Unifying Data-Driven Modeling with Machine Learning to Improve Personalized Treatment of Critical Care Patients

by

Brian McLaverty

Bachelor of Science in Bioengineering, Temple University, 2017

Submitted to the Graduate Faculty of

the Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2023

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Brian McLaverty

It was defended on

January 19th 2023

and approved by

Robert S. Parker, PhD, Professor, Department of Chemical and Petroleum Engineering

Gilles Clermont, MD, Professor, Department of Critical Care Medicine, School of Medicine

Hyo Kyung Lee, PhD, Adjunct Assistant Professor, Department of Industrial Engineering

Jason Shoemaker, PhD, Associate Professor, Department of Chemical and Petroleum Engineering

Dissertation Director: Robert S. Parker, PhD, Professor, Department of Chemical and Petroleum Engineering

Unifying Data-Driven Modeling with Machine Learning to Improve Personalized Treatment of Critical Care Patients

Brian McLaverty, PhD

University of Pittsburgh, 2023

Sepsis is a life-threatening organ dysfunction resulting from a dysregulated host response to infection. Early identification and appropriate management of sepsis has been identified as a global health priority. The development and validation of patient-specific models of acute inflammation offer the opportunity to expand current knowledge of underlying interpatient differences in the fight against infection. Early prediction into identified sepsis subtypes exhibiting differing health outcomes could be used as a guide for clinicians to develop a personalized therapeutic regimen.

A dynamical model of inflammation response in sepsis was derived, leveraging cytokine mediators that play a key role in sepsis pathology. The model was validated against a selected cohort of sepsis patients from the Protocolized Care for Early Septic Shock (ProCESS) trial to create patient-specific models of inflammation. Using the fitted model parameters and unsupervised machine learning, four subtypes of septic patients emerged with distinct inflammation responses and clinical outcomes. A clinical tool was proposed that accurately predicts patient membership into subtypes exhibiting high or low mortality risk using cytokine levels measured within six hours of hospital admission.

Dialysis patients frequently experience intradialytic hypotension (IDH), which is an independent predictor of mortality. Development of a personalized treatment support system for hemodialysis that reduces risk of IDH could potentially improve dialysis patient outcomes. A predictive model of future risk of IDH was developed by training and testing a random forest model using electronic health record data from hemodialysis treatments performed at UPMC acute care facilities. The model forecasted future IDH several hours ahead of occurrence and produced dynamic risk evolution toward instability. An early-warning tool was developed that accurately classifies a patient as high or low risk for IDH using model-derived risk scores available within minutes of dialysis initiation. A risk-based reinforcement

learning algorithm was then developed that recommends preemptive, personalized treatment. The agent-suggested treatment strategy resulted in decreased incidence of hypotension and increased accomplishment of individualized fluid goals *in silico*, at the cost of increased intervention. Clinicians could use the tools presented in this work to provide timely, individualized treatment to patients and potentially improve sepsis and hemodialysis patient outcomes.

Table of Contents

List of Tables

List of Figures

Preface

First, I would like to thank my family for supporting me through the many years of my education. My father, also an engineer, always inspired me growing up as he was able to fix almost anything in front of him. If he didn't initially have a solution to a problem, it seemed as if he would always be able to find one. My mother also worked endlessly to support and care for me and my siblings. Your relentless push to work work hard, stay disciplined, and persevere through adversity remains engrained in me and was essential to get to this point in my graduate studies. I would also like to thank my siblings, Allison and Tracey, as well as extended family and friends for always being there for me and giving me a laugh when it was needed. Nataley, I would like to thank you for putting up with me through many years of graduate studies. You have been an excellent partner, friend and mentor to me throughout my graduate experience. As a bedside nurse, your shared clinical knowledge was invaluable to me as I developed my PhD project. I cannot thank you enough for sticking beside me all of this time. To my current and prior lab members, thank you for welcoming me to the lab with open arms, and providing me guidance and feedback on research problems without hesitation. Bob and Gilles, thank you for your excellent mentorship throughout the years of my graduate education. My engineering skillset has expanded greater than I could of ever dreamed of under your guidance. I have always been impressed and motivated by your impact on the engineering and medical communities.

This dissertation is dedicated to my godfather: John F. McLaverty, who died of prostate cancer in 2011. Your memory has inspired me to pursue a degree and career that focuses on the development of engineering-derived technologies to improve patient outcomes. My hope that the work in this disseration provides meaningful contribution to advancement of personalized therapy in critical care patients.

-Brian McLaverty

1.0 Introduction

1.0.1 Background and Significance

Sepsis is defined as life-threatening organ dysfunction resulting from a dysregulated host response to infection [104]. A patient is clinically diagnosed with sepsis when they present at least two of the following: respiratory rate >= 22 breaths per minute, systolic blood pressure <= 100 mmHg, and altered mental status. Current global estimates of mortality resulting from sepsis range from 20-30% [31, 93]. In the United States alone, sepsis is the leading cause of hospital death and most expensive condition treated, resulting in approximately $24 billion in hospital costs each year [67, 111]. In 2017, the World Health Organization declared sepsis as a global health priority and has encouraged advanced efforts to improve prevention, diagnosis and management of sepsis in response to the syndrome's economic cost and impact on human health.

The acute inflammatory response is the body's non-specific and immediate response to infection and/or tissue damage [38, 1]. Toll-like receptors (TLRs) on resident cells (e.g. macrophages, dendritic cells, mast cells) recognize pathogen-associated molecular factors (PAMPs) and danger associated molecular patterns (DAMPs) presented by microorganisms and damaged cells, respectively, and subsequently release pro-inflammatory mediators (e.g. chemokines, cytokines, vasoactive amines) [1]. Vasoactive mediators encourage vasodilation and increased blood flow to the affected site. Circulating leukocytes (e.g. neutrophils, monocytes) attach to vascular endothelium via weak selectin and ligand interactions and roll along the endothelium wall [1, 19]. Chemokines on vascular endothelium promote leukocyte integrin binding to endothelium cell adhesion molecules, resulting in leukocyte adhesion to endothelium wall. Expression of these chemokines and cell-adhesion molecules is induced by pro-inflammatory mediators (e.g. IL-6) [19]. Leukocytes then migrate through the endothelium, into the tissue and to the site of the infection via chemotaxis. TLRs and opsonic receptors on neutrophil membranes attach themselves to proteins on the pathogen membrane, and the cells subsequently engulf and kill the pathogen using reactive oxygen species (ROS)

[56]. Other recruited leukocytes (e.g. monocytes) differentiate into proinflammatory (M1) macrophages that remove the pathogen via phagocytosis and further encourage production of pro-inflammatory mediators that recruit additional leukocytes [1]. Aged and expended neutrophils are directed to undergo apoptosis and are subsequently removed by M1 macrophages via a phagocytic process called efferocytosis [56]. During the efferocytosis process, M1 macrophages transition into a M2 anti-inflammatory phenotype. M2 macrophages further perform efferocytosis and produce anti-inflammatory cytokines (e.g. IL-10), which function to suppress production of pro-inflammatory mediators (e.g. IL-6) and decrease tissue influx of neutrophils [17, 1]. IL-6, a pleiotropic cytokine, transitions from a neutrophil recruiter to a monocyte recruiter [17, 44]. M2 macrophages transition into (Mres) phenotype that repair damaged tissue via production of anti-fibrotics and anti-oxidants [17]. A successful acute inflammatory response removes the pathogen and repairs damaged tissue, returning the body to equilibrium.

Any dysregulation in the acute inflammatory response, however, can lead to sepsis and result in organ dysfunction and death. For example, during the pro-inflammatory phase of acute inflammation, pro-inflammatory cytokines are upregulated and lead to an influx of neutrophils to kill the pathogen. Nevertheless, toxic ROS and enzymes released by neutrophils can cause excessive damage to local healthy tissue if this phase of inflammation is not properly controlled [44]. Similarly, an imbalanced anti-inflammatory phase could result in persistent inflammation or unsuccessful elimination of the pathogen [112].

The substantial amount of patients affected by sepsis and the financial incentive for improving poor clinical outcomes in the population has led to extensive research efforts in development of treatments and tools for early detection of sepsis. Despite hundreds of clinical trials conducted over the past 30 years that sought to modulate inflammation in sepsis and improve clinical outcomes, none have resulted in FDA-approved treatment that are currently available. In 2001, recombinant human activated protein-C was FDA-approved for its anticoagulant and anti-inflammatory properties that led to an observed reduction in mortality in patients with severe sepsis [8, 43]. However, the observed reduction in mortality was unable to be replicated in a later trial, leading to its withdrawal from the market in 2011 [87, 82]. No other sepsis treatment has successfully completed a phase 3

trial. Other therapeutic strategies to modulate the inflammatory response in sepsis have included: administration of corticosteroids for broad anti-inflammatory effects, inhibition of pro-inflammatory mediators, proteins, enzymes, anti-endotoxins, NO-inhibitory agents, and immunostimulatory agents [97, 72].

Several plausible reasons exist that explain why therapies have been unsuccessful in improving outcomes in humans with sepsis. Challenges exist to replicate the inflammatory response and therapeutic effects observed in preclinical animal models of sepsis in humans. In animals, therapy is administered around an experiment-defined infection time, however, in humans, treatment is given at an unknown time following infection [72]. Inflammation and response to treatment in animals varies depending on the infection procedure and type of infectious agent used [65, 72]. Activated pathways and inflammatory responses to infection are different in animals and humans [98, 118]. Heterogeneity of the septic response in humans also contributes to the difficulty in development of effective therapy. Individual differences in genetics, comorbidities, age, sex, and other attributes influence the human response to infection.

Identification of sepsis endotypes, or subtypes of disease with distinct underlying pathophysiological mechanisms, is a strategy to stratify patients for treatment [70]. Endotype identification differs from phenotyping, or identification of subgroups by observable clinical characteristics, as the latter neglects patient differences in underlying disease process. These distinct underlying mechanistic differences may be important in deciding who should or shouldn't receive a therapy with a specific action. Endotyping has been used to describe subgroups in other disease areas such as asthma and coronavirus [70, 61, 86]. Given the patient heterogeneity observed in sepsis, endotyping is a potential strategy to deliver more targeted therapy and improve clinical outcomes.

Systems engineering tools and mathematical modeling can be leveraged to address some of the challenges in identifying distinct groups of patients that differ in their underlying inflammation response during sepsis. Published mathematical models exist that describe the evolution of inflammation in sepsis after infection onset, yet only a few have been calibrated against human data [60, 88, 91, 23, 125]. A goal of this dissertation is to build patient specific models of inflammation, and then identify sepsis endotypes using inflammation heterogeneity

at the individual level described by the parameterized model. Development of an endotype identification tool that provides quick, accurate patient assignment with readily available clinical measurements is an additional focus.

Hemodialysis, a type of renal replacement therapy (RRT), is a required procedure for patients with renal failure and serves as a replacement for normal nonendocrine kidney function. [78]. In the United States alone, approximately 786,000 people have renal failure, 71% of which undergo regular dialysis treatments [109]. Despite the procedure's efficacy in preventing uremia, patients treated with dialysis have a high risk of mortality [79]. The dialysis population is projected to expand in the future as contributing risk factors to kidney disease, such as diabetes and hypertension, continue to rise [68, 41]. Therefore, there is a critical need for treatment advances that reduce dialysis mortality and improve patient outcomes.

The hemodialysis procedure replaces two main functions of the kidneys: removal of toxins and fluid from the body. The dialysis patient is connected to a dialysis machine through a surgically created vascular access (arteriovenous fistula, graft, or catheter) located in the arm, neck, chest, or leg [94]. Blood is pumped from the patient into a dialyzer (i.e. filter) located within the machine. The dialyzer contains hollow fibers with a semipermeable lining (membrane) that separates the blood from dialysate fluid. The dialysate contains a standardized proportion of water, electrolytes and other solutes found in plasma. Some solutes (i.e. urea, creatinine, magnesium, potassium) are transferred from the blood to the dialysate, while other solutes (e.g. bicarbonate) are transferred from dialysate to the blood [94, 40]. Fluid is removed from the blood through a process called ultrafiltration; dialysate side pressure is reduced, creating a hydrostatic pressure gradient that promotes transfer of fluid from the blood to the dialysate. Some solutes (e.g. sodium) are also transferred into the dialysate via convection. Purified blood is pumped back into the patient and used dialysate fluid is wasted.

Intermittent hemodialysis is typically performed three times per week for three to five hours each treatment to achieve prescribed personalized fluid removal and clearance goals. Fluid goals are calculated by a nephrologist by first estimating a patients dry-weight, or the patient's lowest achievable post-dialysis weight without symptoms or overt signs of

hypotension or fluid overload [45, 69]. The clinician calculates the fluid target for the treatment using the difference between the patient's pre-dialysis weight and the dry weight of the patient, as well as the addition of any fluids scheduled during treatment. Ultrafiltration rate, dialyzer type, blood flow rate, dialysate rate, and type of dialyzer can be tuned by the clinician to accomplish predetermined dialysis targets.

Intradialytic hypotension (IDH) is a complication that occurs in 4-30% of dialysis treatments [16, 59]. The reduction in blood pressure during treatment and its associated symptoms (e.g. cramping, nausea, vomiting, dizziness or fainting) can lead to early discontinuation of therapy, negatively impacting achievement of prescribed fluid removal, clearance goals and renal recovery [4, 122]. Incidence of IDH results in ischemia and cardiac stunning, and increases long-term risk of organ failure and mortality [122, 4, 99, 16, 101]. Several identified patient and treatment-related factors are associated with increased risk of IDH during dialysis treatment. Patient-related risk factors of IDH include older age, diabetes, cardiovascular disease, anemia and poor nutritional status [81]. Normal physiological compensatory mechanisms in response to reductions in blood pressure (e.g. increased heart rate and contractility, increasing in cardiac preload via fluid redistribution) may be compromised in the critically ill [26]. Ultrafiltration rate that exceeds vascular refilling can decrease cardiac preload and thus lead to hypotension. Rapid plasma solute clearance can result in a fluid shift from blood to tissue via osmosis.

Several approaches have been investigated to reduce the incidence of IDH and improve clinical outcomes. Continuous renal replacement therapy (CRRT) removes fluid and toxins 24 hours a day using lower ultrafiltration, blood and dialysate flow rates than intermittent hemodialysis. Thus, CRRT is a preferred modality for ICU patients with history of hemodynamic instability. Nonetheless, evidence of reduced blood pressure instability for patients undergoing CRRT versus intermittent hemodialysis has been mixed [33, 95]. Furthermore, CRRT is more expensive and requires additional hospital resources [40, 7]. Administration of hypertonic solutions (e.g. mannitol, albumin) or high sodium dialysate concentration has been utilized in an attempt to increase oncotic pressure and prevent large osmotic shifts during dialysis that could lead to IDH [74, 71, 55]. Large amounts of fluid administration or significant decreases in ultrafiltration prescription to correct IDH, however, increase risk of fluid overload which

is associated with mortality [103]. Interdialytic weight gain is also observed in patients receiving high sodium dialysate [77]. The limited, conflicting evidence presented in studies indicate a more personalized treatment approach may be necessary to reduce incidence of IDH and optimize dialysis outcomes.

Reinforcement-learning is a subfield of artificial intelligence in which an agent learns, by trial and error, an optimal policy that maximizes expected reward received over time [107]. It has been successfully applied to a variety of domains including autonomous driving, robotic control, gaming and chemical process control [54, 102, 57, 100]. In the healthcare domain, reinforcement learning has been used to optimize treatment for sepsis, cancer, and sedation [58, 83, 84]. Medical decision-making can be described as an iterative process in which a clinician observes a patient's health state and decides whether or not to intervene using their best clinical judgement to maximize the probability of the patient's recovery. The patient response is partly in control of the clinician and partially dictated by patient's underlying biology. A reinforcement learning agent learns the optimal way to act on a patient solely using a large historical dataset of clinician-patient interactions, many of which are suboptimal.

To address clinical challenges observed in the dialysis population, a key goal of this dissertation is to use reinforcement learning to train and test an artificial agent that provides *preemptive*, personalized treatment recommendations for dialysis patients, leveraging patient *risk* of future hypotension. This differs from clinical status quo in which a clinician *reacts* only when overt signs of deterioration in hemodynamic stability and clinical presentation are apparent.

1.0.2 Dissertation Overview

The dissertation is organized as follows. In chapter 2 , patient-specific dynamic models of inflammation in sepsis are developed using key mediators involved in coordinating the response to infection. The model structure describes interpatient variability in infection time, speed, and magnitude of inflammation. The fitted parameters of the model are clustered and reveal endotypes with distinct inflammatory trajectories and clinical outcomes. Baseline

biomarker levels and supervised machine learning were used to quickly classify individual patients belonging to endotypes with high or low observed mortality rates.

In chapter 3, a predictive model of future risk of intradialytic hypotension was developed by training and testing a random forest classifier using electronic health record data available from a large set of patients receiving intermittent hemodialysis. Model-produced risk trajectories demonstrated distinct separation in average risk of future intradialytic hypotension even hours ahead of blood pressure instability. The work was extended into an early warning system which quickly classified a patient as high or low risk of intradialytic hypotension, solely using model-produced risk scores available around dialysis initiation.

In chapter 4, a data-driven algorithm was developed that recommends optimal dialysis treatment to reduce occurrence of intradialytic hypotension and improve dialysis outcomes. A finite Markov decision process (MDP) for dialysis over 15-minute time steps was defined using a small set of states and interventions taken by a clinician. A reinforcement learning agent was used to solve for the optimal policy of the MDP, using predicted future intradialytic hypotension risk and achievement of individualized fluid goals as feedback for its actions. A policy modification was proposed to better align agent-recommended interventions with clinical intuition. Simulated dialysis trajectories were produced using the agent-derived policy to evaluate its treatment efficacy relative to the clinician.

In chapter 5, an overview of the contributions of the work in the areas of dialysis and sepsis are presented. Future directions are included to enhance the current work, with the goal to provide clinicians tools necessary to administer personalized treatment to sepsis and dialysis patients and potentially improve clinical outcomes.

2.0 Identification and Prediction of Sepsis Endotypes

2.1 Background

Sepsis is defined as organ dysfunction resulting from a dysregulated acute inflammatory response to infection [104]. The syndrome is the leading cause of hospital death in the United States and carries an annual economic burden upwards of 24 billion dollars [67, 111]. Hundreds of clinical trials have focused on the development of pharmacological interventions that suppress, neutralize, or stimulate the inflammation process, however, none have resulted in the emergence of a new treatment for sepsis [72]. Heterogeneity in sepsis arising from patient differences in infection site, timing of infection, infection source, comorbidities, and genetic predispositions make treatment of sepsis difficult. Identification of sepsis endotypes, or subtypes of patients defined by distinct underlying pathological mechanisms, could lead to the development of more targeted therapies. For example, several approved treatments for asthma, such as anti-IgE and anti-IL5, have been developed through the identification of asthma endotypes [14, 10, 90, 108]. Research attempts to identify sepsis endotypes have largely focused on identifying patient subgroups from high dimensional gene expression, metabolite, and plasma protein data collected at hospital admission, however, this approach cannot elucidate distinct differences in the dynamic behavior of inflammation between patients [64].

The acute inflammatory response is the body's initial response to pathogen invasion and is regulated by cytokines, chemokines and other inflammatory mediators [96]. Inflammatory cytokines are critical in both the pro- and anti- inflammatory response to infection and their dysregulated production can lead to sepsis [96, 15, 17]. IL-6 is a pro-inflammatory cytokine secreted by M1 macrophages upon infection and plays a role in neutrophil and monocyte recruitment to the site of infection [17, 3, 123, 49]. In healthy volunteers injected with endotoxin, plasma IL-6 concentrations have demonstrated to peak around 3 hours and reach baseline around 8 hours [89]. However, ICU-based studies indicate humans diagnosed with sepsis have recorded plasma IL-6 concentrations that are maximal around hospital admission

and continue to decline 72 hours after clinical presentation. Excessive IL-6 secretion and subsequent neutrophil recruitment can lead to unintended damage to local healthy tissue and organ dysfunction, whereas insufficient cytokine response may result in ineffective removal of the pathogen [113]. Cytokine IL-10 is an anti-inflammatory mediator released by M2 macrophages that suppresses the production of pro-inflammatory mediators such as IL-6 [17, 3, 21]. Serum IL-10 concentrations in humans with sepsis also have demonstrated to have maximum recorded cytokine levels at clinical presentation and which continue to decrease 72 hours after hospital admission [89]. Inappropriate levels and timing of IL-10 secretion has been associated with uncontrollable pathogen growth and excess inflammation [21]. The imbalance of pro-inflammatory and anti-inflammatory cytokines contribute to the underlying pathogenesis of sepsis, and therefore sepsis subtypes may emerge from specific cytokine trajectories.

Sepsis endotype identification may be possible using mechanistic ordinary differential models that describe the human pro- and anti- inflammatory response. Mathematical models have previously been developed to characterize the underlying mechanisms in sepsis, however, few have been validated against human data [60, 88]. This task is challenging due to reasons such as infrequent and missing clinical measurements, and the lack of clinical data describing the dynamics of inflammation prior to clinical presentation. Each patient is infected by the pathogen at an unknown time prior to hospital entrance, making it difficult to determine initial conditions for the inflammation response. In this paper, we introduce a simple mathematical model of the pro- and anti- inflammatory response that initiates at time zero of infection and is evaluated against a cohort of patients from the Protocol-Based Care for Early Septic Shock (ProCESS) trial [124]. This trial investigated the use of early goal-directed therapy to reduce mortality in patients with severe sepsis and septic shock, however, the study resulted in no significant improvement in mortality across treatment arms. The mathematical model is fit to each patient's IL-6 and IL-10 data, representing the pro- and anti- inflammatory response, and the model parameters are clustered. The aim is to identify sepsis endotypes using the clustered parameters, each of which describe groups of patients with distinct inflammatory trajectories that emerge from the mathematical model. The investigation resulted in the discovery of four distinct subtypes of patients that exhibited

different mortality and organ failure rates.

2.2 Methods

2.2.1 Clinical Cohort

A cohort of 316 patients was selected from the 1341 enrolled patients in the ProCESS trial based on the availability of cytokines IL-6 and IL-10 at 0-, 6-, and 24-hour study time, where 0 hours study time was the time of trial enrollment. 46 of these patients also had IL-6 and IL-10 cytokine availability at 72 hours study time. Differing methods of assay validation used at each study site resulted in various limits of quantitation (LLQ), and therefore the smallest, most frequent cytokine measurement across the trial was considered LLQ (IL-6: 28.8 pg/mL, IL-10: 12.64 pg/mL). Each of these subjects had at least one IL-6 and IL-10 measurement $\geq$ LLQ. No significant differences were found in the primary study outcome (60-day in hospital mortality) or secondary outcomes (90-day, 1 year survival, need for organ support) between treatment arms, and thus effects of different methods of intervention were not considered in cohort selection.

2.2.2 Dynamical Model of Inflammation and Model Fitting

IL-6 and IL-10 cytokines were chosen to represent the pro- and anti- inflammatory phases of infection, respectively, due to their crucial role in coordinating the removal of pathogens and resolving inflammation. An initial model of inflammation was developed that described the pro-inflammatory response (IL-6) and anti-inflammatory response (IL-10) as second-order dynamical systems subjected to an infectious input of infinite magnitude and infinitesimal duration (Equation 2.1), where $\tau_{1,cytokine}$ and $\tau_{2,cytokine}$ are time constants describing the rise and fall speeds of the pro- and anti- inflammatory response, $K_{cytokine}$ is the gain describing the magnitude of infectious response, $u(t)$ is the infectious input, $x_{1,cytokine}$ and $x_{2,cytokine}$ are internal system states and $y_{cytokine}$ is the system output or pro- and anti-inflammatory response to infectious input.

$$
\begin{aligned}
\frac{dx_{1,cytokine(t)}}{dt} &= -\frac{1}{\tau_{1,cytokine}} x_{1,cytokine(t)} + \frac{K_{cytokine}}{\tau_{1,cytokine}} u(t) \\
\frac{dx_{2,cytokine(t)}}{dt} &= -\frac{1}{\tau_{2,cytokine}} x_{2,cytokine}(t) + \frac{1}{\tau_{2,cytokine}} x_{1,cytokine}(t) \\
y_{cytokine} &= x_{2,cytokine}(t), u(t) = \delta(t)
\end{aligned}
\tag{2.1}
$$

The initial model describing inflammation suffered from several fundamental issues. The smaller time constants, $\tau_{1,cytokine}$, represent the rise speeds of the responses following an infectious insult. These time constants exhibited identifiability issues as over a third of the patients (n=104) lacked IL-6 and IL-10 data to describe the initial rise of infection. A second issue is the system input is non-physiologic and fails to accurately describe the duration and magnitude of pathogen insult. To address these issues, a reduced model of inflammation was developed which describes the pro- and anti- inflammatory response using two first ordinary differential equations instead of two second-order systems (Equation 2.2).

The pro- and anti- inflammatory response to infection, represented by cytokines IL-6 and IL-10, were modeled as first order system responses to a triangular pulse of infection.

$$
\begin{aligned}
\frac{dx_{IL6}}{dt} &= -\frac{1}{\tau_{IL6}} + \frac{K_{IL6}}{\tau_{IL6}} u(t) \\
y_{IL6} &= x_{IL6}(t) \\
\frac{dx_{IL10}}{dt} &= -\frac{1}{\tau_{IL10}} + \frac{K_{IL10}}{\tau_{IL10}} u(t) \\
y_{IL10} &= x_{IL10}(t)
\end{aligned}
\tag{2.2}
$$

y_{IL6} and y_{IL10} are the output variables which represent the internal states x_{IL6} and x_{IL10} of the inflammation process. Time constants τ_{IL6} and τ_{IL10} represent the speeds of the pro- and anti- inflammatory responses, $u(t)$ represents the triangular pulse infectious input driving the acute inflammatory response, and gains K_{IL6} and K_{IL10} represent the magnitudes of the pro- and anti- inflammatory responses to infection $u(t)$. The triangular system input, $u(t)$, was chosen to have a peak of 6 hours, duration of 48 hours, and area of unity. This guarantees that the model-simulated cytokines peak in ≥ 6 hours, exposure of bloodstream to infection

is physiologically reasonable, and that the magnitude of the simulated inflammatory response is captured by gains K_{IL6} and K_{IL10}.

Cytokine measurements were logged to prevent IL-6 and IL-10 from exhibiting large ranges. Temporal cytokine data is left-censored (clinical data is unavailable prior to trial enrollment), and time of infection is both unknown and variable amongst patients in the cohort. Therefore, measured IL-6 and IL-10 concentrations were uniformly shifted forward in hourly increments [0,30] prior to fitting the model parameters (τ_{IL6}, τ_{IL10}, K_{IL6},K_{IL10}) to the patient cytokine data in order to account for variation in infection time prior to study time initiation. The parameter estimation objective is to minimize the weighted sum of squared error between the clinically measured cytokine concentrations and model-simulated cytokine concentrations (Equation 2.3).

$$
\begin{aligned}
\text{minimize} \quad & F = \sum_{i}^{n}((y_{IL6}(i) - \hat{y}_{IL6}(i))^2 + w(y_{IL10}(i) - \hat{y}_{IL10}(i))^2 + Mz_1 + Mz_2 \\
\text{subject to} \quad & K_{IL10} <= 4.0\tau_{IL10} + 95.2, \\
& K_{IL10} <= 447.3, \\
& K_{IL6} <= 5.8\tau_{IL6} + 138.7, \\
& K_{IL6} <= 593.2, \\
& z_1 = 0 \text{ when } y(i) \geq LLQ \text{ and } \hat{y}(i) >= LLQ, \\
& z_1 = 0 \text{ when } y(i) < LLQ \text{ and } \hat{y}(i) <= LLQ, \\
& z_1 = 1 \text{ when } y(i) < LLQ \text{ and } \hat{y}(i) > LLQ, \\
& z_2 = 0 \text{ when } \hat{y}(168) <= LLQ, \\
& z_2 = 1 \text{ when } \hat{y}(168) > LLQ
\end{aligned}
\tag{2.3}
$$

In Equation 2.3, $\hat{y}_{IL6}$ and $\hat{y}_{IL10}$ are the model-simulated concentrations, y_{IL6} and y_{IL10} are the measured cytokine concentrations, n is the number of datapoints and the weight, $w = \frac{max(y_{IL6})}{max(y_{IL10})}$, is included to scale cytokines to similar magnitudes. The $Mz1$ term is included to drive the model-simulated cytokine concentration below the lower limit of quantitation (LLQ) if the measured cytokine concentration is also below the LLQ, and the $Mz2$ term is included to drive the model-simulation below the LLQ by 168 hours. The inequality constraints

were included to keep the model-simulated cytokine concentrations below the maximum measured concentrations found within the dataset. The optimal patient fit resulted in the lowest weighted sum of squared error across pre-study times. Model parameter estimates for each patient were found using *fmincon* nonlinear programming solver in MATLAB.

2.2.3 Consensus Clustering for Endotype Identification

The next step was to identify subpopulations of patients within the larger cohort population that share similar inflammatory responses, using model fits for each of the patients following the routine described in subsection 2.2.2. K-means clustering, an unsupervised machine learning algorithm, provides the ability to identify k similar clusters of patients in parameter space [115]. Specifically, k-means allocates k centroids in parameter space and allocates each observation, or patient, to the nearest centroid defined by the squared Euclidean distance. The number of clusters must be chosen a priori, and centroids are initialized randomly, resulting in inconsistent clustering assignment across algorithm runs. Consensus clustering addresses the shortcomings of this algorithm by obtaining subsamples of the original dataset via sampling without replacement and then applying a clustering algorithm to many datasets constructed from resampling [76]. If the subsamples represent random samples of patients taken from the subpopulations, then cluster membership and number of clusters should not drastically change when the algorithm is applied across the resampled datasets.

A dataset that included optimal model parameters (τ_{IL6}, τ_{IL10}, K_{IL6},K_{IL10}) and the timeshift (T_{shift}) used for fitting each of the n=316 patients was constructed and used as input for consensus clustering. First, principal component analysis was used to transform the original, correlated parameter space to 4 principal components that describe >90% of data variance. Consensus clustering was then performed by first creating $H = 500$ resampled datasets via sampling without replacement and applying k-means to each of the resampled datasets, where k is chosen a priori. A connectivity matrix, $M(h)$, of size NxN was created for each of the h=1,2...H resampled principal component score datasets to indicate whether pairs of patients (i,j) were clustered together (Equation 2.4).

$$M^{(h)}(i,j) = \begin{cases} 1, & \text{if patient } i \text{ and } j \text{ belong to the same cluster} \\ 0, & \text{otherwise} \end{cases} \tag{2.4}$$

Sampling without replacement creates resampled datasets in which i and j may not both be present. A matrix I, of size NxN was created for each of the h=1,2...H resampled principal component score datasets to indicate if both patients (i,j) were present in the resampled dataset (Equation 2.5).

$$I^{(h)}(i,j) = \begin{cases} 1, & \text{if patient } i \text{ and } j \text{ are present in resampled dataset } h \\ 0, & \text{otherwise} \end{cases} \tag{2.5}$$

A consensus matrix, C, of size NxN was defined to quantify the number of times patients (i,j) are clustered together across each of the resampled datasets and is normalized by the number of times they were both present in the resampled datasets. The consensus matrix, C, consisted of NxN values that ranging from 0-1, where 0 indicated that the patient pair (i,j) was never clustered together and 1 indicated that the patient pair was always clustered together. A larger proportion of 0's and 1's in the consensus matrix indicates cluster membership stability across runs. This analysis was repeated across a set of cluster counts k=1, 2,...K, resulting in the generation of a consensus matrix for each of the k cluster counts (Equation 2.6).

$$C^{(k)}(i,j) = \frac{\sum_{h=1}^{H} M^{(h)}(i,j)}{\sum_{h=1}^{H} I^{(h)}(i,j)} \tag{2.6}$$

Empirical cumulative distributions, $CDF \in [0,1]$, were created for each of the k=1,2,...K, consensus matrices. CDF returns the proportion of patient pair values (i,j) within the consensus matrix for cluster count k less than or equal to a specified value, c (Equation 2.7).

$$CDF^{(k)}(c) = \frac{\sum_{i<j} 1\{C^{(k)}(i,j) <= c\}}{N(N-1)/2} \tag{2.7}$$

An empirical cumulative distribution for each cluster count k was plotted and assessed to determine the cluster count that resulted in the most stable clusters. Stability was further supported by rearranging patient pairs (i,j) in C that were clustered together to be adjacent

to one another and subsequently creating a heatmap display of the rearranged matrix. K-means algorithm was applied to the dataset of principal component scores available for each patient, using the most stable cluster count (k). Then, the parameters of each of the k clusters (endotypes) were averaged to create a representative pro- and anti- inflammation trajectory for each of the k endotypes. Clinical outcomes of each of the endotypes were defined by 14-day mortality.

