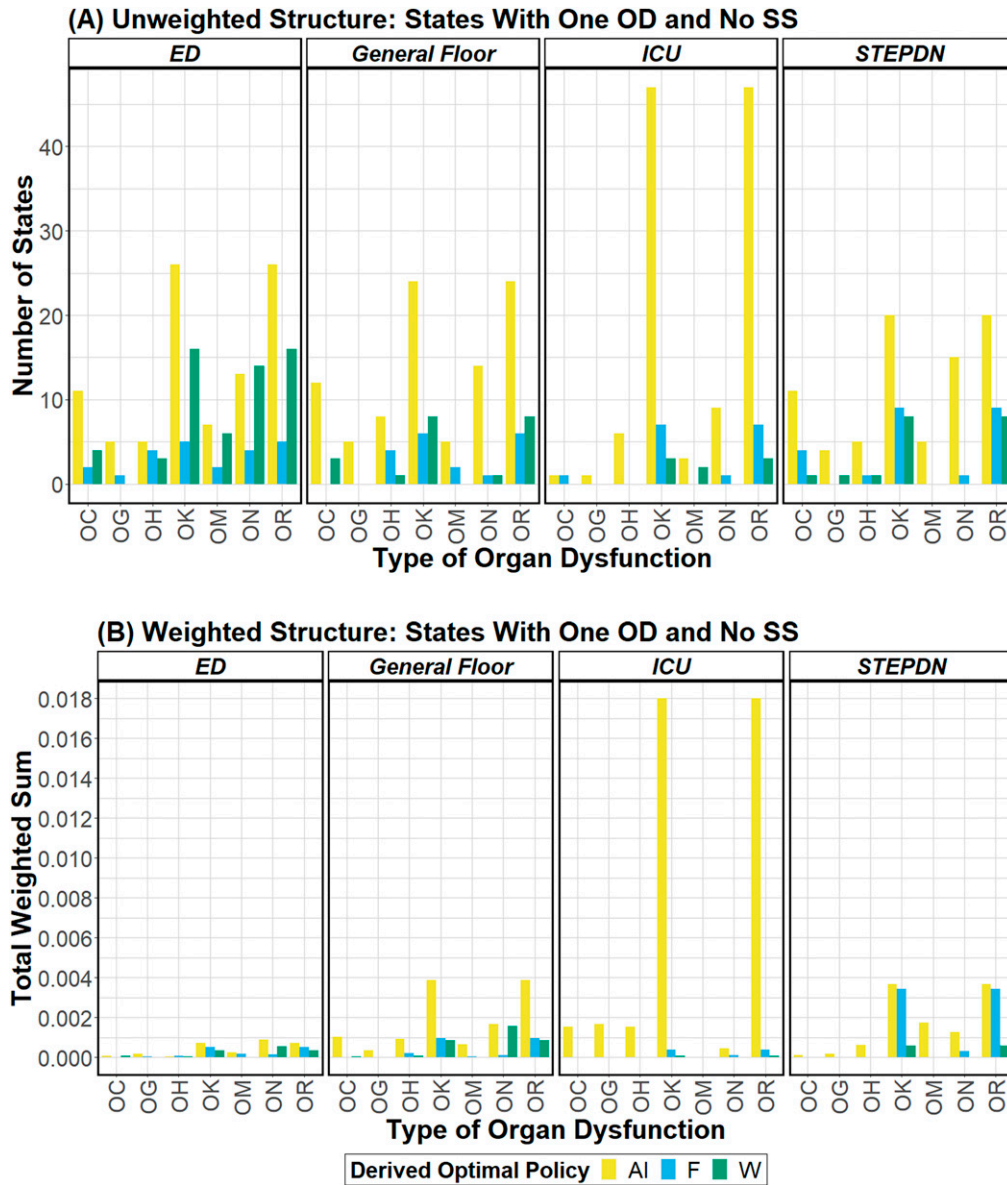


Note. Because SS is present, OC is also present.

5.2.3. States with One Organ System Dysfunction and No Septic Shock. Figure 3 shows the unweighted and weighted distribution of optimal policy action under the baseline cost scenario for states with one organ dysfunction and no SS by location. The weighted and unweighted aggregations of the optimal policy are largely in agreement for this subset of health states.

Once a patient develops a single organ dysfunction in any location, treatment is recommended and the optimal treatment recommendation is primarily anti-infective. The only difference in policy dominance is in the ED for OH. For patients in STEPDN and ICU with or without organ dysfunction, treatment is

Figure 3. (Color online) Unweighted (a) and Weighted (b) Aggregations of States with One Organ Dysfunction and No SS



recommended. In the General Floor and ED, if a patient has at least one organ system dysfunction, anti-infective is recommended. For OK and OR, fluid is a close second for patients in the ED and STEPDN.

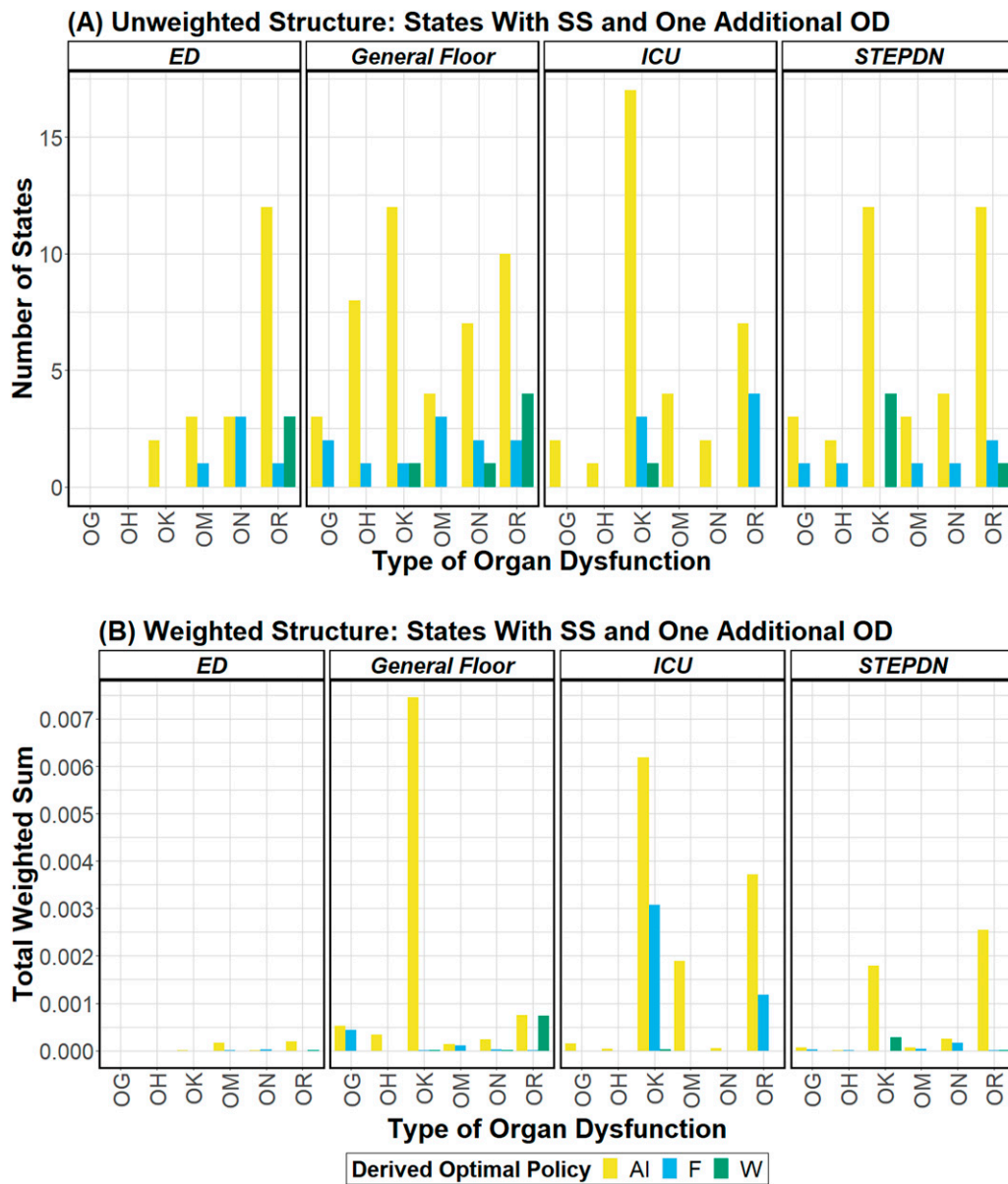
There is one rare instance in which the weighted optimal policy is inconsistent with treatment for less severe health states in the ICU. For OM, the dominant action is wait. Clinical insight suggests that the metabolic organ dysfunction alone does not provide

sufficient insight regarding how the patient’s health will progress. Also, it is important to note that states with only OM are rare, with two unique health states occurring that both transition to another state with probability one. The average time spent in these two states is less than one minute. Because of the small amount of time in the states and certainty regarding where the state will transition next, a trade-off is being made directly in the future expected health cost. This is an example of rare

states impacting the optimal treatment policy through a limited visibility of where these health states can transition to, what health states transition to them, and the time spent in the health state. This rarity of states is largely due to the sample size of data and the complexity of the state space.

