MODEL-BASED DECISION SUPPORT FOR SEPSIS ENDOTYPES

by

Li Ang Zhang

BS, Chemical Engineering and Biomedical Engineering,

Carnegie Mellon University, 2012

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

2017

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This dissertation was presented

by

Li Ang Zhang

It was defended on

August 25th 2017

and approved by

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

Ipsita Banerjee, Ph.D., Associate Professor, Department of Chemical and Petroleum Engineering

Gilles Clermont, M.D., Professor, Department of Critical Care Medicine

David Swigon, Ph.D., Associate Professor, Department of Mathematics

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

MODEL-BASED DECISION SUPPORT FOR SEPSIS ENDOTYPES

Li Ang Zhang, PhD

University of Pittsburgh, 2017

Sepsis is a high mortality syndrome characterized by organ dysfunction due to a severe and dysregulated acute inflammatory response to infection. Research into therapies for this syndrome has historically ended in failure, which has largely been attributed to the elevated levels of subject heterogeneity. What may have been previously attributed to variability in sepsis may be due to mechanistic differences between patients. Endotypes are distinct subtypes of disease, where underlying causes such as mechanistic or pathway related differences manifest into phenotypes of disease.

The lack of mechanistic understanding of immune mediator dynamics and the responses they trigger necessitates a mathematical modeling approach to analyze its complexities. A transfer function model is proposed to describe and cluster the dynamics of key inflammatory mediators. Five sepsis endotypes were discovered and revealed motifs of overwhelming inflammation, various levels of immunosuppression, sustained inflammation, and immunodeficiency. An accurate clinical tool was proposed to classify subjects into endotypes using six-hour trajectories of clinical data.

A physiological ordinary differential equation model of sepsis is proposed that characterizes the interactions of inflammatory signaling molecules, neutrophils, and macrophages across the bone, blood, and tissue compartments of the body. This model used to generate individual subject fits against human sepsis data. Population-level parameter analysis implicated macrophage cell death and cytokine half- dynamics in endotype-level differences.

Several proof-of-concept statistical models were introduced to demonstrate that it is possible to estimate the pre-hospital time of sepsis subjects and to quantify their sepsis-

induced systemic tissue damage. A nearest-neighbor-based method was verified against animal and human data and revealed that identifying infection time-zero of sepsis patients can be quickly estimated with high accuracy using commonly measured clinical features. A logistic regression ensemble model demonstrated revealed early organ dysfunction were significant contributors to systemic damage and mortality. Knowledge of time-zero and systemic damage levels, in combination with an endotype classifier, provides clinicians with a clear depiction of where a subject is located on their sepsis trajectory. Such a decision support system enables therapy timing, early organ support, and targeted therapies to guide personalized treatment and shift patients towards better outcomes in sepsis.

Keywords: mathematical modeling, sepsis, cytokines, ordinary differential equations, machine learning, statistical analysis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I begin by thanking my family for supporting me through this Ph.D. journey. I left my home to pursue higher education in Pittsburgh and somehow ended up staying there for 9 years. I would like to thank my grandfather for introducing me to the clinical world and for inspiring me to help others. My mom and dad deserve more than my thanks for their mentorship, advice, and assistance throughout the years. Thank you for pushing me to do grad school. I promise I will call more often.

Next, I want to thank Annie for putting up with me for many, many years. Somehow, life keeps putting us in a long-distance relationship, but we persevered. You have been a great best friend, sous chef, roommate, travel companion, and life partner. I truly apologize that I had to propose to the Ph.D. committee first.

Matt, thank you for being my "brofessor" and teaching me the secrets to getting swole. I enjoyed shredding powder, brewing Lyapunov, fishing for catfish, and driving the Autobahn with you. Michelle, I enjoyed hanging out with you and discussing our love for dogs and bourbon. Thank you for being my best friend in Pittsburgh (Matt lost that privilege when he decided to graduate first). Ari, thank you for the board game nights and long discussions about MCMC. Tim, thank you for introducing me to Python, ArchOS, and other tangents that effectively distracted me from my research problems. Flo, thank you for being the first to use APT-MCMC in its C++ form. I also enjoyed our lengthy conversations about politics and how we want to improve the world. Megan, I hope that someday we get to run together. I am still searching for a marathon buddy after my last one bailed on me. Christy, thank you for being a good friend, hosting events at your place, and teaching me how to brew. Evan, Monica, and Matt 2.0, it has been great working with you in the lab and I envision great things from all of you in the future.

Bob, thank you for letting me into your lab (on the second try), for tempering my ego, and for tolerating my antics. I appreciate the freedom that you gave me over the years on my research and the unfettered access to your servers. Finally, it was a joy to learn from you and your ideas (that I always think will not work) has been invaluable. I will always enjoy discussing Advanced Applied Bioreactors with you. Gilles, thank you for your guidance in an impressive array of fields. In one meeting with you, I can learn about a clinical topic and then learn about machine learning. Ipsita, thank you for being my mentor since my CMU days and for enduring my early mistakes in the lab. David, it has been a pleasure working with you in MODS and your insights have saved me on many occasions.

I apologize for the people whom I missed who have helped me on this journey. Grad school has helped me grow in many ways and I am eager to make my mark in the world. I believe that in our lifetime, a major, positive shift towards personalized medicine is coming. It is my sincere hope that this work serves as a stepping stone towards that goal.

Li Ang Zhang

张黎昂

August 2017,

Pittsburgh, Pennsylvania

1.0 INTRODUCTION

1.1 BACKGROUND AND SIGNIFICANCE

Each year, sepsis afflicts millions worldwide with extensive morbidity and mortality [1]. This syndrome predominately affects opposite sides of the age spectrum: the pediatric and the elderly, where the incidence and mortality rates increase for each [2]. As such, mortality exceeds 40% for the 85 and older age group [2]. Treatment of sepsis has proven to be a challenge because of the fast-changing dynamics, multiple trajectories, and outcomes of the syndrome [3, 4]. Clinical features of sepsis are highly variable and are dependent on the infection site, type of infection, patient demographics, coexisting illnesses, organ dysfunction, and time since sepsis onset [5]. In the past year, sepsis was redefined for the third time, outdating several diagnosis criteria, such as the Systemic Inflammatory Response Syndrome and the Logistic Organ Dysfunction System. The prevalence of multiple criteria is indicative of the current debate and disagreement among medical professionals [6]. As a result, there is currently no gold standard with which to diagnose sepsis [7].

Advancements in sepsis therapeutics can be generalized as stagnant, where within the past 30 years, the majority of clinical trials for pharmacological interventions in sepsis showed either no effect or a negative effect on mortality [5, 8]. The most promising sepsis therapy, activated protein-c, was briefly approved by the Food and Drug Administration (FDA) following a successful trial in 2001. However, by 2011, a follow-up study revealed that this therapy was ineffective at decreasing mortality and thus the FDA subsequently removed it from the market [1]. To date, there is currently no FDA-approved drug that replaced activated protein-C to treat sepsis [7].

The various pathways to failure as well as the fact that sepsis is often coupled with many comorbidities result in high inter- and intra-patient variabilities of the syndrome. This has led to difficulties in defining, diagnosing, and treating sepsis[7].

Sepsis is a severe acute physiological response that results from the systemic effects of a dysregulated acute inflammatory response to infection and is characterized by organ dysfunction [7]. Acute inflammation represents the first line of defense against infection. First, the body mobilizes immune mediators [3]. These mediators enable the body to mount an antimicrobial response to remove the source of infection, but in doing so damages healthy cells during the process. During severe infections and/or severe damage, overwhelming inflammation may occur [9]. This leads to the systemic inflammation syndrome known as sepsis. Once triggered, three outcomes are possible [3]. The first (i) is the healthy outcome, where the body is able to remove the infection source and inflammatory pathways are properly deactivated. The second (ii) is immunosuppression or immunodeficiency, which can lead to recurrent infection or uncontrolled growth of the invading pathogens [3]. The third case (iii), aseptic sepsis, is persistent inflammation where the dysregulation causes sustained inflammation even after the infection source has been dealt with [3].