Lastly, a clinical tool was developed that predicts a patient into high or low mortality endotype using cytokine IL-6 and IL-10 biomarker levels within 6 hours of study time. Specifically, a logistic regression model was chosen to describe the log odds of an high-mortality outcome as a weighted linear combination of input variables, where the weights are β_j and x_j are the input variables (0 and 6 hour cytokine values) (Equation 2.8). The inverse of the logit function returns the probability high mortality endotype membership using a weighted linear combination of cytokine values as the input (Equation 2.9).

$$logit(P(\mathbf{x})) = log\frac{P(\mathbf{x})}{(1 - P(\mathbf{x}))} = \sum_{j=0}^{k} \beta_j x_j \tag{2.8}$$

$$P(\mathbf{x}) = \frac{1}{1 + exp(-\sum_{j=0}^{k} \beta_j x_j)} \tag{2.9}$$

The parameters (β) of the (binary) logistic regression model were estimated by minimizing the negative log likelihood equation (Equation 2.10), where n is the number of data samples, $\hat{y_i} = Pr(y_i = 1)$ is the predicted probability of the positive class (high mortality) for example i, and y_i is the true positive class probability. Model parameter estimates for each patient were found using L-BFGS solver in Python scikit-learn library.

$$L = \sum_{j=0}^{n} y_i log(\hat{y_i}) + (1 - y_i) log(1 - \hat{y_i}) \tag{2.10}$$

Repeated k-fold cross-validation was used to provide unbiased evaluation of model performance. Specifically, the data was randomly partitioned into k=3 folds. The model is trained using k-1 folds and its prediction performance is evaluated on the remaining fold of data by calculating the area under receiver operator curve (AUC-ROC) and area under precision-recall curve (AUC-PR). ROC curve provided tradeoffs between true positive rate (TPR) and

false positive rate (FPR) at different classification thresholds, while the PR curve provided tradeoffs between precision and recall. This process is repeated until each fold is used once for testing. K-fold cross validation process was repeated 10 times, using different randomization for dataset splitting each repetition. Average model AUC-ROC and AUC-PR across the test folds was reported, along with respective 95% confidence intervals.

2.3 Results

2.4 Identification of Sepsis Endotypes

Optimal parameters resulting from the model fitting routine were used to discover distinct subpopulations of patients within the cohort that share similar inflammatory dynamics. Each n=316 patient's inflammatory dynamics were described by 5 parameters (τ_{IL6}, τ_{IL10}, K_{IL6},K_{IL10},T_{shift}) . Consensus clustering was performed on the matrix of patient principal component scores for a range of cluster counts ($2-10$). The empirical cumulative distributions for each of the consensus matrices are depicted in Figure 2.1. Consensus matrices composed of all 0's and 1's would indicate stability; patients were always clustered together or never clustered together across resamples. Consensus matrices with many fractional entries indicated cluster instability or inconsistent patient clustering. Consensus matrices of cluster counts 2 and 3 were composed of many fractional values, indicated by their increasing cumulative probability across the domain of possible values. The cumulative distribution shape for k=4 clusters steps up at 0, is relatively flat between 0-1, and then takes a second step up around 1. The cumulative distribution functions indicated high cluster stability using k=4 clusters.

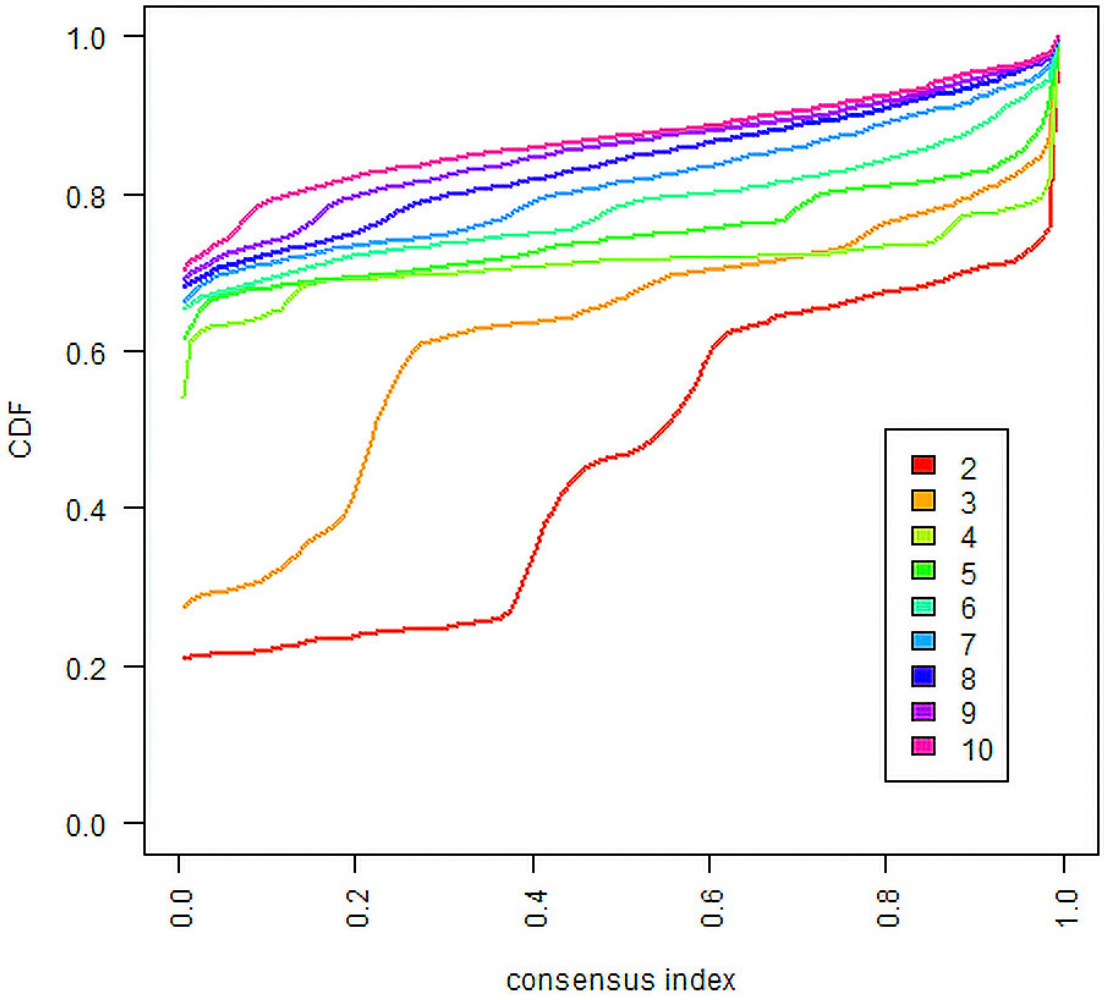

Figure 2.1: Empirical cumulative distribution functions for consensus matrices generated from clusters k=2-10

The consensus matrix that resulted from k=4 clusters is visualized as a colorized heatmap (Figure 2.2), in which a well-defined 4 block structure emerges along the diagonal. Heatmaps for consensus matrices for cluster counts prior to and succeeding k=4 were visualized and demonstrated instability (Figure 2.3, Figure 2.4).

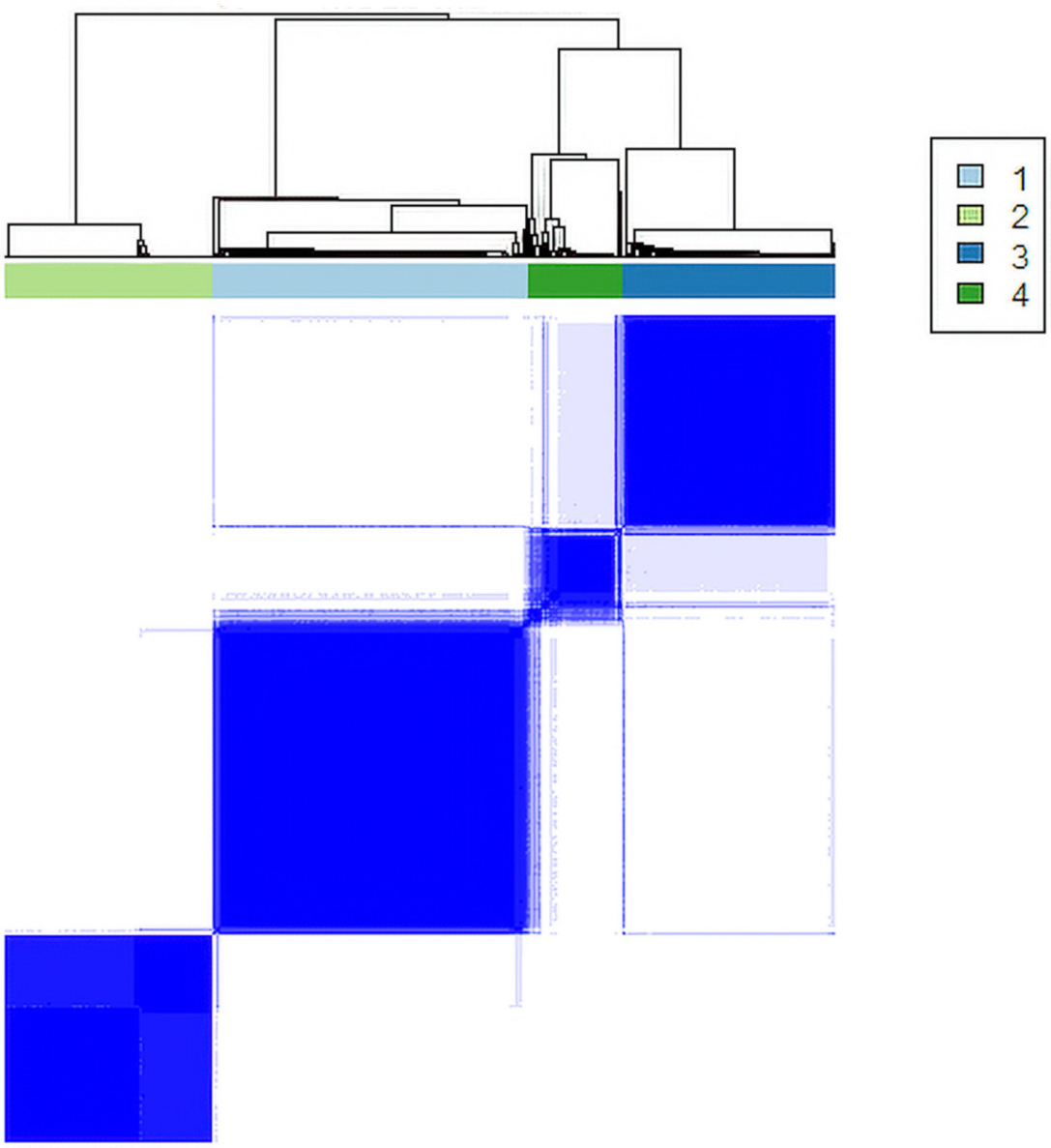

Figure 2.2: Heatmap visualization of consensus matrix stability for optimal k=4 clusters. Heatmap created by arranging consensus matrix values so that items clustered together are adjacent to each other. Consensus matrix values range from 0 (never clustered together) to 1 (always clustered together), where white represents 0, and dark blue represents 1. Relative proportion of patients in each cluster depicted by dendrogram.

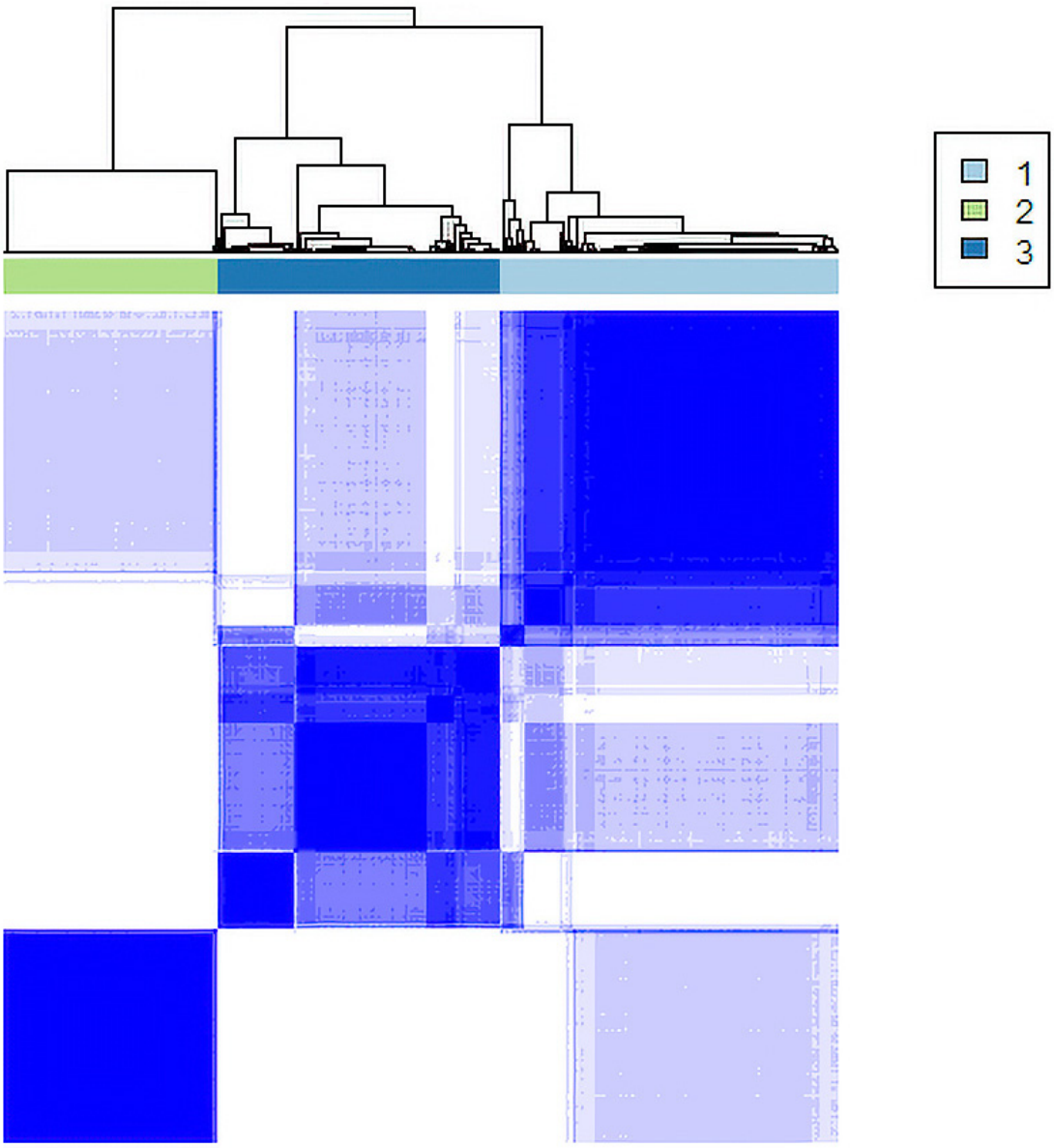

Figure 2.3: Heatmap visualization of consensus matrix stability for k=3 clusters. Heatmap created by arranging consensus matrix values so that items clustered together are adjacent to each other. Consensus matrix values range from 0 (never clustered together) to 1 (always clustered together), where white represents 0, and dark blue represents 1.Relative proportion of patients in each cluster depicted by dendrogram.

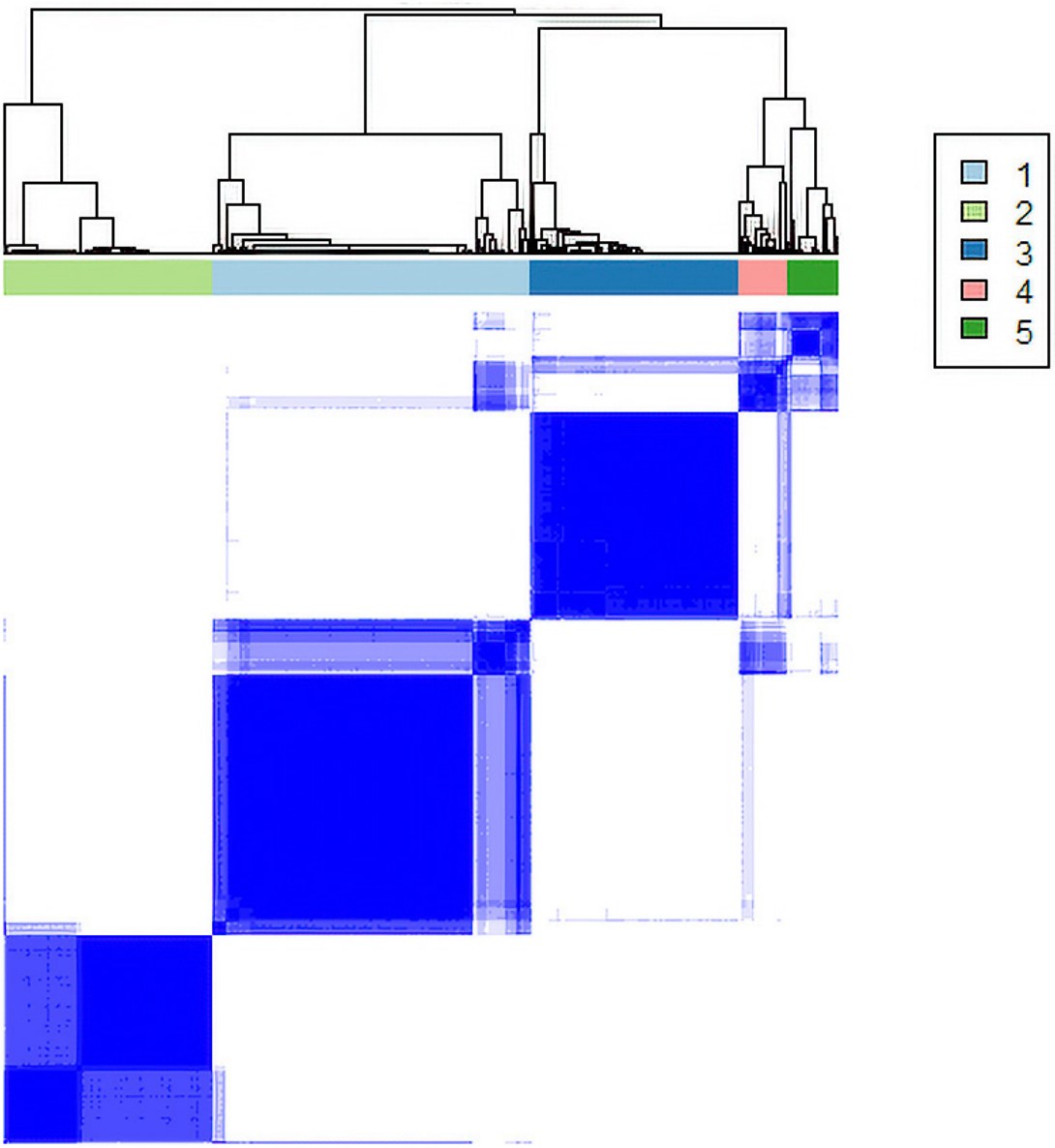

Figure 2.4: Heatmap visualization of consensus matrix stability for k=5 clusters. Heatmap created by arranging consensus matrix values so that items clustered together are adjacent to each other. Consensus matrix values range from 0 (never clustered together) to 1 (always clustered together), where white represents 0, and dark blue represents 1. Relative proportion of patients in each cluster depicted by dendrogram.

Geometric visualization of clustered principal component scores depicts four distinct patient

subpopulations of patients defined by their model parameters (Figure 2.5). Correlation between the principal components and dynamic model parameters can be described by the principle component coefficients (Table 2.1). Principal component 1 was positively correlated with all model parameters, principal component 2 was positively correlated with τ_{IL10} and K_{IL10} and negatively correlated with τ_{IL6} and K_{IL6}, and principal component 3 was strongly correlated with the time shift or pre-study time.

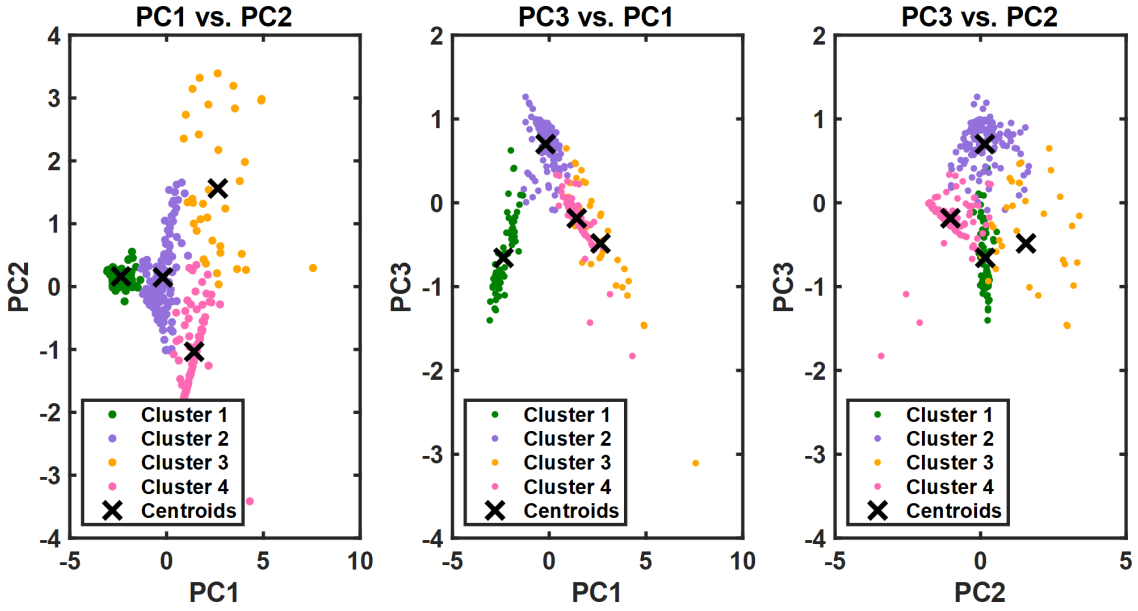

Figure 2.5: Visualization of k=4 clusters in PCA space

Table 2.1: Correlation between Principal Components and Model Parameters

	Principal Component 1	**Principal Component 2**	**Principal Component 3**
τ_{IL6}	.47	-.32	-.19
K_{IL6}	.45	-.54	-.36
τ_{IL10}	.44	.65	-.01
K_{IL10}	.47	.39	-.22
T_{shift}	.41	-.18	.89

Inflammatory trajectories for each patient subpopulation were represented by clusters'

mean model parameters (Figure 2.6). An average patient from cluster 1 had a quick inflammatory response that was initially significantly elevated after infection but quickly peaked and returned to below quantifiable levels. This quick inflammatory trajectory is reflective of their smaller time constants and gains compared to other patient subtypes. An average patient in cluster 2 had an inflammatory response that peaked nearly 15 hours after subjects in cluster 1 and reached LLQ much later. Inflammatory responses of patients in cluster 3 also peaked much later than subjects in cluster 1 and had a long-protracted decline in both IL-6 and IL-10. Subjects in cluster 2 and 3 had increasingly larger time constants and gains compared to cluster 1. Patients in cluster 4 had a similar IL-6 dynamic profile to patients in cluster 2, although lower in amplitude. These subjects also experienced a long, protracted decline in IL-10 at lower magnitudes.

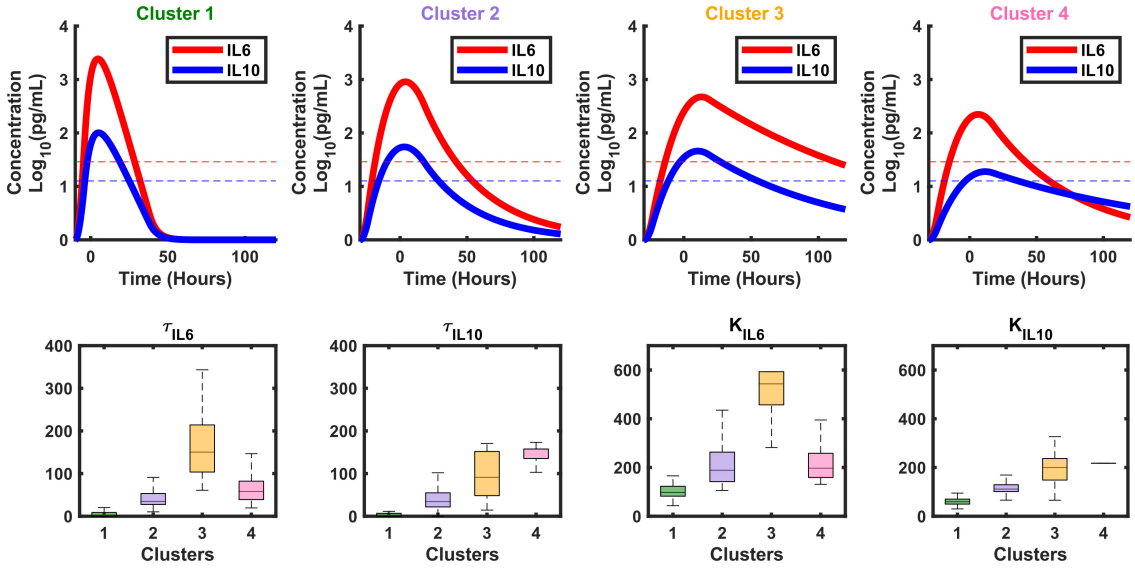

Figure 2.6: Cytokine trajectories generated from mean model parameters of each cluster and model parameter distributions

Overall, 42 (13%) of the patients in the cohort died within 14 days of hospitalization. Clinical outcomes of endotypes were significantly different ($p<.05$, chi-squared test), with the highest 14- day mortality rate and respiratory failure rate after 1 week of hospitalization experienced by patients belonging to cluster 3 (Table 2.2). Subjects in cluster 1 and cluster 4 experienced lower 14-day mortality and respiratory failure rates than patients belonging

to cluster 2 and cluster 3.

Table 2.2: Mortality and Respiratory Failure Rates for Sepsis Subtypes

	Patients (count)	**14 Day Mortality (%)**	**Respiratory Failure (1 Week) (%)**
Cluster 1	79	6.3	30.4
Cluster 2	120	16.7	43.3
Cluster 3	35	28.6	51.4
Cluster 4	82	8.5	25.6

2.5 Early Warning Tool for Prediction Into High/Low Mortality Inflammation Subtypes

A clinical tool was developed that predicts whether a patient belongs to a high or low mortality endotype, solely using cytokine IL-6 and IL-10 biomarker levels within 6 hours of study time. Patients in clusters 2 and 3 were considered the positive class (high mortality), whereas patients in clusters 1 and 4 were considered the negative class (low mortality rate) ($p<.05$). Logistic regression produced the probability of high-mortality class membership, and binary classification performance was evaluated at different probability thresholds using a ROC curve and a PR curve (Figure 2.7). As the probability threshold is decreased, more patients are classified as positive, increasing TPR but concurrently increasing FPR and decreasing precision. The prediction model demonstrates skill; AUC-ROC (.72 vs. .50) and AUC-PR (.71 vs. .49) are greater than a no-skill baseline classifier.

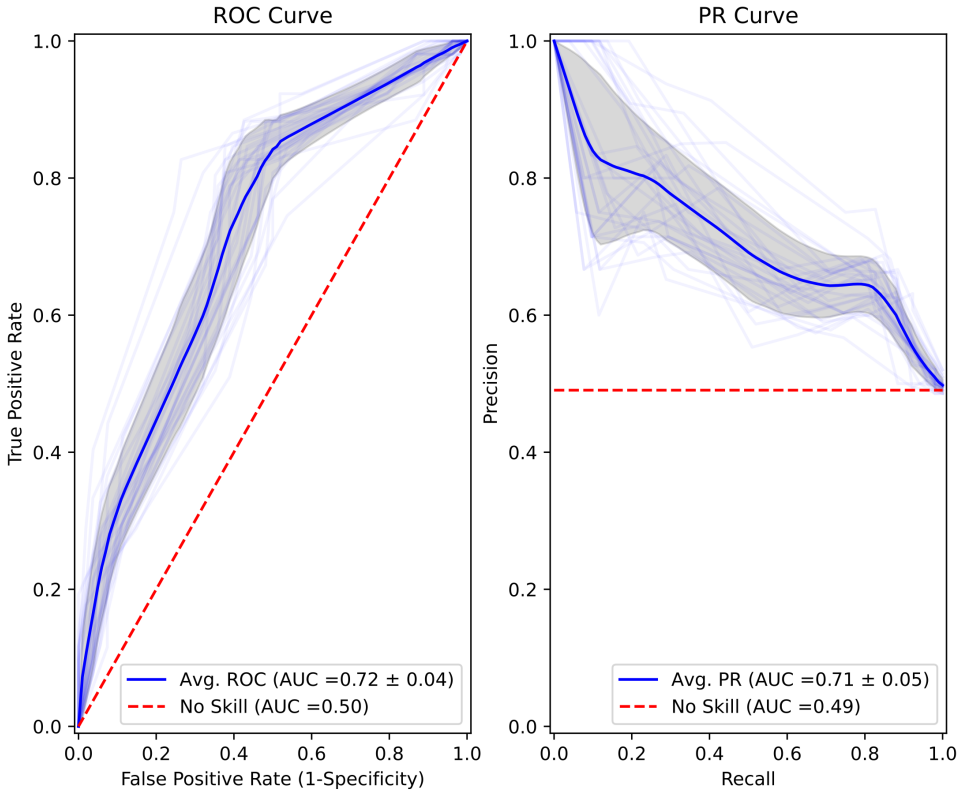

Figure 2.7: ROC curve (left) and PR Curve (right) for early detection tool

2.6 Discussion

A mathematical model developed describing inflammation using only 5 parameters was evaluated against human IL-6 and IL-10 data, both of which are critical inflammatory mediators in the pro- and anti- inflammatory response to infection. The parameters from patient specific model fits were clustered using k-means, and a consensus across multiple runs of the algorithm was used to identify four distinct endotypes of patients that differ in their underlying mechanisms that lead to sepsis. Patients from cluster 1 had a robust

inflammation response that quickly and aggressively responded to infection and then returned to their baseline state, reflected by the subtype's low 14-day mortality and respiratory failure rate. Although larger cytokine magnitudes have been associated with negative clinical outcomes, this endotype's inflammatory dynamics and clinical outcome further support that the duration of serum exposure to cytokines plays a critical role in the body's ability to properly respond to infection [9, 13, 42, 85]. Subjects from cluster 4 have a less aggressive inflammatory response, indicated by their low magnitude cytokine trajectories, and their immune system requires a lower amount of energy to remove the source of infection and return to immunological equilibrium. Subjects from cluster 2 and 3 exhibit an aggressive inflammatory response to infection over a prolonged period and have higher 14-day mortality and respiratory failure rates than patients in clusters 1 and 4. One explanation is that the patients prolonged exposure to excessive hyperinflammation cannot be successfully balanced by anti-inflammatory response, leading to consequences such as persistent damage to local healthy tissue, organ failure, and high risk of mortality [96, 15, 113].