5.2.4. States with Septic Shock and One Additional Organ System Dysfunction. Figure 4 shows the unweighted and weighted distribution of optimal policy action under the baseline cost scenario for states with two organ system dysfunctions and SS by location. If a patient has SS, they also have OC, so this

Figure 4. (Color online) Unweighted (a) and Weighted (b) Aggregations of States Septic Shock and One Additional Organ System Dysfunction



Note. Because SS is present, OC is present but not included in the x-axis.

analysis focuses on the second organ system dysfunction. For this set of states, the treatment policy is dominated by anti-infective in the ED, General Floor, ICU, and STEPDN. In the ED, SS is less likely to occur.

5.2.5. Initial Treatment Heuristic Policy. The structure of the optimal policy can be characterized as follows:

1. If the patient has one or more organ dysfunctions or SS, treat
 - a. If the patient only has OC and SS on the General Floor, treat with fluid.
 - b. Otherwise, treat with anti-infective.
2. If the patient has no organ dysfunction and no SS,
 - a. in the ED or General Floor, wait
 - b. in ICU or STEPDN, treat with anti-infective.

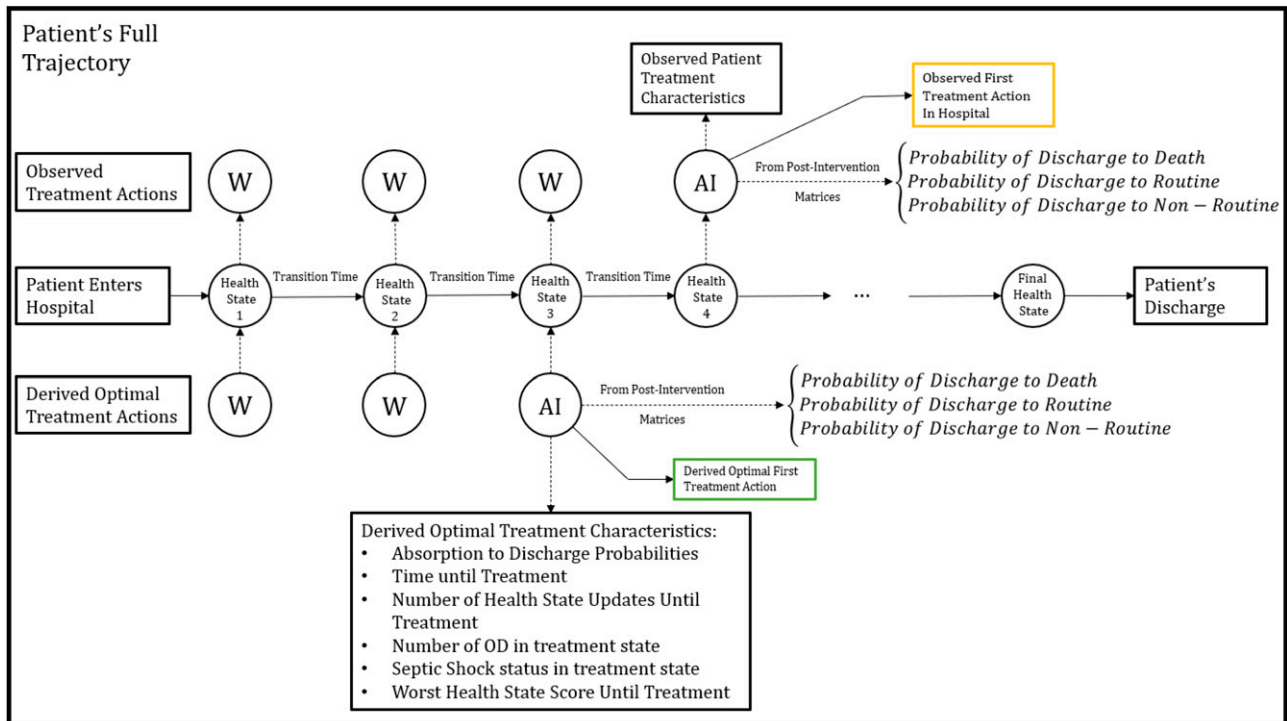
This policy is close to a heuristic policy of always treating a patient with at least one organ dysfunction or SS. However, it is more nuanced by providing a treatment recommendation and considering organ dysfunction type and location.

5.3. Application of Optimal Treatment Policy

In this section, we evaluate the performance of the baseline optimal policy on a sample of patients. As shown in Figure 5, given a patient’s full health state trajectory and the optimal first treatment policy derived from the MDP, when the patient reaches a health state in which the policy recommends treatment, the state and treatment are recorded. With this information, (i) the treatment location, (ii) the absorption probabilities of each discharge disposition, (iii) the time until treatment, (iv) the number of organ dysfunctions at the time of treatment, (v) the septic shock status at the time of treatment, and (vi) the proportion of optimal treatment actions are evaluated. The absorption probabilities are calculated from the postintervention TPMs, as discussed in Section 4.5.

The optimal policy is tested on 30 samples of 5,000 randomly selected patient trajectories from the NDTP. We compare the treatment characteristics of the optimal policy’s first treatment recommendation to the first treatment the patient actually received. Subsequently, we calculate the differences in each treatment characteristic and determine if the differences between treatments are

Figure 5. (Color online) Application of Optimal Policy to a Patient’s Health Trajectory



Downloaded from informs.org by [152.7.224.31] on 16 November 2022, at 04:02. For personal use only, all rights reserved.

Table 10. Comparison of Optimal Treatment Policy and Observed Treatment

Panel A: Mean count by treatment location				
Optimal treatment location	Observed treatment location			
	ED	General floor	STEPDN	ICU
ED	2,912.5 (12.3)	665.1 (9.1)	187.1 (4.6)	94.0 (3.2)
General Floor	199.3 (5.3)	564.6 (9.0)	7.0 (0.9)	9.0 (1.1)
STEPDN	38.7 (1.8)	21.3 (2.1)	87.0 (2.9)	13.6 (1.1)
ICU	36.9 (2.4)	21.1 (2.3)	6.7 (1.0)	136.1 (4.6)
Panel B: Mean difference in absorption to death probability				
ED	-0.003*** (0.0003)	-0.008*** (0.001)	0.016*** (0.002)	-0.105*** (0.005)
General Floor	-0.008*** (0.001)	-0.004*** (0.001)	0.008** (0.008)	-0.097*** (0.012)
STEPDN	-0.021*** (0.006)	-0.004*** (0.002)	-0.008*** (0.002)	-0.135*** (0.007)
ICU	0.010** (0.008)	0.028*** (0.005)	0.071*** (0.019)	-0.046*** (0.003)
Panel C: Mean difference in time until treatment				
ED	-49*** (1.10)	-1,157*** (23.7)	-1,408*** (42.7)	-1,385*** (83.0)
General Floor	1,381*** (51.0)	898*** (44.1)	-1,578*** (674.7)	398 (963.7)
STEPDN	823*** (82.6)	1,109*** (188.3)	-466*** (66.1)	-1,117** (918.5)
ICU	704*** (64.5)	-2,541*** (529.9)	-4,351*** (1,007.2)	-14 (61.5)
Panel D: Mean difference in number of organ dysfunction at time of treatment				
ED	-0.30*** (0.005)	-0.22*** (0.009)	-0.40*** (0.023)	-1.15*** (0.039)
General Floor	0.40*** (0.021)	0.50*** (0.014)	0.03 (0.133)	-0.48*** (0.102)
STEPDN	-0.18*** (0.039)	-0.00 (0.032)	-0.28*** (0.025)	-0.91*** (0.094)
ICU	0.28*** (0.059)	-0.26*** (0.082)	-0.45*** (0.166)	-0.40*** (0.029)
Panel E: Mean difference in proportion with septic shock at time of treatment				
ED	0.017*** (0.001)	0.006** (0.003)	-0.008*** (0.005)	0.020*** (0.009)
General Floor	0.019*** (0.004)	0.010*** (0.003)	0.000 (0.000)	0.084*** (0.032)
STEPDN	0.083*** (0.019)	0.014** (0.010)	-0.007* (0.007)	0.040*** (0.026)
ICU	0.228*** (0.027)	-0.000 (0.017)	-0.022 (0.061)	0.056*** (0.008)
Panel F: Proportion of states that use anti-infection define stars				
ED	0.80 (0.003)	0.79 (0.007)	0.83 (0.011)	0.82 (0.013)
General Floor	0.80 (0.011)	0.84 (0.005)	0.90 (0.048)	0.99 (0.013)
STEPDN	0.87 (0.021)	0.56 (0.038)	0.55 (0.021)	0.59 (0.047)
ICU	0.86 (0.020)	0.90 (0.026)	0.92 (0.048)	0.85 (0.010)

Notes. Each panel compares optimal treatment policy (row) to the treatment the patient received in the hospital (column) by location. Each value has a significance level and 95% confidence interval half width. The statistics are calculated from randomly sampling 5,000 patients from the NDTP for 30 replications. Panel A: The distribution of patients' treatment locations under the optimal treatment policy and their observed treatment. The mean difference in (panel B) absorption to death probability, (panel C) time until treatment, (panel D) number of organ dysfunctions in the treatment state, and (panel E) proportion of patients with septic shock in their treatment state between the optimal first treatment policy and the observed treatment. Panel F: the proportion of states in which the optimal policy recommends anti-infective treatment.