During an acute inflammatory response, there are three phases: initiation, the inflammation response, and resolution. Embedded in almost all tissue, resident macrophages are responsible for the recognizing nearby pathogens and initiating the inflammatory response by producing a variety of inflammatory mediators including chemokines, cytokines, and vasoactive agents [10]. Important cytokines such as Tumor Necrosis Factor-α (TNF-α) and Interleukin-6 (IL-6) recruit neutrophils towards the site of infection by effecting change on the cell membrane, causing neutrophils to physically roll along endothelial walls [11]. The Interleukin-8 (IL-8) chemokine stops the rolling and allows neutrophils to extravasate through endothelial walls without erythrocyte leakage [10, 11]. Neutrophils migrate through the tissue and, upon reaching the infection site, become activated and release toxic substances to eliminate the pathogen. Most of these substances are reactive oxygen species. During this stage, M1 (pro-inflammatory monocyte-derived) macrophages are simultaneously recruited to the infection site via similar cytokine and chemokine interactions in order to aid the elimination effort and to produce additional mediators to sustain inflammation

[12, 13]. Upon removing the pathogen, macrophages shift production from pro-inflammatory mediators to anti-inflammatory mediators and begin the resolution phase [10, 13]. IL-6 transitions its effect from a neutrophil recruiter to a monocyte recruiter [14]. During this stage, recruited monocytes exhibit an anti-inflammatory phenotype and may differentiate into M2 macrophages [13]. M2 macrophages become more prevalent in tissue and are responsible for producing Interleukin-10 (IL-10), a master anti-inflammatory cytokine [13, 15]. Additionally, macrophages remove dead tissue and promote repair [10, 12]. Finally, the inflammatory cascade is turned off due to the presence of IL-10 and by the lack of IL-6 and TNF-α producing M1 macrophages. This healthy scenario represents case (i). This process is illustrated by the simplified diagram shown in Figure 1.1. The healthy scenario, case (i), occurs when pathogens are fully eliminated and the anti-inflammatory compartment successfully resolves the inflammatory compartment.

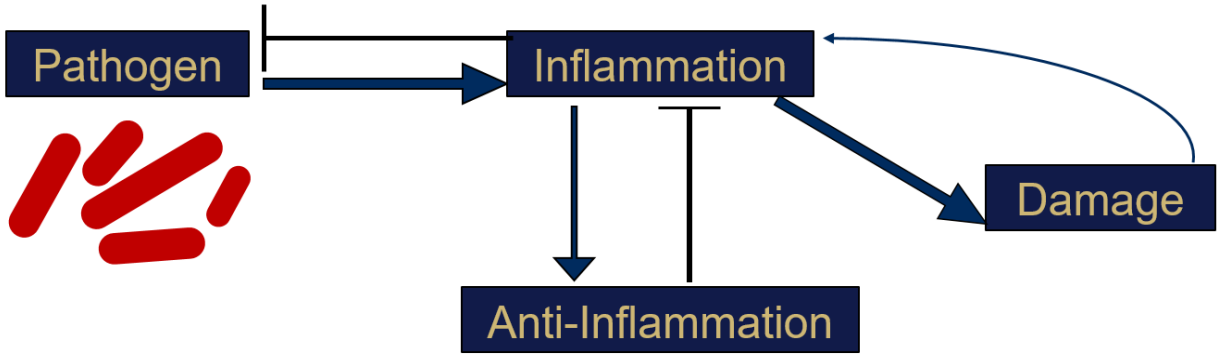


Figure 1.1: Simplified schematic of the inflammatory response. Pathogens invading the host triggers an inflammatory response. This response eliminates pathogens, but causes unavoidable collateral damage to healthy tissue, thereby causing damage. This damage further initiates inflammation. Inflammation also triggers an anti-inflammatory response. The anti-inflammatory response inhibits the inflammatory response. Adapted from Reynolds, et al, 2006 [4].

Sepsis may result when one or more aspects of acute inflammatory response become dysregulated. This often occurs due to high pathogen load or within the immunocompromised elderly. During the neutrophil-dominant stage, high levels of neutrophil activity may trigger the aseptic sepsis case (iii). The substances released by neutrophils are highly toxic and cause unavoidable collateral damage to nearby tissue [10]. This damage triggers further inflammation and leads to a persistent inflammatory response, which in turn causes more tis-

Table 1.1: Overview of Select Sepsis Therapeutic Agents and Their Efficacy in Animal Models and in Human Trials.

Therapeutic Agent		**Mortality Benefit**	
Name	**Description**	**Positive**	**None or Negative**
IL1-RA	IL-1 receptor antagonist	Mouse[16], Baboon[17]	Human[18]
MPSS	Glucocorticoid	Baboon[19]	Human [20]
Lenercept	TNF-α receptor antagonist	Baboon[21]	Human [22]
CDP571	anti-TNF-α antibody	Baboon[23]	Human [24]
BN 5021	PAF antagonist	Mouse [25]	Human [26]
BB-882	PAF receptor antagonist	Mouse [27]	Human [28]
TCV-309	PAF antagonist	Mouse [29]	Human [30]
TAK-242	TLR-4 inhibitor	Mouse [31]	Human [32]
Tifacogin	Preventing coagulation	Mouse [33], Baboon [34]	Human [35]

(IL1-RA: Interleukin-1 receptor antagonist, MPSS: Methylprednisolone sodium succinate, TNF-α: Tumor necrosis factor alpha, PAF: Platelet activating factor, TLR-4: Toll-like receptor 4)

sue damage. This cycle persists even if the pathogen is fully eliminated. Organ dysfunction and failure, prolonged ICU stays, metabolic abnormalities, and/or death are common consequences of this case [7]. Another dysregulation occurs if the anti-inflammatory pathway is triggered prior to the full elimination of pathogens, triggering case (ii). A high inflammatory response may trigger a similarly high anti-inflammatory response too early. In this scenario, the host immune response is unable to mount a sufficient attack on the pathogen and the host may succumb to the infection.

1.2 PROBLEMS WITH CURRENT APPROACHES IN SEPSIS THERAPY

Given the complexities of the inflammatory pathways, much of sepsis research and experiments are derived from animal models. Promising therapies in animal models have often failed in human trials due to a lack of efficacy or due to safety concerns [36, 37]. Table 1.1 outlines several promising sepsis therapeutic agents over the years that demonstrated significant mortality benefits in animal trials but not human trials. The historical lack of

successful clinical trials has resulted in pharmaceutical companies referring to sepsis as a "graveyard" [38]. The sheer amount of inter- and intra-patient variability in septic humans contribute to this translational disconnect. For example, this disconnect may result from patient differences in pre-existing disease burden or the timing of applied therapeutics [39–42]. Animal models typically use identical animals with uniform levels of inflammation (low genetic variability and controlled induction of sepsis) and controlled therapy timing (applied exactly x hours after experimentally induced sepsis). Applying this level of control in human trials is currently infeasible because of the inability of clinicians to quantify such variability for each human patient.

Mechanistic understanding of sepsis may hold the answers behind sepsis heterogeneity and differences in inflammatory pathways. The mechanistic differences in inflammatory pathways can result from immunosenescence, age-related differences in immunity [43]. Analysis of genomic data from human sepsis subjects suggests that certain populations are predisposed to sepsis due to mechanistic implications of certain gene mutations [44]. Furthermore, a fixed-effect analysis of this data, a statistical approach to identify confounding effects on these gene mutations, revealed that age, ethnicity, and comorbidities may be important factors. These findings lend evidence to the idea that sepsis heterogeneity can be attributed to pathological mechanistic differences between certain populations.

Given the long historical failure of sepsis therapies and the issues with patient variability, the preconception of a "one-size-fits-all" therapy may be inappropriate [41, 45]. Post-hoc analyses of several past human trials suggest several cohorts of sepsis patients may have benefited from the experimental therapies (despite the overall conclusion of no mortality benefit across trial subjects) [18, 46]. A targeted therapeutic approach to sepsis may be a promising outlook for treatment. Targeted therapeutics can be developed by understanding and use the disease subtypes to formulate a therapy. This has been shown to be an effective approach in the treatment of asthma. Asthmatic research using disease subtypes to develop targeted therapies have been successful, which lends support to the hypothesis that a singular therapy for sepsis is inappropriate [47, 48]. An "endotype" is a disease subtype characterized by pathological and mechanistic differences. Sepsis subtype identification has immense potential to improve clinical outcomes (such as lowering mortality or lowering

multiple organ failure rates), which is why it is surprising that there is a lack of existing literature on the topic. The existence of subtypes in sepsis has been hypothesized over the years, but few have tried to identify and characterize them [45, 49, 50].

1.3 DECISION SUPPORT SYSTEMS

To address the clinical challenges with sepsis, mathematical model-based decision support may provide the solution. Mathematics and *in silico* studies in sepsis have offered descriptive models that can aid clinicians in understanding the ongoing dynamics of sepsis and ascertain the effects of potential treatments [3, 4, 51–55]. Some of these studies have analyzed parameter-based outcome bifurcations, revealing that mechanistic differences in the parameter space may lead to death or survival states [3, 4]. However, few models in existing literature have been calibrated against human sepsis data, limiting the translatability of its results to the clinic. The goal of this dissertation is to use mathematical models, calibrated on human sepsis data, to produce insights into how and why sepsis endotypes manifest and to provide clinicians with the appropriate quantitative tools to deal with the sepsis heterogeneity.