These results open the possibility of using targeted therapy in the acute care setting. To illustrate, if a subject was identified to have a subtype 2 or 3 inflammatory response, a clinician may decide to prescribe an immunomodulatory agent that suppresses the patient's inflammation. Discovery of sepsis endotypes could also enhance clinical trial design by identifying which patients would successfully respond to a given therapeutic. IL-6 and IL-10 play a critical role in the dysregulated immune response in sepsis; however, the dynamics of these biomarkers describe a small portion of the underlying pathophysiology of the syndrome. The patient cohort analyzed in this study represents only a sample of the 1341 subjects in the ProCESS trial and therefore other sepsis endotypes may exist. Future research will focus on using an alternative modeling approach to identify additional sepsis subtypes present within the clinical trial.

3.0 Dynamic Risk Modeling of Intradialytic Hypotension

3.1 Background

Intradialytic hypotension (IDH) occurs in up to 30% of hemodialysis sessions [59]. The Kidney Disease Outcomes Quality Initiative (KDOQI) defines IDH as a decrease in systolic blood pressure (SBP) ≥ 20 mmHg or a decrease in mean arterial pressure (MAP) ≥ 10 mmHg accompanied by clinical symptoms, however, significant variation in the definition of IDH exists [122, 2]. There are a variety of patient-related and treatment-related contributors to IDH pathology. Identified patient-related factors that increase risk of IDH include, but are not limited to, autonomic and cardiac dysfunction, cardiovascular disease, age >65, diabetes, lower body weight, female sex, and Hispanic origin [106, 81]. Hemodialysis reduces plasma volume via ultrafiltration [26]. Solute clearance reduces plasma osmolality, resulting in osmotic shifts in fluid from the intravascular space to the interstitial space. Inadequate physiological compensatory response (e.g. increased heart rate, recruitment of unstressed blood volume) to reduced blood volume induced by hemodialysis can lead to IDH. Hypoperfusion and consequent organ damage induced by recurrent IDH mitigates kidney recovery. In human studies, occurrence of IDH has demonstrated to slow kidney recovery, measured by early discontinuation of hemodialysis prior to hospital discharge [4]. In animals models of acute kidney injury (AKI), hypotension significantly reduced autoregulation of renal blood flow and kidney clearance [50]. Furthermore, IDH has been found to be independently associated with mortality, even when statistically adjusting for other prognostic factors [99, 101]. Early prediction of IDH would allow clinicians to take preventative measures that could potentially reduce incidence of IDH and improve clinical outcomes.

Machine-learning algorithms exist that use featurized electronic health record (EHR) data to predict hypotension. Hatib et al. developed a machine learning algorithm that predicted hypotension with an AUC-ROC of 0.94 fifteen minutes ahead of instability, leveraging millions of features derived from high-resolution arterial waveform data [39]. The practical

use of this hypotension prediction algorithm was expanded in a separate study on patients that underwent elective, noncardiac surgery [121]. A clinician-derived interpretation of the underlying physiological cause of instability was produced when the algorithm-predicted risk of hypotension exceeded a devised threshold. The algorithm predicted hypotension with an AUC-ROC of 0.81 at 15 minutes ahead of instability, and alarms produced by the system incited clinical action that reduced incidence of hypotension in the surgical population. Yoon et. al developed a hypotension prediction model using featurized, multigranular intensive care unit (ICU) data. The model had an AUC-ROC of 0.93 fifteen minutes prior to hypotension [127]. The model-produced risk scores were utilized within a continuous alert system that was designed to minimize alarm fatigue in the ICU setting. Other machine learning algorithms have been devised to accurately predict adverse events such as tachycardia and generalized instability in critical care settings using featurized, multigranular EHR data [128, 18]. Machine learning could potentially be used to forecast IDH risk real time, and model-produced risk prediction could be used to devise an early warning system for IDH. In this study, a prediction model of future IDH risk is developed, specifically by training and testing a random forest classifier using EHR data from a population of patients receiving hemodialysis at UPMC. Model-produced risk scores available within a small time frame around dialysis initiation were utilized to derive an early warning system for IDH. The system is conceptualized to provide an alert of impending IDH to a clinician, who could then administer treatment preemptively to prevent occurrence of IDH.

3.2 Methods

3.2.1 Source Data and Data Processing

The study used EHR information collected from patients receiving hemodialysis at University of Pittsburgh Medical Center (UPMC) acute care facilities from June 2016-June 2020. The multigranular EHR data, which included vital signs, labs, intake and output, administered medications, and clinical notes, was used to develop features for a

predictive model of future risk of IDH. A cohort of 276 patients that underwent a total of 1685 hemodialysis sessions were selected that had adequate availability of EHR information for risk model featurization. A clinical definition of IDH, defined as SBP<90mmHg and MAP<65mm Hg, was used to flag a total of 95 (34.4%) patients and 253 (15.0 %) treatments that experienced IDH. Table 3.1 and Table 3.2 describe the study demographics and treatment specific information, respectively. Overall, the patients were mostly white males and had non-Hispanic or Latino ethnicity. Patients that experienced IDH were older in age (67 vs. 62, p=.0072), had longer hospital stays (17.5 days vs. 12 days, p=.0015) and exhibited higher in-hospital mortality (30.5% vs. 15.5%, p=.0033). Ethnicity was also found to be significantly different between the two groups, however, over 15% of the patients declined or failed to specify their ethnicity. The median time of treatment and IDH occurrence was 210 minutes and 54 minutes after dialysis initiation, respectively.

Table 3.1: Patient Demographics

Characteristics		**All Patients**	**IDH (n=95)**	**No IDH (n=181)**	**P-values**
Age, years [median (IQR)]		64 [53-71]	67 [59,73]	62 [49,70]	.0072
Male [count (%)]		159 (57.6)	51 (53.7)	108 (59.7)	.34
Admission Weight, kg [median (IQR)]		80.3 (68-105)	79.0 [66.0,104.0]	82 [70.0-105.0]	.71
Race [count (%)]					.13
	White	183 (66.3)	70 (73.7)	113 (62.4)	
	African American	61 (22.1)	18 (18.9)	43 (23.8)	
	Other	32 (11.6)	7 (7.4)	25 (13.8)	
Ethnicity [count (%)]					.0096
	Hispanic or Latino	3 (1.1)	0 (0.0)	3 (1.7)	
	Not Hispanic or Latino	230 (83.3)	89 (93.7)	141 (77.9)	
	Declined	16 (5.8)	2 (2.1)	14 (7.7)	
	Not Specified	27 (9.8)	4 (4.2)	23 (12.7)	
Hospital length of stay, days [median (IQR)]		14.0 (7.0-27.5)	17.5 (11-34)	12 (5-21)	.0015
In Hospital Mortality [count (%)]		57 (20.7)	29 (30.5)	28 (15.5)	.0033

Table 3.2: Treatment Information

Characteristic	**Value**
Treatments with Incidence of IDH [count (%)]	253 (15.0)
Time of IDH After Treatment Start, minutes [median(IQR)]	54.2 [28.0,113.5]
Treatments Originating from ICU [count (%)]	558 (33.1)
Length of Treatment, minutes [median(IQR)]	210.0 [176.0, 235.0]
Total Ultrafiltrate Removed, mL [median (IQR)]	1997.0 [1135.5,2549.0]
Estimated Dialysis Fluid Removal Goal , mL [median (IQR)]	1825.8 [1111.0,2430.0]
Treatments with Fluid Removal Goal Achieved, mL [count (%)]	1082 (64.2)
Initial Ultrafiltration Rate, mL/min [median (IQR)]	8.2 [2.7,12.4]
Increase Ultrafiltrate, mL/min [median (IQR)]	4.5 [2.6, 8.1]
Decrease Ultrafiltrate, mL/min [median (IQR)]	4.5 [2.7,7.2]
Treatments with Vasopressors Administered [count(%)]	76 (4.5)
Treatments with Mannitol Administered [count (%)]	153 (9.1)
Initial Vasopressors Dosage, mcg/kg/min [median (IQR)]	.05 [.02,.08]
Increase Vasopressors, mcg/kg/min [median (IQR)]	.05 [.02,.16]
Mannitol Dosage, gm/mL	.25

Waveform data from heart monitors (telemetry data) and arterial lines was available for a subset of dialysis treatments but was not used in the study. Laboratory measurements used for analysis were chosen based on frequency and availability. These measurements result from basic metabolic panel (BMP) and comprehensive metabolic panel (CMP) blood tests. Vital signs, recorded by clinicians throughout treatment, included MAP, SBP, diastolic blood pressure (DBP), oxygen saturation, and heart rate. Physiologically implausible vital sign values were removed that met the following definitions: SBP, MAP, and DBP <10 mmHg and >300mmHg, heart rate<10 beats/min or >300 beats/min; these removed measurements were subsequently replaced using the average of the two nearest measurements. Additional vital sign features, such as minimum, maximum, slope, and linear weighted moving average

(LWMA), were derived using the past 30 minutes of vitals time series data. Clinical notes included past medical history, reason for admission, current patient state and an ongoing care plan; these notes were used to derive features describing underlying disease of patient. A total of n=89 static and dynamic patient features were obtained or computed each minute; when data availability was limited, the last observation carried forward (LOCF) approach was used to impute missing feature values. A complete list of features computed and used for risk prediction are available in Table A.1.

3.2.2 Random Forest Algorithm

A random forest classifier uses collection of decision trees to make a class prediction for a sample in a dataset [12]. A single decision tree recursively partitions a dataset into subsets of data using entropy, E, to maximally separate samples belonging to k possible classes [53]. The entropy of the data at the parent tree node, S, is calculated (Equation 3.1). Information gain, IG, resulting from splitting the set of data at a parent tree node, S, into subsets (children) S_1 and S_2 using feature F is evaluated using each feature and all possible split points for the features (Equation 3.3) [30, 11]. $P(C_i, S)$ is the proportion of samples in S belonging to class C_i, and $E(F, T; S)$ is net entropy resulting from splitting S using threshold T and feature F (Equation 3.2). The optimal decision threshold at a parent tree node is chosen based as the one that maximizes IG.

$$E(S) = -\sum_{i=1}^{k} P(C_i, S) log P(C_i, S) \tag{3.1}$$

$$E(F, T; S) = \frac{|S_1|}{|S|} E(S_1) + \frac{|S_2|}{|S|} E(S_2) \tag{3.2}$$

$$IG = E(S) - E(F, T; S) \tag{3.3}$$

To prevent overfitting, decision tree partitioning is halted based on a predefined criterion (e.g. maximum depth of tree) . Alternatively, the decision tree is post-pruned after partitioning. A sample is classified to one of the k classes by first traversing the trained tree from the root node (origin) to a terminal node (leaf) using the optimized decision thresholds at each

node. Then, the class majority of the samples in the leaf is used to assign the class of the sample. Alternatively, predicted class probability can be deduced using the proportion of samples in the leaf belonging to the class. In a random forest, a set of decision trees grown are grown using bootstrapped datasets derived from the original dataset, and only a random subset of features are considered for splitting each tree node. These modifications reduce high variance exhibited by a single decision tree and create an uncorrelated set of trees for prediction. The predicted class probability by a random forest is the average predicted class probability returned by the trees in the forest.

3.2.3 IDH Prediction Model

A predictive model of future risk of IDH was developed by training and testing a random forest classification model using n=253 hypotensive hemodialysis sessions (HS) and n=1432 nonhypotensive hemodialysis sessions (NHS). Scikit-learn package 0.24.1 in Python 3.8.5 was used to train and test random forest models for hypotension prediction. Featurized data samples, available at each minute of hemodialysis, were randomly selected from NHS and prior to IDH onset in HS to train and test the model (Figure 3.1).

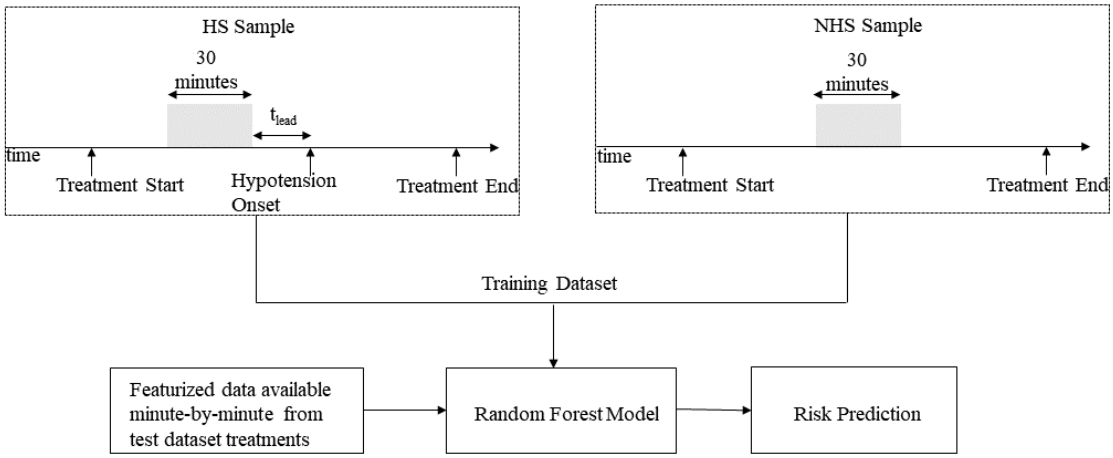

Figure 3.1: Risk model training design and production of risk scores

Selected HS (n=253) and NHS samples (n=1432) were split into a train (67%) and test set (33%) in stratified fashioned. Model parameters were tuned via stratified k-fold

cross validation (k=3) using the train dataset, and the average area under the receiving operating curve (AUC-ROC) and precision recall curve (AUC-PR) was computed across the validation folds. Briefly, stratified k-fold cross-validation splits the train dataset into k subsets, retaining the class distribution of the original dataset (stratified) during partitioning. The model is trained using k-1 folds, and the model performance is evaluated using the remaining (validation) fold. This process is repeated k times so each of the folds are used for validation once. Model parameters of the random forest were tuned by repeating stratified k-fold cross-validation (k=3) on the train dataset using a set of possible parameter values. Tuned random-forest model parameters included maximum depth of tree, number of features considered to split each tree node, and number of trees in the forest. Entropy was chosen to evaluate the split quality at each tree node. The performance of the tuned model that exhibited maximum average AUC-ROC on validation folds was further evaluated on the held out test set. To evaluate the effect of lead time on model prediction performance, training and testing was repeated using HS samples with lead times of 0, 5, 15, and 30 minutes to IDH onset. The importance of candidate features in IDH prediction was derived with the information gain that results from using the feature for tree partitioning, averaged across all of the decision trees in the forest.

The trained risk model produced an absolute risk score (AR) between 0 and 1 and represents the predicted probability of future IDH utilizing the prior 30 minutes of featurized clinical data. The probability of IDH of a featurized sample produced by a single tree in the forest was calculated as the proportion of the samples belonging to the hypotensive class in the assigned leaf over the total number of samples in the tree leaf. The probability of future IDH produced by the random forest model was calculated as the average probability produced by the trees in the forest. The model's ability to produce clinically relevant dynamic risk patterns was determined by producing minute-to-minute risk trajectories in the two-hour timespan leading up to IDH for HS (n=84) and NHS (n=473) from the test dataset. Specifically, the trained random forest model was applied to minute-to-minute featurized raw data for hemodialysis sessions in the test dataset (n=557), providing dynamic absolute risk (AR) of IDH. Time course AR were transformed to relative risk scores (RRS) that reflect IDH risk relative to the average AR of a NHS (Equation 3.4). By this definition, the average

RRS of a NHS is equal to 1.

$$RRS = \frac{AR}{Average\,AR\,in\,NHS} \tag{3.4}$$

The hypothetical IDH onset time for NHS was chosen such that the distribution of IDH onset times, relative to the start of hemodialysis, matched the distribution of real IDH onset times from HS. Additional risk trajectories were produced for HS and NHS to evaluate the evolution of risk beginning 30 minutes before the start of hemodialysis until the time of IDH. The model's capacity to produce evolving risk in response to clinical features was assessed on an individual treatment-by-treatment basis.

3.2.4 Derivation of Early Warning System for IDH

An early warning system was developed using the projected risk of IDH produced by the random forest model. A baseline early warning system was derived using RRS calculated at time of initiation of IDH. A ROC curve was produced to evaluate the tradeoffs between the true positive rate (TPR) and false positive rate (FPR) using different decision thresholds for classification. An optimal threshold was selected using the maximum Youden's J Statistic, where c is the threshold that maximizes the statistic (Equation 3.5) [129]. The statistic is maximal when sensitivity (TPR) and specificity or true negative rate (TNR) are 1 and assumes the two measures are equally weighted.

$$J = max_c(TPR_c - FPR_c) \tag{3.5}$$

The optimal threshold c was then used to classify a patient into high or low IDH risk class at hemodialysis initiation. Specifically, if the $RRS < c$, the patient was classified as low IDH risk and if $RRS \geq c$, the patient was classified as high IDH risk. TPR and FPR were calculated to evaluate the performance of the baseline classifier at the optimized threshold. A second early warning system was developed to evaluate whether additional risk score features, derived from risk trends within 15 minutes of dialysis initiation, could improve classification performance. Treatments with risk score availability up to 15 minutes after treatment initiation were included for analysis. Risk score features derived included

minimum, maximum, mean, and standard deviation of RRS available within 15 minutes of dialysis initiation. A small decision tree was applied to derived risk score features and utilized to classify patients as high or low IDH risk, where the threshold that maximizes Youden's statistic was used to make a classification decision.

3.3 Results

3.3.1 IDH Prediction Model

The prediction performance and importance of features in IDH prediction were evaluated for each of the candidate models with differing lead-times to IDH. Table 3.3 describes the top features of derived risk models as a function of the lead time parameter. Most important features across each of the models were derived from blood pressure measurements. Labs such as white blood cell count (WBC), prothrombin time (PT), red cell distribution width (RDW), and prior incidence of hypotension became increasingly important with longer lead time to hypotension.

Table 3.3: Top Features of Random Forest Model as Function of Lead Time

Lead Time (minutes)	**0**	**5**	**15**	**30**
Features	MAP	SBP_min_30	SBP_minimum_30	SBP_LWMA_30
	SBP	SBP_LWMA_30	SBP_LWMA_30	SBP_minimum_30
	SBP_minimum_30	MAP_minimum_30	SBP_max_30	SBP
	MAP_LWMA_30	SBP	SBP	MAP_minimum_30
	SBP_LWMA_30	SBP_max_30	MAP_minimum_30	SBP_maximum_30
	MAP_minimum_30	MAP_LWMA_30	MAP_LWMA_30	MAP_LWMA_30
	DBP	MAP	MAP	MAP
	SBP_wslope_30	MAP_max_30	MAP_maximum_30	MAP_maximum_30
	MAP_wslope_30	DBP_minimum_30	DBP_minimum_30	DBP_minimum_30
	SBP_uwslope_30	DBP_LWMA_30	DBP_LWMA_30	DBP_LWMA_30
	DBP_wslope_30	Prior_ID_Hypotension	Prior_ID_Hypotension	DBP
	DBP_minimum_30	DBP_maximum_30	WBC	DBP_maximum_30
	DBP_LWMA_30	DBP	PT	Prior_ID_Hypotension
	SBP_cv_30	WBC	DBP	WBC
	MAP_uwslope_30	RDW	DBP_maximum_30	RDW

*abbreviations: _30: of past thirty minutes of measurements, LWMA: linearly weighted moving average, wslope: slope from weighted linear regression, uwslope: slope from unweighted linear regression, Prior_ID_hypotension: prior incidence of intradialytic hypotension

The performance of each candidate model was evaluated on the test set using AUC-ROC and AUC-PR metrics. The model with a lead time of 0 minutes (lead-0) performed best out of all the tested models, with AUC-ROC=1.0 and AUC-PR=1.0. The performance of trained models with longer lead-times to IDH varied. To describe, trained lead-5 and lead-30 model had AUC-ROC of 0.90 and 0.89 and AUC-PR of 0.68 and 0.64. The lead-0 model was chosen for further evaluation. The trained lead-0 model was applied to featurized data available every minute leading to IDH onset, and model's ability to discriminate between future IDH and non-IDH over time was evaluated. The AUC-ROC evolved from 0.74 two

hours prior to IDH to 1.0 at IDH onset; the model could discriminate between impending IDH and non-IDH far ahead of blood pressure instability, and discrimination increased as the model was applied closer to IDH onset (Figure 3.2). The final, tuned random forest model (lead-0) parameters were max tree depth=2, number of features considered to split each tree node=17, number of trees in random forest=100.

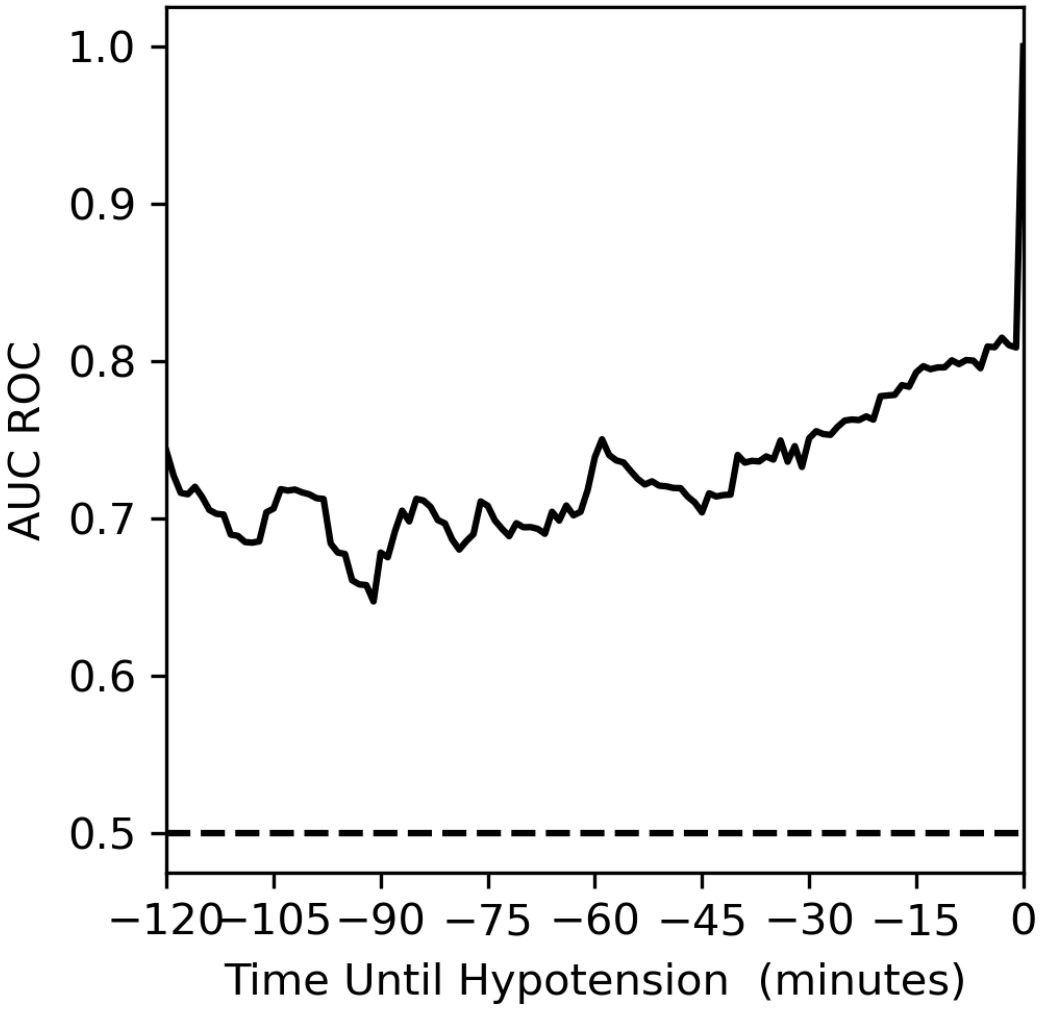

Figure 3.2: AUC-ROC evolution over time for IDH prediction model

Minute-to-minute risk trajectories were generated by applying the trained random forest model on the hemodialysis sessions in the test dataset (n=557). Figure 3.3 depicts the evolution of *RRS* for HS and NHS beginning two hours ahead of IDH onset. There is

distinct separation in average *RRS* (p<.05, two-sample t-test) between the HS and NHS even 2 hours ahead of event, where the average *RRS* of HS is 4 times higher than NHS. The risk separation between the two groups increases leading up to the time of IDH. On average, the *RRS* of HS elevates beginning around 90 minutes prior to IDH and escalates again around 30 minutes prior to IDH. The *RRS* of NHS remains low for the 2-hour period leading up to hypothetical IDH onset.

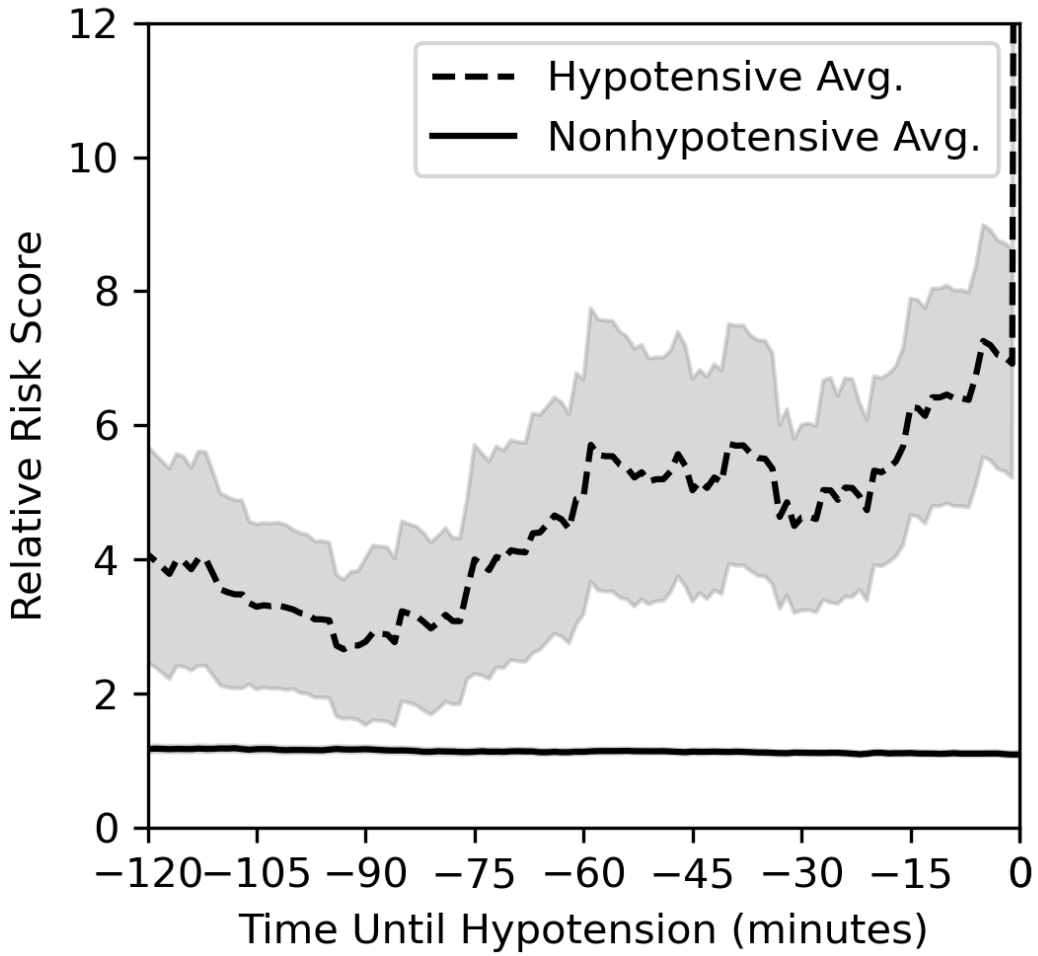

Figure 3.3: Mean relative risk trajectory before event shown for n=84 hypotensive dialysis sessions (dashed line) and n=473 nonhypotensive dialysis sessions (solid line). Shaded gray areas represent 95% confidence interval

Risk trajectories were generated beginning 30 minutes prior to hemodialysis initiation for HS and NHS in Figure 3.4 . There is distinct separation in average RRS ($p<.05$, two-sample t-test) between the HS and NHS throughout the observed period. HS and NHS experience distinct risk evolution from hemodialysis baseline. For HS, average RRS evolves from baseline and elevates over the observed period as continuous fluid removal induces hemodynamic instability on the average HS. Interestingly, on average, NHS experience a slight elevation in RRS at the start of hemodialysis; however, RRS remains relatively stable afterward. This may reflect the physiological stress induced on all patients upon ultrafiltration initiation. The risk trajectories in Figure 3.3 and Figure 3.4 are clinically relevant descriptions of the evolving risk of the average hemodialysis patient with and without impending IDH. Hence, there is a distinct difference between the model generated risk trajectories for HS and NHS.

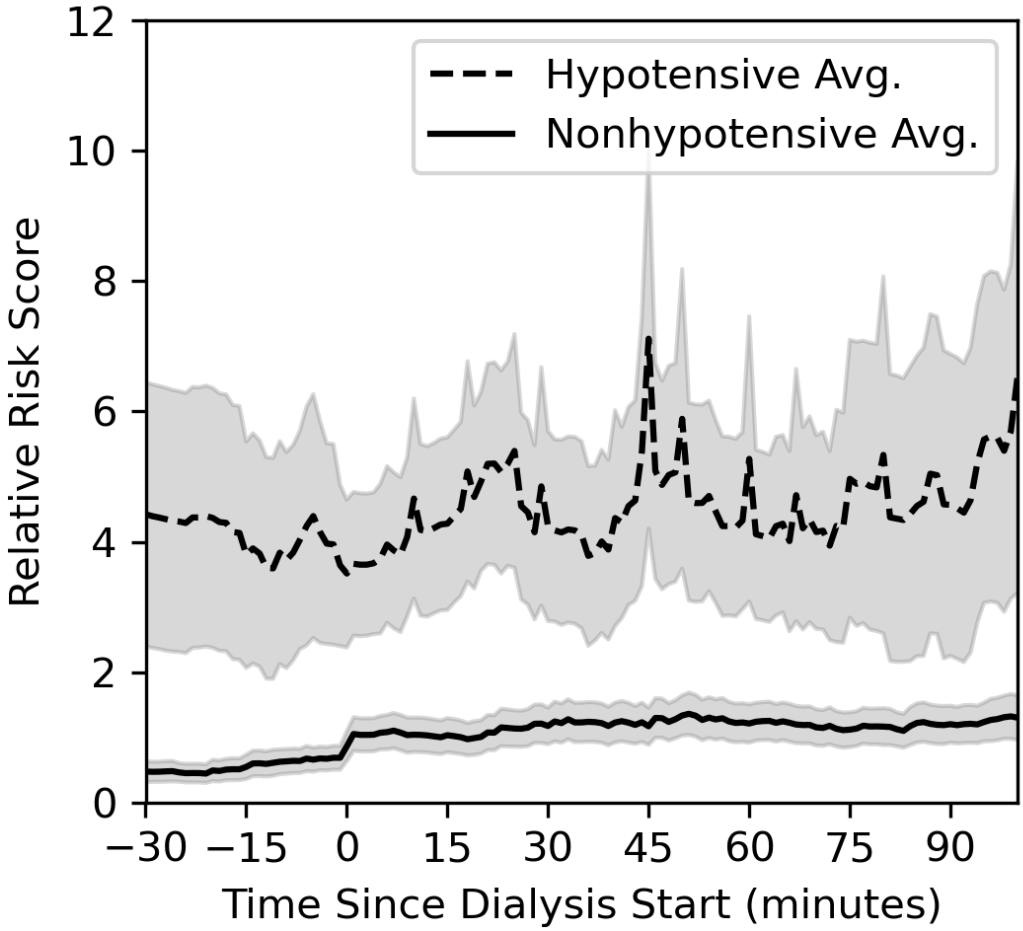

Figure 3.4: Mean relative risk evolution beginning 30 minutes prior to start of dialysis shown for n=84 hypotensive dialysis sessions (dashed line) and n=473 nonhypotensive dialysis sessions (solid line). Gray areas represent 95% confidence intervals.