***<0.01; ** <0.05; * <0.1.

significant. Table 10, panels A–F show the results of the six key performance metrics categorized by the treatment location under the optimal first treatment policy and the location of the observed first treatment. The average value across the 30 replications of each key performance metric is shown with the 95% confidence interval half-width and the significance level. Location combinations are defined in coordinate pairs, for example, (ED, GF) is the matrix element where the optimal policy

recommends treatment in the ED and the observed treatment location is the General Floor.

Table 10, panel A shows the average number of patients categorized by their treatment location under the optimal first treatment policy and the observed first treatment. The upper triangular portion of this matrix shows regions in which the optimal policy recommends treating the patient in less or equally severe locations as the observed treatment location.

On average, 93% of patients are in this area with the most in (ED, ED). If a patient receives treatment before entering a more severe location, they might be able to prevent a visit to the location.

Table 10, panel B shows the optimal first treatment policy significantly reduces the average absorption to death probability by 0.007 ($p = 1E-27$) or 2.2% across the population. In the four most common location combinations, (ED, ED), (ED, GF), (GF, ED), and (GF, GF), the optimal policy significantly reduces the absorption to death probabilities. Table 10, panel C shows the optimal policy reduces the time until treatment by 106 minutes ($p = 2E-16$) and 98.2% of the patients are treated within 72 hours on average.

Table 10, panel D shows the average difference in the number of organ system dysfunctions between the optimal first treatment policy and the observed first treatment. The number of organ system dysfunctions at the time of treatment is an indicator of the severity of patient health before treatment. Under the optimal policy, approximately 85% of the patients are treated in less severe health states and the average number of organ system dysfunction is reduced by 0.19 ($p = 9E-35$).

Table 10, panel E shows the average difference in the number of patients with septic shock at the time of treatment. Across the 30 replications of 5,000 patients each, there is only a significant reduction in the number of patients with septic shock at the time of treatment under the optimal policy in the (ED, SD), (ED, ICU), (SD, SD), and (SD, ICU). Table 10, panel F shows that the majority of patients are treated optimally with anti-infective first.

When Table 10, panel A–F are examined together, we observe that the optimal policy treats patients significantly quicker and in less severe health states than their observed treatment. The timing of the treatment is critical as delays in treatment are associated with increased mortality risk (Kumar et al. 2006, Andersson et al. 2019). Treating in less severe health states is also important as the time with organ system dysfunction can impact patient outcomes. The optimal treatment policy performs significantly better for these metrics, and it treats patients in states with lower absorption to death probabilities. For all location coordinates, the optimal policy performs better than the actual treatment in at least one of the key performance indicators.

This analysis highlights the potential trade-offs in the treatment strategies. For example, when the optimal

policy treats in the ED but the observed treatment location is in the STEP DN, the observed treatment had a lower absorption to death probability but the optimal policy treats 23 hours sooner and in a less severe health state. A similar trade-off can be examined for (ED, GF) and (GF, GF). Under the optimal policy, the absorption to death probability is lower; however, the treatment is delayed, and there is an increase in the severity of the health state.

There are some counterintuitive results. For example, a natural question arises regarding why, under the optimal policy, treatment does not occur until the patient gets to the ICU when the patient was actually treated in the ED. Intuitively, an optimal policy should treat before a patient enters a more severe health state. This occurs when health states in the patient trajectory have no known optimal policy. The health state only has a known optimal policy if it is in the filtered states of the DTP. States that are not in this set do not have a known optimal policy. Without a known optimal policy, we are forced to wait until a state with a known optimal policy is reached. Full information about the impact of the treatment is required to solve for an optimal treatment decision; an optimal policy could not be identified under partial information. Potential ways to address this issue are to use the optimal policy for a different location if it is available or to use the heuristic developed in Section 5.2.

5.3.1. Performance of the Optimal Policy on the DTP.

We perform the same analysis for the DTP. Among the four mean difference performance metrics, all are significantly lower under the optimal policy than the actual treatment they received.

1. The mean difference in the absorption to death probability for the DTP is -0.027 ($p = 8.5E-11$).
2. The mean difference in time until treatment is $-2,326$ minutes or 38.7 hours ($p = 2.2E-16$).
3. The mean difference in the number of organ system dysfunctions is -0.61 ($p = 2.2E-16$).
4. The mean difference in the proportion of patients treated with SS is -0.103 ($p = 8.05E-11$).

5.3.2. Performance of Optimal Policy Heuristic.

Additionally, we perform the same analysis for the NDTP using the heuristic defined from the optimal policy structure in Section 5.2. The mean difference in the absorption to death probability for the NDTP is -0.008 ($p = 4E-27$); the mean difference in time until treatment is -38

minutes ($p = 6.4E-9$); the mean difference in the number of organ system dysfunctions is -0.27 ($p = 2.2E-27$), and the mean difference in the proportion of patients treated with SS is -0.003 ($p = 5.2E-7$). The heuristic is defined for all states, and the optimal policy is only defined for the filtered states of the DTP. This can lead to situations in which the treatment is to wait because optimal policy is not known for a state, but the heuristic can recommend a treatment action. Under the optimal policy, treatment recommendations are made 1.13 hours earlier than the heuristic on average, which translates to lower mortality risk (Kumar et al. 2006). The heuristic provides a good solution for the initial treatment of patients if the optimal policy cannot be referenced.

5.4. Sensitivity Analysis on Cost Structure

We test 39 different cost scenarios shown in Table B.1 in Appendix B. We vary the terminal cost, the cost of treatment relative to the terminal cost, and the cost of anti-infective relative fluid. We find that the structure of the optimal policy is similar across cost scenarios. The primary difference between the structures is the optimal type of treatment; the general conditions of states for which treatment is optimal remain the same. We find that for all cost structures with terminal cost less than 10,000, each policy performs significantly better than the observed treatment policy for the absorption to death probability, time until treatment, and health state severity. When the terminal cost becomes large, the policy sacrifices the time until treatment and health state severity metrics for a greater reduction in the absorption to death probability.

5.5. Limitations

There are some limitations to our analysis. Given the nature of sepsis, patients with cardiovascular conditions are more difficult to treat, more likely to have worse health conditions, and more likely to suffer worse outcomes. To ensure that the results of the optimal policies are not biased by the underlying health conditions of the patient population, a CTMC for the subset of patients who received delayed treatment and had either congestive heart failure or valvular disease was developed. The state space of this subpopulation's CTMC was compared with the DTP states grouped by each optimal action. We found that

there were no significant differences in the states in the cardiovascular comorbidity DTP state space and entire DTP state space.

One key challenge in a model like this is finding a comparison group without treatment to understand the counterfactual. Untreated patients with infection or sepsis are extremely rare among hospitalized patients. As discussed in the methods, our solution was to identify a group of patients with delayed treatment and use their clinical course as a proxy for an untreated or natural history group. However, a limitation is that this cohort of patients significantly differs from the NDTP in age and gender, which could bias the comparison. Additionally, there could be unobserved differences between the groups that may bias the results. This is a somewhat unavoidable situation, given that it would be unethical to leave patients with infection or sepsis knowingly untreated.

The size of the DTP and NDTP data sets is another limiting factor in this work that may have reduced the number of possible states represented and resulted in data sparseness. The size of the state space limits the number of states for which an optimal first treatment action can be found. Because it is unethical to delay treatment, alternative approaches (such as combining data from multiple hospital systems) should be explored to more fully characterize the optimal first treatment action for the entire possible state space.

The data sparseness, driven by the state space complexity and the data set size, creates instances of state transitions that occur a small number of times across all patient visits. This phenomenon results in some average transition times calculated with small sample sizes; approximately 20.4% of the states are visited 10 or more times. This may cause the average time until a transition to be uncertain. These transitions are not removed as removal would severely reduce the connectivity of the Markovian network. In general, hypothesis testing of the proportion of various attributes reveals that the proportion of low transition counts is greatest in the ED and originates in states with more missingness than higher transition count instances. These findings are intuitive as patients never reenter the ED and never lose information once gained. In future work, we will quantify the value of the stochastic solution gained by addressing the uncertainty in the transition rates.

One limitation of our approach is that our model does not have an option to recommend that both treatments be given at the same time. This option may be the optimal scenario in many clinical situations in which this is possible. However, as we mentioned in the introduction, it is not always possible because of limited nursing time and competing priorities with other patients.

The final limitation is the estimation of the costs for each state. It is difficult to quantify the costs of our parameters from the literature, so we rely on expert opinion. The cost assumptions are relative. Additionally, we perform extensive sensitivity analysis to explore the effect of different cost assumptions on the optimal policy. For future work, this framework could be applied to the SOFA for comparison. In addition, survival curves of the organ system dysfunctions could be used to determine the relative value of cost rates of the organ system dysfunction and septic shock.

6. Conclusion and Future Work

Sepsis is considered a medical emergency, and delays in initial treatment are associated with increased morbidity and mortality. Because there is no gold standard for diagnosis and treatment of sepsis, we present a generalizable framework to optimally design first treatment policies for patients with suspected infection who are at risk for sepsis. This framework addresses the uncertainty of timing and type of first treatment and provides an optimal first treatment policy that minimizes patients' expected morbidity and mortality. This framework models the stochastic progression of patient health from hospital admission using a CTMC constructed from a population of patients that received delayed treatment upon entering the hospital or during their stay. This CTMC is used to develop a stopping problem-based MDP to identify an optimal first treatment policy.