1.4 DISSERTATION OVERVIEW

The focus of this dissertation is to identify, characterize, and analyze the subtypes of adult sepsis. The underlying hypothesis is that providing clinicians with the ability to classify patients into sepsis subtypes will improve sepsis therapy and improve clinical outcomes. Endotype-driven approaches may address major components of the translational disconnect between animal models and human trials. Identifying sepsis endotypes may capture the root cause of much of the variability associated with the syndrome. Analysis of specific subtypes can identify how mechanistic differences cause the dysregulations in Section 1.1. Knowledge

of such mechanistic differences can enable the discovery of new pharmaceutical targets for therapy or reveal cohorts that benefit from specific treatments.

The secondary focus of this dissertation is the creation of translatable clinical tools to support clinicians in their decision-making process during the treatment of sepsis patients. These tools aim to provide early sepsis subtype classification and enable therapy timing and were designed with (i) low measurement burden (relatively easy to measure biomarkers) and (ii) low temporal burden (short sequential measurements only).

The dissertation is organized as follows. Chapter 2 identifies clusters of sepsis patients with similar clinical features using a statistical approach. Hierarchical clustering revealed that inflammatory mediators predominately drove the formation of the clusters and that they are associated with sepsis-related mortality and organ dysfunction.

Chapter 3 studies how the dynamics of cytokines, the aforementioned inflammatory mediators, reveal distinct patterns. These behaviors were characterized by five distinct groups. This work was extended as a proposed clinical tool, which quickly classified a subject's endotype within 6 hours of clinical presentation.

Chapter 4 focuses on the importance of cytokines and their dynamical relationship to endotyping in sepsis. Cytokines are responsible for recruiting white blood cells, which ultimately carry out the relevant antimicrobial actions during acute inflammation. To analyze the discovered endotypes in the broader context of sepsis, a tissue and blood compartmental mechanistic ordinary differential equation model was developed to capture both white blood cell and cytokine dynamics. This model was used to generate endotype-specific fits, which was used to analyze endotype pathology.

Chapter 5 describes exploratory models that address other aspects of sepsis variability that can improve the translatability of animal research to humans. To address the tight experimental controls with regards to therapy timing and controlled injury levels seen in animal research, machine-learning approaches were taken to identify sepsis time zero in patients and to quantify global tissue damage during sepsis. With the knowledge of disease time zero and systemic damage, a clinician can be provided with a grasp of where a septic patient is located along an endotype's disease trajectory. This combined knowledge can be

used to finely select suitable patients for clinical trials and may offer improvements to the status quo of care.

Chapter 6 describes several notable mathematical tools developed for this project. A network optimization tool is first introduced as a method to detect inflammatory pathways in data. This tool identifies the best boolean rule network that best describes a dataset to aid in the analysis of immune pathways and model structure design. A parameter fitting toolkit, APT-MCMC, is then introduced as an efficient Monte Carlo method to fit such models. APT-MCMC contains features for advanced fitting scenarios pertaining to mathematical models in medicine, including fitting initial conditions, handling of infusions, and fitting left-censored data.

Chapter 7 provides perspective on the contributions this dissertation makes in the field of sepsis. Short and long-term future improvements are proposed, with the ultimate goal envisioned as a simple to use bedside tool to enable clinicians to personalize treatment for each septic patient under their care.

2.0 HIERARCHICAL CLUSTERING APPROACH TO IDENTIFY SEPSIS ENDOPHENOTYPES

Clinical features of sepsis vary greatly among patients and are dependent on pathogen factors, host factors, time elapsed between onset of infection and clinical presentation, and interventions [5]. Failure to understand the vital interplay between these factors has contributed to many failed sepsis clinical trials in the past few decades, and there is a clinical need to revisit the current approach to sepsis research [8, 56, 57]. Sepsis research may benefit by borrowing a promising technique from another disease area that may be the key to capturing physiologic variability: the study of disease endotypes.

Asthma research has demonstrated success in applying clinical endotypes for the purposes of clinical diagnosis and classification, predictions, and therapy development [47]. Endotypes are disease subtypes that are caused by pathobiological differences and are often expressed through phenotypic variability [48]. Several severe asthma endotypes and their indicative biomarkers have been identified, enabling the research and development of targeted therapeutics [48]. For instance, glucocorticoids are the gold standard for asthma treatment, but their efficacy varies, sometimes to the point of no benefit, among asthma endotypes [47]. Furthermore, asthma endotypes are being used to guide clinical study design [47]. Consideration of endotypes in sepsis, which dispels the notion of a single therapeutic approach, may lead to better targeted and more effective treatment [45].

The use of endotypes in sepsis may reveal key underlying pathobiological differences between different host responses to sepsis and may address much of the clinical variability observed in septic patients. In the field of asthma, researchers have demonstrated that endotypes can be derived from clusters of phenotypes despite differences in clinical presentations and disease severity [47].

Several groups have previously identified sepsis clusters with differing mortality rates [49, 50]. Fjell, et al., used an unsupervised clustering approach to segregate patients using cytokines and other signaling molecules [49]. They identified three groups of varying cytokine magnitudes. The highest cytokine group was associated with elevated mortality and morbidity such as renal failure and coagulopathy. Knox, et al., used a self-organizing map approach to identify four clusters with differing clinical outcomes [50]. This approach trained a map of nodes to represent a dimensionally-reduced version of a clinical data set (with variables such as temperature, white blood cell count, etc.). The identified node clusters were re-colored to visually demonstrate differences between clusters in sepsis-related organ failure assessment (SOFA) scores and sepsis severity [58].

The ideal endotype classification system for sepsis would involve a combination of a few highly informative signaling molecules (cytokines) and intuitive clinical biomarkers, such as vital signs, to be practical in a clinical setting. The work presented by Fjell, et al., utilizes only signaling molecules, many of which are rarely collected in an ICU. The work by Knox, et al., takes a dimension reduction approach which renders it difficult to interpret the clinical features of each endophenotype.

The underlying hypothesis for this work was sepsis endotypes were identifiable and separable via clinical and diagnostic biomarkers in a multi-dimensional manner. Because endotypes are generally explained by genetic differences (or some other underlying mechanistic distinction), the stratified groups in this work are better described as endophenotypes. The hierarchical dendrogram clustering and heatmap approach from Fjell, et al. was adapted for use with a rich and longitudinal human sepsis data set. This approach allowed the visualization of defining features of each endophenotype that would be lost in a dimension-reduction approach. Finally, an important end goal was to establish endophenotypes with distinct clinical features, including 14-day all-cause mortality and risk of multiple organ failure.

2.1 METHODS

A retrospective analysis of the Protocol-Based Care for Early Septic Shock (ProCESS) trial was conducted [59]. This trial was propelled by the seminal 2001 journal article by Rivers, et al., which demonstrated mortality benefits from early goal-directed therapy (EGDT) in sepsis [60]. EGDT involves monitoring hemodynamics using a central venous catheter and controlling arterial pressures and hematocrit to within pre-set therapeutic zones via fluids and vasopressors [60]. The ProCESS trial encompassed 31 emergency departments in the United states and enrolled 1341 subjects into one of three treatment arms to compare the clinical outcomes of EGDT, protocol-based care (combination of EGDT and standard care), and standard care. The primary clinical outcome was 60-day in-hospital mortality. At the conclusion of the study in 2014, the ProCESS investigators determined that there were no mortality differences among any of the three treatment arms. Furthermore, a retrospective analysis of all EGDT studies (including ProCESS) in 2015 concluded that EGDT had no effect on mortality outcomes in sepsis patients [61].

In light of this result, the ProCESS trial amassed a rich data set from its subjects. Longitudinal and high fidelity clinical data were collected for many of the 1341 subjects. This data was retrospectively analyzed for this chapter. 84 biomarkers were measured temporally. The variance among different biomarkers varied in scale and those with high variance were selected for analysis. Biomarkers were converted to the z-score (zero-mean and one standard deviation). The biomarkers were sorted by the range in z-scores (highest range to lowest range to represent the amount of variance within a biomarker). The top 50% of biomarkers were selected for inclusion in the heatmap analysis.

An availability analysis at baseline, as defined by trial enrollment, was performed on the remaining 42 biomarkers. To improve biomarker availability, missing baseline values were substituted by a measurement within 3 hours of time-zero, when necessary. This yielded a convenience cohort of 493 patients with the following biomarkers fully measured at baseline: glucose, platelet count, systolic and diastolic blood pressure, urine output, temperature, heart rate, Tumor Necrosis Factor-*alpha* (TNF), Interleukin-10 (IL-10), Interleukin-6 (IL-6), lactate, blood urea nitrogen, creatinine, white blood cell count, respiratory rate, and

Table 2.1: Comparison of convenience cohort for hierarchical clustering against the rest of ProCESS database.