Figure 3.3 and Figure 3.4 depict average *RRS* evolution of NHS and HS towards IDH and within an observed timeframe around dialysis initiation. The evolution of *RRS* with respect to clinical variables trends was evaluated on an individual treatment basis. The patient receiving dialysis in Figure 3.5 has relatively low *RRS* at baseline and experienced IDH over 3 hours after initiation. This model-produced *RRS* evolution is low and relatively stable until about two hours into dialysis, after which *RRS* begins to escalate toward the

IDH event in response to declining blood pressure trends.

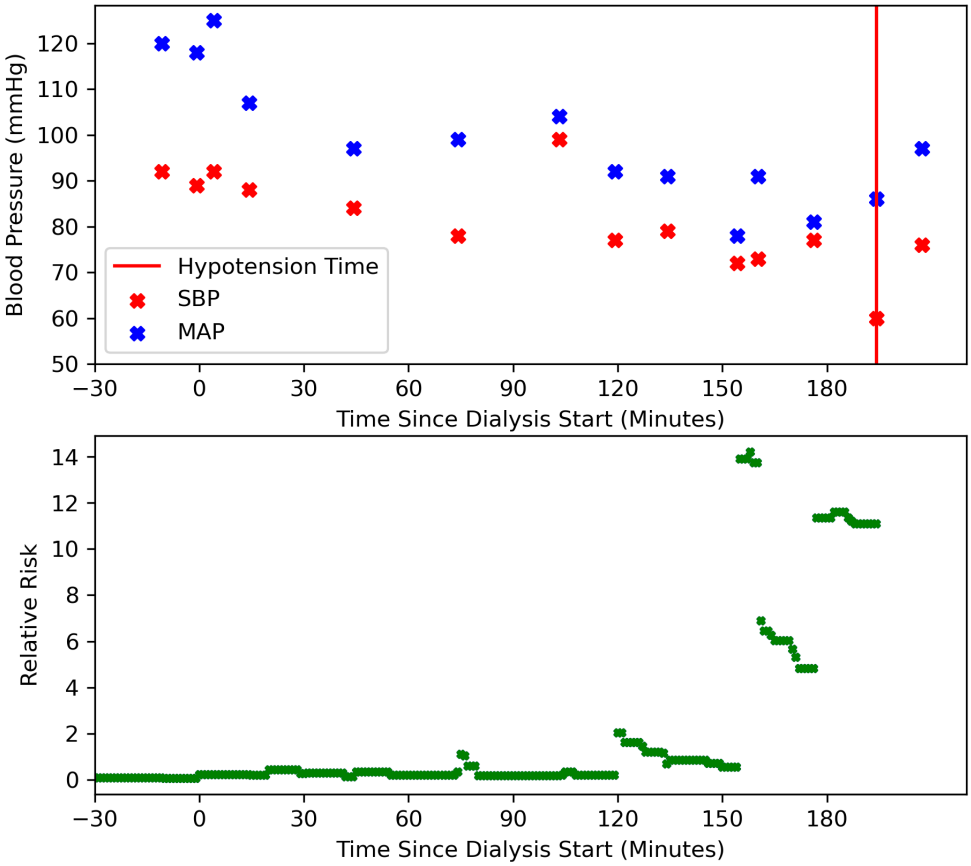

Figure 3.5: Example risk evolution during dialysis session with IDH Event

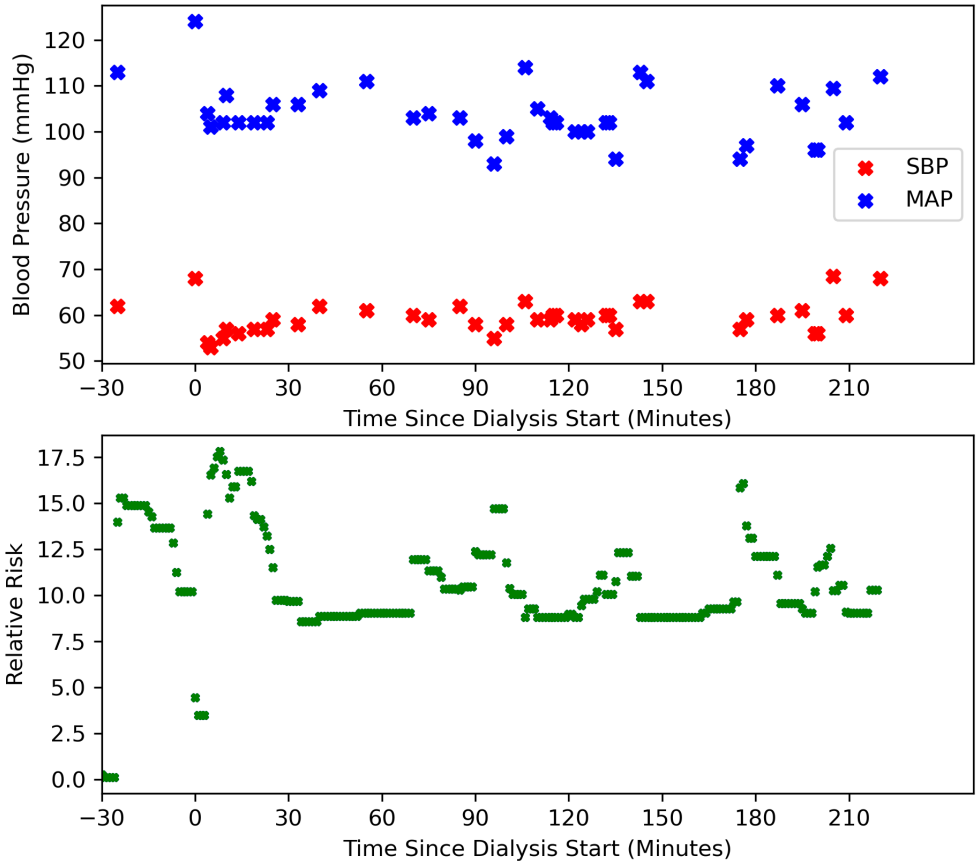

Figure 3.6: Example risk evolution during dialysis session without IDH Event

The patient receiving dialysis in Figure 3.6 is high risk at baseline yet does not experience an IDH event. The patient's *RRS* remains elevated throughout treatment. Nevertheless, the patient's blood pressure rapidly changed at times, and the model-produced risk trends reflect these dynamic changes in blood pressure.

3.3.2 Early Warning System

Features derived from risk trajectory information available within 15 minutes of dialysis initiation were used to develop an early warning classification system for impending IDH

during treatment. The baseline early warning system, using *RRS* available immediately prior to dialysis initiation, produced an AUC-ROC of 0.71 and AUC-PR of 0.41. The optimized threshold derived from maximizing Youden's J Statistic was used as a decision point to classify patients as high or low IDH risk; classification predictions producing a TPR of 0.58 and FPR of 0.21. The second early warning system used decision points derived from risk trajectory information available within 15 minutes of dialysis initiation Figure 3.7. This early warning system produced an AUC-ROC of 0.81 and AUC-PR of 0.54, indicating additional risk score dynamics improved the ability to distinguish between patients with and without impending IDH. Using the optimized threshold as a decision point, the classifier produced a TPR of 0.80 and a FPR of 0.29.

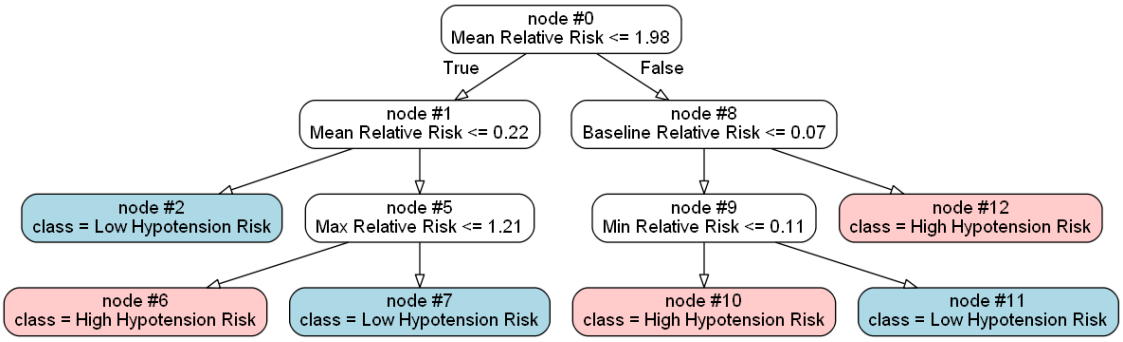

Figure 3.7: Early warning classifier using risk scores available within 15 minutes of dialysis initiation

3.4 Discussion

A data-driven prediction model of future risk of IDH was developed using machine learning, and a risk-score triggered early warning system was proposed that could be acted upon by clinicians at the bedside. The machine learning model produced risk of future IDH using the prior 30 minutes of featurized EHR data, and risk trajectories were produced by applying the model to collected minute-by-minute featurized data from HS and NHS available up to IDH onset. A warning threshold was derived using risk scores available

around hemodialysis initiation, and the threshold subsequently was used to classify patients as high or low risk of future IDH during the current treatment.

Risk trajectory analysis provided an interpretable description of the model's ability to generate IDH risk over time. While static variables, such as current laboratory measurements and vital signs, can solely be used to derive a risk score and predict future patient instability, these variables fail to capture dynamic physiological changes that precede hemodynamic instability; this could potentially result in inaccurate prediction and an inappropriate proceeding clinical response to the calculated score. For example, severity illness scoring systems have demonstrated to potentially misclassify patient illness severity as the calculated scores fail to utilize clinical variable evolution [24]. In this study, the risk prediction model leveraged variables describing vital sign dynamics over an interval of time (30 minutes) preceding prediction, and these variables revealed to be important IDH predictors. Produced risk trajectories illustrate significant separation in average IDH risk between NHS and HS throughout the observed two-hour interval preceding IDH, as well as the observed window around hemodialysis initiation. Separation in average IDH risk between the two groups became wider approaching IDH onset. The distinct separation in average risk scores between the two groups, even 2 hours ahead of event, further motivated the development of an early warning system for IDH that leveraged the model-generated risk scores. The early warning system was able to capture $> 80\%$ of future IDH within minutes of dialysis initiation. In response to a high IDH risk indication by the early warning system, a clinician could take preventative measures such as decreasing dialysate temperature, which acts to elevate blood pressure by increasing vascular resistance [27, 26]. Administration of hypertonic fluids such as mannitol [74] can prevent net fluid shifts from intravascular to intercellular space. Other preemptive measures such as sodium variation modeling and ultrafiltration profiling could also be used to reduce incidence of IDH [26].

Nevertheless, this study has several limitations. First, absent validation of the risk model using an external cohort of dialysis patients in a major limitation. Internal validation of the risk model was performed using a randomly separated test dataset. Features most important in predicting IDH were derived from blood pressure. While blood pressure is heavily utilized clinically for hemodynamic support decisions, other features derived from

high frequency, filtered waveform data may provide better risk prediction at the cost of reduced translation to patient settings where waveform data is infrequently acquired, i.e. outpatient dialysis facilities. It was assumed there was insignificant delay between recorded time of vital signs and the real measurement time. Clinical data is irregularly recorded; therefore, it is possible vital signs and laboratory dynamics leading to IDH or real IDH events could have been missed. In this study, a conventional definition of IDH used at UPMC was utilized to establish predicted binary classes of the model. Significant variation exists in IDH definition, using alternative blood pressure thresholds as well as existence of accompanied symptoms or use of nursing interventions [2]. Using an alternative definition of IDH could alter its associated significance with downstream consequences of IDH [32]. Furthermore, in the context of this study, a change in hypotension definition would adjust the predicted class labels and effect both IDH risk prediction and the clinical utility of the early warning system.

It is worth mentioning several neural network models have been recently developed that also provide accurate real-time IDH risk prediction [63, 5]. However, these studies provide insufficient guidance in regard to the implementation of the predictions to reduce incidence of IDH. The risk trajectories produced by the model in this study demonstrated the ability to forecast IDH far ahead of instability. The utility of the model-produced risk was extended by developing an early warning system for IDH that is simple and interpretable. The early warning system developed leveraged risk score trends available within 15 minutes of dialysis initiation and demonstrated improvement in prediction performance over the baseline early warning system. Nonetheless, 10% of the original hypotensive events were removed as these treatments exhibited IDH immediately proceeding dialysis initiation. In addition, we selected an operating threshold for classification that considered specificity and sensitivity of equal importance. This decision threshold can be tuned according to clinical interests or facility demands.

Variation in IDH risk dynamics at treatment-by-treatment level exists that is outside the common band of its respective NHS and HS subgroup. It is possible a larger set of IDH risk trajectory subtypes exist, and identification of these subgroups could create an enhanced and more personalized IDH detection system. Although the early warning

system supports the use of preventative treatments, study data evaluating the effect of these interventions on reducing IDH is limited and has mixed results [36, 73]. A higher degree of treatment personalization may be required to reduce incidence of IDH and its downstream consequences. In chapter 4, a reinforcement learning-based algorithm is developed that recommends preemptive, personalized treatment for hemodialysis patients based on the patient's risk trajectory.

4.0 Risk-Based Reinforcement Learning Algorithm for Hemodialysis

4.1 Background

Hypotension occurs in up to 30% of hemodialysis treatments and has been associated with increased mortality in various studies, including the one conducted in chapter 3 [59]. Several machine learning models have been developed to provide accurate real-time risk of future hypotension with clinical utility [128, 121, 63, 5]. Yoon et. al developed a machine-learning derived hypotension prediction model using multigranular ICU monitor data and designed a warning system that was conceptualized to provide clinicians real-time hypotension alerts while mitigating alarm fatigue. The Hypotension Prediction (HYPE) trial investigated the use of a hypotension recommendation system in a surgical setting; this consisted of a machine learning algorithm that utilizes arterial waveform data to predict hypotension and a clinician-derived interpretation of the underlying physiological cause of blood pressure instability [121]. The use of the early-warning system incited clinician action that resulted in decreased incidence of hypotension compared to standard care. In chapter 3, a machine learning model was developed that forecasted future risk of IDH real time, and an early warning system for IDH was developed using model-produced risk available around dialysis initiation. Still, no tool exists that prescribes dialysis treatment real-time based on patient risk of future IDH. The development of an autonomous control system that predicts future IDH and suggests *optimal*, *personalized* therapy could improve outcomes in the dialysis setting.

Reinforcement learning (RL) is an attractive methodology to learn an optimal policy, or way to act on a patient, to maximize the possibility of achieving desired patient outcomes. RL has been applied to several healthcare contexts, including treatment and management of sepsis, anesthesia, and cancer [58, 83, 84]. The RL agent (controller) receives feedback for its interactions with the environment (plant) in the form of rewards, with the goal to learn how to act to maximize expected reward over time [130]. It is both difficult and time consuming to accurately represent complex, diverse physiological responses to treatment with a mathematical model. Using RL, an optimal treatment policy is learned by the agent

without a model of the environment. An RL agent can learn an optimal policy using a dataset of sequential interventions and their respective responses that originated from a vast number of patients and treatments.

In this section, an RL-based algorithm is introduced in which the agent learns an optimal treatment policy for dialysis patients, utilizing the evolving patient risk of IDH and long-term achievement of individualized ultrafiltration goals as feedback for its actions. A model of future IDH risk was trained and evaluated specifically on a cohort of patients receiving intermittent dialysis in chapter 3. The model output was utilized to reward the agent based on alterations in patient IDH risk in response to dialysis interventions, and a set of states with a projected risk were derived to describe environment of the Markov Decision Process (MDP). Q-learning, a reinforcement learning algorithm, was used to solve for the optimal policy of the MDP, and microsimulations of dialysis trajectories were produced to evaluate the efficacy of the agent-derived treatment policy relative to the clinician policy.

4.2 Methods

4.2.1 Theoretical Framework

4.2.1.1 Markov Decision Process

A Markov Decision Process (MDP) is a mathematical formalism of a discrete time, stochastic decision-making process. Specifically, a finite Markov decision process defined by a set of states, S, a set of actions $A(s_k)$ available for the states in S, a reward function R, and a probability transition function $P(s_{k+1}|s_k, a_k)$, which represents the environment dynamics or probability of transitioning from state $s_k \in S$ at timestep k to state s_{k+1} at timestep $k+1$ when action $a_k \in A$ is taken [107]. The dynamics of the MDP, $P(s_{k+1}|s_k, a_k)$, satisfy the Markov Property: the future depends current state and action and is conditionally independent of the previous states and actions. The current state is assumed to contain all relevant information needed to define the future.

Reinforcement learning-based algorithms solve a MDP, where an agent is tasked to learn

the optimal way to interact with the stochastic MDP environment in order to achieve a goal. The agent-environment interaction described by a MDP is demonstrated by Figure 4.1.

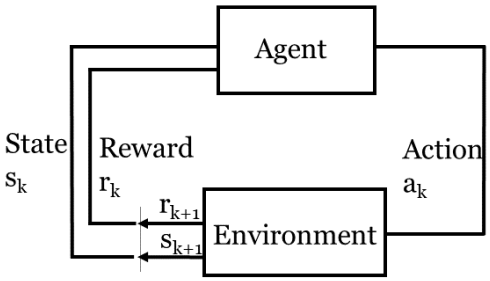

Figure 4.1: Reinforcement learning schematic

At each time step k, the agent interacts with the environment taking an action $a_k \in A$ based on the state $s_k \in S$ of the environment. The environment transitions to a new state s_{k+1} and emits a reward r_{k+1} that is used to provide feedback to the agent for the quality of a_k in s_k. The agent is tasked to maximize cumulative (discounted) reward over time (Equation 4.1).

$$G_k = r_{k+1} + \gamma r_{k+2} + \gamma^2 r_{k+3} + ... = \sum_{k=0}^{\infty} \gamma^k r_{k+1} \tag{4.1}$$

The discount factor, $\gamma \in [0, 1]$, describes present value of future reward or the horizon of the agent; a reward received k timesteps in the future is worth γ^k times less than if it were received immediately.

There is a theoretical relationship between reward discounting and the way an agent acts in an environment [105]. Future reward value is weighted by the probability of agent "surviving" at time t without experiencing a hazard (Equation 4.2).

$$v(r_t) = S(t)r(t) \tag{4.2}$$

Formally, the probability of hazard time T occurring at time t is described by the density function $f(t)$ (Equation 4.3), and probability of the hazard time T occurring at some time prior or at time t is $F(t)$ (Equation 4.4).

$$f(t) = lim_{\Delta t \to 0} \frac{P(t <= T < t + \Delta t)}{\Delta t} \tag{4.3}$$

$$F(t) = P(T <= t) = \int_0^t f(u)du \tag{4.4}$$

The survival function, $S(t)$, describes the probability that T exceeds time t and is the complement of $F(t)$ (Equation 4.5).

$$S(t) = P(T > t) = 1 - F(t) \tag{4.5}$$

The hazard function, $h(t)$, is described as the instantaneous probability of the hazard occurring at time t given that the hazard has not occurred (Equation 4.6).

$$\begin{aligned} h(t) &= \frac{f(t)}{S(t)} = \frac{1}{S(t)} \frac{dF(t)}{dt} \\ &= \frac{1}{S(t)} \frac{d(1 - S(t))}{dt} \\ &= \frac{-d \log(S(t))}{dt} = \frac{-ds(t)}{dt} \frac{1}{s(t)} \end{aligned} \tag{4.6}$$

Under the assumption that the hazard rate (λ) is constant, the survival function can be expressed in terms of the hazard rate, effective horizon time (τ) and discount factor (γ) (Equation 4.7). Equation 4.7 demonstrates that as hazard rate $\lambda \to \infty$, $\gamma \to 0$, and as effective horizon time $\tau \to 0$, $\gamma \to 0$.

$$v(r_t) = S(t)r(t) = e^{-\lambda t} = e^{-\frac{t}{\tau}} = \gamma^t \tag{4.7}$$

The value of a reward decays exponentially over time (Figure 4.2). As $\gamma \to 0$, the agent desires to maximize immediate reward, whereas the agent increasingly weighs future reward received as $\gamma \to 1$.

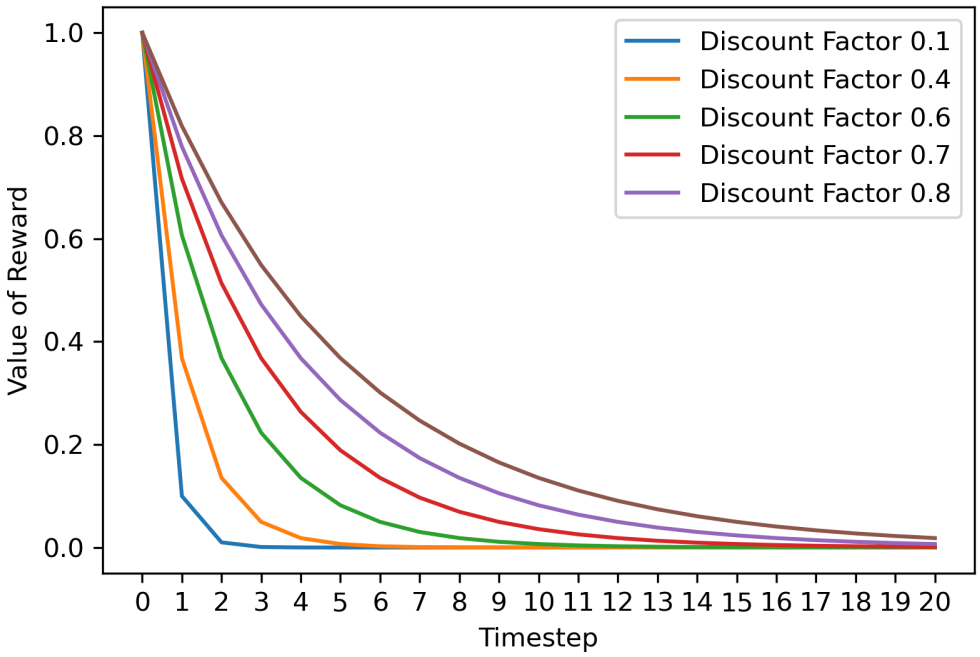

Figure 4.2: Value of Reward Over Time as Function of Discount Factor

4.2.1.2 Learning an Optimal Policy Using a Reinforcement Learning Agent

The relationship between cumulative (discounted) returns at successive timesteps can be written recursively (Equation 4.8).

$$\begin{aligned} G_k &= r_{k+1} + \gamma r_{k+2} + \gamma^2 r_{k+3} + ... \\ &= r_{k+1} + \gamma(r_{k+2} + \gamma r_{k+3} + ...) \\ &= r_{k+1} + \gamma G_{k+1} \end{aligned} \tag{4.8}$$

A policy, $(\pi(a|s))$, is a mapping of states to actions or the probability that $a_k = a$ when $s_k = s$. The relationship between the value of a state and its successor states under a policy

is described by (Equation 4.9), also known as the Bellman Equation. The value of a state is the summation of the expected discounted value of the next state and the reward received along the way.

$$\begin{aligned} v_\pi(s) &= \mathbb{E}_\pi[G_k|s_k = s] \\ &= \mathbb{E}_\pi[r_{k+1} + \gamma G_{k+1}|s_k = s] \\ &= \sum\nolimits_a \pi(a|s) \sum_{s',r} p(s', r|s, a)[r + \gamma v_\pi(s')] \end{aligned} \tag{4.9}$$

Similarly, state-action value function calculates the expected reward taking action a in state s and following the policy (π) afterwards (Equation 4.10). The value of a state-action pair is the summation of the expected discounted value of the next state-action pair and the reward received along the way.

$$\begin{aligned} Q_\pi(s, a) &= \mathbb{E}_\pi[G_k|s_k = s, a_k = a] \\ &= \mathbb{E}_\pi[r_{k+1} + \gamma G_{k+1}|s_k = s, a_k = a] \\ &= \sum_{s',r} p(s', r|s, a)[r + \gamma \sum_{a'} \pi(a'|s') q_\pi(s', a')] \end{aligned} \tag{4.10}$$

An optimal policy, π^*, is one in which the expected return is greater than any other policy for all $s \in S$. The value function and state-action value function for an optimal policy is called an optimal value function, $v^*(s)$, and optimal state-action value function, $Q^*(s, a)$, respectively. The value of a state under the optimal policy, $\pi^*(s)$, is equal to the expected return for the best action from that state (Equation 4.11).

$$\begin{aligned} v^*(s) &= \max_a q_{\pi*}(s, a) \\ &= \max_a \mathbb{E}[r_{k+1} + \gamma v*(s_{k+1})|s_k = s, a_k = a] \\ &= \max_a \sum_{s',r} p(s', r|s, a)[r + \gamma v^*(s')] \end{aligned} \tag{4.11}$$

Therefore, the bellman optimality equation for the state-action value function can be written as (Equation 4.12).

$$
\begin{aligned}
Q^*(s,a) &= \mathbb{E}[r_{k+1} + \gamma \max_{a_{k+1}} Q^*(s_{k+1}, a_{k+1}) | s_k = s, a_k = a] \\
&= \sum_{s',r} p(s', r|s, a)[r + \gamma \max_{a'} Q^*(s', a')]
\end{aligned}
\tag{4.12}
$$

$Q*$ is the cumulative discounted reward taking action a_k in state s_k and following the optimal policy afterward. If accurate description of the environment dynamics ($p(s', r|s, a)$) exists, dynamic programming techniques such as value iteration or policy iteration can be used to solve for $Q*$.

Q-learning, an off-policy reinforcement learning algorithm, can be used used to solve for the optimal policy of the MDP [120, 83, 62, 29]. The algorithm solves the bellman equation of optimality (Equation 4.12) without a model of the environment. As depicted in Figure 4.12, the reinforcement learning agent is external to the environment. The agent iteratively improves its policy by acting on the environment and observing the subsequent response of the environment to its actions in the form of rewards. In the case that agent learning cannot occur by interacting with the environment online, learning can occur offline using experience stored in a previously collected dataset. Specifically, Q is first initialized to zero. Then, tuples of experience are sampled from a static dataset $D = \{(s_k, a_k, s_{k+1}, r_{k+1})\}$ and the algorithm is updated using a running (weighted) average of current (Q_c) and new (Q_n) state action value information (Equation 4.13)

$$
\begin{aligned}
Q_c &= Q_{k-1}(s_k, a_k) \\
Q_n &= r_{k+1} + \gamma \max_{a_{k+1}} Q_{k-1}(s_{k+1}, a_{k+1}) \\
Q_k(s_k, a_k) &\leftarrow Q_c + \alpha(Q_n - Q_c)
\end{aligned}
\tag{4.13}
$$

The step size or learning rate, $\alpha \in [0, 1]$, weighs how much new information overrides old information during the algorithm update. As $\alpha \to 1$, the agent places larger weight on new experience during its policy update. Theoretically, Q converges to the optimal state value function $Q*$ if the state-action values are visited infinitely often (Equation 4.14) and the learning rate α is decreased appropriately (Equation 4.15).

$$\sum_{k=0}^{\infty} \alpha_k(s_k, a_k) = \infty, \forall(s, a) \tag{4.14}$$

$$\sum_{k=0}^{\infty} \alpha_k^2(s_k, a_k) < \infty, \forall(s, a) \tag{4.15}$$

Upon convergence, the optimal policy can be extracted by $Q*$ by taking the action that results in the maximum expected cumulative reward from the current state (Equation 4.16)

$$\pi^*(s_k) = \arg\max_{a_k} Q^*(s_k, a_k) \tag{4.16}$$

4.2.1.3 Q-learning Based Closed-Loop Risk Control of Intermittent Hemodialysis

Q-learning was used to develop an algorithm that recommends preemptive, personalized treatment for hemodialysis patients using the patient's IDH risk trajectory. The agent-derived policy identifies the optimal way to act on a patient to maximize the probability of achieving individualized fluid removal goals without incidence of IDH. In this section, the dialysis MDP and RL-based methods used to solve for the optimal policy are described.

4.2.2 MDP States

The states of the MDP were derived using a data-driven approach, using the patients' physiological features and their projected absolute risk (AR) calculated by the random forest model. As described in Source Data and Data Processing, the relative importance of candidate features in IDH prediction was derived with the information gain that results from using the feature for tree partitioning, averaged across all of the decision trees in the forest. Feature importances were sorted in descending order to evaluate the relative contribution of variables in IDH prediction (Figure 4.3). The features (n=8) with highest importance in IDH prediction were chosen to derive the state space, as feature importance decreased slowly after n=8 with the use of additional features. The variables with highest feature importance included: current-time SBP, MAP, and diastolic blood pressure (DBP), minimum and past

30 minute linear-weighted moving average of systolic blood pressure (SBP) and mean arterial pressure (MAP) measurements, and the slope of SBP over the past 30 minutes.

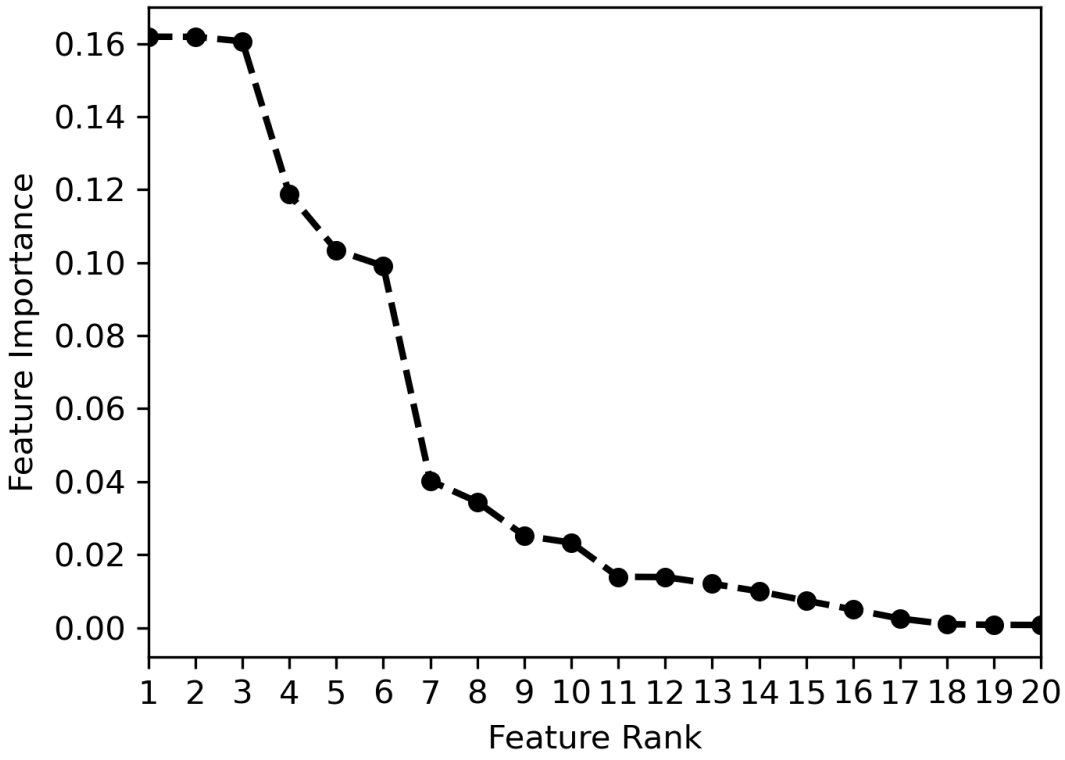

Figure 4.3: Feature Importance vs. Feature Rank

Relative risk (RR) and personalized risk (PR) were derived from the absolute risk (AR) projected from the model and used as additional variables to define the state space. RR represents the IDH risk relative to average IDH risk in NHS (Equation 3.4), while PR describes deviation in RR from patients own RR at hemodialysis initiation (RR_0) (Equation 4.17).

$$PR = RR - RR_0 \tag{4.17}$$

These 10 features were extracted from each treatment every 15 minutes from dialysis initiation and up to, but not including, the time of IDH onset.