The optimal first treatment policy acts as an alert system for treatment that also provides a treatment recommendation. By examining the structural characteristics of the optimal policy, we find the dominant patterns for treatment in health states that are common amongst patients entering the hospital to develop a heuristic rule for initial treatment. This heuristic focuses on location, the number of organ system dysfunctions, and treatment type to identify rules for initial treatment. We find that having at least

one organ dysfunction triggers treatment and that anti-infective is generally the best treatment option. We find when patients have no additional organ system dysfunction and septic shock that fluid is the best first treatment and that when patients are in the ICU or STEP DN, even without organ dysfunction, it is best to treat them. This heuristic structure of the optimal policy provides easily interpretable guidelines for the optimal first treatment. In addition, the treatment characteristics of the heuristic are robust to changes in cost structure.

When the optimal policy is compared with observed treatment, we find that the optimal policy significantly reduces the absorption probability to death, time until treatment, and the number of organ dysfunctions at the time of treatment. This is critical as delays in treatment lead to increased mortality risk, and untreated organ dysfunction can have lifelong health implications. When treatments are disaggregated by location, we also observe that the optimal treatment policy treats most patients before they enter more severe locations. This is critical for preventing patients from entering the ICU, which is often a constrained resource. These metrics suggest that implementable first treatment policies can be identified that outperform observed treatment.

We use data to formulate this problem from a metropolitan regional Level 1 trauma center. The framework we develop to identify the improved first treatment policy is generalizable to other hospital systems. This system could act as an alert system for treatment that could be integrated into the EHR.

This framework can be personalized to the attributes of the patients. In future work, this framework will be applied to populations with comorbidity. This would facilitate personalized treatment for patients who enter the hospital with suspected infection, are at risk for sepsis, and have underlying health conditions. Another area for future work is developing a finite horizon model to find the first treatment that considers time-dependent thresholds for action. CMS time limits identify bounds within which intervention must occur after health state criteria are met. It would be possible to establish time limits informed by the infinite horizon policy and compare with the current CMS recommendations.

Appendix A. Policy Structure For States with Greater Than Three Organ Dysfunction

Figures A.1 and A.2 show states aggregated by the number of organ system dysfunctions, location, and by optimal treatment action. In general, there is an agreement between unweighted and weighted dominance. There are three location/number of organ dysfunctions combinations that have differing dominant actions by weighting.

Under the weighted aggregation, states in the ED with three and four organ dysfunctions have dominant actions of waiting and fluid, respectively. Note that the weights in the ED are very small relative to the count of states. Because the states do not enter the ED once they have left, these states are transient with respect to changing location

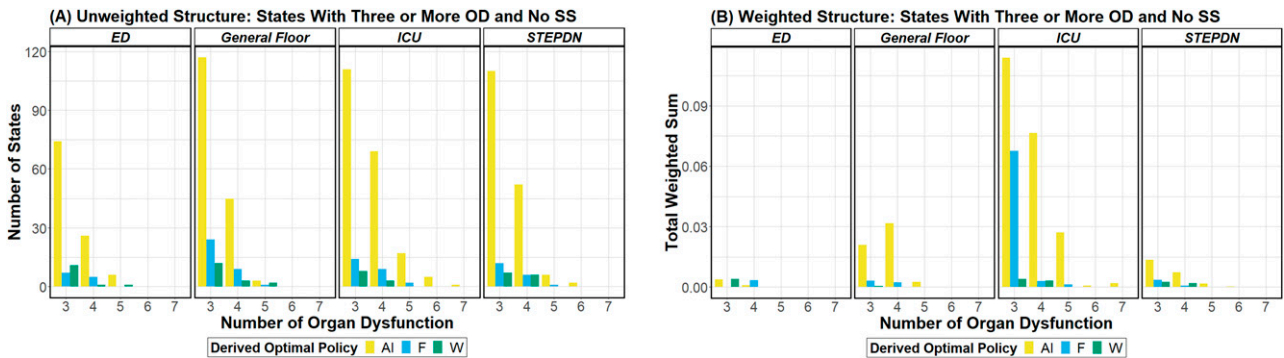
away from the ED, creating volatility compared with the unweighted.

Under both the weighted and unweighted aggregation, the state with seven organ system dysfunctions and septic shock in the ICU optimally waits. This state is rare in the DTP, occurring for one patient and transitioning to death with probability one. Waiting is due to an end of horizon effect.

Appendix B. Sensitivity Analysis on Cost Structure

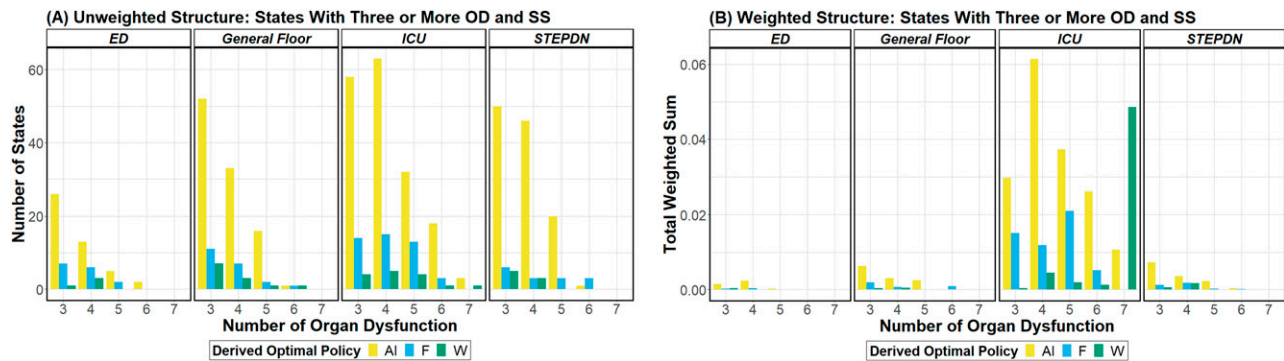
Table B.1 shows 39 cost scenarios used for sensitivity analysis on the cost structure. For all scenarios, location costs and the routine discharge cost are held constant. We test three sets of terminal cost values; for each terminal cost

Figure A.1. (Color online) Unweighted and Weighted Aggregations of States with Greater than Three Organ Dysfunction and No Septic Shock



Note. States are separated by location and color-coded by optimal action.

Figure A.2. (Color online) Unweighted and Weighted Aggregations of States with Greater than Three Organ Dysfunction and Septic Shock



Note. States are separated by location and color-coded by optimal action.

Table B.1 Cost Scenarios for Sensitivity Analysis

Discharge Costs	Scenario Number	Cost of Fluid	Cost of Anti-Infective
Routine: 0; Non-routine: 1,000; Death: 2,000 Routine: 0; Non-routine: 250; Death: 500	Baseline	200	200
	1	0	0
	2	25	12.5
	3	12.5	25
	4	25	25
	5	50	25
	6	25	50
	7	50	50
	8	75	37.5
	9	37.5	75
	10	75	75
	11	100	50
	12	50	100
Routine: 0; Non-routine: 1,000; Death: 2,000	13	100	100
	14	0	0
	15	100	50
	16	50	100
	17	100	100
	18	200	100
	19	100	200
	20	200	200
	21	300	150
	22	150	300
	23	300	300
Routine: 0; Non-routine: 5,000; Death: 10,000	24	400	200
	25	200	400
	26	400	400
	27	0	0
	28	500	250
	29	250	500
	30	500	500
	31	1,000	500
	32	500	1,000
	33	1,000	1,000
	34	1,500	750
	35	750	1,500
	36	1,500	1,500
	37	2,000	1,000
	38	1,000	2,000
	39	2,000	1,000

Notes. The location are constant for all scenarios. They are 0, 1, 3, and 5 for the General Floor, STEPDN, ED, and ICU, respectively.

value, we test 13 action cost combinations. The action costs ranged from 0%–20% of the terminal cost value. For each 5% increase in both fluid and anti-infective action cost, we test combinations in which fluid or anti-infective is 50% of the alternative. For each cost scenario, the MDP is formulated with the corresponding cost values, and solved, generating 39 optimal treatment policies.

Because the costs relativistically weigh different attributes of patient health, the cost scenarios take on interpretable

meaning. When the terminal cost of death is lower, the cost rate associated with each health state has a larger relative weight; this emphasizes avoiding detrimental health states. Similarly, when the cost of death is greater, the terminal disposition has a greater weight compared with the cost rates and location cost of a health state; this emphasizes the discharge disposition of the patient above the health states they visit. By testing action costs that increase relative to the terminal cost, the value of treatment is observed. In

scenarios in which the cost of treatment is large, the sensitivity analysis shows if the high cost of treatment is worth the benefits of the future expected health compared with the future of waiting. We explore if there is a threshold cost for initiating treatment. By testing action costs that vary relative to each other, we explore if there is a threshold for which a treatment is preferred and if there are conditions for which a treatment is always optimal. Establishing these thresholds informs treatment when treatment risk is uncertain.

We examine the policies generated by the different cost scenarios from the same perspective of policy dominance used in Section 5.2. We use the weighted aggregation for this analysis, as it accounts for the importance of the state in the DTP Markov chain. The objective is to identify where the policies are in agreement and where there are differences. The following sections compare the policies for the different sets of states as discussed in Section 5.2.