Characteristic	ProCESS Database n=847	Whole Cohort n=493	
Age	61.5±17.9	60.0±17.2	
Male sex – no. (%)	457 (54.0%)	290 (58.8%)	
Race – no. (%)			
White	575 (67.9%)	340 (69.0%)	
Black	224 (26.4%)	109 (22.1%)	
Other	48 (5.7%)	44 (8.9%)	*
Source of sepsis – no. (%)			
Pneumonia	272 (32.1%)	171 (34.7%)	
Urinary Tract Infection	160 (18.9%)	124 (25.2%)	*
Intraabdominal infection	127 (15.0%)	49 (9.9%)	*
Infection of unknown source	112 (13.2%)	58 (11.8%)	
Skin or soft-tissue infection	63 (7.4%)	33 (6.7%)	
Catheter-related infection	27 (3.2%)	11 (2.2%)	
Central nervous system infection	7 (0.8%)	3 (0.6%)	
Endocarditis	4 (0.5%)	3 (0.6%)	
Other	52 (6.1%)	33 (6.7%)	
Determined not to have infection	23 (2.7%)	8 (1.6%)	
Positive blood culture – no. (%)	249 (29.4%)	147 (29.8%)	
Injury Severity			
APACHE II	20.5±7.5	21.2±7.8	
APACHE III	60.8±22.3	61.7±23.2	
Charlson	2.6±2.6	2.7±2.7	
SOFA	7.1±3.5	7.3±3.6	
Physiological variables			
Systolic blood pressure – mmHG	100.6±23.9	98.7±25.0	*
Heart rate – beats/min	103.3±20.6	103.0±20.1	
Temperature – C	37.3±1.3	37.3±1.4	
Respiratory rate – breaths/min	23.0±6.6	22.6±7.1	
Total bilirubin – mg/deciliter	1.5±2.2	1.4±1.8	
Serum lactate – mmol/liter	2.9±2.5	3.0±3.0	
Outcome			
14 All-cause Mortality – no. (%)	149 (17.6%)	90 (18.3%)	
Mortality – days	48.6±71.8	58.3±84.7	
Multiple Organ Failure – no. (%)			
Baseline	468 (55.3%)	282 (57.2%)	
Within 14 Days (Any)	473 (55.8%)	276 (56.0%)	
New Organ Failure in the first week – no. (%)			
Cardiovascular	506 (59.7%)	303 (61.5%)	
Respiratory	289 (34.1%)	183 (37.1%)	
Renal	27 (3.2%)	20 (4.1%)	
Stay in hospital – days	12.1±11.6	10.8±10.0	
ICU	5.3±7.3	4.5±4.6	
Serious Adverse Events – no. (%)	53 (6.3%)	29 (5.9%)	

$*$ denotes $p \leq 0.05$ from Dunn' or Chi-Squared tests to determine differences between continuous or categorical variables, respectively.

potassium. This convenience cohort was compared against the rest of the ProCESS subjects and the statistical results are presented in Table 2.1. Very few cohort characteristics or clinical outcomes were statistically different, indicating that this convenience cohort was representative of the ProCESS patients as a whole.

Missing baseline biomarkers were imputed in order to increase the number of analyzable biomarkers within the convenience cohort. Biomarkers with less than 5% missing at baseline were included after imputation using the predictive mean matching program from the *mice* package in R. Imputed biomarkers were sodium, chloride, hemoglobin, thrombin-antithrombin (TAT) complex, D-Dimer, and calcium. Finally, bilirubin, with 18% missing baseline values, was imputed using a linear model, where other baseline biomarkers served as predictors (see Equation 2.1).

$$X_{bilirubin}(0) = \beta_0 + \sum_{i=1}^{N} \beta_i X_i(0) \tag{2.1}$$

Fitting the β coefficients were performed via the linear regression function from the *Scipy* package in Python. The result was a complete dataset consisting of 493 subjects and 18 biomarkers at baseline. In order to analyze changes within this cohort over time, 6 and 24-hour post-trial enrollment data were considered as well. Missing measurements at these time points were imputed using a last measured value carried forward approach. Data values for patients that died by 6 or 24 hours were not imputed and were left as black columns within the heatmaps.

The final step of data processing was to normalize the data. The natural log of the data was taken to prevent biomarkers from exhibiting large ranges (such as cytokine measurements). An empirical cumulative distribution function (eCDF) was calculated for each biomarker across the entire convenience cohort using measurement values at 0, 6, 24, and 72 hours. Measurements at 72 hours were included to provide more data points for the eCDF but were not used to generate heatmaps due to the excessive amounts of missing measurements. Each biomarker was evaluated against their respective eCDF to obtain a normalized value between 0 and 1. This normalized value was linearly transformed to boundaries of [-10, 10]. A value of zero (corresponding to white on a heatmap) represents the population median for a biomarker. An important note is that a value of zero makes no indication

of a "healthy" state. Given the large number of biomarkers included for analysis and the severely ill cohort, no consideration was performed to account for each biomarker's criteria for healthy levels. This normalization scheme step ensured comparability of biomarkers among endophenotypes and time points.

2.1.1 Heatmap

Identifying patterns of distinct septic responses is nontrivial and necessitates unsupervised clustering because there is no currently accepted classification of sepsis in clinical practice. Hierarchical clustering was selected as the clustering algorithm because clusters merges can be visualized intuitively via a dendrogram tree. Hierarchical clustering and subsequent heat map visualization were performed using the *heatmap3* package in R. Clustering options was set to Euclidean distance and to the complete linkage function for bottom-up clustering. Clustering was performed on the baseline dataset with both biomarker and patient clustering enabled. Endophenotype identification was performed by analyzing the resulting heatmap and dendrogram. Biomarker differences in the heatmap, the heights of the patient-level dendrogram tree, and clinical intuition informed the creation of similar groups within the heatmap. Hourly visualizations of patient data at 6 and 24 hours were performed with the heatmap technique and ordered using the hierarchical clustering results of the baseline dataset to preserve the identified endophenotype groups.

Clinical outcomes were defined by 14-day all-cause mortality and 14-day multiple organ failure rates within each endotype. Unlike the clinical outcome in the ProCESS database (60-day in-hospital mortality), 14-day all-cause mortality isolated mortality likely due to the septic shock rather than downstream complications of sepsis, which is consistent with previous literature [62]. Furthermore, sepsis severity definitions were rendered obsolete by the 2016 redefinition of sepsis and multiple organ failure rates served as a proxy for sepsis severity for each endotype [7]. SOFA values were calculated for each of the 493 patients using the appropriate values at baseline and then on days 1 through 14 using daily maximum or minimum values. A distinction was made between baseline MOF and 14-day MOF in order to distinguish patients that were sick upon clinical presentation (with baseline MOF) or

patients who worsened (no baseline MOF, but developed within 14-days). Baseline MOF was defined by baseline SOFA ≥ 2. 14-day MOF was defined by any daily SOFA ≥ 2.

2.1.2 Statistical Analysis

Kruskal-Wallis and Chi-squared tests were performed on continuous and discrete variables, respectively, test for statistically significant differences in variables among each endophenotype ($p< 0.05$). Patient characteristics such as infection source, demographics, injury severity, and baseline physiological variables were analyzed. Outcomes such as mortality, multiple organ failure, and hospital length of stay were also evaluated. Significance testing was used to identify comorbidity differences between each endophenotype.

Further statistical testing was performed to test if any generated endophenotype was a time-shift of another (for example, is group A at hour 0 the same as group B at hour 6?). Pair-wise testing was performed between each time point of each endophenotype. The energy-distance test statistic was chosen due to its ability to compare two multivariate distributions and generate a distance between them [63]. This statistic was used to generate a multivariate distribution of clinical biomarkers (generated on raw values prior to log and CDF transformation) for each endophenotype at each time point. Using these distributions, the energy-distance algorithm tested for any statistical differences between each time point of each endotype. The R package *energy* was used for these computations. To account for dependency issues arising from multiple comparisons, permutation testing (10,000 replicates for each pair-wise test, significance at $p< 0.05$) was conducted.