Principal Component Analysis (PCA) was used to reduce the set of n=10 features to a new set of uncorrelated variables (principal components) that capture the majority of

the variance described by the original dataset [48]. Figure 4.4 demonstrates additional cumulative variance described by principal components diminishes after the use of four principal components. Four principal components were chosen to describe the data, as the components captured 90% of the original dataset variance.

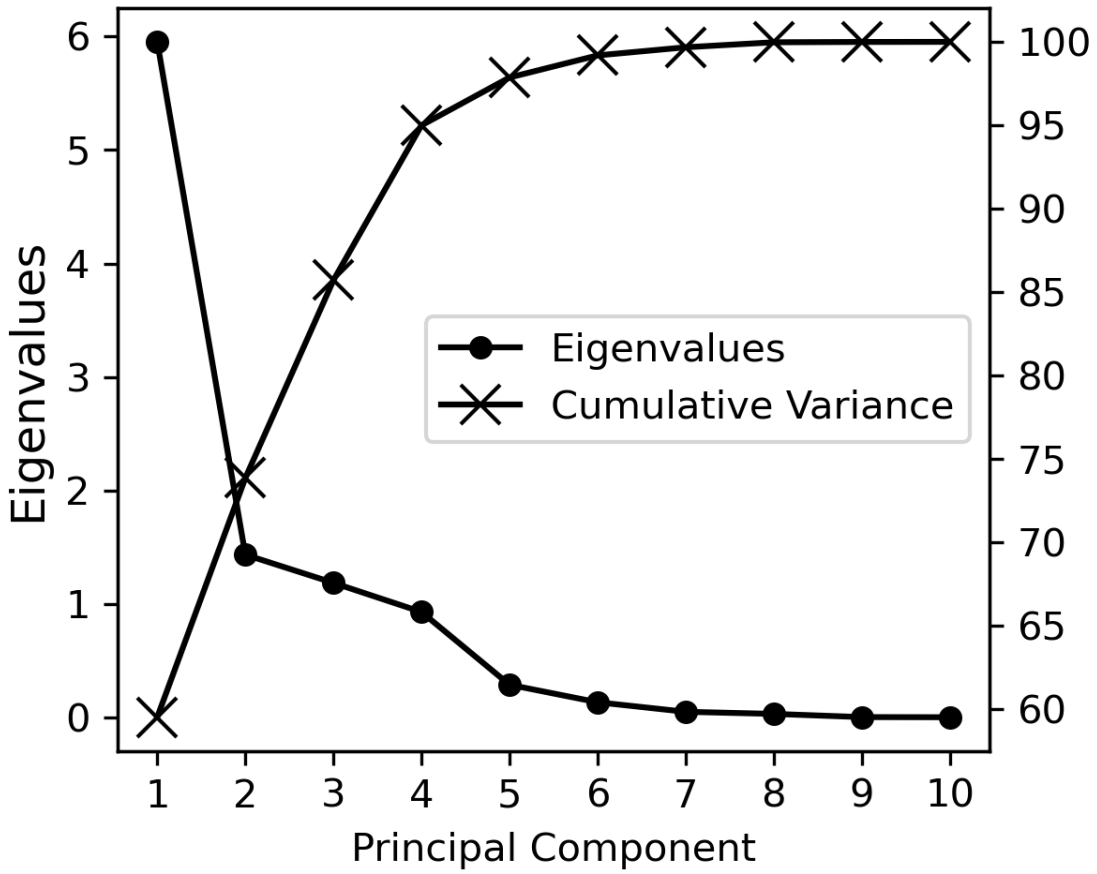

Figure 4.4: Eigenvalues and cumulative variance explained by the principal components.

A discrete set of states are required for a finite MDP, however, the selected variables are continuous. A decision tree was applied to the reduced-dimensionality dataset to derive boundaries between sampled variables with and without IDH experienced in one timestep (15 minutes)(Figure 4.5).

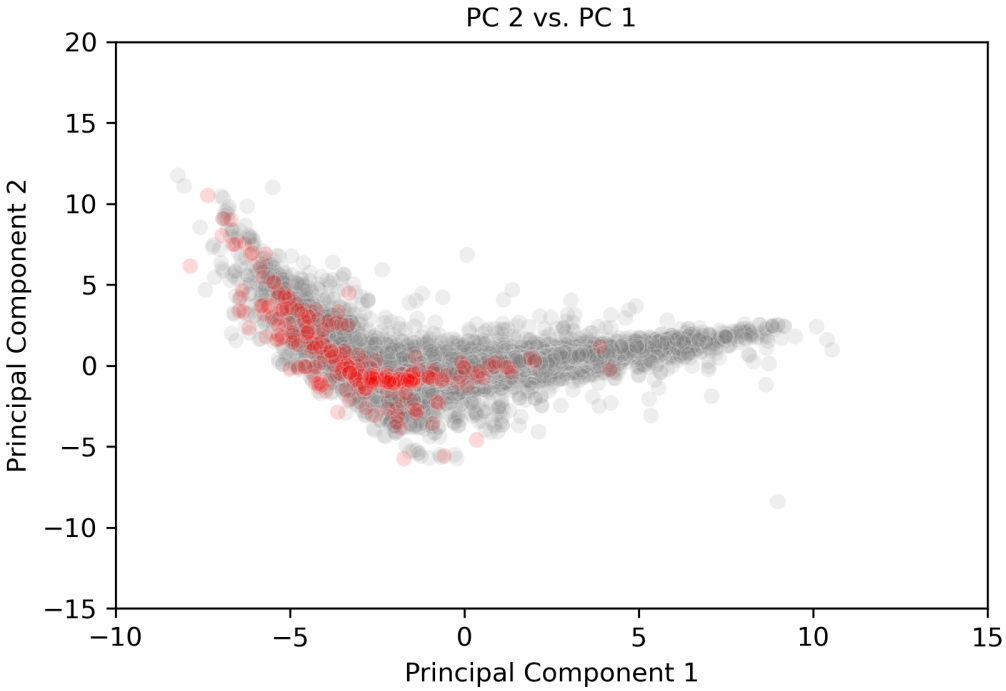

Figure 4.5: Projection of sampled features on first two principal components. Red and gray samples are samples with and without experienced IDH in one timestep, respectively

Specifically, a decision tree was grown with a max depth, or maximum path length from root node to leaves of 3. Then, the decision tree was simplified by pruning the tree from the bottom up, removing split points that resulted in lowest information gain (Equation 3.3). The boundaries created six regions with distinct clinical features and projected risk of hypotension (Figure 4.6).

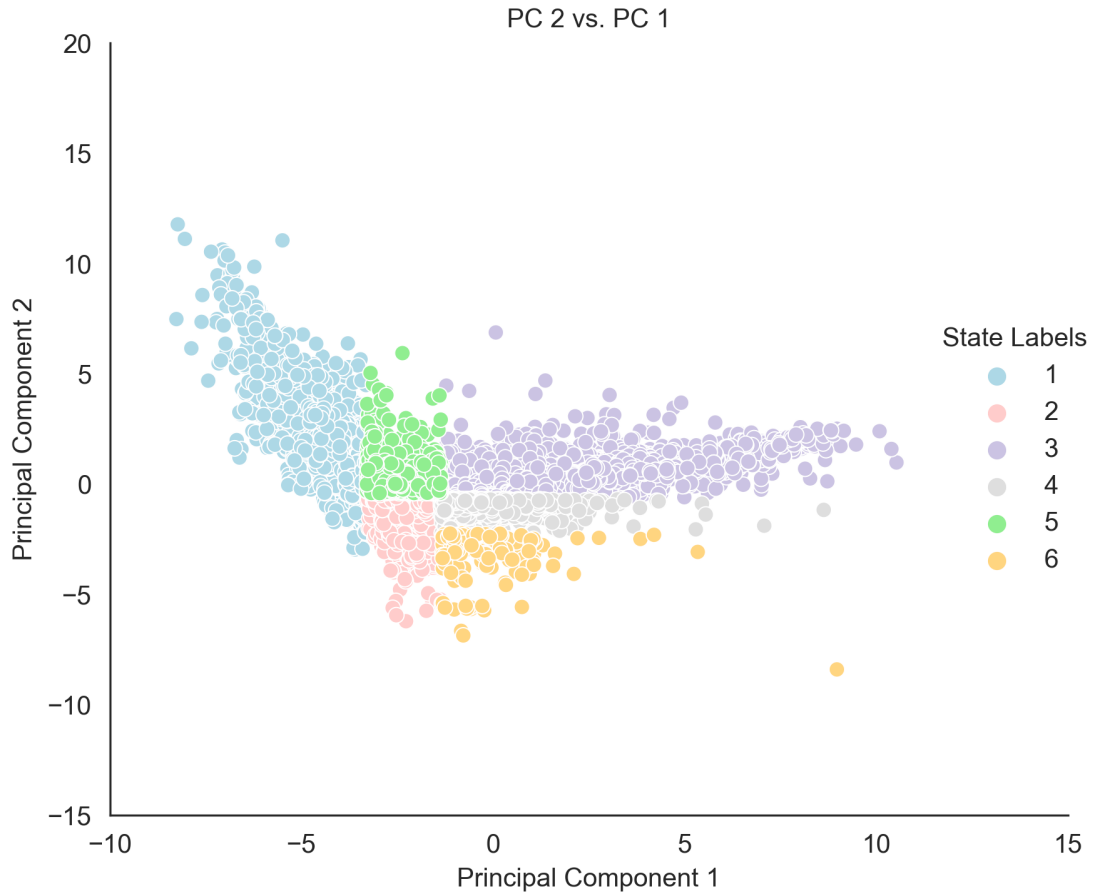

Figure 4.6: Projection of sampled features on first two principal components. Red and gray samples are samples with and without experienced IDH in one timestep, respectively

A hemodialysis patient was ascribed to a risk state (n=6) each timestep of dialysis using a PCA projection of their clinical variables and associated risk of future IDH. (Figure 4.7) .

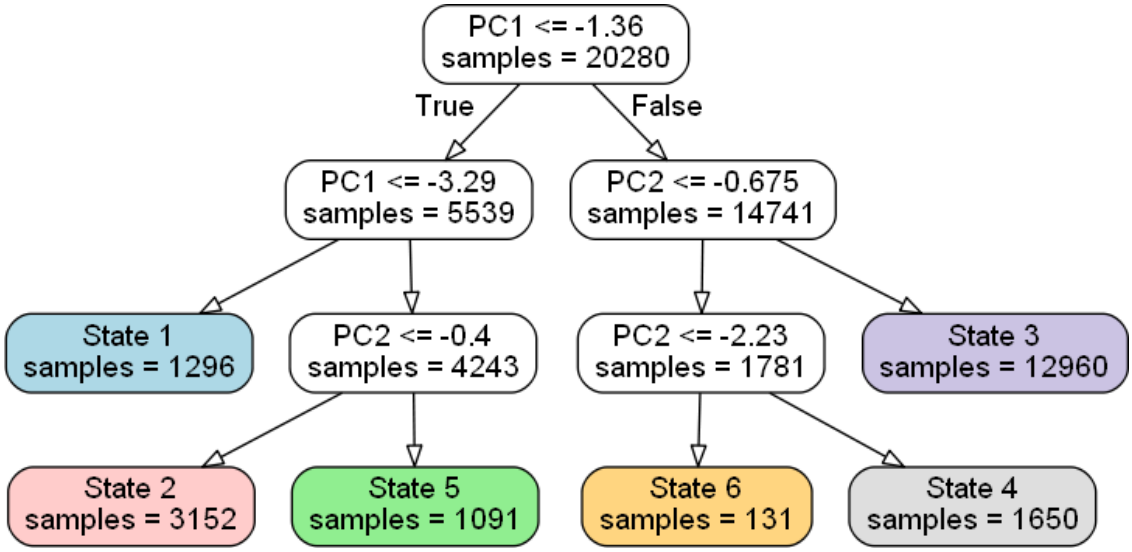

Figure 4.7: Decision Tree Description of Risk State Assignment. Patient is assigned a risk state in real time by traversing the root node (origin) to a terminal node (leaf) using decision thresholds at each node.

The state space was expanded (n=12) to include a discrete indicator of personalized fluid removal status. Fluid removal goals were estimated on a treatment-by-treatment basis using the initial ultrafiltration rate set by the clinician and calculating the cumulative fluid pulled if the rate was held constant for three hours (Figure 4.8). Fluid removal rate was integrated over time to determine the total volume removed from the patient up to the current treatment time. At each timestep during dialysis, a patient was ascribed a discrete volume status of [0-50%) or [50-100%) as a function of their personalized fluid removal goal. Fluid goals were satisfied if the volume removed exceeded the estimated fluid removal goal. Two terminal states were added to the state space that indicated if IDH occurred or fluid goals were achieved.

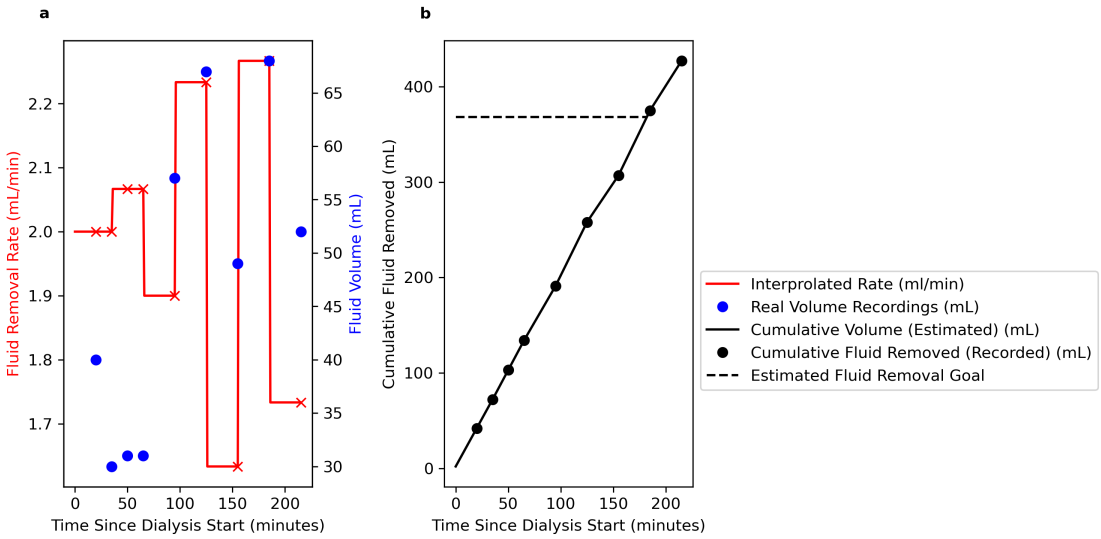

Figure 4.8: **a**. Fluid removal vs. time since dialysis initiation for individual patient. Blue and red depict actual recorded fluid volume and fluid removal rate over time, respectively. **b** Cumulative fluid removal vs. time. Dashed line added to depict estimated fluid removal goal during treatment for patient.

4.2.3 MDP Rewards

The objective is for the agent to learn preemptive, personalized dialysis treatment that migates IDH risk and maximizes possibility of achieving individualized fluid removal goals. The random forest model developed in chapter 3 provided projected risk of IDH real-time using clinical measurements and prior medical history. Personalized risk of IDH (Equation 4.17) provided change in risk relative to the patient's own baseline. An effective reward function should reward the agent for decreases in personalized risk of IDH and penalize the agent otherwise. Specifically, the agent is rewarded at timestep k if a patient's personalized risk of IDH decreased after taking action a_k on a patient in state s_k (Equation 4.18).

$$r_{k+1} = -RR_0 * [|PR_{k+1}| - |PR_k|] \tag{4.18}$$

At terminal timesteps of a patient's dialysis trajectory, the agent was rewarded if dialysis

fluid goals were achieved without incidence of IDH and penalized otherwise. The terminal reward magnitude was selected to be greater than intermediate rewards received at any given timestep.

4.2.4 MDP Actions

The set of actions taken by the agent included no intervention, increase ultrafiltration rate, decrease ultrafiltration rate, increase vasopressor dose, and administer mannitol or albumin. These clinical interventions were collected from EHR data available for each dialysis treatment (n=1685) in 15-minute intervals from dialysis initiation. Vasopressors used by clinicians included epinephrine, norepinephrine, dopamine, phenylephrine, and vasopressin: these were converted to their equivalent norepinephrine dosage (mcg/kg/min) [51]. 25% mannitol and 25% albumin were assumed to be bolused over 5 minutes from the time of recording. Albumin and mannitol administration was merged as a single action as both increase plasma osmolality and encourage vascular refill via osmosis. Interventions by pump and clinician were recorded as volumes in EHR; rate was assumed to be constant between two recordings and determined using time between previous and current recording. An increase or decrease was indicated by $\geq 25\%$ change in flow rate.

4.2.5 Offline Q-Learning to Learn Optimal Treatment

Q-learning was used to solve for the optimal policy of the MDP. Q was initially set to zero. Transitions from real dialysis trajectories were stored in a static dataset $D = \{(s_k, a_k, s_{k+1}, r_{k+1})\}$, and tuples of $(s_k, a_k, s_{k+1}, r_{k+1})$ were sampled to update the Q (Equation 4.13). The learning rate was decayed throughout the learning process using a linear learning rate schedule [29]. Discount factor is related to the horizon of the agent; as $\tau \to 0$, $\gamma \to 0$, less distant reward is used to calculate the optimal action in a state. Thus, discount factor(γ) was set to 0.1 to reflect a short agent horizon and urgency of the agent (clinician) to decrease patient risk of hypotension. Theoretically, Q-learning converges under the assumptions of Equation 4.14 and Equation 4.15. In practice, the algorithm was updated until the rewards in Q stabilized ($\Delta Q < 10^{-7}$) . Upon convergence, the optimal policy, $\pi^*(s_k)$, was extracted

from Q^* using (Equation 4.16).

4.2.6 Microsimulations

Microsimulations of the MDP using the RL-generated optimal policy were produced to evaluate its efficacy in reducing incidence of IDH while simultaneously achieving individualized fluid removal goals. Specifically, a dialysis patient was initialized one of 6 possible initial risk states according to the distribution of initial risk states found within the dataset. Simulation parameters were defined using real dialysis parameters set by clinicians during treatments in the dataset. The initial ultrafiltration rate of each simulated treatment was set as the median rate in the dataset (median=6.4 mL/kg/hr (IQR: 3.5 mL/kg/hr-9.6 mL/kg/min)). Patient weight was assumed as the median patient weight in the dataset (median=80.3 kg(IQR=68.0 kg-105.0 kg), and an increase or decrease in rate was set to 50% change in flow rate (median=43.2% (IQR: 32.8%-62.1%)). The maximum ultrafiltration rate allowable during simulation was set to the maximum rate observed in the dataset (44.09 mL/kg/hr). A total of 25,000 simulated dialysis treatments were produced by applying the optimal policy in Equation 4.16 and producing subsequent state transitions according to $P(s_{k+1}|s_k, a_k)$ until a terminal state was reached or a maximum of five hours of simulated treatment time was observed.

The optimal policy in Equation 4.16 is deterministic, and the optimal intervention in a state is the one with maximum expected reward. In cases in which vasopressors or mannitol was recommended as the optimal intervention, the optimal action and an alternate, less rewarding treatment intervention were chosen probabilistically according to a softmax distribution (Equation 4.19).

$$\frac{p_s(a_2)}{p_s(a_1)} = e^{(Q(s,a_2)+b-Q(s,a_1))} = e^{\Delta Q_{mod}} \tag{4.19}$$

$Q^*(s, a_2) < Q^*(s, a_1)$, and a_1 is the optimal action chosen in a state according to Equation 4.16. $P_s(a_2)$ and $P_s(a_1)$ are the probabilities of choosing a_1 or a_2 in state s. The bias term, b, was added so the alternative, less rewarding action in the state was chosen more frequently than the optimal action if vasopressor or mannitol was initially the recommended optimal

action. The parameter was tuned by sweeping b and creating n=25000 dialysis trajectories with each b value. Dosing behavior was evaluated as a function of increased bias toward use of alternative interventions. Simulated dosing patterns and dialysis outcomes of the optimal agent-suggested policy (Equation 4.16) and modified agent policy (Equation 4.19) used *in silico* were compared to the real clinical dosing patterns and dialysis outcomes of the clinician.

The analysis was repeated to evaluate the effect of agent-derived therapy on specific subpopulations (ICU and non-ICU based treatments). Specifically, probability transition functions ($P(s_{k+1}|s_k, a_k)$) were reconstructed from dataset transitions for both subgroups, and the simulation and analysis procedure defined above for the general population of patients was repeated for each subgroup. This analysis was performed to evaluate how the RL policy trained using all patient data performs on specific subpopulations.

4.2.7 Statistical Analysis

Normality of continuous data distributions was calculated via Shapiro-Wilk test. Two-sample t-test was performed for statistical comparisons of distributions that were both continuous and normal, while the two-sample Kolmogorov-Smirnov test was used to compare continuous, non-normal distributions. Normal, continuous distributions were reported with mean and interquartile range (IQR), while non-normal continuous distributions were reported with median and interquartile range. For categorical variables, chi-squared test was used for statistical comparison when <20% of cells had expected frequencies <5, otherwise, Fisher's exact test was implemented. Categorical variables were reported with counts and frequencies. P< .05 chosen as the threshold for significance. In text, significant comparisons with p< .0001 were indicated as such, but exact values can be found in Appendix .

4.3 Results

4.3.1 Optimal Agent-Suggested Treatment Policy vs. Clinician

Figure 4.9 depicts several example simulated trajectories following the optimal policy. The agent acts to maximize the probability of a dialysis patient reaching his or her fluid removal goal without incidence of IDH.

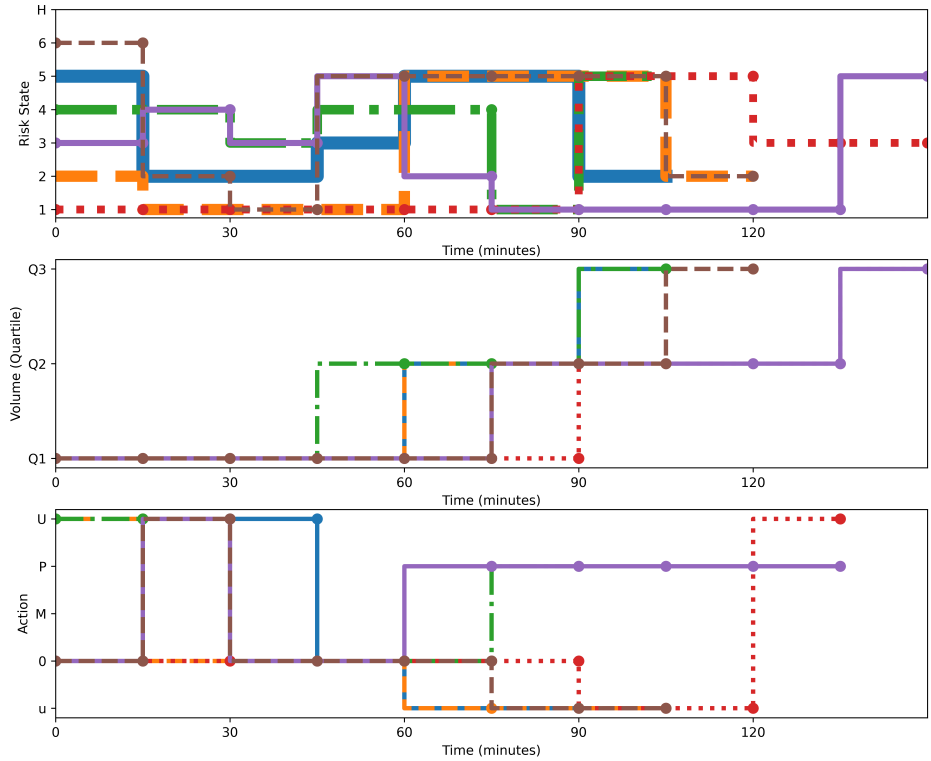

Figure 4.9: Example Simulated Dialysis Trajectories. Top Graph: Risk state of patient each simulated timestep. Middle Graph: % of personalized fluid goal. Q1:[0-50%), Q2:[50-100%), Q3: 100%. Bottom Graph: Actions recommended by agent each timestep. 0: no intervention, U: increase ultrafiltrate, u: decrease ultrafiltrate, action M: administer albumin or mannitol, P: increase vasopressors

The clinician and optimal interventions were collected each timestep of real and simulated dialysis sessions, respectively (Figure 4.10). As depicted in Figure 4.10, the clinician most frequently made no change in ultrafiltration rate, followed by increasing and decreasing ultrafiltration rate. Administration of albumin or mannitol and increase of vasopressor dose were infrequent interventions taken by clinicians during hemodialysis. The frequency of applied interventions was increased significantly during dialysis sessions *in-silico* relative to the clinician. To illustrate, the clinician and agent intervened 18.4% and 48.7 % of treatment timesteps, respectively. The agent altered ultrafiltration rate more routinely than the clinician; ultrafiltration was changed 32% of timesteps by the RL and 17.78% of timesteps by the clinician. Rare and more aggressive interventions (i.e administration of albumin or mannitol bolus or increasing vasopressor dose) were recommended more often by the RL agent relative to the clinician. Specifically, albumin/mannitol administration and escalation of vasopressor dosage were taken 28 times and 25 times more frequently *in-silico* than the clinician, respectively.

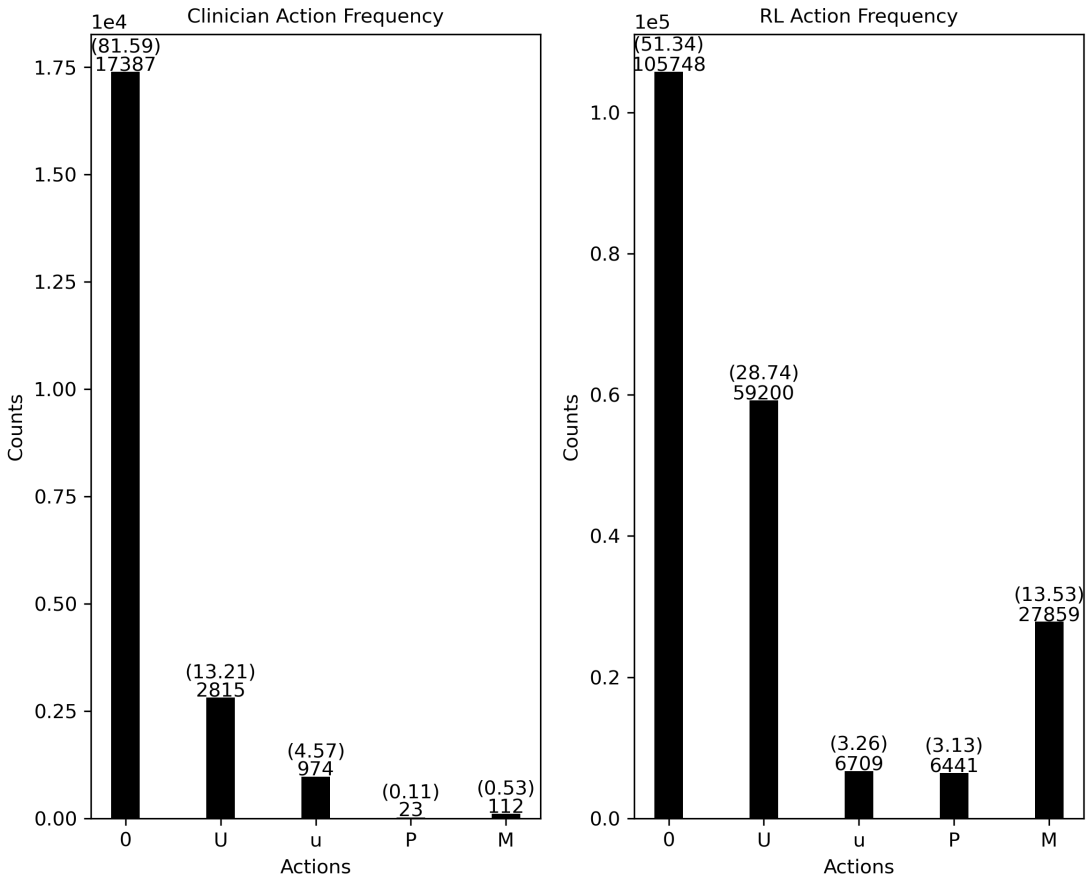

Figure 4.10: Clinician vs. Agent Recommended Interventions (optimal policy) . Actions selected by clinician every 15 minutes across all treatments (left). Actions recommended by agent every timestep during simulated dialysis treatments (right). Action 0: no intervention, action U: increase ultrafiltrate, action u: decrease ultrafiltrate, action M: administer albumin or mannitol, P: increase vasopressors

The dosing behavior of the agent (per treatment) following the optimal policy is depicted in Figure 4.11. Increase in vasopressor dose and administration of mannitol or albumin were actions taken a maximum of 12 and 11 times per treatment, respectively *in silico* (Figure 4.11). In comparison, increase of vasopressor dose and administration of mannitol or albumin were interventions taken a maximum of 6 and 4 times by the clinician per treatment of the dialysis sessions available in the dataset, respectively. Despite apparent overuse of aggressive treatments by the agent, the proportion of treatments with experienced

hypotension was reduced by 63.3% (p< .0001), and the proportion of treatments with achievement of fluid removal goals was increased by 47.4% *in silico* (p< .0001). While not a primary study outcome, the deployed agent-recommended policy resulted in accomplishment of fluid removal goals without incidence of hemodynamic instability faster than the clinician treatment policy (Table 4.1).