The figures in this section that describe which action is dominant are presented through a series of grouped scatter plots. The X and Y axes on each individual scatter plot represent the cost of fluid treatment and the cost of anti-infective treatment, respectively. The scatter plots are grouped by their organ dysfunction status, terminal cost, and location. Anti-infective or fluid is the dominant action if the sum of states' weights is greater than the other treatment option and waiting. Waiting is the dominant action if the sum of states' weights is greater than the sum of anti-infective and fluid. Otherwise, the sum of anti-infective and fluid is greater than waiting and an alert for treatment is dominant, which is referred to as action. Within the action region, the optimal policy should be used to identify the type of treatment.

B.1. States With No Organ System Dysfunction and No Septic Shock

States with no organ system dysfunction and no septic shock are the healthiest states in the state space. Figure B.1 shows the dominant action for each of the 39 optimal policies for this set of states. The optimal actions are equivalent and correspond to the dominance characteristics discussed in Section 5.2. Specifically, if the patient is on the General Floor or in the ED with no organ dysfunction, wait; if the patient is in the ICU or STEPDN with no organ dysfunction, treat with anti-infective.

B.2. States With Cardiovascular Organ System Dysfunction and Septic Shock

Figure B.2 shows the dominant action for the 39 cost scenarios for this set of states. In this figure, as would be expected, we observe that a treatment action is always dominant. This supports the results in Figure 2 and the heuristic in Section 5.2. Figure 2 shows that optimal treatment varies primarily by location and terminal cost. In the

STEPDN location, anti-infective is the dominant optimal treatment for all cost scenarios. For the ED and ICU, we observe that as the terminal cost increases and fluid becomes the preferred treatment action. On the General Floor, we observe a similar pattern but with anti-infective and fluid reversed.

B.3 States With One Organ System Dysfunction and No Septic Shock

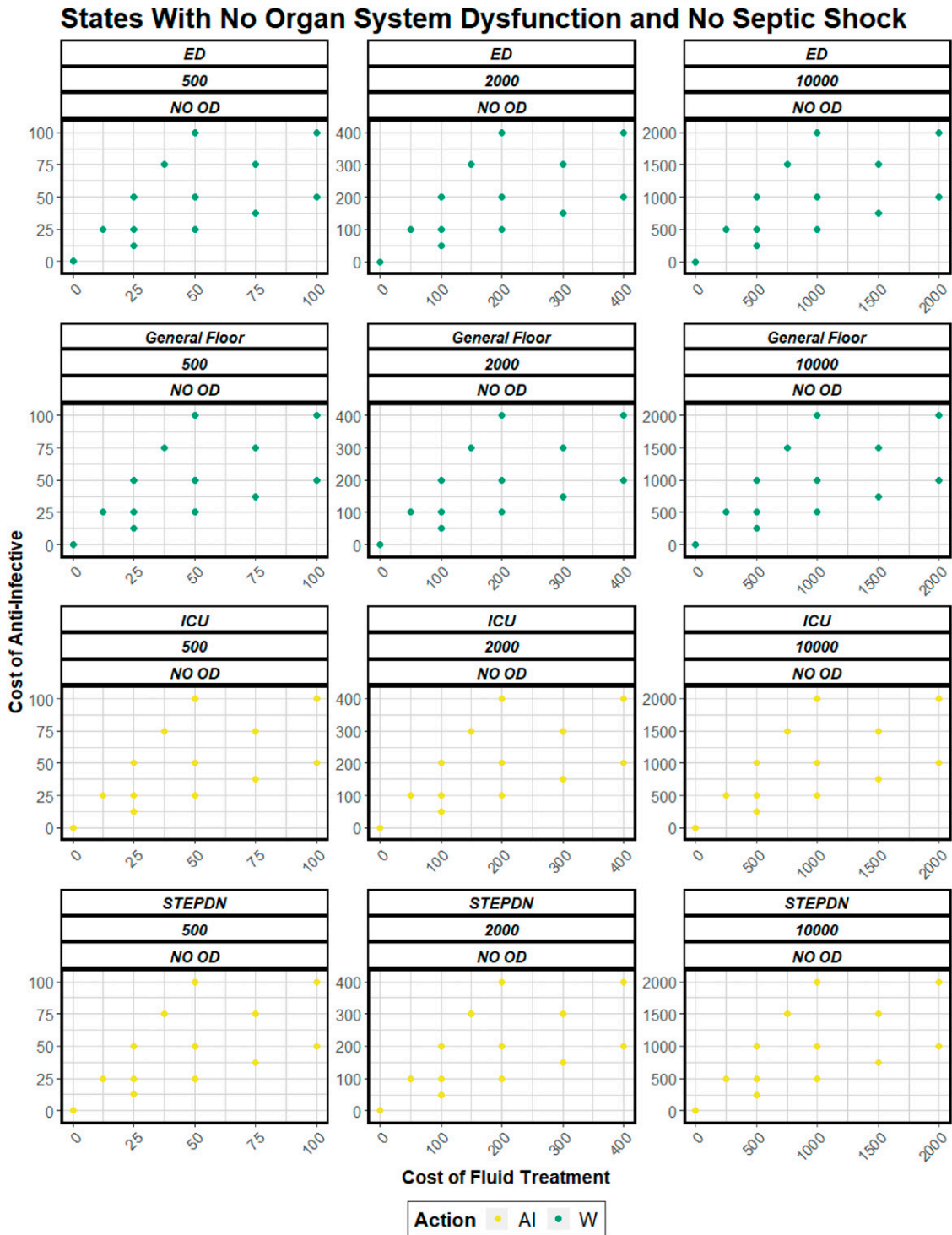
We extend the graph in the previous section to examine multiple types of organ system dysfunction simultaneously. Figure B.3 shows the dominant optimal policy for states with one organ system dysfunction and no septic shock. The columns of the matrix of graphs represent the type of organ system dysfunction present for the set of states. The rows correspond to the combination of terminal cost of death and location. For each subgraph, the cost of anti-infective and fluid varies according to the cost scenarios corresponding to the terminal costs. For example, the graph in the top left corner of the matrix represents a set of states with OC in the ED where the cost of death is 500.

First, we examine how the policy structure changes in the ED. For terminal costs of 500 and 2,000, for all organ system dysfunction except OH, anti-infective is the dominant action. There is no organ system dysfunction with a single dominant action across all 39 cost scenarios. Once the cost of death is 10,000, and the discharge disposition is weighted more, the policy changes. Fluid and anti-infective become more competitive for the dominant as regions in which fluid is less costly are optimally treated by fluid for OG, OH, OM, OK, and OR. Except for the ON, the heuristic developed in Section 5.2 holds, that is, treatment should be given if the patient has an organ system dysfunction.

On the General Floor, anti-infective is the dominant action for OC, OG, OH, and OM organ system dysfunction for all 39 cost scenarios. We observe a similar trend as in the ED for the other three organ system dysfunctions for which anti-infective is largely dominant over all cost scenarios with terminal costs of 500 and 2,000. For ON, fluid is favored for some cost scenarios. Under the terminal cost of 10,000, when the discharge disposition is primarily weighted in the minimization problem, waiting becomes a more competitive action for neurological, renal, and respiratory organ system dysfunction. For OK and OR, if the treatment cost is a larger proportion of the terminal cost, the dominant action becomes waiting. This implies if the cost of treatment is expected to be high, treatment might not be warranted. In general, these policies support the heuristic developed in Section 5.2. For the ON, OK, and OR treatment, the potential cost of treatment should be carefully considered prior to treatment.