2.2 RESULTS

Figure 2.1 shows the results of the clustering algorithm and the groups are visualized via heatmap. White (zero in the normalized scale) represented the population median for a biomarker. Red and blue colors indicate biomarker levels relative to the population; dark red indicated that a patient's biomarker was elevated with respect to the population me-

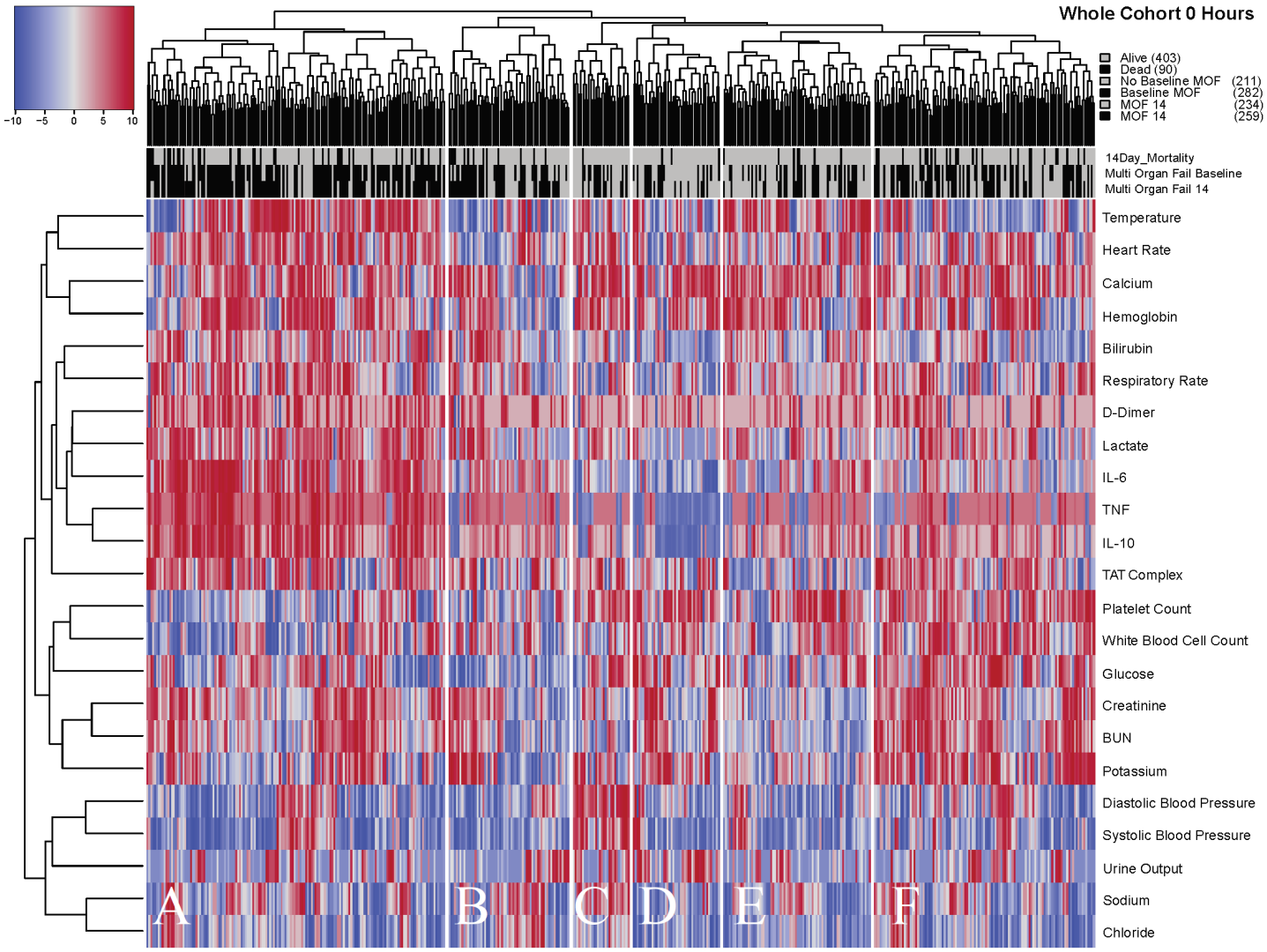


Figure 2.1: Heatmap of clustered baseline biomarkers in 493 patients with 6 identified subclusters. Biomarkers are individually normalized to be between -10 and 10, with 0 (white) representing the biomarker median. Dendrograms illustrate the clustering performed on biomarkers (left) and on patients (top). Subgroups were identified by analyzing the patient dendrogram. Group A is characterized by high cytokinemia, high lactate, and elevated mortality and multiple organ failure rates. Group B is characterized by medium cytokinemia and lower lactate. Group C is characterized by elevated blood pressure and low mortality rates. Group D is characterized by low cytokinemia and low mortality rates. Group E is characterized by a low IL-6 response and low values of creatinine and BUN, which is indicative of renal health. Group F is characterized by high IL-10 and TNF responses along with elevated creatinine and BUN, which is indicative of poor renal health. (BUN = blood urea nitrogen, IL-# = Interleukin-#, TNF = Tumor Necrosis Factor-α)

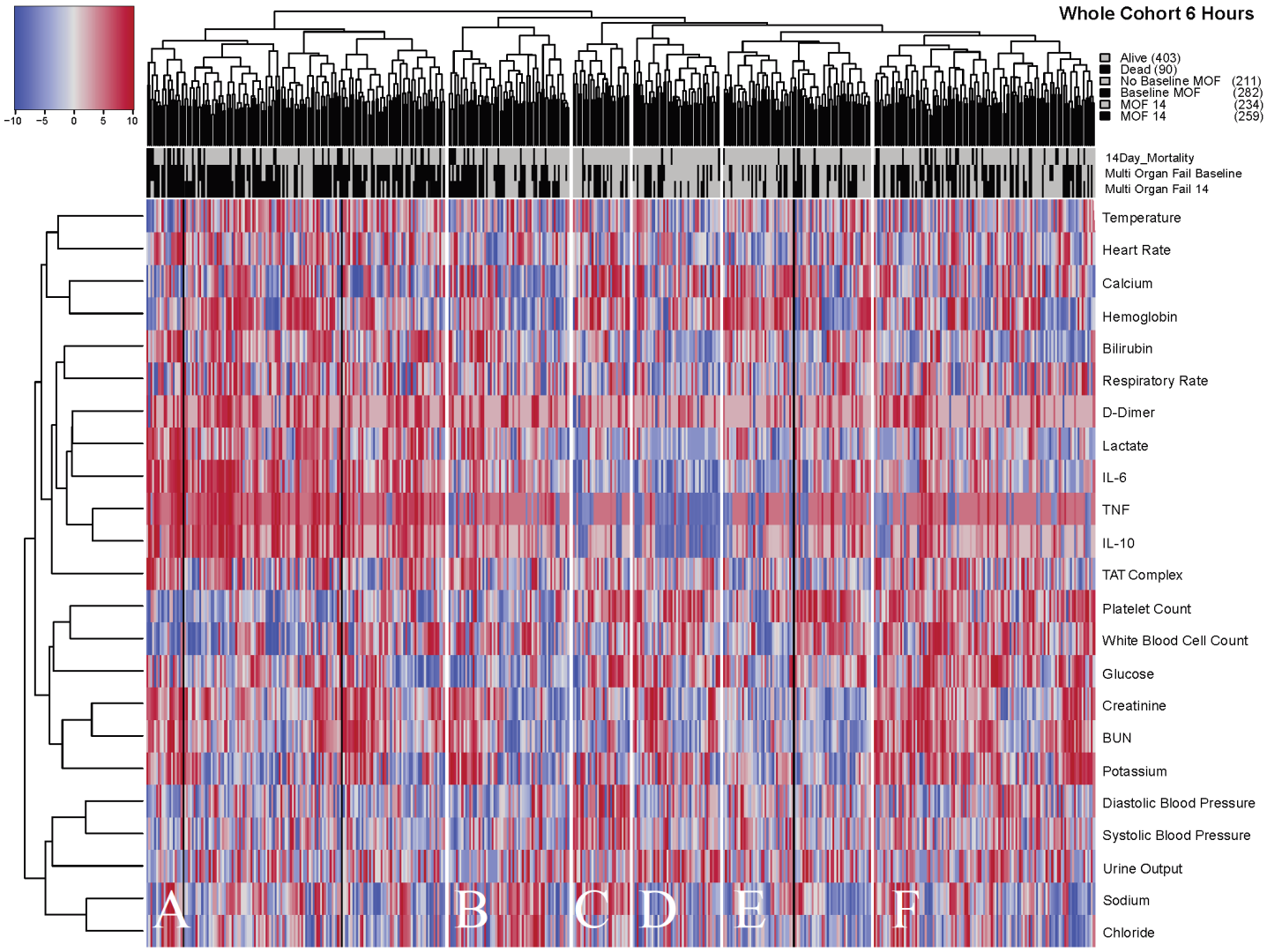


Figure 2.2: Heatmap of 6-hour biomarkers in 493 patients with 6 identified subclusters. Hierarchical clustering was not performed. Cluster ordering and Groups A-F were preserved from Figure 2.1. This heatmap demonstrates the progression of the 493 patients since baseline. Generally, there are improvements in all groups. Patients with elevated cytokinemia begin to taper down (return toward baseline). Blood pressure and urine output generally increased. Those with elevated temperatures at baseline tended to abate by 6 hours.