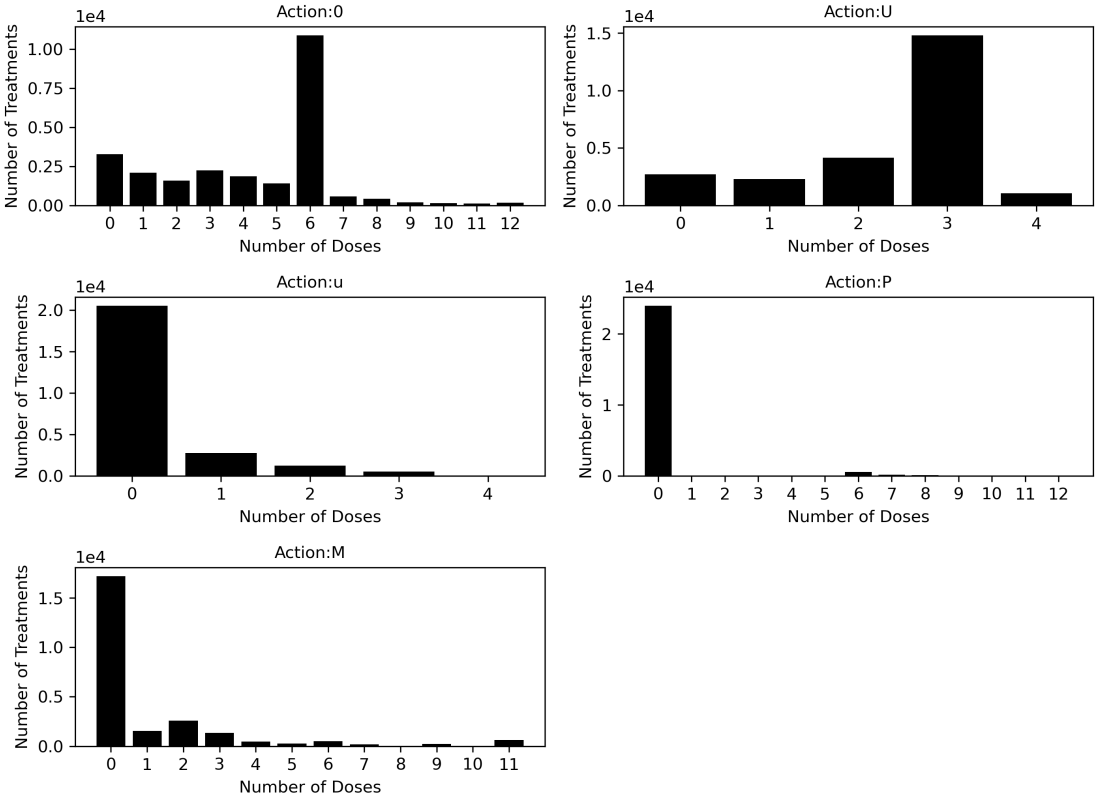

Figure 4.11: RL Agent intervention frequency per dialysis treatment (optimal policy). Action 0: no intervention, action U: increase ultrafiltrate, action u: decrease ultrafiltrate, action M: administer albumin or mannitol, P: increase vasopressors

Table 4.1: Outcomes of optimal treatment policy vs. clinician Policy

	IDH [count (%)]	**Fluid Goals Achieved [count (%)]**	**Time to Goal State (minutes) [median (IQR)]**
RL Treatment Policy	1364 (5.5)	23664 (94.6)	135 [105,140]
Clinician Treatment Policy	253 (15.0)	1082 (64.2)	176 [164,182]

4.3.2 Modified Agent-Suggested Treatment Policy vs. Clinician

The agent-derived (optimal) policy could potentially lead to improved patient outcomes at the expense of increased use of aggressive treatments. However, these rare and aggressive interventions are unlikely to be applied as frequently if recommended to a clinician. To illustrate, vasopressor support is only used for dialysis patients that originate from the intensive care unit (ICU) and have a history of blood pressure instability. Furthermore, there is evidence excessive vasopressor usage can cause ischemia and necrosis in limbs and digits [80, 46]. The agent should select interventions with high expected value that lead to improved dialysis outcomes, however, these actions must also meet clinical and operational expectation. Dosing behavior and dialysis goal achievement of the modified agent-policy was evaluated via simulation. Figure 4.12 demonstrates that as bias increases, the proportion of treatments with recommendation of vasopressor escalation or mannitol administration decreases. The bias closest to observed clinical dosing ($\Delta Q_{mod} \rightarrow 2.25$) was chosen to reduce overdosing and better align RL dosing with clinical intuition.

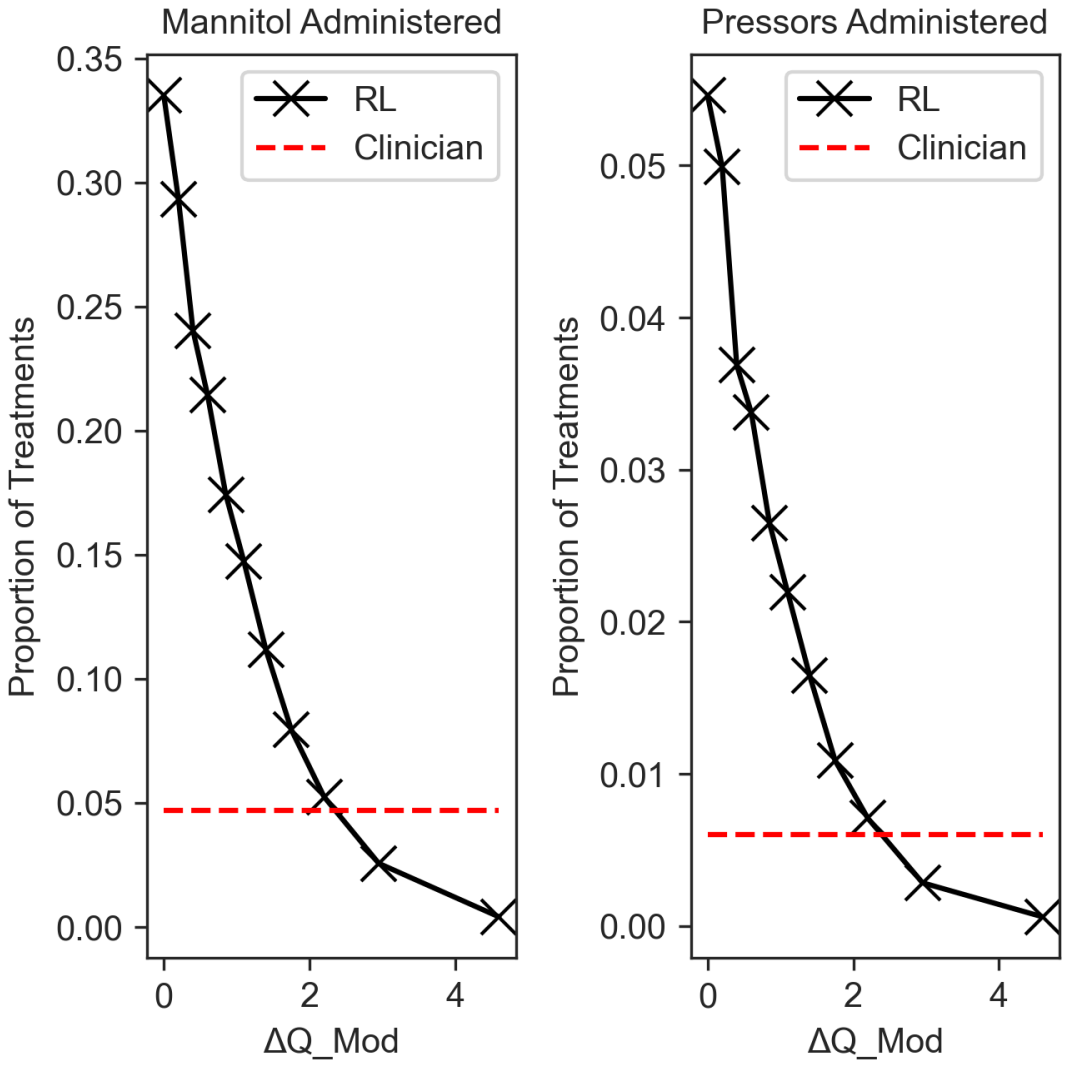

Figure 4.12: Effect of Bias on Dosing

The dosing behavior of the modified agent policy was evaluated by collecting the interventions taken at each timestep during simulated treatments (Figure 4.13). In comparison to dosing using the optimal agent policy (Figure 4.10), use of aggressive treatments are reduced at the expense of increased alteration of ultrafiltration. Specifically, ultrafiltration was changed 32% and 41% of the time following the optimal agent policy and the modified agent policy *in silico*, respectively. Ultrafiltration is decreased >3x more often following the modified agent policy than the optimal agent policy *in silico*. As depicted in Figure 4.13, the clinician

and agent most frequently made no change in ultrafiltration rate, followed by increasing and decreasing ultrafiltration rate.

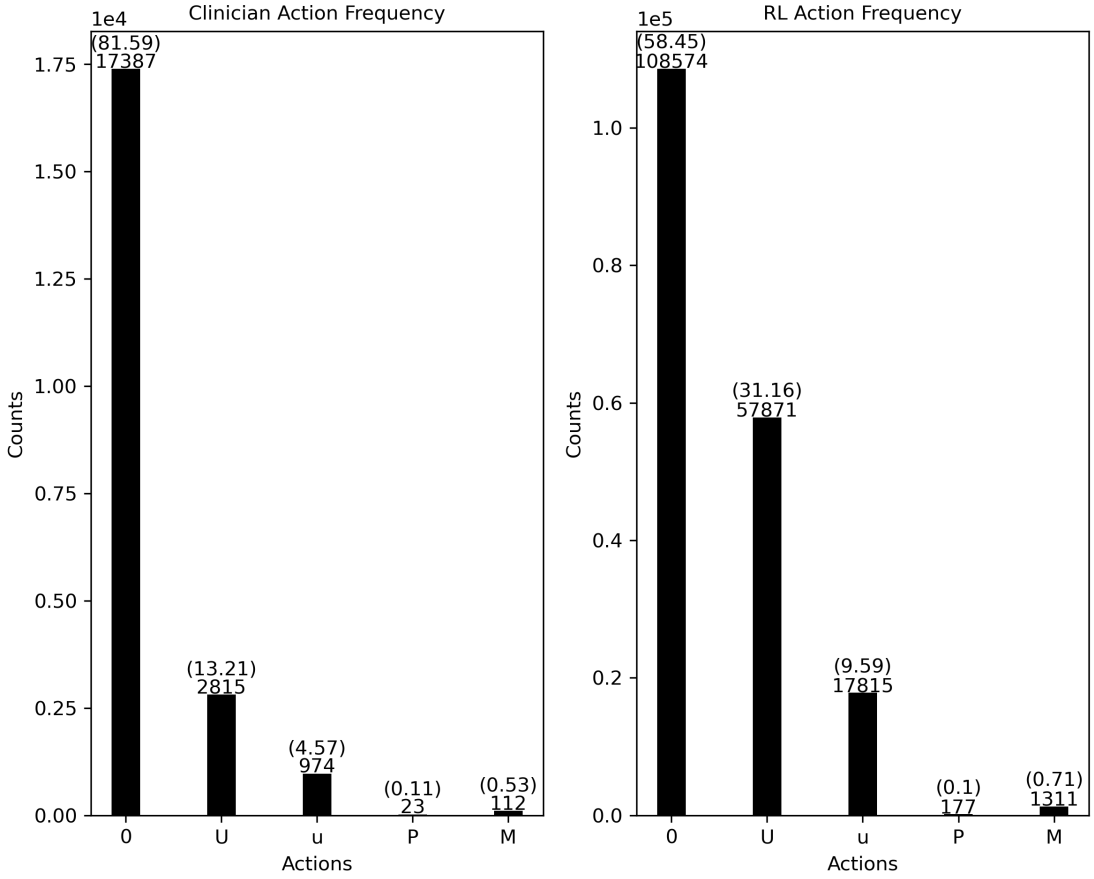

Figure 4.13: Clinician vs. Agent Recommended Interventions (modified policy). Action 0: no intervention, action U: increase ultrafiltrate, action u: decrease ultrafiltrate, action M: administer albumin or mannitol, P: increase vasopressors

The dosing behavior of the agent (per treatment) following the modified agent policy is depicted in Figure 4.14. Aggressive treatments measures, i.e. increasing vasopressor support and administration of mannitol or albumin, were applied a maximum of 3 times *in silico* and below the maximum dosing observed in the dataset. Figure 4.14 and Table 4.2 elucidate that the RL recommends changes in ultrafiltration during treatment more frequently than the clinician.

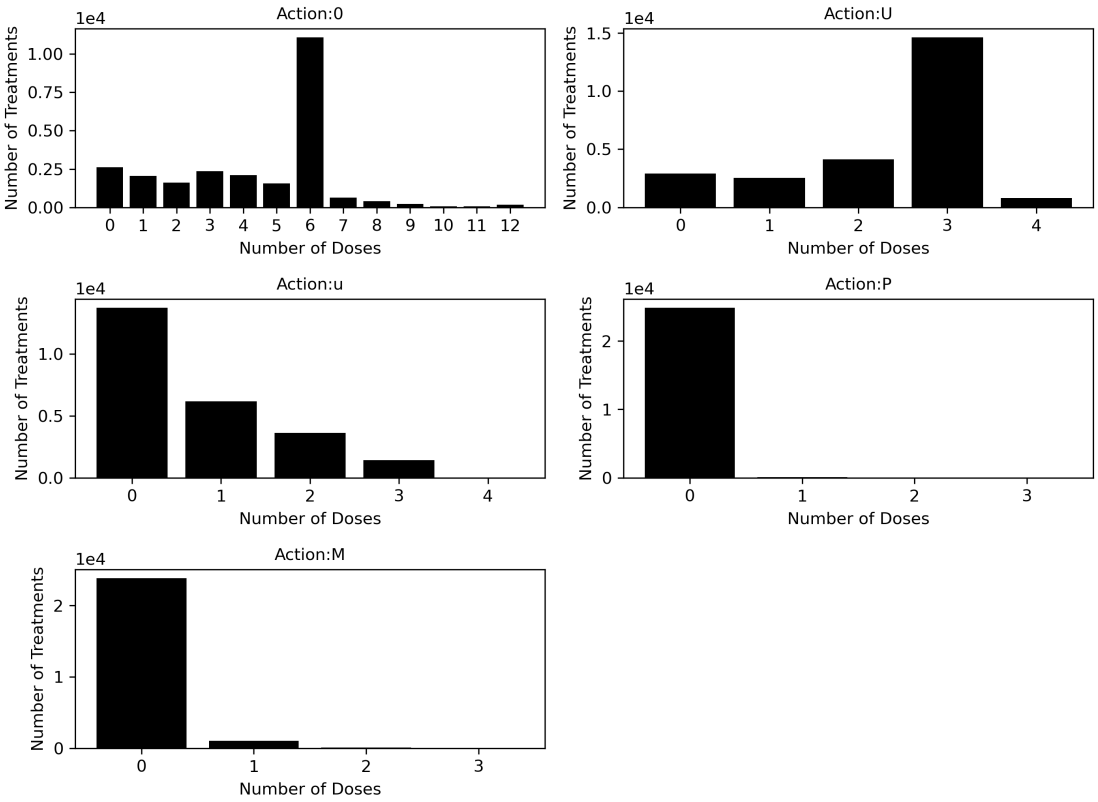

Figure 4.14: RL Agent intervention frequency per dialysis treatment (modified policy). Action 0: no intervention, action U: increase ultrafiltrate, action u: decrease ultrafiltrate, action M: administer albumin or mannitol, P: increase vasopressors

Table 4.2: Per Treatment Intervention Information for agent policy vs. clinician policy

	0	**u**	**U**	**P**	**M**
Treatments Action Taken (RL) [counts (%)]	22392 (89.6)	11264 (45.1)	22087 (88.3)	145 (.6)	1181 (4.7)
Treatments Action Taken Multiple Times (RL) [counts (%)]	20326 (81.3)	5072 (20.3)	19556 (78.2)	29 (.12)	122 (.49)
Treatments Action Taken (Clinician) [counts (%)]	1661 (98.6)	675 (40.1)	1653 (98.1)	19 (1.1)	83 (4.9)
Treatments Action Taken Multiple Times (Clinician) [counts (%)]	1607 (94.8)	229 (13.5)	782 (46.4)	8 (.5)	32 (1.9)

A comparison of primary outcomes resulting from agent and clinician-prescribed treatment are summarized in Table 4.3. The proportion of treatments with experienced hypotension was reduced by 51.3% ($p<.0001$), and the proportion of treatments with achievement of fluid

removal goals was increased by 45.0% ($p < .0001$) when the agent-recommended treatment policy was deployed *in silico*. There is a cost to using the modified agent policy over the optimal policy; IDH occurred in 7.3% and 5.5% of treatments following the modified and optimal agent policy, respectively.

Table 4.3: Outcomes of modified treatment policy vs. clinician policy

	IDH [count (%)]	**Fluid Goals Achieved [count (%)]**	**Time to Goal State (minutes) [median (IQR)]**
RL Treatment Policy	1817 (7.3)	23265 (93.1)	135 [90,140]
Clinician Treatment Policy	253 (15.0)	1082 (64.2)	176 [164,182]

4.3.3 Effect of Agent-Suggested Treatment Policy on Patient Subpopulations

4.3.3.1 Agent-Suggested Policy on ICU vs. Non-ICU

It was hypothesized that the general agent-suggested policy derived in subsection 4.3.2 may experience contrasting performance on different subpopulations of patients. 558 (33.1%) of the dialysis treatments were ICU-based, and 261 (94.6%) of patients received ICU-based treatment at least once during their hospital stay. Table 4.4 describe differences in clinical characteristics between the two subpopulations. ICU-based treatments had a higher frequency of IDH (18.1 % vs. 13.5 %, p=.0013) and satisfaction of systemic inflammatory response syndrome (SIRS) at dialysis initiation (75.1 % vs. 37.2 %, p<.0001). Absolute neutrophil count, anion gap, and phosphate were abnormally elevated and significantly higher at baseline for ICU-based treatments. Calcium levels were abnormally low and significantly lower than non-ICU based treatments at dialysis initiation. Prothrombin time and blood urea nitrogen levels were abnormally high in both subgroups, yet both laboratory measurements were significantly higher at baseline for ICU-based treatment in comparison to non-ICU based treatment. At dialysis initiation, red blood cell count was abnormally low for both subgroups and significantly lower for ICU-based treatment in comparison to non-ICU based treatment.

Table 4.4: Clinical characteristics of ICU and non-ICU based treatment

Characteristic	**ICU**	**non-ICU**	**P-values**	**Reference Range**
IDH [count(%)]	101 (18.1)	152 (13.5)	1.30E-02	---
SIRS [count (%)]	419 (75.1)	572 (37.2)	1.30E-21	---
Calcium, mg/dL	8.3 (7.7-8.9)	8.6 (8.1-9.1)	5.9E-11	8.4-10.2
Anion gap, mEq/L	17.0 (14.0-20.0)	15.0 (13.0-18.0)	1.80E-08	7.0-15.0
Absolute neutrophil count, count*10^9/L	7.8 (5.0-12.4)	6.4 (4.0-9.7)	2.20E-07	2.2-7.7
Phosphate, mg/dL	4.7 (3.6-6.2)	4.1 (3.3-5.4)	1.50E-05	2.5-4.6
Prothrombin time, seconds	15.8 (14.6-17.4)	15.2 (13.9-16.5)	1.90E-11	11.2-14.7
Blood urea nitrogen, mg/dL	47.0 (31.0-72.5)	42.0 (29.0-60.0)	5.80E-05	8.0-26.0
Red blood cell count, count*10^12/L	2.8 (2.5-3.3)	2.9 (2.8-3.2)	2.00E-02	3.7-4.9

The dosing behavior of the modified agent policy on ICU and non-ICU treatments was assessed by collecting the interventions taken at each timestep during simulated dialysis sessions. As depicted in Figure 4.15, both the clinician and agent most frequently made no change in ultrafiltration rate, followed by increasing and decreasing filtration rate during non-ICU treatment. Mannitol was infrequently administered relative to the remaining set of actions taken by the clinician and agent. Vasopressor support is only available for dialysis patients that originate from the intensive care unit (ICU), therefore, increasing vasopressor support is not part of the set of interventions taken by the clinician or agent. Despite these similarities, the agent altered ultrafiltration rate more routinely than the clinician; ultrafiltration was altered 41.8% of the time by the RL agent and 17.1% of the time by the clinician.

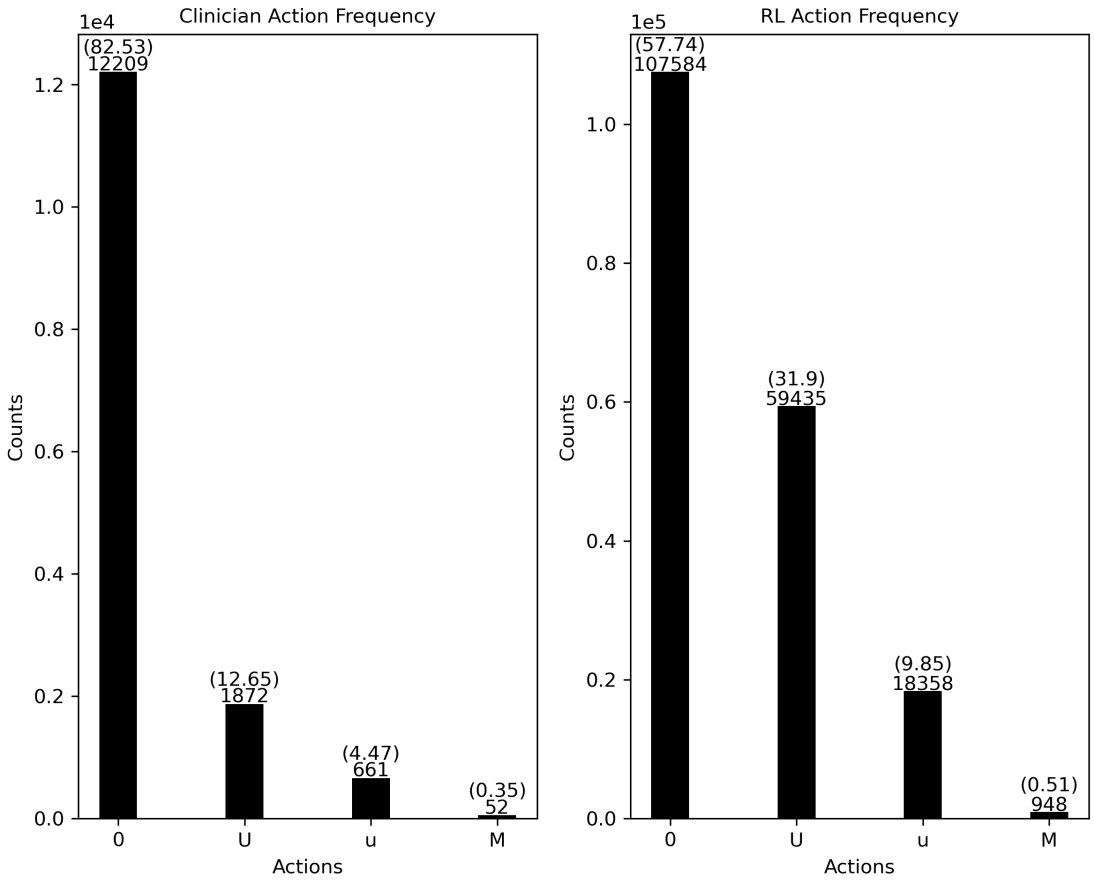

Figure 4.15: Clinician vs. Agent Recommended Interventions (nonICU) . Action 0: no intervention, action U: increase ultrafiltrate, action u: decrease ultrafiltrate, action M: administer albumin or mannitol, P: increase vasopressors

Similarly, both the clinician and agent most frequently made no change in ultrafiltration rate, followed by increasing and decreasing filtration rate during ICU treatments (Figure 4.16). Aggressive treatments (i.e. administration of albumin or mannitol bolus or increasing vasopressor dose) were infrequently taken relative to the remaining set of actions taken by the RL agent and clinician.

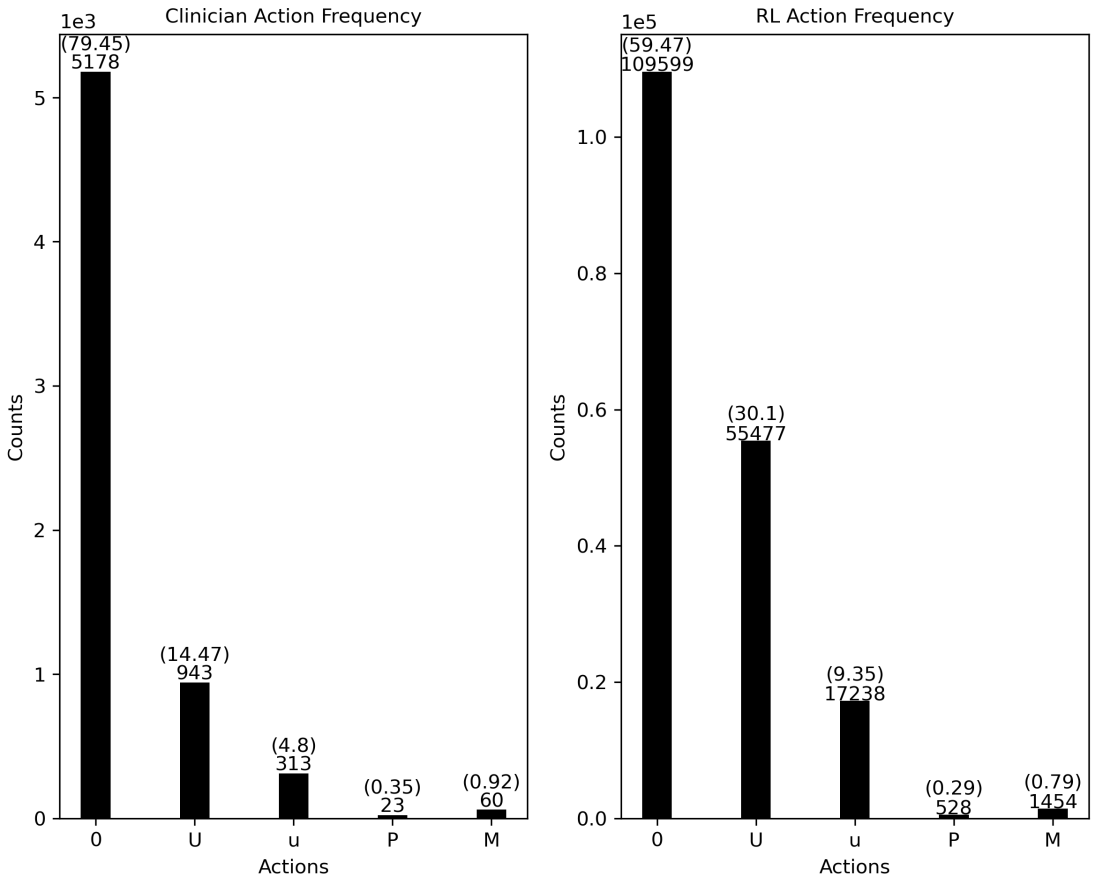

Figure 4.16: Clinician vs. Agent Recommended Interventions (ICU). Action 0: no intervention, action U: increase ultrafiltrate, action u: decrease ultrafiltrate, action M: administer albumin or mannitol, P: increase vasopressors

Table 4.5 compares dialysis outcomes of agent and clinician-prescribed treatment on ICU and nonICU subgroups. As previously mentioned, IDH incidence was higher in ICU-based treatment relative to non-ICU based treatment (18.1% vs. 13.5%,p=.013) performed at UPMC. Similarly, incidence of IDH was higher in ICU-based treatment vs. non-ICU based treatment *in silico* using the agent-derived treatment (9.9% vs. 5.7%). Nevertheless, use of the agent-derived treatment policy *in silico* resulted in a significant reduction (p<.0001) in incidence of IDH and increased achievement of fluid removal goals (p<.0001) in both subgroups. Specifically, IDH occurrence was reduced by 45.3 % (p<.0001) and 58.4% (p<.0001) in ICU and non-ICU groups, respectively. Accomplishment of fluid goals was

increased by 56.9% (p<.0001) and 40.5% (p<.0001) in ICU and non-ICU groups, respectively.

Table 4.5: Outcomes of clinician and agent treatment policy for ICU and non-ICU subpopulations

	IDH [count (%)]	**Fluid Goals Achieved [count (%)]**
RL Treatment Policy (ICU)	2472 (9.9)	22819 (91.3)
Clinician Treatment Policy (ICU)	101 (18.1)	323 (57.9)
RL Treatment Policy (non-ICU)	1420 (5.7)	23580 (94.3)
Clinician Treatment Policy (non-ICU)	152 (13.5)	758 (67.3)

4.4 Discussion

This study demonstrates the utility of using reinforcement learning to derive an optimal treatment policy, solely from a vast number of sequential intervention decisions stored in an offline dataset. A model of future IDH risk was trained and validated using a cohort of patients that received hemodialysis at UPMC acute care facilities in chapter 3. In this study, the RL agent learned the value of taking actions in each health state derived from predictive clinical variables, and a projected future risk of IDH from the model developed in chapter 3. Reward structure is a critical element in the learning design as dialysis treatment decisions need to consider alterations in patient IDH risk while accounting for longer-term goals of dialysis. In the healthcare setting, RL agents have been predominantly rewarded using a single reward at the end of treatment or using a mixture of short-term and terminal rewards with a universal goal for all patients [58, 83]. The risk-reward structure used by the agent in this study directly addresses this issue as the reward signal is derived from a mathematical mapping of patient-specific health characteristics to the probability of an undesired health outcome. Without dialysis, occurrence of uremia and death is inevitable in patients with end-stage renal disease (ESRD). Nevertheless, hypoperfusion and consequent organ damage induced by occurrence of IDH contributes to poor clinical outcomes observed in the dialysis

population, including decreased kidney recovery and death [106, 26]. The agent in this study is rewarded as a function of forecasted blood pressure instability and thus learns how to act preemptively in order to reduce incidence of IDH *in silico*.

Simulated trajectories of the MDP following the optimal policy revealed a 63.3% decrease (p<.0001) in incidence of IDH and 47.4% increase (p<.0001) in fluid goal achievement *in silico* relative to the clinician. Nevertheless, the optimal policy recommended rare and aggressive interventions, e.g. increase vasopressor support and administration of mannitol, at a significantly greater extent than would be utilized clinically. The optimal policy returns the single best action with maximum expected reward. A tunable function was utilized to provide flexibility in selecting alternative, less-aggressive interventions that were sufficiently rewarding, and its implementation produced dosing behavior that met clinical expectation. The modified agent-suggested policy resulted in 51.3% (p<.0001) decrease in incidence of IDH and 43.0% increase (p<.0001) in fluid goal achievement *in silico* relative to the clinician, at the expense of increased actuation of ultrafiltration rate. ICU-based treatments experienced higher incidence of IDH (18.1 vs. 13.5,p=.013) and achievement of fluid removal goals (67.3 vs. 57.9 , p=1.6e-04) compared to non-ICU based treatments performed at UPMC. At baseline, patients receiving ICU-based treatment presented evidence of being sicker, which could lead to greater incidence of blood pressure instability. The agent-derived policy led to a decreased incidence of IDH and increased achievement of fluid removal goals in both subpopulations *in silico*. Nevertheless, incidence of IDH was higher in ICU-based treatment in comparison non-ICU based treatment *in silico*. This algorithm framework is envisioned to be integrated into a closed-loop control system used in the dialysis setting. Alternatively, the system could be utilized semi-closed loop as a recommendation system during the dialysis procedure.

The study in this section has several limitations. First, absent validation of the risk model and RL policy performance using an external cohort of dialysis patients is a major limitation. The trained IDH risk model performance was evaluated on a randomly selected internal validation dataset, and performance of the RL agent was evaluated using dialysis treatments performed internally at UPMC acute care facilities. Future research will focus on collecting a large set of EHR information from dialysis treatments performed externally and

using the collected data to further evaluate RL treatment efficacy. Currently, over 60,000 hours of data was collected from 1685 hemodialysis treatments performed at UPMC; this dataset is continously expanding with additional non-ICU and ICU-based treatments. The agent will be retrained using the expanded dataset, generating increased confidence in the agent-suggested interventions, especially in regard to rare actions (e.g. administration of mannitol, increase in vasopressor administration rate).

Q-learning is an off-policy algorithm and therefore the (optimal) policy learned is different than the behavior (clinician) policy, nevertheless, the learned policy must still meet clinical expectation. Extending state description to include mechanism, clinical history, or intervention history could better align optimal agent-recommended interventions with clinical reasoning. Embedding additional state information and constraints (e.g. mechanism, dosing), however, can greatly increase state-space dimensionality and the necessary data to accurately calculate state-action pair values. Alternatively, to address the issue of overabuse of mannitol and vasopressor support following the optimal policy *in silico*, a modified policy was developed to allow for probabilistic selection of an less aggressive treatment in a state. Undoubtedly, this action selection method diverges from the optimal policy returned from Q-learning that is a single (greedy) action that maximizes value in a given state. Alternative methods, such as penalizing the use of aggressive treatments (e.g. mannitol and vasopressors), could reduce overdosing without divergence from mathematical optimality, however, these may be difficult to appropriately tune to better meet clinical intuition while improving clinical outcomes *in silico*.