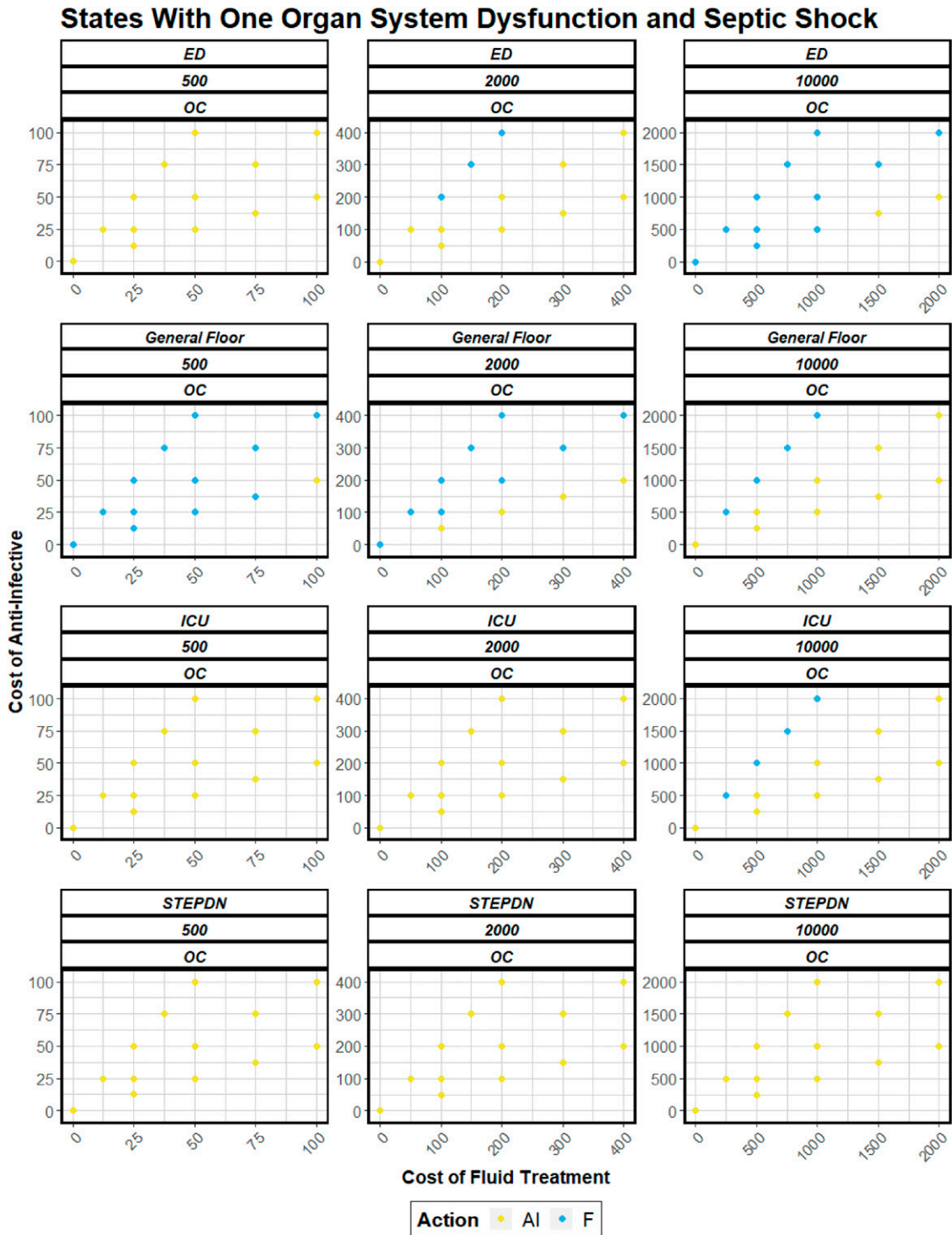
For the ICU, we observe that OG, OM, and neurological organ system dysfunction (ON) have a constant dominant optimal action across all 39 cost scenarios. We again observe that anti-infective is dominant for all organ system

Figure B.1. (Color online) Dominant Action Under the Weighted Aggregation of States with No Organ Dysfunction and No Septic Shock



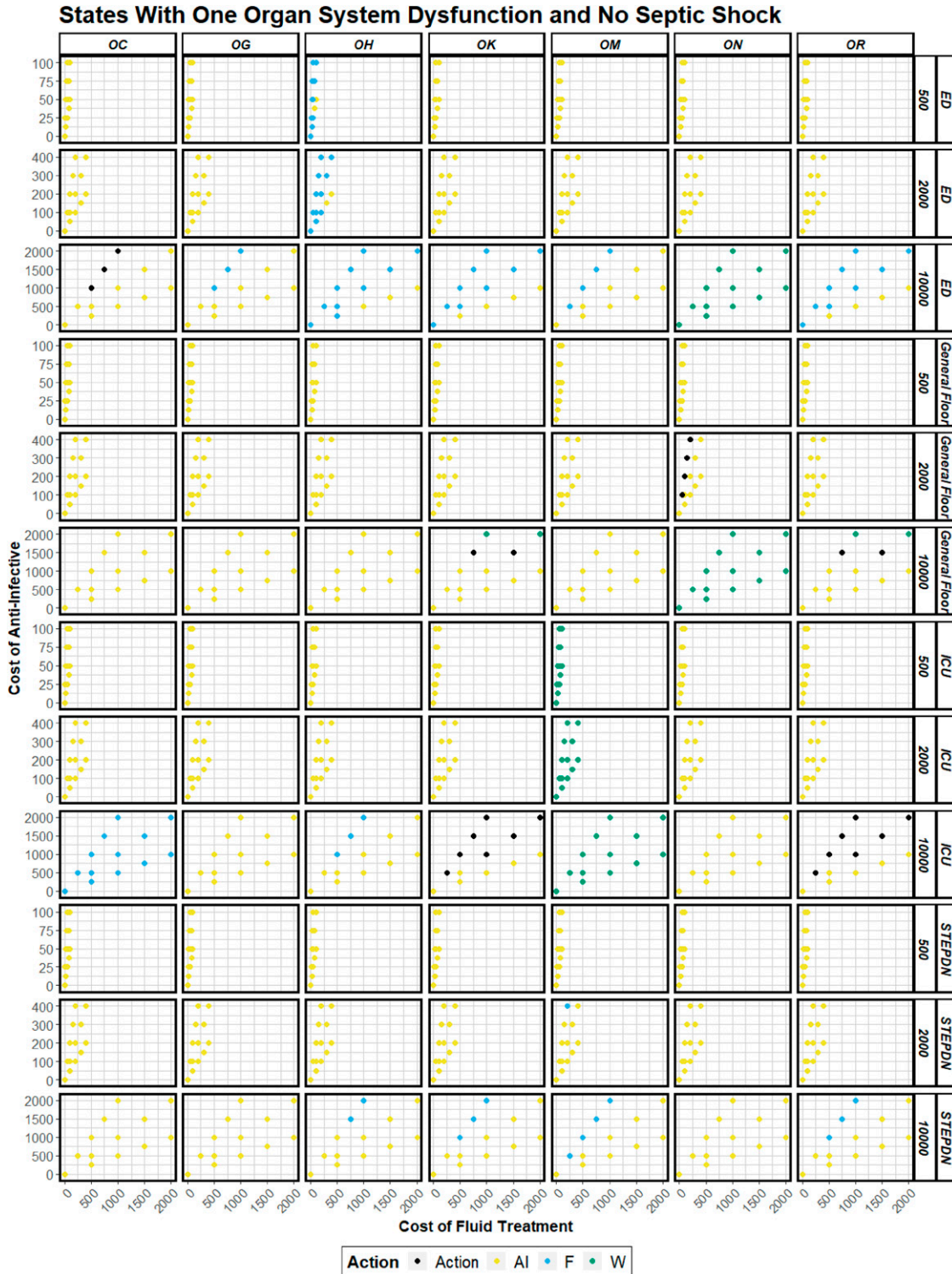
Notes. Each subgraph shows the 13 cost scenarios for each terminal cost parameter. Subgraphs are separated by terminal cost parameters and location.

Figure B.2. (Color online) Dominant Action Under the Weighted Aggregation of States with One Organ Dysfunction and Septic Shock



Notes. Each subgraph shows the 13 cost scenarios for each terminal cost parameter. Subgraphs are separated by terminal cost parameters and location. Because septic shock is present, cardiovascular organ dysfunction is the only possible single organ dysfunction.

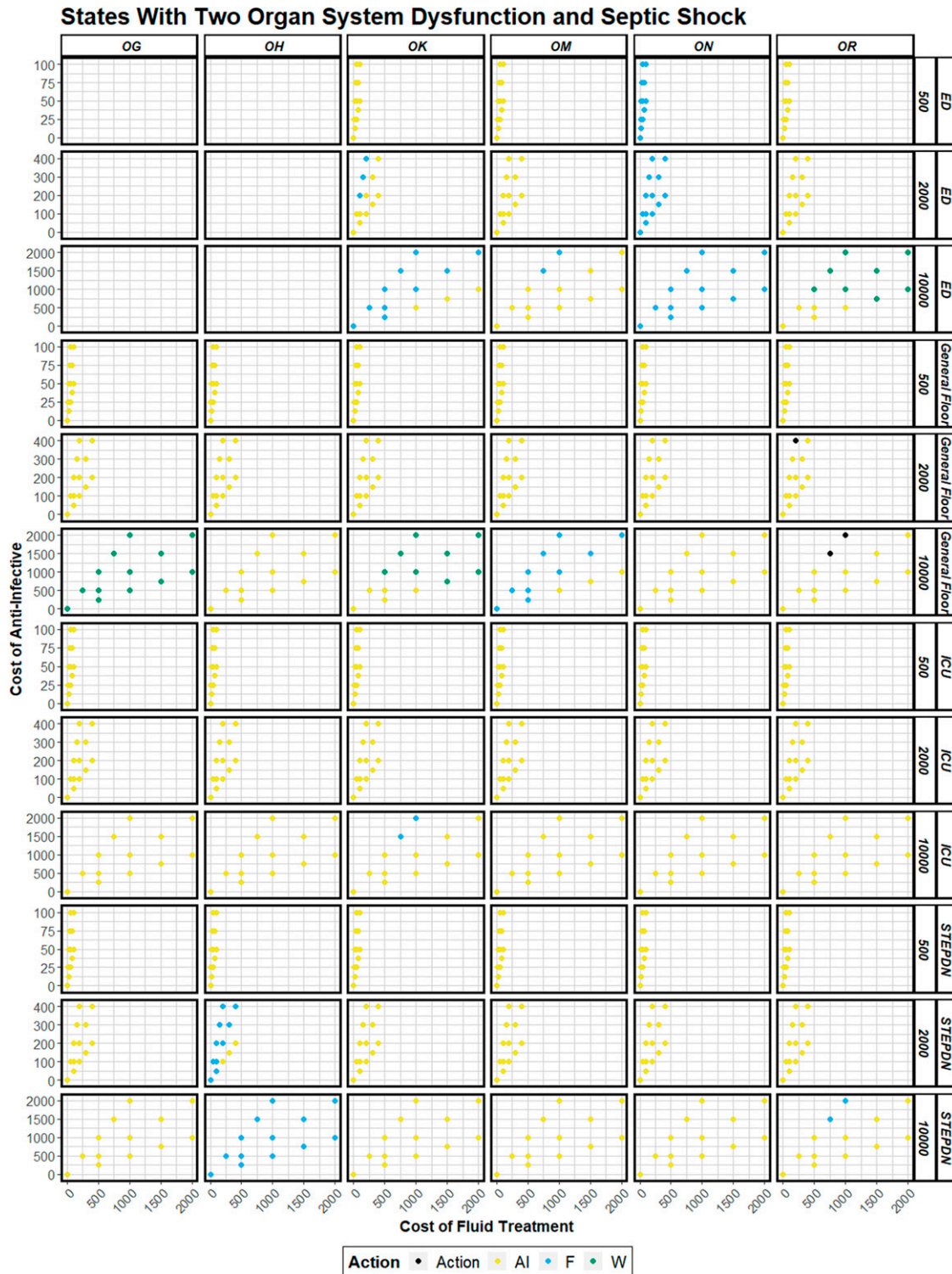
Figure B.3. (Color online) Dominant Action Under the Weighted Aggregation of States with One Organ Dysfunction and No Septic Shock