dian and dark blue indicated reduced levels. The analysis of the dendrogram yielded six endophenotypes, labeled A-F. Figures 2.2 and 2.3 demonstrate patient progression over the next 6 and 24 hours, respectively, with patient ordering preserved (using the ordering in-

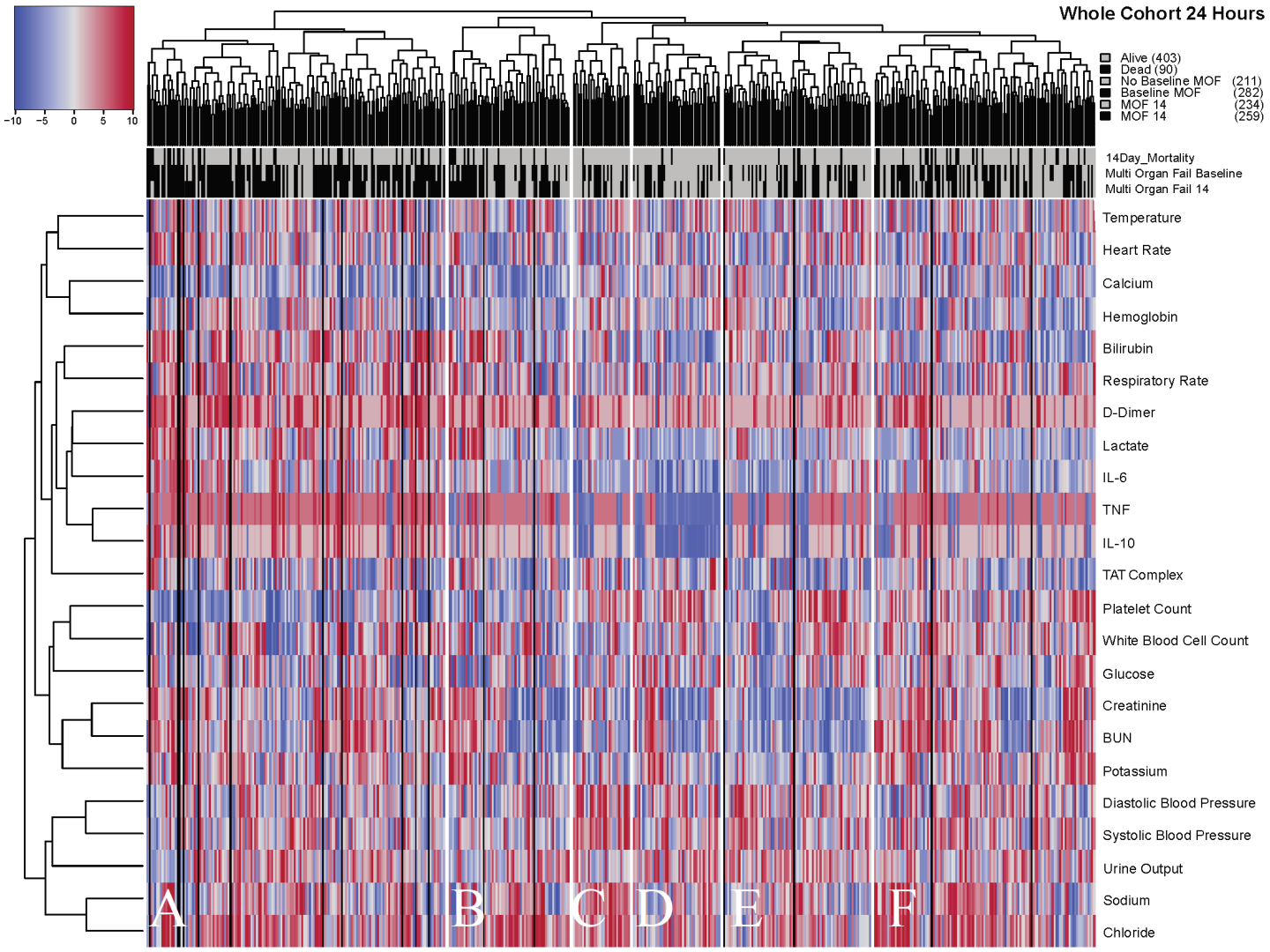


Figure 2.3: Heatmap of 24-hour biomarkers in 493 patients with 6 identified subclusters. Hierarchical clustering was not performed. Cluster ordering and Groups A-F were preserved from Figure 2.1. This heatmap demonstrates the progression of the 493 patients since baseline. There are more white biomarkers in this heatmap than in Figures 2.1 and 2.2, indicating a tendency of biomarkers to move away from their extremes. Abatement of cytokinemia is demonstrated in Group A, but the majority of patient IL-6 and IL-10 levels were still in the red, which portend worse prognoses. Group D continues to have a low cytokine response.

formed from baseline clustering). Table 2.2 shows the patient characteristics and outcome differences between each endophenotype, including significant differences in mortality, organ failure events, and the length of ICU stay.

Table 2.2: Comparison of endophenotypes identified via hierarchical clustering.

	Endophenotype						
Characteristic	A n=158	B n=64	C n=30	D n=46	E n=78	F n=117	
Age	61.2±17.2	58.6±16.4	57.8±18.7	61.5±15.6	58.7±16.1	59.8±18.2	
Male sex – no. (%)	100 (63.3%)	32 (50.0%)	15 (50.0%)	24 (52.2%)	48 (61.5%)	71 (60.7%)	
Race – no. (%)							
White	107 (67.7%)	47 (73.4%)	15 (50.0%)	31 (67.4%)	52 (66.7%)	88 (75.2%)	
Black	38 (24.1%)	9 (14.1%)	13 (43.3%)	10 (21.7%)	21 (26.9%)	18 (15.4%)	*
Other	13 (8.2%)	8 (12.5%)	2 (6.7%)	5 (10.9%)	5 (6.4%)	11 (9.4%)	
Source of sepsis – no. (%)							
Pneumonia	56 (35.4%)	23 (35.9%)	11 (36.7%)	14 (30.4%)	32 (41.0%)	35 (29.9%)	
Urinary Tract Infection	42 (26.6%)	17 (26.6%)	6 (20.0%)	11 (23.9%)	14 (17.9%)	34 (29.1%)	
Intraabdominal infection	15 (9.5%)	8 (12.5%)	1 (3.3%)	4 (8.7%)	12 (15.4%)	9 (7.7%)	
Infection of unknown source	17 (10.8%)	6 (9.4%)	5 (16.7%)	2 (4.3%)	5 (6.4%)	23 (19.7%)	*
Skin or soft-tissue infection	8 (5.1%)	3 (4.7%)	2 (6.7%)	5 (10.9%)	11 (14.1%)	4 (3.4%)	*
Catheter-related infection	5 (3.2%)	1 (1.6%)	0 (0.0%)	2 (4.3%)	0 (0.0%)	3 (2.6%)	
Central nervous system infection	1 (0.6%)	0 (0.0%)	0 (0.0%)	1 (2.2%)	0 (0.0%)	1 (0.9%)	
Endocarditis	1 (0.6%)	2 (3.1%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
Other	13 (8.2%)	3 (4.7%)	2 (6.7%)	7 (15.2%)	3 (3.8%)	5 (4.3%)	
Determined not to have infection	0 (0.0%)	1 (1.6%)	3 (10.0%)	0 (0.0%)	1 (1.3%)	3 (2.6%)	*
Positive blood culture – no. (%)	75 (47.5%)	20 (31.2%)	1 (3.3%)	8 (17.4%)	18 (23.1%)	25 (21.4%)	*
Injury Severity							
APACHE II	23.1±8.0	21.6±8.0	20.0±8.6	18.7±6.2	17.9±7.1	21.8±7.3	
APACHE III	63.4±23.5	67.8±22.0	57.4±21.8	57.4±19.9	56.1±23.8	62.5±23.1	
Charlson	2.9±2.8	3.0±2.7	2.7±2.7	2.4±2.2	2.5±2.3	2.8±2.9	
SOFA	8.9±3.7	7.9±3.9	5.5±3.3	5.5±2.8	6.0±2.6	6.8±3.2	*
CNS	0.8±1.2	0.8±1.2	1.3±1.4	0.5±1.1	0.7±1.2	0.9±1.3	
Cardiac	2.5±1.6	2.6±1.5	0.9±1.4	1.8±1.4	2.1±1.6	1.8±1.5	*
Coagulation	1.0±1.2	0.9±1.1	0.1±0.2	0.2±0.7	0.4±0.9	0.2±0.7	*
Liver	0.8±1.0	0.8±1.1	0.4±0.8	0.2±0.5	0.5±0.9	0.3±0.7	*
Renal	1.7±1.2	1.2±1.3	0.9±1.2	1.4±1.5	0.5±0.7	1.7±1.2	*
Respiratory	2.1±1.2	1.6±1.1	1.9±1.0	1.4±1.0	1.7±1.0	1.8±1.1	*
Physiological variables							
Systolic blood pressure – mmHG	95.6±25.2	90.8±13.0	134.7±28.3	97.1±28.3	97.2±22.1	99.4±21.3	*
Heart rate – beats/min	110.2±17.9	93.7±19.7	108.9±18.5	95.5±21.5	102.2±18.1	100.3±20.2	*
Temperature – C	37.8±1.4	36.4±1.4	37.4±1.0	36.9±1.1	37.8±1.2	36.7±1.2	*
Respiratory rate – breaths/min	24.7±7.9	21.4±5.9	21.2±7.9	18.4±5.3	21.9±4.3	22.7±7.4	*
Total bilirubin – mg/deciliter	1.8±2.0	2.5±3.3	0.8±0.3	1.2±1.2	1.0±0.7	1.0±1.1	*
Serum lactate – mmol/liter	4.0±3.2	3.8±4.6	2.5±2.2	1.5±1.2	2.3±1.9	2.9±2.6	*
Outcome							
14 Day All-cause Mortality – no.(%)	41 (25.9%)	13 (20.3%)	1 (3.3%)	3 (6.5%)	7 (9.0%)	25 (21.4%)	*
Mortality – days	35.9±67.9	62.4±76.9	78.0±73.1	116.4±103.8	84.8±110.9	56.1±82.8	*
Multiple Organ Failure – no. (%)							
Baseline	120 (75.9%)	36 (56.2%)	13 (43.3%)	15 (32.6%)	32 (41.0%)	66 (56.4%)	*
Within 14 Days (Any)	119 (75.3%)	40 (62.5%)	12 (40.0%)	16 (34.8%)	29 (37.2%)	60 (51.3%)	*
New Organ Failure in the first week – no. (%)							
Cardiovascular	118 (74.7%)	43 (67.2%)	9 (30.0%)	23 (50.0%)	42 (53.8%)	68 (58.1%)	*
Respiratory	77 (48.7%)	23 (35.9%)	12 (40.0%)	9 (19.6%)	14 (17.9%)	48 (41.0%)	*
Renal	14 (8.9%)	1 (1.6%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	5 (4.3%)	*
Stay in hospital – days	11.2±11.9	10.4±7.7	11.4±9.4	8.3±4.3	10.3±9.0	11.5±10.5	
ICU	5.3±5.4	4.5±3.2	3.8±5.0	2.9±2.8	3.3±3.2	4.9±5.1	*
Serious Adverse Events – no. (%)	19 (12.0%)	1 (1.6%)	0 (0.0%)	0 (0.0%)	2 (2.6%)	7 (6.0%)	