Dialysis settings outside UPMC may use alternative interventions in response to blood pressure instability, such as decreasing dialysate temperature. These can be incorporated into the action space, or alternative, rewarding interventions could be recommended if the current recommended action is not deployed at the site. As mentioned in chapter 3, features derived from high-frequency, filtered waveform data may improve risk prediction at the cost of reduced translation to patient settings where waveform data is infrequently acquired, such as outpatient facilities. High-frequency data opens the possibility of using a smaller agent (controller) sample time if implemented as a closed-loop control system.

Clinical data-related issues amplify the challenge of using an offline dataset to derive a

optimal treatment policy for dialysis. The RL-derived policy is predicated on the fact that all sequential treatment-response information necessary for learning is available. Interventions and vital signs are manually recorded by clinicians with varying frequency. Laboratory measurements are gathered routinely but on a longer timeframe. Missing clinical measurements were imputed via last observation carried forward approach to create a complete set of inputs for risk score generation on a 15 minute timestep basis. These imputed clinical measurements were utilized to develop the states of the hemodialysis reinforcement learning problem. Any missing administration rate information from intake/output data was imputed with the last available rate recorded; these rates were then leveraged to create a discrete set of actions for the Q-learning algorithm. Thus, the real patient response and clinical intervention may not always be accurately reflected by the data. In practice, a clinician may intervene not only using instrument measurements, but the clinical presentation of the patient, which is neither recorded by the clinician or available to the agent during the learning process. Adequate clearance of toxins, the second primary goal of dialysis, was not included in the reward scheme for the agent as personalized clearance goals depend on a complex combination of factors such as blood flow rate, time on dialysis, and dialyzer type and were not easily derived from available data. The utilization of estimated dialysis goals, i.e. fluid targets, instead of the use of real prescribed dialysis parameters is another study limitation, and this choice was based on limited data availability. Despite the promising results demonstrated in this section in regard to the use of reinforcement learning to improve clinical outcomes in hemodialysis setting, any trained reinforcement learning policy would need to be deployed on hemodialysis patients in a clinical trial to truly support (or reject) its clinical utility.

5.0 Summary and Outlook

Sepsis is a dysregulated host response to infection; if the syndrome is not properly managed it can lead to organ failure and death. Despite many clinical trials that aimed to modulate inflammation and improve clinical outcomes in sepsis patients, none have resulted in the emergence of a FDA-approved therapy that is currently available. The poor outcomes observed in clinical trials could be partially attributed to patient heterogeneity in the fight against infection. A mathematical model was developed that captured interpatient variability in infection time, magnitude, and speed of the pro- and anti-inflammatory response. Sepsis endotypes, subtypes of patients that differ in their underlying mechanisms to fight infection, were identified using a robust clustering method (consensus clustering). These endotypes differed in their pro- and anti- inflammatory dynamics in response to infection and their respective organ failure and mortality rates. Machine learning was used to predict patient membership into high or low mortality endotypes using IL-6 and IL-10 cytokine levels at study enrollment.

Hemodialysis is a treatment that removes excess fluid and toxic wastes in patients with kidney failure. IDH commonly occurs during hemodialysis treatment and is associated with early discontinuation of treatment, cardiovascular events and mortality. Early prediction of IDH could allow for preemptive administration of treatment and subsequently reduce the incidence and poor outcomes associated with IDH. A machine learning model that predicted future IDH risk was trained and tested using featurized EHR data collected from n=276 patients and n=1685 dialysis treatments conducted in acute care facilities at UPMC. The model provided continuous future IDH risk prediction. An early warning classification system was developed to assign a patient as high or low IDH risk at dialysis initiation, utilizing the IDH risk projection from the model.

Reinforcement learning was leveraged to bridge future IDH risk prediction with an optimized, therapeutic strategy. A finite Markov decision process (MDP) for dialysis over 15-minute time steps was constructed from sequences of patient states, actions, and projected IDH risk during dialysis treatments. The RL agent solved for the optimal policy of the MDP,

using projected risk of IDH as feedback for its interventions. The optimal policy provides a recommended series of interventions to take based on the evolving state of the patient. The optimal policy was tuned to enhance its clinical utility.

5.0.1 Utility of Developed Clinical Tools

5.0.1.1 Sepsis Endotype Identification and Early Classification Tool

In chapter 2, a model was developed that describes inflammation dynamics using five parameters. The model was validated against time course IL-6 and IL-10 data available for n=316 sepsis patients to acquire patient-specific models of inflammation. The sepsis endotype identification approach developed in chapter 2 accounted for patient variation in pre-hospital time and inflammation dynamics by clustering fitted model parameters. This differs from common endotyping approaches that cluster genomic, transcriptomic, metabolite and plasma protein data collected at a single, variable timepoint during the patient's inflammation trajectory. Consensus clustering identified four stable clusters (endotypes) of patients, and distinct inflammation trajectories were produced using average model parameters of each cluster. Patients from cluster 1 demonstrated a quick aggressive response to infection, while patients from cluster 4 required less energy to remove the pathogen and return to a basal state. The patients of these endotypes experienced lower mortality and organ failure rate ($p<.05$) relative to patients in clusters 2 and 3, who demonstrated high and persistent inflammation. This demonstrates the value of leveraging knowledge of inflammation dynamics in addition to absolute levels of mediators for patient prognosis. Endotype identification could be used to identify patients who would or wouldn't be responsive to a therapy with specific action. Prediction from the classification tool developed in chapter 2 could be used by a clinician to develop a treatment regimen that tailors patient outcome. For example, a clinician may prescribe immunomodulatory agents to a patient predicted with membership to clusters 2 or 3 that exhibit high, persistent inflammation and worse clinical outcomes.

5.0.1.2 IDH Risk Prediction Model and Early Warning System for IDH

In chapter 3, a predictive model of future IDH risk was developed by training and testing a random forest classifier on EHR data collected during hemodialysis treatments performed at UPMC. When a held out historical stream of EHR data was fed to the model, distinct separation and evolution of risk between HS and NHS leading up to the IDH event was readily apparent. The AUC-ROC of the model evolved from .74 two hours from IDH occurrence to 1.0 at the time of IDH. Most existing IDH prediction algorithms lack utility as the end user (clinician) needs to know when, who and how to act based on the risk continuously generated by the model. A early warning system was proposed that classified a patient as high or low risk for future IDH during treatment using risk scores around dialysis initiation and an optimized risk threshold. Classifier predictions resulted in a TPR=0.80 and a FPR=0.29 using the chosen risk threshold. The classification system could be leveraged by a clinician to modify dialysis prescription (e.g. modification of dialysate solution, decreasing dialysate temperature, infusion of hypertonic solutions) in a way that could potentially improve a patient's blood pressure stability during hemodialysis.

5.0.1.3 Risk-Based Reinforcement Learning Algorithm for Hemodialysis

The risk model in chapter 3 produced accurate prediction of future IDH risk, however, the tool alone cannot provide patient-specific dialysis prescription recommendations. In chapter 4, a reinforcement learning agent was tasked to learn an optimal treatment policy to decrease incidence of IDH in patients receiving hemodialysis, using projected risk of future IDH as feedback for its actions. Dialysis trajectories were simulated using the optimal policy and demonstrated a significant decrease in IDH (5.5% vs. 15.0%) and increase in achievement of fluid goals (94.6% vs. 64.2 %). Despite this, use of a strictly optimal policy led to over abuse of aggressive treatments (i.e. vasopressors), limiting its clinical utility. A policy modification was implemented that decreased aggressive interventions at the cost of increased actuation of ultrafiltration and a slight increase in incidence of hypotension (+1.8%) . Nevertheless, the modified policy resulted in a significant reduction in IDH occurrence (7.3% vs. 15.0%) and increase in fluid goal achievement (93.1% vs. 64.2%) *in*

silico. The treatment policy recommends preemptive, personalized dialysis interventions based on the evolving state of the patient; this tool could be leveraged by a clinician for treatment decision-making during the dialysis procedure and could potentially improve clinical outcomes.

5.0.2 Scientific Dissemination

The work throughout this dissertation was presented orally to wide audiences consisting of clinicians, scientists, and engineers. Dynamic modeling and sepsis subtype identification work in chapter 2 was presented orally at the *Annual American Institute of Chemical Engineers (AIChE) Meeting* in 2019. Risk modeling development and evaluation work in chapter 3 was orally presented at *International Conference on Complex Acute Illness (ICCAI) 2021: AI in Critical Illness: Emergence and Emergent Issues* and is to be published in Journal of Critical Care as an extended abstract. The risk-based reinforcement learning algorithm development and evaluation work discussed in chapter 4 was presented orally at *International Conference on Complex Acute Illness (ICCAI) 2022: Impactful AI in Critical Care* and *Foundations of Computer Aided Process Operations(FOCAPO)/Chemical Process Control (CPC) Conference 2023*. The work presented at *International Conference on Complex Acute Illness (ICCAI) 2022: Impactful AI in Critical Care* is to be published in Journal of Critical Care as an extended abstract, and the conference paper from *Foundations of Computer Aided Process Operations(FOCAPO)/Chemical Process Control (CPC) Conference 2023* will be published in Computers & Chemical Engineering.

5.0.3 Future Directions

5.0.3.1 Sepsis Endotype Identification and Early Classification Tool

The clinical tool developed in chapter 2 demonstrated skill in predicting patient membership into high/low mortality endotypes (AUC-ROC=.72, AUC-PR=.71). Additional features could be derived from the ProCESS dataset to improve class prediction. In chapter 2, four endotypes were identified by clustering fitted model parameters for n=316 subjects. The

cohort represents 24% of the total patients enrolled in the ProCESS trial and therefore other sepsis endotypes may exist. These subjects were leveraged based on full availability of IL-6 and IL-10 data at 0-, 6-, and 24-hour study time. Modeling of other biomarker trajectories (e.g. lactate) could be used to identify interpatient differences in inflammation during sepsis. Altered microcirculation in sepsis leads to decreased tissue perfusion; cells leverage anaerobic respiration for energy production in the face of inadequate oxygen supply, resulting in increased production of lactate [116]. Normal clearance of lactate by the liver may also be impacted in sepsis patients [116, 6]. Hyperlactatemia is associated with mortality [117, 34]. A mathematical model that accurately captures lactate dynamics in response to infectious input, validated against patient data available in ProCESS dataset, could be used to identify additional sepsis subtypes with differing outcomes. Based on limitations in data availability, using two compartment model subject to an infectious input could be a starting point to capture the biexponential trajectory of lactate observed in humans following a controlled acute LPS challenge [75].

5.0.3.2 IDH Risk Prediction Model and Early Warning System for IDH

The Kidney Disease Outcomes Quality Initiative (KDOQI) defines IDH as a decrease in systolic blood pressure (SBP) $\geq$ 20 mmHg or a decrease in mean arterial pressure (MAP) $\geq$ 10 mmHg accompanied by clinical symptoms [122], however, significant variation in IDH definition exists [2]. For the IDH risk prediction model in chapter 3, drops in blood pressure were not included in the definition. Furthermore, IDH risk prediction models recently developed provide higher real-time IDH risk prediction accuracy than the one developed in chapter 3 [63]. Notable predictive features from these models not leveraged in the current IDH risk prediction model include statistical features derived from time-varying dialysate composition, ultrafiltration and blood flow rate [63]. The size of the UPMC cohort is continuously expanding, and filtered, high-frequency (waveform) data could be leveraged to develop additional features that may improve prediction accuracy [126]. The risk prediction framework also should be trained and tested using multiple IDH definitions to evaluate its utility for different clinical settings.

IDH is only one form of instability experienced by hemodialysis patients. Another potential extension of this work is to create a general dialysis instability prediction model that also encompasses prediction of tachycardia (heart rate $\geq$ 130 beats/min) and cardiac arrythmias (e.g. atrial fibrillation). Arrythmias and tachycardia are associated with increased risk of stroke, heart failure, and death [92, 52, 119]. Results from a recent six-month study conducted on n=66 outpatient maintenance hemodialysis patients demonstrated clinically significant arrythmias (CSA) occurred in two-thirds of the population, with an average of 4.5 CSA events per patient per month [92]. Incidence of CSA events significantly increased risk of cardiovascular hospitalization in this population. The use of sodium variate modeling, high dialysate calcium, or higher dialysis temperatures significantly increased risk of CSA event during treatment.

Early united prediction and detection of tachycardia, cardiac arrythmias, and IDH could allow for preemptive treatment to be administered and subsequently decrease incidence of poor clinical outcomes associated with these events. A separate machine learning model could be trained and tested to produce the evolution of future risk of tachycardia, cardiac arrythmias, and IDH. A risk score triggered early warning system, similar to the one developed in chapter 3, could be developed to classify a patient as high or low risk for each of events at dialysis initiation. The clinician could use the risk classification returned by the system to provide more targeted treatment to the patient. To describe how a total instability prediction could influence decision making, beta-blockers and antiarrhythmics are classifications of medications that reduce heart rate and arrythmias, however, there is evidence the latter may lead to lower risk of blood pressure instability [114, 37, 25, 20]. If a patient was classified as high-risk for IDH and tachycardia/arrythmias with this system, this may challenge the administration of beta-blockers.

5.0.3.3 Risk-Based Reinforcement Learning Algorithm for Hemodialysis

A primary extension of the work in chapter 4 is to test the RL agent on an external cohort of ICU dialysis patients. In chapter 4, EHR information was collected from n=276 patients and n=1685 treatments performed at University of Pittsburgh Medical Center (UPMC) acute

care facilities from June 2016-June 2020. Of these treatments, n=558 (33.1%) were ICU-based treatments. A MDP representation of the dialysis process was derived from patient data, using predicted IDH risk and achievement of individualized fluid goals as feedback for its actions. The agent-suggested policy was deployed *in silico* to evaluate its treatment efficacy, which resulted in significant reduction in IDH and an increase in accomplishment of fluid removal goals. As a preliminary investigation, an agent was trained separately on ICU-based (n=558) and non-ICU based treatments (n=1127) available in the internal UPMC dataset, producing two distinct agent-derived treatment policies. The policies were deployed *in silico* to evaluate treatment efficacy and clinical plausibility.

The dosing behavior of the agent and clinician during ICU and non-ICU based treatments was assessed by collecting interventions taken at each timestep during treatment. As depicted in Figure 5.1, both the clinician and agent most frequently made no change in ultrafiltration rate, followed by increasing and decreasing ultrafiltration rate during non-ICU treatments. Mannitol was infrequently administered relative to the remaining set of actions taken by the clinician and agent. Despite these similarities, the agent altered ultrafiltration rate more routinely than the clinician; ultrafiltration was altered 44.6% and 17.1% of timesteps by the RL and clinician, respectively.

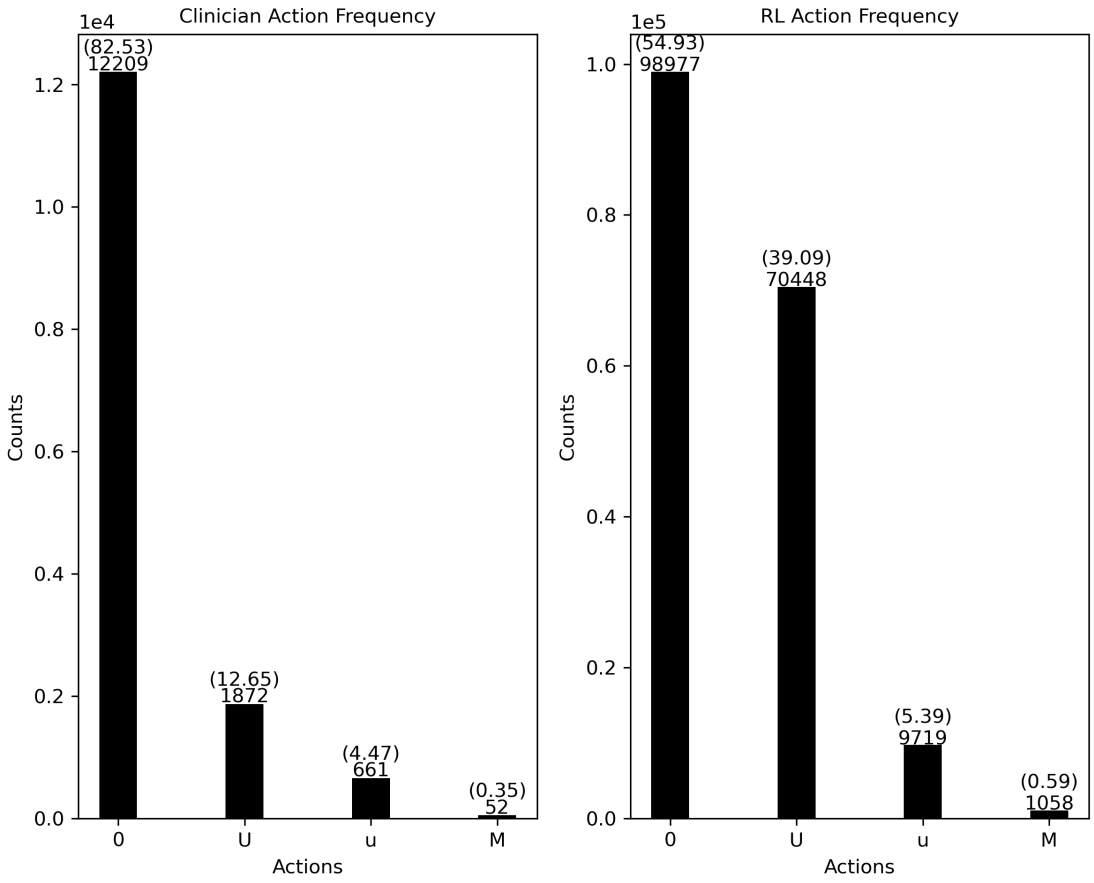

Figure 5.1: Clinician vs. Agent Recommended Interventions (NonICU). Action 0: no intervention, action U: increase ultrafiltrate, action u: decrease ultrafiltrate, action M: administer albumin or mannitol, P: increase vasopressors

Similarly, both the clinician and agent most frequently made no change in ultrafiltration rate, followed by increasing and decreasing ultrafiltration rate during ICU treatments (Figure 5.2). Aggressive treatments (i.e. administration of albumin or mannitol or increasing vasopressor dose) were infrequently taken relative to the remaining set of actions taken by the RL agent and clinician. Interestingly, the RL recommended increasing ultrafiltration less for ICU patients than non-ICU patients (39.05% of timesteps vs. 23.58 % of timesteps) , and frequency of decreases in ultrafiltration rate were similar (5.98 % of timesteps vs. 5.51% of timesteps).

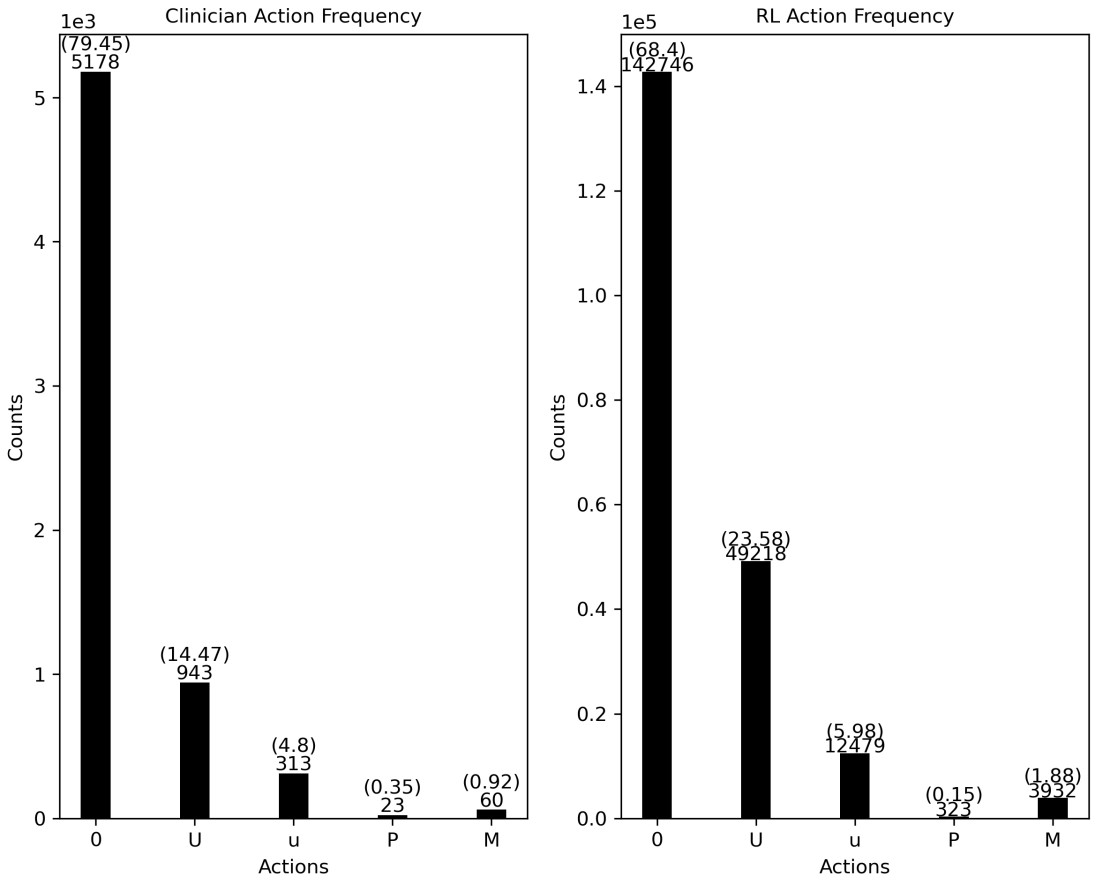

Figure 5.2: Clinician vs. Agent Recommended Interventions (ICU) . Action 0: no intervention, action U: increase ultrafiltrate, action u: decrease ultrafiltrate, action M: administer albumin or mannitol, P: increase vasopressors

Dialysis outcomes of the clinician policy and RL policy were compared (Table 5.1). The agent-derived treatment policy *in silico* resulted in a significant reduction ($p<.0001$) in incidence of IDH and increased achievement of fluid removal goals ($p<.0001$) in both subgroups. IDH occurrence was reduced by 34.8% ($p<.0001$) and 69.6% ($p<.0001$) in ICU and non-ICU groups, respectively. Accomplishment of fluid goals was increased by 53.7% ($p<.0001$) and 42.9% ($p<.0001$) in ICU and non-ICU groups, respectively.

Table 5.1: Outcomes of clinician and agent treatment policies for ICU and non-ICU subpopulations

	Treatments (count)	**IDH (%)**	**Fluid Goals Achieved (%)**
RL Treatment Policy (ICU)	25000	11.8	89
Clinician Treatment Policy (ICU)	558	18.1	57.9
RL Treatment Policy (non-ICU)	25000	4.1	95.9
Clinician Treatment Policy (non-ICU)	1127	13.5	67.3

The simulation results demonstrate the RL framework described in chapter 4 improves both fluid goal achievement and decreases incidence of IDH *in silico* when an RL agent is separately trained and tested on ICU and non-ICU based treatments. Although preliminary results are promising, inter-site differences in dialysis may exist; clinicians may use different dialysis interventions, and patients may experience different responses to treatment and exhibit different dialysis outcomes. In future work, the RL framework will be tested using EHR data from the publicly available Medical Information Mart for Intensive Care (MIMIC)-IV database. Specifically, a subset of the database contains EHR data for n=1044 patients and n=2140 ICU-based dialysis treatments conduced at Beth Israel Deaconess Medical Center (BIDMC) in Boston, MA between 2008-2019 [47]. The database contains elements necessary to reconstruct a testbed for the RL agent developed chapter 4, including vital signs, medications, intake and output, laboratory measurements, and clinical notes. A comparison of demographic and treatment information between the UPMC dialysis cohort and selected dialysis treatment cohort from Beth Israel Deaconess Medical Center is shown in Table 5.2.

Table 5.2: Demographics/Treatment Information for Pitt ICU vs. BIDMC ICU Dialysis Cohorts

Demographics			
		UPMC ICU (n=261)	BIDMC (n=1044)
Age, years [median (IQR)]		63 (52-71)	63 (52-72)
Male [count (%)]		149 (57.1)	627 (60.1)
Admission weight, kg [median (IQR)]		80.3 (68.0-105.0)	79.8 (66.5-94.28)
Race [count (%)]			
	White	174 (66.7)	575 (55.1)
	African-American	57 (21.8)	213 (20.4)
	Other	30 (11.5)	256 (24.5)
Hospital length of stay, days [median (IQR)]		15.0 (8.0-29.0)	15.4 (7.9-29.2)
In Hospital Mortality [count(%)]		56 (21.5)	100 (9.7)
Treatment Information			
		UPMC ICU (n=558)	BIDMC (n=2140)
Incidence of IDH [count (%)]		101 (18.1)	730 (34.1)
Mannitol/Albumin Administration [count (%)]		50 (9.0)	38 (1.8)
Vasopressor Support [count (%)]		76 (13.6)	299 (14.0)

BIDMC and UPMC dialysis cohorts are mostly white, male and are similar in age. In hospital-mortality is higher in the UPMC ICU dialysis patients (21.5% vs. 9.6%), while IDH incidence is higher in BIDMC ICU-based dialysis treatments (34.1% vs. 18.1%). The proportion of treatments with administration of mannitol is higher in the UPMC cohort (9.0% vs. 1.8%), while vasopressor use is similar in both cohorts (13.6% vs. 14.0%)

Two sets of analysis could be engaged with the acquired external dialysis dataset. The RL policy devised for UPMC ICU patients can be tested on BIDMC patients *in silico* by first projecting patient EHR data onto the state space constructed in chapter 4. Initial collected treatment information from the BIDMC cohort (Table 5.2) indicates that the interventions used on the BIDMC cohort and UPMC cohort overlap, thus the action set is likely the same. Dialysis trajectories can be produced by applying the RL treatment policy and producing

state transitions via a probability transition function constructed from BIDMC patient data. Achievement of fluid goals and incidence of IDH using the RL policy *in silico* can be evaluated against the BIDMC clinician. The RL framework in chapter 4 can also be used to train and test a new RL policy specific to the BIDMC cohort population.

Adequate clearance of toxins, the second primary goal of dialysis, was not included in the reward scheme for the agent as personalized clearance goals were not easily derived from the available data. In practice, Kt/V is used to measure dialysis dose adequacy, where K is the dialyzer clearance (mL/minute), t is time on dialysis (minutes) and V is the urea distribution volume (mL). Kt/V is a dimensionless number representing volume of urea cleared during treatment relative to the urea distribution volume [35]. Single-pool urea kinetic models are primarily used in dialysis units to estimate Kt/V (also known as $SpKt/V$). Current Kidney Disease Outcomes Quality Initiative (KDOQI) guidelines recommend a target $SpKt/V$ >1.2 for patients with low residual kidney clearance receiving intermittent hemodialysis three times per week [22]. The single-pool kinetic model can overestimate Kt/V as it assumes urea is confined to one compartment (pool) of the body and fails to account for the increase in urea concentration immediately following treatment, known as the urea rebound effect [110]. Two compartment models have been developed to address limitations of the single-pool kinetic model. In the two-pool kinetic model, the human body is split into an intracellular compartment and extracellular fluid compartment. The mass and volume balance for this process is shown in Equation 5.1.

$$\begin{aligned} \frac{d}{dt}(\alpha_1 V C_1) &= G + X(C_2 - C_1) - KC_1 \\ \frac{d}{dt}(\alpha_2 V C_2) &= X(C_1 - C_2) \\ \frac{dV}{dt} &= F_{in}(t) - F_{out}(t) \end{aligned} \tag{5.1}$$

The total of volume of fluid in the body is V, and $\alpha_1 V$ and $\alpha_2 V$ are the fluid volumes of the extracellular and intracellular compartments of the body distinguished by constants α_1 and α_2, respectively. Urea is generated in the extracellular compartment with rate G and exchanged between the two compartments at a rate proportional to the concentration difference between two compartments ($X(C_2 - C_1)$). Urea is cleared from the extracellular

compartment at rate K governed by dialyzer clearance with rate K_d and residual renal clearance K_r (K=K_d+K_r). Change in urea distribution volume (V) over time is governed by the rate of fluid administration (F_{in}) and fluid removal rate (F_{out}). Some of these model variables and parameters can be defined based on known physiological values (V, α_1, α_2, F_{in}, F_{out}, X). Remaining parameters vary on a per treatment basis (e.g. G, K_r,K_d) and can be determined by fitting the model parameters to time course BUN data via constrained nonlinear least squares. Significant modifications of this second-order compartment model to better ground estimation of renal clearance Kr on a individual basis is ongoing work of our research team [66]. The fully defined two compartment model produces BUN trajectories over time, and patient-specific dialysis dose adequacy (eKt/V) can be estimated using the model [28]. Thus, an additional requirement to receive a reward at terminal timesteps of dialysis could be if treatment $eKt/V > 1.0$, where the threshold for eKt/V is set according to current clinical guidelines [28]. More specific conditional terminal rewards can be designed and weighted according to clinical importance of clearance, fluid removal goals, and IDH.

Appendix

Table A.1: Predictive Variables Used in IDH Prediction Model

Category	**Items**	**Units**
Vital Signs	mean arterial pressure (MAP)	mmHg
	systolic blood pressure (SBP)	mmHg
	heart rate	beats/minute
	diastolic blood pressure	mmHg
	respiratory rate	breaths/minute
	oxygen saturation	%
Featurized Vital Signs	LWMA	
	Minimum (min)	
	maximum (max)	
	wslope	
	uwslope	
	coefficient of variation (cv)	
	hematocrit	%
	carbon dioxide	mmol/L
	basophils	%
	glucose	mg/dL
	absolute monocytes	count/L
	RDW	%
	chloride	mmol/L
	absolute eosinophils	count/L
	absolute basophils	count/L
	calcium	mg/dL

Table A.1 (continued)

Category	Items	Units
	anion cap	mEq/L
	mean corpuscular volume	fL
	absolute lymphocytes	count/L
	sodium	mmol/L
	potassium	mmol/L
	phosphate	mg/dL
	monocytes	%
	lymphocytes	%
	mean platelet volume	fL
	eosinophils	%
	mean corpuscular hemoglobin	pg
	red blood cells	count/L
	platelets	count/L
	blood urea nitrogen	mg/dL
	absolute neutrophils	count/L
	white blood cells	count/L
	neutrophils	%
	MCHC	gm/dL
	creatinine	mg/dL
	hemoglobin	gm/dL
	aspartate aminotransferase	IU/L
	bilirubin	mg/dL
	total protein	gm/dL
	albumin	gm/dL
	alkaline phosphatase	IU/L
	alanine transaminase	IU/L
	international normalized ratio	
	prothrombin time (PT)	sec

Table A.1 (continued)

Category	Items	Units
	partial thromboplastin time	sec
	lactate (whole blood)	mmol/L
	magnesium	mg/dL
Medical History	liver disease or dysfunction	
	heart disease or dysfunction	
	diabetes	
	prior incidence of IDH	
	prior incidence of hypotension	

Bibliography

[1] Ana L Alessandri, Lirlandia P Sousa, Christopher D Lucas, Adriano G Rossi, Vanessa Pinho, and Mauro M Teixeira. Resolution of inflammation: mechanisms and opportunity for drug development. *Pharmacology therapeutics*, 139(2):189–212, 2013.

[2] Magdalene M Assimon and Jennifer E Flythe. Definitions of intradialytic hypotension. In *Seminars in dialysis*, volume 30, pages 464–472. Wiley Online Library.

[3] Chiraz Atri, Fatma Z Guerfali, and Dhafer Laouini. Role of human macrophage polarization in inflammation during infectious diseases. *International journal of molecular sciences*, 19(6):1801, 2018.

[4] Joshua J Augustine, Diane Sandy, Tracy H Seifert, and Emil P Paganini. A randomized controlled trial comparing intermittent with continuous dialysis in patients with arf. *American journal of kidney diseases*, 44(6):1000–1007, 2004.