Notes. Each subgraph shows the 13 cost scenarios for each terminal cost parameter. Subgraphs are separated by terminal cost parameters and location.

Downloaded from informs.org by [152.7.224.31] on 16 November 2022, at 04:02 . For personal use only, all rights reserved.

Figure B.4. (Color online) Dominant Action Under the Weighted Aggregation of States with Two Organ Dysfunction and Septic Shock



Notes. Each subgraph shows the 13 cost scenarios for each terminal cost parameter. Subgraphs are separated by terminal cost parameters and location. Because septic shock is present, cardiovascular is present in all states but not listed in the organ dysfunction columns.

Downloaded from informs.org by [152.7.224.31] on 16 November 2022, at 04:02 . For personal use only, all rights reserved.

dysfunction at terminal costs of 500 and 2,000, except for OM. As discussed in Section 5.2, the ICU with OM is a rare state. We observe that fluid becomes the recommended optimal treatment when the terminal cost is 10,000. This supports the heuristic developed in Section 5.2 that one organ system dysfunction in the ICU warrants treatment. As the terminal cost increases, the discharge disposition is more heavily weighted. Because patients are at greater risk for death while in the ICU, the optimal policy with the highest terminal cost should be used. Fluid should be carefully considered for optimal treatment.

Finally, in the STEP DN location, we observe that anti-infective is the dominant action for almost all cost scenarios and organ dysfunctions. For OC, OG, and ON, the policy structure is the same for all terminal costs indicating that anti-infective is always the dominant action. We observe that fluid becomes the recommended optimal treatment for OH, OM, OK, and OR under the 10,000 terminal cost when the cost of fluid is less than the cost of anti-infective. In general, the optimal policies support the heuristic developed in Section 5.2.

B.4. States With Two Organ System Dysfunction and Septic Shock

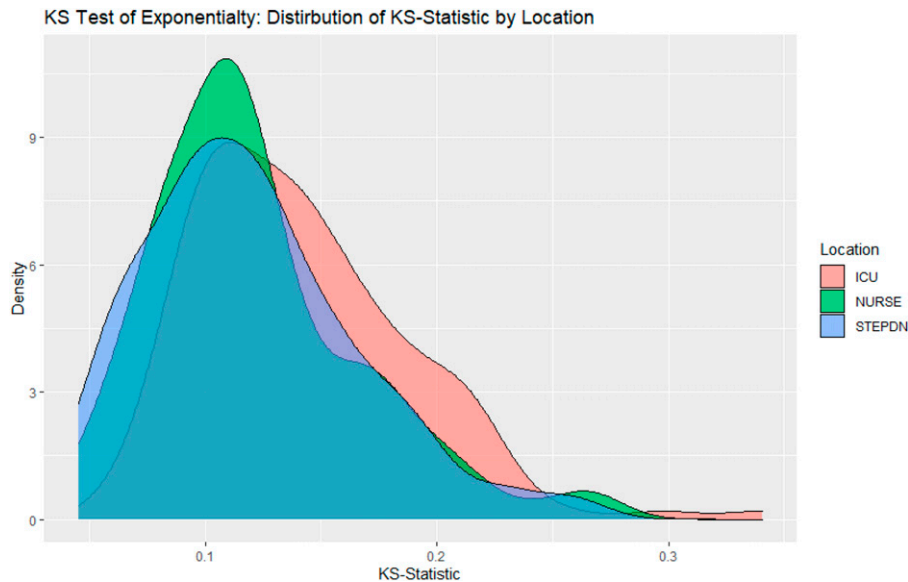
For states with two organ system dysfunctions, Figure B.4 shows the effect of the cost structure on the optimal

policy similar to Figure B.3. For all cost/location/organ system dysfunction combinations except for the combinations of (10,000/ED/OR), (10,000/General Floor/OG), and (10,000/General Floor/OK), all cost scenarios support the heuristic developed in Section 5.2, once a patient has at least one organ system dysfunction and septic shock treatment needs to be administered; in most cases, treatment with anti-infective is the optimal first treatment. For the combinations listed, waiting becomes optimal at the terminal cost of 10,000.

B.5. Distribution of the Sojourn Time

We test the appropriateness of the assumption that the sojourn time follows an exponential distribution using the Kolmogorov-Smirnov (KS) hypothesis test for exponentiality for states that are visited a sufficiently large ($n \geq 30$) number of times in the DTP. Figure B.5 shows a histogram of the KS-statistics for these states. Across all locations, 52% of states' time-until-transition distribution is statistically similar to an exponential distribution at alpha of 0.05. By location, the ICU had the fewest states with a time-until-transition distribution that is similar to an exponential distribution. This may be related to the censoring of the ICU data.

Figure 12. (Color online) Distribution of the KS-Statistics of Exponential Goodness of Fit Tests on the Time to Leave a State



Note. States are shown by location and must have $n \geq 30$ instances of the state being visited in EHRs.