$*$ denotes $p \leq 0.05$ from Dunn' or Chi-Squared tests to determine differences between continuous or categorical variables, respectively.

2.2.1 High-Risk Endophenotypes

Endophenotypes A, B, and F were at high risk of cardiovascular, respiratory, and renal failure. 14-day mortality was $> 20\%$ and 60-day mortality was $> 30\%$. Together, these endophenotypes represent 339 of 493 patients analyzed. Endophenotype A was highly distinctive due to the presence of elevated cytokines (Interleukin-6, Tumor Necrosis Factor-α, and Interleukin-10) across all patients. D-Dimer, TAT-complex, lactate, and bilirubin were elevated as well, indicating coagulation, tissue perfusion, and hepatic issues. This group had the highest incidence of MOF (both baseline and within 14-days) and 14-day mortality. Table 2.2 shows that this group also had the highest incidence of positive blood cultures, 28 and 60-day mortality, cardiovascular, respiratory, and renal failure, and number of serious adverse events. This endophenotype also had the largest number of early deaths for any endophenotype (out of 158 patients, 2 died within 6 hours and 11 died within 24 hours). Patients with elevated creatinine and blood urea nitrogen (BUN), which is indicative of renal deficiency or failure, were at particularly high risk, with 4 deaths prior to 24 hours. Patients with elevated diastolic and systolic blood pressure at baseline had lower rates of MOF and mortality. Overall, patients in this group demonstrated a severe and sustained systemic inflammatory response to infection. At baseline, a typical patient had high cytokinemia and multiple biomarkers indicating the beginning of multiple organ system failure. This endophenotype was most likely descriptive of early-stage sepsis patients.

Endophenotype F patients demonstrated low baseline cytokinemia, high white blood cell counts, platelet counts, and elevated potassium levels. In addition, TAT complex decreased in most patients within 24 hours. Despite the apparently protective profile of these clinical biomarkers, this endophenotype had the second highest levels of MOF and mortality. Defining feature were high levels of creatinine, BUN, and lactate. Temperature, heart rate, hemoglobin, respiratory rate, white blood cell count, platelet count, and potassium tended to drop among patients. Two patients died within 24 hours. Table 2.2 showed that this group had the highest incidence of infection from an unknown source. Furthermore, this group was at high risk of cardiovascular and respiratory failure. Overall, biomarker progressions indicate that this group was most likely succumbing to infection during late-stage

sepsis. A possible mechanism could relate to the emergence of an immunosuppressed profile, as indicated by an elevated IL-10 and decreased IL-6 among most patients. Furthermore, most endophenotypes did not exhibit clear or consistent trends in white blood cell counts (WBC), while group F patients uniformly exhibited elevated levels at baseline that depleted over a 24-hour period.

A portion of Endophenotype B patients presented with baseline cytokinemia, but cytokine magnitudes and clinical outcomes were more favorable than that of A or F. Glucose and systolic blood pressure were uniformly decreased in this group. Half of the patients exhibited elevated lactate, creatinine, and BUN (Figure 2.1, group B, left, highest branch of the vertical dendrogram), while half did not (Figure 2.1, group B, right, highest branch of the vertical dendrogram). There were no significant biomarker changes over 24 hours except for decreasing IL-6. Table 2.2 showed that this group had the second highest incidences of positive blood culture and second highest 60-day all-cause mortality. Additionally, this group had the highest levels of baseline APACHE III severity of illness score and Charlson burden of chronic illness score [64, 65]. This suggests that this group of patients may have been enrolled later in the course of illness, possibly after peak cytokine levels had been reached [66]. This endophenotype launched a (relatively) appropriate inflammatory response to infection. However, the late complications of sepsis, rather than the initial insult, were responsible for the high mortality.

2.2.2 Low-Risk Endophenotypes

Endophenotypes C, D, and E demonstrated lower rates of MOF and mortality. These groups were characterized by low to no cytokinemia and low lactate. Furthermore, the 154 patients from these groups presented with low injury severity scores and low SOFA scores.

Endophenotype C contained patients exhibiting elevated baseline heart rates, hemoglobin, platelet count, lactate, and diastolic and systolic blood pressures. This group was not afflicted with cytokinemia: despite varying levels of TNF and IL-10, IL-6 was uniformly low among this cohort. Except for blood pressure, biomarkers had decreasing trends within 24 hours. While only 1 of 30 patients died within 14-days, Table 1 revealed elevated rates of res-

piratory failure and 60-day mortality. This suggested that deaths within this endophenotype were caused by later complications rather than the initial sepsis insult.

Endophenotype D contained patients with low or nonexistent baseline cytokinemia, low lactate values, and elevated platelet counts. This group had the lowest risk of MOF and 14-day mortality. Table 1 showed that this group had the lowest 60-day mortality as well, indicating fewer late-stage complications.

Endophenotype E contained patients with elevated temperatures, and low creatinine and BUN values. Few patients had baseline cytokinemia. This group had slightly higher rates of MOF and mortality than endophenotypes C and D. One patient died within 6 hours and another died within 24 hours. Over time, IL-6 decreased faster than IL-10. This group appeared to have a normal inflammatory response and baseline biomarkers were not indicative of any organ deficiency.

2.2.3 Similarity between Endophenotype

Figure 2.4 illustrates the results of the multivariate distance matrix to compare the similarity between each endophenotype and their time points. Many of the 0-vs-6 hour comparisons and the 0-vs-24 hour comparisons yielded significant differences, which further support the notion that the identified endophenotypes were not mere time-shifts of one another. Endophenotype A at 0 and 6 hours were significantly different from most other endophenotypes and time points. Furthermore, none of the pair-wise comparisons yielded low dissimilarity statistics between any test (no dark colors). The darkest region occurs within the 24-vs-24 hour comparisons, indicating that there were some similarities among endophenotypes at 24 hours. This demonstrated that patients that segregate into different baseline endophenotypes may progress to similar clinical states after a period, but their baseline clinical values and time-series progression may still play a larger role in determining outcomes.