[5] Tae Wuk Bae, Min Seong Kim, Jong Won Park, Kee Koo Kwon, and Kyu Hyung Kim. Multilayer perceptron-based real-time intradialytic hypotension prediction using patient baseline information and heart-rate variation. *International journal of environmental research and public health*, 19(16):10373, 2022.

[6] Jan Bakker. Lactate is the target for early resuscitation in sepsis. *Revista Brasileira de Terapia Intensiva*, 29:124–127, 2017.

[7] William Beaubien-Souligny, Yifan Yang, Karen EA Burns, Jan O Friedrich, Alejandro Meraz-Muñoz, Edward G Clark, Neill K Adhikari, Sean M Bagshaw, and Ron Wald. Intra-dialytic hypotension following the transition from continuous to intermittent renal replacement therapy. *Annals of intensive care*, 11(1):1–10, 2021.

[8] Gordon R Bernard, Jean-Louis Vincent, Pierre-Francois Laterre, Steven P LaRosa, Jean-Francois Dhainaut, Angel Lopez-Rodriguez, Jay S Steingrub, Gary E Garber, Jeffrey D Helterbrand, and E Wesley Ely. Efficacy and safety of recombinant human activated protein c for severe sepsis. *New England journal of medicine*, 344(10):699–709, 2001.

[9] F. A. Bozza, J. I. Salluh, A. M. Japiassu, M. Soares, E. F. Assis, R. N. Gomes, M. T. Bozza, H. C. Castro-Faria-Neto, and P. T. Bozza. Cytokine profiles as markers of disease severity in sepsis: a multiplex analysis. *Critical Care*, 11(2):8, 2007.

[10] F. Braido, A. Tiotiu, K. Kowal, S. Mihaicuta, P. Novakova, and I. K. Oguzulgen. Phenotypes/endotypes-driven treatment in asthma. *Current Opinion in Allergy and Clinical Immunology*, 18(3):184–189, 2018.

[11] Leo Breiman. Some properties of splitting criteria. *Machine learning*, 24(1):41–47, 1996.

[12] Leoa Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[13] T. Calandra, J. Gerain, D. Heumann, J. D. Baumgartner, and M. P. Glauser. High circulating levels of interleukin-6 in patients with septic shock - evolution during sepsis, prognostic value, and interplay with other cytokines. *American Journal of Medicine*, 91(1):23–29, 1991.

[14] P. Campo, F. Rodriguez, S. Sanchez-Garcia, P. Barranco, S. Quirce, C. Perez-Frances, E. Gomez-Torrijos, R. Cardenas, J. M. Olaguibel, J. Delgado, and Seaic Asthma Severe Asthma Workgrp. Phenotypes and endotypes of uncontrolled severe asthma: New treatments. *Journal of Investigational Allergology and Clinical Immunology*, 23(2):76–88, 2013.

[15] J. M. Cavaillon, M. Adib-Conquy, C. Fitting, C. Adrie, and D. Payen. Cytokine cascade in sepsis. *Scandinavian Journal of Infectious Diseases*, 35(9):535–544, 2003.

[16] Chia-Ter Chao, Jenq-Wen Huang, and Chung-Jen Yen. Intradialytic hypotension and cardiac remodeling: a vicious cycle. *BioMed research international*, 2015, 2015.

[17] Hina Chaudhry, Juhua Zhou, YIN Zhong, Mir Mustafa Ali, Franklin McGuire, Prakash S Nagarkatti, and Mitzi Nagarkatti. Role of cytokines as a double-edged sword in sepsis. *In Vivo*, 27(6):669–684, 2013.

[18] Lujie Chen, Olufunmilayo Ogundele, Gilles Clermont, Marilyn Hravnak, Michael R Pinsky, and Artur W Dubrawski. Dynamic and personalized risk forecast in step-down units. implications for monitoring paradigms. *Annals of the American Thoracic Society*, 14(3):384–391, 2017.

[19] Thomas Clahsen and Fred Schaper. Interleukin-6 acts in the fashion of a classical chemokine on monocytic cells by inducing integrin activation, cell adhesion, actin polymerization, chemotaxis, and transmigration. *Journal of leukocyte biology*, 84(6):1521–1529, 2008.

[20] Henry F Clemo, Mark A Wood, David M Gilligan, and Kenneth A Ellenbogen. Intravenous amiodarone for acute heart rate control in the critically ill patient with atrial tachyarrhythmias. *The American journal of cardiology*, 81(5):594–598, 1998.

[21] Kevin N Couper, Daniel G Blount, and Eleanor M Riley. Il-10: the master regulator of immunity to infection. *The Journal of Immunology*, 180(9):5771–5777, 2008.

[22] John T Daugirdas, Thomas A Depner, Jula Inrig, Rajnish Mehrotra, Michael V Rocco, Rita S Suri, Daniel E Weiner, Nancy Greer, Areef Ishani, and Roderick MacDonald. Kdoqi clinical practice guideline for hemodialysis adequacy: 2015 update. *American Journal of Kidney Diseases*, 66(5):884–930, 2015.

[23] Judy Day, Jonathan Rubin, Yoram Vodovotz, Carson C Chow, Angela Reynolds, and Gilles Clermont. A reduced mathematical model of the acute inflammatory response ii. capturing scenarios of repeated endotoxin administration. *Journal of theoretical biology*, 242(1):237–256, 2006.

[24] Rodrigo Octávio Deliberato, Stephanie Ko, Matthieu Komorowski, MA Armengol de La Hoz, Maria P Frushicheva, Jesse D Raffa, Alistair EW Johnson, Leo Anthony Celi, and David J Stone. Severity of illness scores may misclassify critically ill obese patients. *Critical care medicine*, 46(3):394–400, 2018.

[25] Georg Delle Karth, Alexander Geppert, Thomas Neunteufl, Ute Priglinger, Markus Haumer, Michael Gschwandtner, Peter Siostrzonek, and Gottfried Heinz. Amiodarone versus diltiazem for rate control in critically ill patients with atrial tachyarrhythmias. *Critical care medicine*, 29(6):1149–1153, 2001.

[26] Adrianna Douvris, Khalid Zeid, Swapnil Hiremath, Sean M Bagshaw, Ron Wald, William Beaubien-Souligny, Jennifer Kong, Claudio Ronco, and Edward G Clark. Mechanisms for hemodynamic instability related to renal replacement therapy: a narrative review. *Intensive care medicine*, 45(10):1333–1346, 2019.

[27] Fahad Y Edrees, Sreelatha Katari, Jack D Baty, and Anitha Vijayan. A pilot study evaluating the effect of cooler dialysate temperature on hemodynamic stability during prolonged intermittent renal replacement therapy in acute kidney injury. *Critical care medicine*, 47(2):e74–e80, 2019.

[28] Garabed Eknoyan, Gerald J Beck, Alfred K Cheung, John T Daugirdas, Tom Greene, John W Kusek, Michael Allon, James Bailey, James A Delmez, and Thomas A Depner. Effect of dialysis dose and membrane flux in maintenance hemodialysis. *New England Journal of Medicine*, 347(25):2010–2019, 2002.

[29] Eyal Even-Dar, Yishay Mansour, and Peter Bartlett. Learning rates for q-learning. *Journal of machine learning Research*, 5(1), 2003.

[30] Usama Fayyad and Keki Irani. Multi-interval discretization of continuous-valued attributes for classification learning. 1993.

[31] Carolin Fleischmann, André Scherag, Neill KJ Adhikari, Christiane S Hartog, Thomas Tsaganos, Peter Schlattmann, Derek C Angus, and Konrad Reinhart. Assessment of global incidence and mortality of hospital-treated sepsis. current estimates and limitations. *American journal of respiratory and critical care medicine*, 193(3):259–272, 2016.

[32] Jennifer E Flythe, Hui Xue, Katherine E Lynch, Gary C Curhan, and Steven M Brunelli. Association of mortality risk with various definitions of intradialytic hypotension. *Journal of the American Society of Nephrology*, 26(3):724–734, 2015.

[33] Vladimir Gašparović, Ina Filipović-Grčić, Marijan Merkler, and Zoran Pišl. Continuous renal replacement therapy (crrt) or intermittent hemodialysis (ihd)—what is the procedure of choice in critically ill patients? *Renal failure*, 25(5):855–862, 2003.

[34] Amina Godinjak, Selma Jusufovic, Admir Rama, Amer Iglica, Faris Zvizdic, Adis Kukuljac, Ira Tancica, and Sejla Rozajac. Hyperlactatemia and the importance of repeated lactate measurements in critically ill patients. *Medical Archives*, 71(6):404, 2017.

[35] Frank A Gotch and John A Sargent. A mechanistic analysis of the national cooperative dialysis study (ncds). *Kidney international*, 28(3):526–534, 1985.

[36] Kathryn S Gray, Dena E Cohen, and Steven M Brunelli. Dialysate temperature of 36 c: Association with clinical outcomes. *Journal of Nephrology*, 31(1):129–136, 2018.

[37] Jari Halonen, Pertti Loponen, Otso Järvinen, Jari Karjalainen, Ilkka Parviainen, Pirjo Halonen, Jarkko Magga, Anu Turpeinen, Mikko Hippeläinen, and Juha Hartikainen. Metoprolol versus amiodarone in the prevention of atrial fibrillation after cardiac surgery: a randomized trial. *Annals of internal medicine*, 153(11):703–709, 2010.

[38] Sally Hannoodee and Dian N Nasuruddin. *Acute inflammatory response*. StatPearls Publishing, 2021.

[39] Feras Hatib, Zhongping Jian, Sai Buddi, Christine Lee, Jos Settels, Karen Sibert, Joseph Rinehart, and Maxime Cannesson. Machine-learning algorithm to predict hypotension based on high-fidelity arterial pressure waveform analysis. *Anesthesiology*, 129(4):663–674, 2018.

[40] Jonathan Himmelfarb and T Alp Ikizler. Hemodialysis. *New England Journal of Medicine*, 363(19):1833–1845, 2010.

[41] Jonathan Himmelfarb, Raymond Vanholder, Rajnish Mehrotra, and Marcello Tonelli. The current and future landscape of dialysis. *Nature Reviews Nephrology*, 16(10):573–585, 2020.

[42] P. M. Honore, E. Hoste, Z. Molnar, R. Jacobs, O. Joannes-Boyau, Mlng Mabrain, and L. G. Forni. Cytokine removal in human septic shock: Where are we and where are we going? *Annals of Intensive Care*, 9:13, 2019.

[43] Anita Marie Hosac. Drotrecogin alfa (activated): the first fda-approved treatment for severe sepsis. In *Baylor University Medical Center Proceedings*, volume 15, pages 224–227. Taylor Francis.

[44] Suzanne M Hurst, Thomas S Wilkinson, Rachel M McLoughlin, Suzanne Jones, Sankichi Horiuchi, Naoki Yamamoto, Stefan Rose-John, Gerald M Fuller, Nicholas Topley, and Simon A Jones. Il-6 and its soluble receptor orchestrate a temporal switch in the pattern of leukocyte recruitment seen during acute inflammation. *Immunity*, 14(6):705–714, 2001.

[45] Jack Q Jaeger and Ravindra L Mehta. Assessment of dry weight in hemodialysis: an overview. *Journal of the American Society of Nephrology*, 10(2):392–403, 1999.

[46] Shruti Jesani, Sherif Elkattawy, Muhammad Atif Masood Noori, Sarah Ayad, Suha Abuaita, Kirolos Gergis, Omar Elkattawy, and Vipin Garg. Vasopressor-induced digital ischemia. *Cureus*, 13(7), 2021.

[47] Alistair Johnson, Lucas Bulgarelli, Tom Pollard, Steven Horng, Leo Anthony Celi, and Roger Mark. Mimic-iv. *PhysioNet. Available online at: https://physionet.org/content/mimiciv/1.0/(accessed August 23, 2021)*, 2020.

[48] Ian T Jolliffe and Jorge Cadima. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202, 2016.

[49] Gilles Kaplanski, Valérie Marin, Félix Montero-Julian, Alberto Mantovani, and Catherine Farnarier. Il-6: a regulator of the transition from neutrophil to monocyte recruitment during inflammation. *Trends in immunology*, 24(1):25–29, 2003.

[50] Stephen P Kelleher, John B Robinette, Frederick Miller, and John D Conger. Effect of hemorrhagic reduction in blood pressure on recovery from acute renal failure. *Kidney international*, 31(3):725–730, 1987.

[51] Ashish Khanna, Shane W English, Xueyuan S Wang, Kealy Ham, James Tumlin, Harold Szerlip, Laurence W Busse, Laith Altaweel, Timothy E Albertson, and Caleb Mackey. Angiotensin ii for the treatment of vasodilatory shock. *New England Journal of Medicine*, 377(5):419–430, 2017.

[52] Yazan Khouri, Tiona Stephens, Gloria Ayuba, Hazim AlAmeri, Nour Juratli, and Peter A McCullough. Understanding and managing atrial fibrillation in patients with kidney disease. *Journal of Atrial Fibrillation*, 7(6), 2015.

[53] Carl Kingsford and Steven L Salzberg. What are decision trees? *Nature biotechnology*, 26(9):1011–1013, 2008.

[54] B Ravi Kiran, Ibrahim Sobh, Victor Talpaert, Patrick Mannion, Ahmad A Al Sallab, Senthil Yogamani, and Patrick Pérez. Deep reinforcement learning for autonomous driving: A survey. *IEEE Transactions on Intelligent Transportation Systems*, 2021.

[55] Greg A Knoll, Jenny A Grabowski, Geoffrey F Dervin, and Keith O'Rourke. A randomized, controlled trial of albumin versus saline for the treatment of intradialytic hypotension. *Journal of the American Society of Nephrology*, 15(2):487–492, 2004.

[56] Scott D Kobayashi, Natalia Malachowa, and Frank R DeLeo. Neutrophils and bacterial immune evasion. *Journal of Innate Immunity*, 10(5-6):432–441, 2018.

[57] Jens Kober, J Andrew Bagnell, and Jan Peters. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, 32(11):1238–1274, 2013.

[58] Matthieu Komorowski, Leo A Celi, Omar Badawi, Anthony C Gordon, and A Aldo Faisal. The artificial intelligence clinician learns optimal treatment strategies for sepsis in intensive care. *Nature medicine*, 24(11):1716–1720, 2018.

[59] Johanna Kuipers, Loes M Verboom, Karin JR Ipema, Wolter Paans, Wim P Krijnen, Carlo AJM Gaillard, Ralf Westerhuis, and Casper FM Franssen. The prevalence of intradialytic hypotension in patients on conventional hemodialysis: a systematic review with meta-analysis. *American journal of nephrology*, 49(6):497–506, 2019.

[60] R. Kumar, G. Clermont, Y. Vodovotz, and C. C. Chow. The dynamics of acute inflammation. *Journal of Theoretical Biology*, 230(2):145–155, 2004.

[61] Merin E Kuruvilla, F Lee, and Gerald B Lee. Understanding asthma phenotypes, endotypes, and mechanisms of disease. *Clinical reviews in allergy immunology*, 56(2):219–233, 2019.

[62] Sascha Lange, Thomas Gabel, and Martin Riedmiller. *Batch reinforcement learning*, pages 45–73. Springer, 2012.

[63] Hojun Lee, Donghwan Yun, Jayeon Yoo, Kiyoon Yoo, Yong Chul Kim, Dong Ki Kim, Kook-Hwan Oh, Kwon Wook Joo, Yon Su Kim, and Nojun Kwak. Deep learning model for real-time prediction of intradialytic hypotension. *Clinical Journal of the American Society of Nephrology*, 16(3):396–406, 2021.

[64] A. Leligdowicz and M. A. Matthay. Heterogeneity in sepsis: new biological evidence with clinical applications. *Critical Care*, 23, 2019.

[65] Zhang Lin, Xinru Liu, Lulu Sun, Jinbao Li, Zhenglin Hu, Haisheng Xie, Xianpeng Zu, Xiaoming Deng, and Weidong Zhang. Comparison of sepsis rats induced by caecal ligation puncture or staphylococcus aureus using a lc-qtof-ms metabolomics approach. *Infection, Genetics and Evolution*, 43:86–93, 2016.

[66] Annabelle Lint, Robert S. Parker, and G. Clermont. Modeling blood urea nitrogen (bun) dynamics to estimate dialysis adequacy. *Journal of critical care*, 2023.

[67] V. Liu, G. J. Escobar, J. D. Greene, J. Soule, A. Whippy, D. C. Angus, and T. J. Iwashyna. Hospital deaths in patients with sepsis from 2 independent cohorts. *Jama-Journal of the American Medical Association*, 312(1):90–92, 2014.

[68] Thaminda Liyanage, Toshiharu Ninomiya, Vivekanand Jha, Bruce Neal, Halle Marie Patrice, Ikechi Okpechi, Ming-hui Zhao, Jicheng Lv, Amit X Garg, and John Knight. Worldwide access to treatment for end-stage kidney disease: a systematic review. *The Lancet*, 385(9981):1975–1982, 2015.

[69] Francesco Locatelli, Andrea Cavalli, Benedetta Tucci, Sara Viganò, and Salvatore Di Filippo. Opinion: Can chronic volume overload be recognized and prevented in hemodialysis patients? newer methods to assess volume status. In *Seminars in Dialysis*, volume 22, pages 492–494. Blackwell Publishing Ltd Oxford, UK.

[70] Jan Lötvall, Cezmi A Akdis, Leonard B Bacharier, Leif Bjermer, Thomas B Casale, Adnan Custovic, Robert F Lemanske Jr, Andrew J Wardlaw, Sally E Wenzel, and Paul A Greenberger. Asthma endotypes: a new approach to classification of disease entities within the asthma syndrome. *Journal of Allergy and Clinical Immunology*, 127(2):355–360, 2011.

[71] Etienne Macedo, Bethany Karl, Euyhyun Lee, and Ravindra L Mehta. A randomized trial of albumin infusion to prevent intradialytic hypotension in hospitalized hypoalbuminemic patients. *Critical care*, 25(1):1–8, 2021.

[72] J. C. Marshall. Special issue: Sepsis why have clinical trials in sepsis failed? *Trends in Molecular Medicine*, 20(4):195–203, 2014.

[73] Finnian R Mc Causland, Brian Claggett, Venkata S Sabbisetti, Petr Jarolim, and Sushrut S Waikar. Hypertonic mannitol for the prevention of intradialytic hypotension: a randomized controlled trial. *American Journal of Kidney Diseases*, 74(4):483–490, 2019.

[74] Finnian R Mc Causland, Lisa M Prior, Eliot Heher, and Sushrut S Waikar. Preservation of blood pressure stability with hypertonic mannitol during hemodialysis initiation. *American journal of nephrology*, 36(2):168–174, 2012.

[75] Burkhard Michaeli, Alexandre Martinez, Jean-Pierre Revelly, Marie-Christine Cayeux, René L Chioléro, Luc Tappy, and Mette M Berger. Effects of endotoxin on lactate metabolism in humans. *Critical care*, 16(4):1–11, 2012.

[76] Stefano Monti, Pablo Tamayo, Jill Mesirov, and Todd Golub. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Machine learning*, 52(1):91–118, 2003.

[77] Jair Munoz Mendoza, Sumi Sun, Glenn M Chertow, John Moran, Sheila Doss, and Brigitte Schiller. Dialysate sodium and sodium gradient in maintenance hemodialysis: a neglected sodium restriction approach? *Nephrology Dialysis Transplantation*, 26(4):1281–1287, 2011.

[78] Himani N Murdeshwar and Fatima Anjum. Hemodialysis. 2020.

[79] Kyla L Naylor, S Joseph Kim, Eric McArthur, Amit X Garg, Megan K McCallum, and Gregory A Knoll. Mortality in incident maintenance dialysis patients versus incident solid organ cancer patients: a population-based cohort. *American Journal of Kidney Diseases*, 73(6):765–776, 2019.

[80] Anastassia Newbury, Katharine D Harper, Arianna Trionfo, Frederick V Ramsey, and Joseph J Thoder. Why not life and limb? vasopressor use in intensive care unit patients the cause of acute limb ischemia. *Hand*, 15(2):177–184, 2020.

[81] Ogochukwu Chinedum Okoye, Henry Enyinmisan Slater, and Nilum Rajora. Prevalence and risk factors of intra-dialytic hypotension: a 5 year retrospective report from a single nigerian centre. *Pan African Medical Journal*, 28(1), 2017.

[82] Steven M Opal, R Phillip Dellinger, Jean-Louis Vincent, Henry Masur, and Derek C Angus. The next generation of sepsis trials: What's next after the demise of recombinant human activated protein c? *Critical care medicine*, 42(7):1714, 2014.

[83] Regina Padmanabhan, Nader Meskin, and Wassim M Haddad. Closed-loop control of anesthesia and mean arterial pressure using reinforcement learning. *Biomedical Signal Processing and Control*, 22:54–64, 2015.

[84] Regina Padmanabhan, Nader Meskin, and Wassim M Haddad. Learning-based control of cancer chemotherapy treatment. *IFAC-PapersOnLine*, 50(1):15127–15132, 2017.

[85] Michael R Pinsky, Jean-Louis Vincent, Jacques Deviere, Marisa Alegre, Robert J Kahn, and Etienne Dupont. Serum cytokine levels in human septic shock: relation to multiple-system organ failure and mortality. *Chest*, 103(2):565–575, 1993.

[86] Benjamin L Ranard, Murad Megjhani, Kalijah Terilli, Kevin Doyle, Jan Claassen, Michael R Pinsky, Gilles Clermont, Yoram Vodovotz, Shadnaz Asgari, and Soojin Park. Identification of endotypes of hospitalized covid-19 patients. *Frontiers in medicine*, 8, 2021.

[87] V Marco Ranieri, B Taylor Thompson, Philip S Barie, Jean-François Dhainaut, Ivor S Douglas, Simon Finfer, Bengt Gårdlund, John C Marshall, Andrew Rhodes, and Antonio Artigas. Drotrecogin alfa (activated) in adults with septic shock. *New England Journal of Medicine*, 366(22):2055–2064, 2012.

[88] A. Reynolds, J. Rubin, G. Clermont, J. Day, Y. Vodovotz, and G. B. Ermentrout. A reduced mathematical model of the acute inflammatory response: I. derivation of model and analysis of anti-inflammation. *Journal of Theoretical Biology*, 242(1):220–236, 2006.

[89] Emanuel P Rivers, Anja Kathrin Jaehne, H Bryant Nguyen, Demosthenes G Papamatheakis, Daniel Singer, James J Yang, Samantha Brown, and Howard Klausner. Early biomarker activity in severe sepsis and septic shock and a contemporary review of immunotherapy trials: not a time to give up, but to give it earlier. *Shock*, 39(2):127–137, 2013.

[90] F. Roufosse. Targeting the interleukin-5 pathway for treatment of eosinophilic conditions other than asthma. *Frontiers in Medicine*, 5, 2018.

[91] Anirban Roy, Gilles Clermont, Silvia Daun, and Robert S Parker. A mathematical model of acute inflammatory response to endotoxin challenge. In *AIChE Annual Meeting, Salt Lake City, UT, 538g*.

[92] Prabir Roy-Chaudhury, Jim A Tumlin, Bruce A Koplan, Alexandru I Costea, Vijay Kher, Don Williamson, Saurabh Pokhariyal, David M Charytan, James Tumlin, and Vikranth Reddy. Primary outcomes of the monitoring in dialysis study indicate that clinically significant arrhythmias are common in hemodialysis patients and related to dialytic cycle. *Kidney international*, 93(4):941–951, 2018.

[93] Kristina E Rudd, Sarah Charlotte Johnson, Kareha M Agesa, Katya Anne Shackelford, Derrick Tsoi, Daniel Rhodes Kievlan, Danny V Colombara, Kevin S Ikuta, Niranjan Kissoon, and Simon Finfer. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the global burden of disease study. *The Lancet*, 395(10219):200–211, 2020.

[94] Ramin Sam. Hemodialysis: diffusion and ultrafiltration. *Austin J Nephrol Hypertens*, 1(2):1010, 2014.

[95] Tonio Schoenfelder, Xiaoyu Chen, and Hans-Holger Bleß. Effects of continuous and intermittent renal replacement therapies among adult patients with acute kidney injury. *GMS health technology assessment*, 13, 2017.

[96] W. Schulte, J. Bernhagen, and R. Bucala. Cytokines in sepsis: Potent immunoregulators and potential therapeutic targets-an updated view. *Mediators of Inflammation*, 2013, 2013.

[97] Nitin Seam and Anthony F Suffredini. Mechanisms of sepsis and insights from clinical trials. *Drug Discovery Today: Disease Mechanisms*, 4(2):83–93, 2007.

[98] Junhee Seok, H Shaw Warren, Alex G Cuenca, Michael N Mindrinos, Henry V Baker, Weihong Xu, Daniel R Richards, Grace P McDonald-Smith, Hong Gao, and Laura Hennessy. Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proceedings of the National Academy of Sciences*, 110(9):3507–3512, 2013.

[99] Khaled Shawwa, Panagiotis Kompotiatis, Jacob C Jentzer, Brandon M Wiley, Amy W Williams, John J Dillon, Robert C Albright, and Kianoush B Kashani. Hypotension within one-hour from starting crrt is associated with in-hospital mortality. *Journal of Critical Care*, 54:7–13, 2019.

[100] Joohyun Shin, Thomas A Badgwell, Kuang-Hung Liu, and Jay H Lee. Reinforcement learning–overview of recent progress and implications for process control. *Computers Chemical Engineering*, 127:282–294, 2019.

[101] Tatsuya Shoji, Yoshiharu Tsubakihara, Masamitsu Fujii, and Enyu Imai. Hemodialysis-associated hypotension as an independent risk factor for two-year mortality in hemodialysis patients. *Kidney international*, 66(3):1212–1220, 2004.

[102] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, and Marc Lanctot. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

[103] Jonathan A Silversides, Ruxandra Pinto, Rottem Kuint, Ron Wald, Michelle A Hladunewich, Stephen E Lapinsky, and Neill KJ Adhikari. Fluid balance, intradialytic hypotension, and outcomes in critically ill patients undergoing renal replacement therapy: a cohort study. *Critical care*, 18(6):1–10, 2014.

[104] Mervyn Singer, Clifford S Deutschman, Christopher Warren Seymour, Manu Shankar-Hari, Djillali Annane, Michael Bauer, Rinaldo Bellomo, Gordon R Bernard, Jean-Daniel Chiche, and Craig M Coopersmith. The third international consensus definitions for sepsis and septic shock (sepsis-3). *Jama*, 315(8):801–810, 2016.

[105] Peter D Sozou. On hyperbolic discounting and uncertain hazard rates. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1409):2015–2020, 1998.

[106] W Sulowicz and A Radziszewski. Pathogenesis and treatment of dialysis hypotension. *Kidney International*, 70:S36–S39, 2006.

[107] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[108] S. Svenningsen and P. Nair. Asthma endotypes and an overview of targeted therapy for asthma. *Frontiers in Medicine*, 4, 2017.

[109] United States Renal Data System. 2020 usrds annual data report: Epidemiology of kidney disease in the united states. Report, , 2020.

[110] James E Tattersall, Dominic DeTakats, Paul Chamney, Roger N Greenwood, and Ken Farrington. The post-hemodialysis rebound: predicting and quantifying its effect on kt/v. *Kidney international*, 50(6):2094–2102, 1996.

[111] Celeste M Torio and Roxanne M Andrews. Statistical brief 160. *Healthcare Cost and Utilization Project (HCUP) Statistical Briefs*, 2013.

[112] Panagiotis Tsirigotis, Spiros Chondropoulos, Konstantinos Gkirkas, Josef Meletiadis, and Ioanna Dimopoulou. Balanced control of both hyper and hypo-inflammatory phases as a new treatment paradigm in sepsis. *Journal of Thoracic Disease*, 8(5):E312, 2016.

[113] Yelena V Lerman and Minsoo Kim. Neutrophil migration under normal and sepsis conditions. *Cardiovascular Haematological Disorders-Drug Targets (Formerly Current Drug Targets-Cardiovascular Hematological Disorders)*, 15(1):19–28, 2015.

[114] Patricia Vassallo and Richard G Trohman. Prescribing amiodarone: an evidence-based review of clinical indications. *Jama*, 298(11):1312–1322, 2007.

[115] Sergei Vassilvitskii and David Arthur. k-means++: The advantages of careful seeding. In *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1027–1035.

[116] Jean-Louis Vincent and Jan Bakker. Blood lactate levels in sepsis: in 8 questions. *Current Opinion in Critical Care*, 27(3):298–302, 2021.

[117] Jean-Louis Vincent, Amanda Quintairos e Silva, Lúcio Couto, and Fabio S Taccone. The value of blood lactate kinetics in critically ill patients: a systematic review. *Critical care*, 20(1):1–14, 2016.

[118] H Shaw Warren, Catherine Fitting, Eva Hoff, Minou Adib-Conquy, Laura Beasley-Topliffe, Brenda Tesini, Xueya Liang, Catherine Valentine, Judith Hellman, and Douglas Hayden. Resilience to bacterial infection: difference between species could be due to proteins in serum. *The Journal of infectious diseases*, 201(2):223–232, 2010.

[119] Hiroshi Watanabe and Tohru Minamino. Atrial fibrillation in patients with end-stage kidney disease on dialysis. *Internal Medicine*, pages 0735–17, 2018.

[120] Christopher J. C. H. Watkins and Peter Dayan. Q-learning. *Machine Learning*, 8(3):279–292, 1992.

[121] Marije Wijnberge, Bart F Geerts, Liselotte Hol, Nikki Lemmers, Marijn P Mulder, Patrick Berge, Jimmy Schenk, Lotte E Terwindt, Markus W Hollmann, and Alexander P Vlaar. Effect of a machine learning–derived early warning system for intraoperative hypotension vs standard care on depth and duration of intraoperative hypotension during elective noncardiac surgery: the hype randomized clinical trial. *Jama*, 323(11):1052–1060, 2020.

[122] K DOQI Workgroup. K/doqi clinical practice guidelines for cardiovascular disease in dialysis patients. *Am J Kidney Dis*, 45:S1–S153, 2005.

[123] Helen L Wright, Robert J Moots, Roger C Bucknall, and Steven W Edwards. Neutrophil function in inflammation and inflammatory diseases. *Rheumatology*, 49(9):1618–1631, 2010.

[124] D. M. Yealy, J. A. Kellum, D. T. Huang, A. E. Barnato, L. A. Weissfeld, F. Pike, T. Terndrup, H. E. Wang, P. C. Hou, F. LoVecchio, M. R. Filbin, N. I. Shapiro, D. C. Angus, and Cess Investigators Pro. A randomized trial of protocol-based care for early septic shock. *New England Journal of Medicine*, 370(18):1683–1693, 2014.

[125] Hao Hong Yiu, Andrea L Graham, and Robert F Stengel. Dynamics of a cytokine storm. 2012.

[126] JH Yoon, S Malakouti, M Hauskrecht, MR Pinsky, and G Clermont. *Machine Learning Driven Prediction of Hypotension Using Real World Multigranular Data*, pages A5615–A5615. American Thoracic Society, 2022.

[127] Joo Heung Yoon, Vincent Jeanselme, Artur Dubrawski, Marilyn Hravnak, Michael R Pinsky, and Gilles Clermont. Prediction of hypotension events with physiologic vital sign signatures in the intensive care unit. *Critical Care*, 24(1):1–9, 2020.

[128] Joo Heung Yoon, Lidan Mu, Lujie Chen, Artur Dubrawski, Marilyn Hravnak, Michael R Pinsky, and Gilles Clermont. Predicting tachycardia as a surrogate for instability in the intensive care unit. *Journal of Clinical Monitoring and Computing*, 33(6):973–985, 2019.

[129] William J Youden. Index for rating diagnostic tests. *Cancer*, 3(1):32–35, 1950.

[130] Chao Yu, Jiming Liu, and Shamim Nemati. Reinforcement learning in healthcare: A survey. *arXiv preprint arXiv:1908.08796*, 2019.