References

- Abraham EP (2015) The antibiotics resistance crisis. *Comprehensive Biochemistry* 11(4):181–224.
- Alberti C, Brun-Buisson C, Chevret S, Antonelli M, Goodman SV, Martin C, Moreno R, et al. (2005) Systemic inflammatory response and progression to severe sepsis in critically ill infected patients. *Amer. J. Respiratory Critical Care Medicine* 171(5):461–468.
- Andersson M, Östholm-Balkhed Å, Fredrikson M, Holmbom M, Hällgren A, Berg S, Hanberger H (2019) Delay of appropriate antibiotic treatment is associated with high mortality in patients with community-onset sepsis in a Swedish setting. *Eur. J. Clinical Microbiology Infectious Diseases* 38(7):1223–1234.
- Annane D, Bellissant E, Bollaert PE, Briegel J, Keh D, Kupfer Y (2004) Corticosteroids for severe sepsis and septic shock: A systematic review and meta-analysis. *BMJ* 329(7464):480.
- Ayvaci M, Mobini Z, Özer Z (2021) To catch a killer: A data-driven personalized and compliance-aware sepsis alert system. Working paper, The University of Texas, Dallas, TX.
- Bauer M, Gerlach H, Vogelmann T, Preissing F, Stiefel J, Adam D (2020) Mortality in sepsis and septic shock in Europe, North America and Australia between 2009 and 2019: Results from a systematic review and meta-analysis. *Critical Care* 24(1):239.
- Bone RC, Fisher CJ Jr, Clemmer TP, Slotman GJ, Metz CA, Balk RA (1989) Sepsis syndrome: A valid clinical entity. *Critical Care Medicine* 17(5):389–393.
- Buchman TG, Simpson SQ, Sciarretta KL, Finne KP, Sowers N, Collier M, Chavan S, et al. (2020) Sepsis among Medicare beneficiaries: 1. The burdens of sepsis, 2012–2018. *Critical Care Medicine* 48(3):276–288.
- Capan M, Ivy JS, Wilson JR, Huddleston JM (2017) A stochastic model of acute-care decisions based on patient and provider heterogeneity. *Health Care Management Sci.* 20(2):187–206.
- CDC (2019) Centers for Disease Control and Prevention. Accessed July 22, 2019, <https://www.cdc.gov/sepsis/datareports/index.html>.
- CDC (2021) How antibiotic resistance happens. Accessed March 25, 2022, <https://www.cdc.gov/drugresistance/about/how-resistance-happens.html>.
- Chen KF, Liu SH, Li CH, Wu CC, Chaou CH, Tzeng IS, Hsieh YH, et al. (2017) Development and validation of a parsimonious and pragmatic CHARM score to predict mortality in patients with suspected sepsis. *Amer. J. Emergency Medicine* 35(4):640–646.
- Ghassemi M, Wu M, Hughes MC, Szolovits P, Doshi-Velez F (2017) Predicting intervention onset in the ICU with switching state space models. *AMIA Joint Summits Translational Sci. Proc.* 2017: 82–91.
- Ghosh S, Li J, Cao L, Ramamohanarao K (2017) Septic shock prediction for ICU patients via coupled HMM walking on sequential contrast patterns. *J. Biomedical Inform.* 66:19–31.
- Girard TD, Opal SM, Ely EW (2005) Insights into severe sepsis in older patients: From epidemiology to evidence-based management. *Clinical Infectious Diseases* 40(5):719–727.
- Gupta A, Liu T, Crick C (2020) Utilizing time series data embedded in electronic health records to develop continuous mortality risk prediction models using hidden Markov models: A sepsis case study. *Statist. Methods Medical Res.* 29(11):3409–3423.
- Hotchkiss RS, Moldawer LL, Opal SM, Reinhart K, Turnbull IR, Vincent JL (2016) Sepsis and septic shock. *Nature Rev. Disease Primers* 2:16045.
- Huang DT, Clermont G, Dremsizov TT, Angus DC, ProCESS Investigators (2007) Implementation of early goal-directed therapy for severe sepsis and septic shock: A decision analysis. *Critical Care Medicine* 35(9):2090–2100.
- Johnson AE, Pollard TJ, Shen L, Lehman LW, Feng M, Ghassemi M, Moody B, Szolovits P, Celi LA, Mark RG (2016) MIMIC-III, a freely accessible critical care database. *Sci. Data* 3:160035.
- Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA (2018) The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nature Medicine* 24(11):1716–1720.
- Kumar A, Roberts D, Wood KE, Light B, Parrillo JE, Sharma S, Suppes R, et al. (2006) Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Critical Care Medicine* 34(6): 1589–1596.
- Levy MM, Macias WL, Vincent JL, Russell JA, Silva E, Trzaskoma B, Williams MD (2005) Early changes in organ function predict eventual survival in severe sepsis. *Critical Care Medicine* 33(10): 2194–2201.
- Magni P, Quaglini S, Marchetti M, Barosi G (2000) Deciding when to intervene: A Markov decision process approach. *Internat. J. Medical Inform.* 60(3):237–253.
- Marshall JC, Deitch E, Moldawer LL, Opal S, Redl H, van der Poll T (2005) Preclinical models of shock and sepsis: What can they tell us? *Shock* 24(Suppl 1):1–6.
- Martin GS, Mannino DM, Moss M (2006) The effect of age on the development and outcome of adult sepsis. *Critical Care Medicine* 34(1):15–21.
- NIH (2019) Sepsis. Accessed July 22, 2019, https://www.nigms.nih.gov/education/pages/factsheet_sepsis.aspx.
- Papathanassoglou E, Middleton N, Benbenishty J, Williams G, Christofi MD, Hegadoren K (2017) Systematic review of gender-dependent outcomes in sepsis. *Nursing Critical Care* 22(5):284–292.
- Parente JD, Chase JG, Moeller K, Shaw GM (2021) High inter-patient variability in sepsis evolution: A hidden Markov model analysis. *Comput. Methods Programs Biomedicine* 201:105956.
- Paxton C, Niculescu-Mizil A, Saria S (2013) Developing predictive models using electronic medical records: Challenges and pitfalls. *AMIA Annual Symp. Proc.* 2013:1109–1115.
- Puterman ML (2014) *Markov Decision Processes: Discrete Stochastic Dynamic Programming* (John Wiley & Sons, Hoboken, NJ).
- Raghu A, Komorowski M, Celi LA, Szolovits P, Ghassemi M (2017) Continuous state-space models for optimal sepsis treatment: A deep reinforcement learning approach. Doshi-Velez F, Fackler J, Kale D, Ranganath R, Wallace B, Wiens J, eds. *Proc. 2nd Machine Learn. Healthcare Conf., Proceedings of Machine Learning Research*, vol. 68 (PMLR, Cambridge, MA), 147–163.
- Rangel-Frausto MS, Pittet D, Hwang T, Woolson RF, Wenzel RP (1998) The dynamics of disease progression in sepsis: Markov modeling describing the natural history and the likely impact of effective antiseptic agents. *Clinical Infectious Diseases* 27(1):185–190.
- Rhodes A, Evans LE, Alhazzani W, Levy MM, Antonelli M, Ferrer R, Kumar A, et al. (2017) Surviving sepsis campaign: International guidelines for management of sepsis and septic shock: 2016. *Critical Care Medicine* 45(3):486–552.
- Ribas VJ, Vellido A, Ruiz-Rodríguez JC, Rello J (2012) Severe sepsis mortality prediction with logistic regression over latent factors. *Expert Systems Appl.* 39(2):1937–1943.

- Rubulotta F, Marshall JC, Ramsay G, Nelson D, Levy M, Williams M (2009) Predisposition, insult/infection, response, and organ dysfunction: A new model for staging severe sepsis. *Critical Care Medicine* 37(4):1329–1335.
- Saka G, Kreke JE, Schaefer AJ, Chang CCH, Roberts MS, Angus DC; GenIMS Investigators (2007) Use of dynamic microsimulation to predict disease progression in patients with pneumonia-related sepsis. *Critical Care* 11(3):R65.
- Schoenberg MH, Weiss M, Radermacher P (1998) Outcome of patients with sepsis and septic shock after ICU treatment. *Langenbecks Arch. Surgery* 383(1):44–48.
- Shankar-Hari M, Phillips GS, Levy ML, Seymour CW, Liu VX, Deutschman CS, Angus DC, Rubenfeld GD, Singer M, Sepsis Definitions Task Force (2016) Developing a new definition and assessing new clinical criteria for septic shock: For the third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* 315(8):775–787.
- Shapiro NI, Howell MD, Talmor D, Donnino M, Ngo L, Bates DW (2007) Mortality in emergency department sepsis (MEDS) score predicts 1-year mortality. *Critical Care Medicine* 35(1):192–198.
- Shechter SM, Bailey MD, Schaefer AJ, Roberts MS (2008) The optimal time to initiate HIV therapy under ordered health states. *Oper. Res.* 56(1):20–33.
- Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, Bellomo R, et al. (2016) The third international consensus definitions for sepsis and septic shock (sepsis-3). *JAMA* 315(8):801–810.
- Sonnenberg FA, Beck JR (1993) Markov models in medical decision making: A practical guide. *Medical Decision Making* 13(4):322–338.
- SSC (2016) SSC guidelines. Accessed July 22, 2019, <http://www.survivingsepsis.org/Guidelines/Pages/default.aspx>.
- Tsoukalas A, Albertson T, Tagkopoulou I (2015) From data to optimal decision making: A data-driven, probabilistic machine learning approach to decision support for patients with sepsis. *JMIR Medical Inform.* 3(1):e11.
- van Wyk F, Khojandi A, Mohammed A, Begoli E, Davis RL, Kamaleswaran R (2019) A minimal set of physiologic markers in continuous high frequency data streams predict adult sepsis onset earlier. *Internat. J. Medical Inform.* 122:55–62.
- Yealy DM, Huang DT, Delaney A, Knight M, Randolph AG, Daniels R, Nutbeam T (2015) Recognizing and managing sepsis: What needs to be done? *BMC Medicine* 13(1):98.

Erik Rosenstrom is an operation research PhD student at North Carolina State University. He received his BS in industrial and systems engineering with a focus on operations research from Georgia Institute of Technology. His research interests include decision making under uncertainty as applied to improving health policy. He has particular interest in agent-based simulation, Markov decision process, and stochastic optimization methodologies.

Sareh Meshkinfam is an operation research data Scientist in Dynamic Ideas LLC (Waltham, Massachusetts). She earned her PhD in industrial and systems engineering from North Carolina State University in 2020. Her work and research focus on large-scale data analysis and optimization in areas such as healthcare, transportation, supply chain, and production. Her PhD dissertation involved analyzing K-12 students' performance and sepsis patient health conditions using statistical and optimization models.

Julie Simmons Ivy is a professor in the Department of Industrial and Systems Engineering at North Carolina State University. She received her PhD in industrial and operations engineering from the University of Michigan. Her research interests are mathematical modeling of stochastic dynamic systems with an emphasis on statistics and decision analysis as applied to healthcare, public health, and humanitarian logistics.

Shadi Hassani Goodarzi is a data scientist at Citrix. She earned her PhD in industrial and systems engineering from the North Carolina State University in 2018. Her work and research focus on mathematical modeling of large-scale problems. In her PhD dissertation, she studied the treatment and natural history of breast cancer treatment using Markovian models and large-scale data analysis paired with simulation techniques, respectively.

Muge Capan is an associate clinical professor at the decision sciences and management information systems department at LeBow College of Business. She received her PhD in industrial and systems engineering at North Carolina State University with a minor in operations research. Her research focuses on data science, statistical analysis, and decision modeling in healthcare to develop smart and connected clinical decision support systems.

Jeanne Huddleston, MD, MS is a physician and professor of medicine at Mayo Clinic, Rochester. She led the integration of engineering principles into clinical care, creating the first Clinical Engineering Learning Laboratory with the Department of Emergency Medicine at Mayo Clinic. Her research goal is the translation of engineering principles to healthcare delivery to improve the safety, efficiency, and reliability of the healthcare experience for patients, their families, and providers.

Santiago Romero-Brufau, MD, PhD, MS is an assistant professor of healthcare systems engineering and medicine at Mayo Clinic, Rochester. He is also part-time instructor in health data science at the Harvard School of Public Health. His main areas of interest are in the application of machine learning for clinical decision support and bridging the implementation gap of machine learning models into clinical practice.