Interestingly, endophenotypes A and F had similar 14-day mortalities, demonstrating that distinct baseline profiles and sepsis trajectories may have similar risks of death. The largest notable difference between these endophenotypes was the level of cytokinemia present at baseline. Furthermore, white blood cell counts within endophenotype F were uniformly

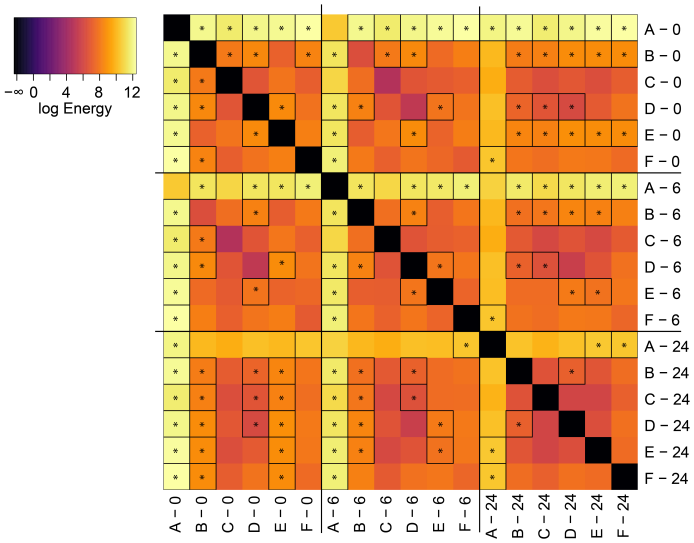


Figure 2.4: Multivariate energy-distances between endophenotypes A-F and their time points. The distances are displayed in a heatmap where darker colors represent similar groups and lighter colors represent differing groups. Most group-vs-group comparisons do not reveal any similarities. Some similarities are present between different groups at time 24 hours. Cells with an asterisk indicate that the null hypothesis (of equal multivariate distributions, $p < 0.05$) is rejected.

elevated at baseline and decreased over time, while patients within endophenotype A split (half were elevated and half decreased at baseline) and remained unchanged over time.

2.3 DISCUSSION

Despite improving understanding of sepsis pathology, one of the critical issues afflicting sepsis therapies and research is high inter-patient variability and the need for better patient stratification. The goal of this study was to explore clinical values to identify the types of subgroups that may exist in clinical sepsis populations. Outside of differing clinical outcomes, there were clear differences between each endotype presented in Figure 2.1. Most notable were the differences in cytokinemia and lactate. However, endotype F demonstrated that even

low baseline cytokinemia may lead to high mortality and MOF and that biomarker progression over time is worthy of additional study. This exploratory work supports the notion of endotype-specific therapies. Heat map clustering of patient cytokines and other laboratory and clinical features is a promising method to group patients into distinct endotypes. Availability of additional domains of data would enhance the richness of their description, offer further segregation of endotypes, and potentially offer insight as to underlying mechanisms. Classification of these endotypes can act as a basis for more precise therapies even at the rather coarse level of the description presented. The high-risk endophenotypes A, B, and F are distinctive enough to warrant specific treatment or therapy options. Immunosuppression drugs would likely be harmful to the slightly immunosuppressed patients in endophenotype F, but could be beneficial for the high cytokinemia patients in endophenotype A. Knowledge of sepsis endotypes has the clear potential to shift the current paradigm of research and trial design towards better outcomes. For example, clinical trials can benefit from better patient characterization at enrollment, to better match intervention and presumed pathophysiology, or later, to interpret subgroup benefit of such interventions.

The identified groups in this work were dubbed endophenotypes due to their discovery using observable clinical biomarkers (phenotypes). However, these biomarkers were hypothesized to be clinical manifestations of distinct underlying pathophysiological differences (endotypes). Along the spectrum between endotypes and phenotypes, the identified groups fall somewhere in the middle. The visualization of clinical data as heatmaps revealed areas of similarity and dissimilarity between groups, which indicated that clinical biomarkers alone cannot fully explain the endotypes present in sepsis. The ability to discriminate the underlying endotype from the clinical phenotype with good accuracy is a strength of the current approach. Clinical decisions are based on what clinicians can readily observe and measure, and many of the markers characterizing endotypes, even at the coarse level we present, are not readily available in a timely fashion. This issue will presumably be amplified as endotype enriching work proceeds. Thus, there is a trade-off between extensive knowledge of endotypes with their implied underlying biology, and the pragmatic decisions clinicians are facing when treating sepsis. Thus, endotype characterization allowing precise therapies to

be applied in a timely fashion is key to the translational relevance of detailed endotyping work.

In previous reports on sepsis clustering, Kellum, et al., studied cytokine (IL-6 and IL-10) longitudinal profiles in the GenIMS cohort of patients admitted with community-acquired pneumonia. These authors identified cytokine-based clusters with different outcomes. Specifically, high cytokinemia trajectories in IL-6 and IL-10 corresponded to a higher overall burden of inflammation and mortality, which corroborates our findings with endophenotype A. Fjell, et al., analyzes signaling molecules and cytokines using a similar hierarchical clustering method, but their biomarkers were pre-selected and did not include any clinical biomarkers [49]. Our approach using a non-biased feature selection (based on standard deviations and subject to data availability) demonstrated biomarker progressions over time within each endophenotype. Knox, et al., used similar clinical variables, but underwent a dimension reduction technique, rendering it difficult to interpret the resulting clusters in a clinically meaningful way [50]. The resulting clusters had to be recolored to present one variable at a time, while the approach we took presented all clinical information in a multivariate way.

It has been demonstrated that the baseline time point in human sepsis data may not be meaningful due to the variability associated with each patient's pre-hospital time, but this time could be recovered using clinical measurements [41]. Thus, the biomarker similarity within each endophenotype suggests that the pre-hospital time (or true sepsis time zero) for grouped subjects appears to be close, and patients are at similar points along the endotype-specific sepsis trajectory. A dissimilarity test indicated that each heatmap was not similar enough to be grouped together at any time point or cluster. The implication is that classifying a patient into an endotype may be sufficient to estimate his or her pre-hospital time (true sepsis time zero) and may be a one-step procedure for tailoring a suitable therapy.

Comorbidities played an influential role in each endophenotype. High-risk endophenotypes A and F had the highest rates of renal failure as a comorbidity. Diabetes was most prevalent in endophenotype F and lowest in B and E. Endophenotype B had the largest population of AIDS patients (10%), which partially explains why B also had the largest population of immune-compromised patients. Endophenotypes E had the highest incidence of unknown infection source, which may suggest that these patients may not have received

appropriate or timely antibiotic coverage. The lowest 14-day mortality groups, endophenotype C and D, were predominately characterized by the hypertension comorbidity (70%+ within either cohort). This suggested that hypertension may be somewhat protective during the initial onset of sepsis.

2.3.1 Study Limitations

This work can be improved via validation using a larger dataset. Despite the small cohort sizes used in this study, the ProCESS database is still one of the largest of its kind. A validation database would enable the exploration of other clinically interesting variables because many had to be eliminated for this work due to low sample sizes or rates within the ProCESS population. Furthermore, there is an element of subjectivity in the current approach. Determining the height to "cut" the dendrogram tree in Figure 1 affected the resulting groups, ranging from 2 groups at the top level to 493 groups at the bottom-most level. While endophenotype A was clearly distinct from the rest (largest height on dendrogram tree), the division of the remaining groups was based on cutting the dendrogram to yield visually distinctive regions. As a result, segregating the data required a degree of human intervention. Despite the subjectivity, Table 2.2 revealed many differences between the endophenotypes discovered.

2.3.2 Application into Temporal qSOFA Trajectories

The hierarchical clustering approach outlined in Section 2.1 was extended to explore the predictive capability of sequential measurements of the quick Sequential Organ Failure Assessment (qSOFA) score with respect to mortality. The qSOFA score was recommended by the recent sepsis redefinition (Sepsis-3) as a fast method of assessing a patient with suspected infection without the need of laboratory tests [7]. The qSOFA score consisted of measuring respiratory rate, systolic blood pressure, and the Glasgow Coma Score (GCS). Each clinical feature may contribute 0 or 1 to the overall qSOFA score for a total of 3 points (worst score). The manuscript for this work is pending co-author review (Kievlan, Zhang, et al., Evalu-

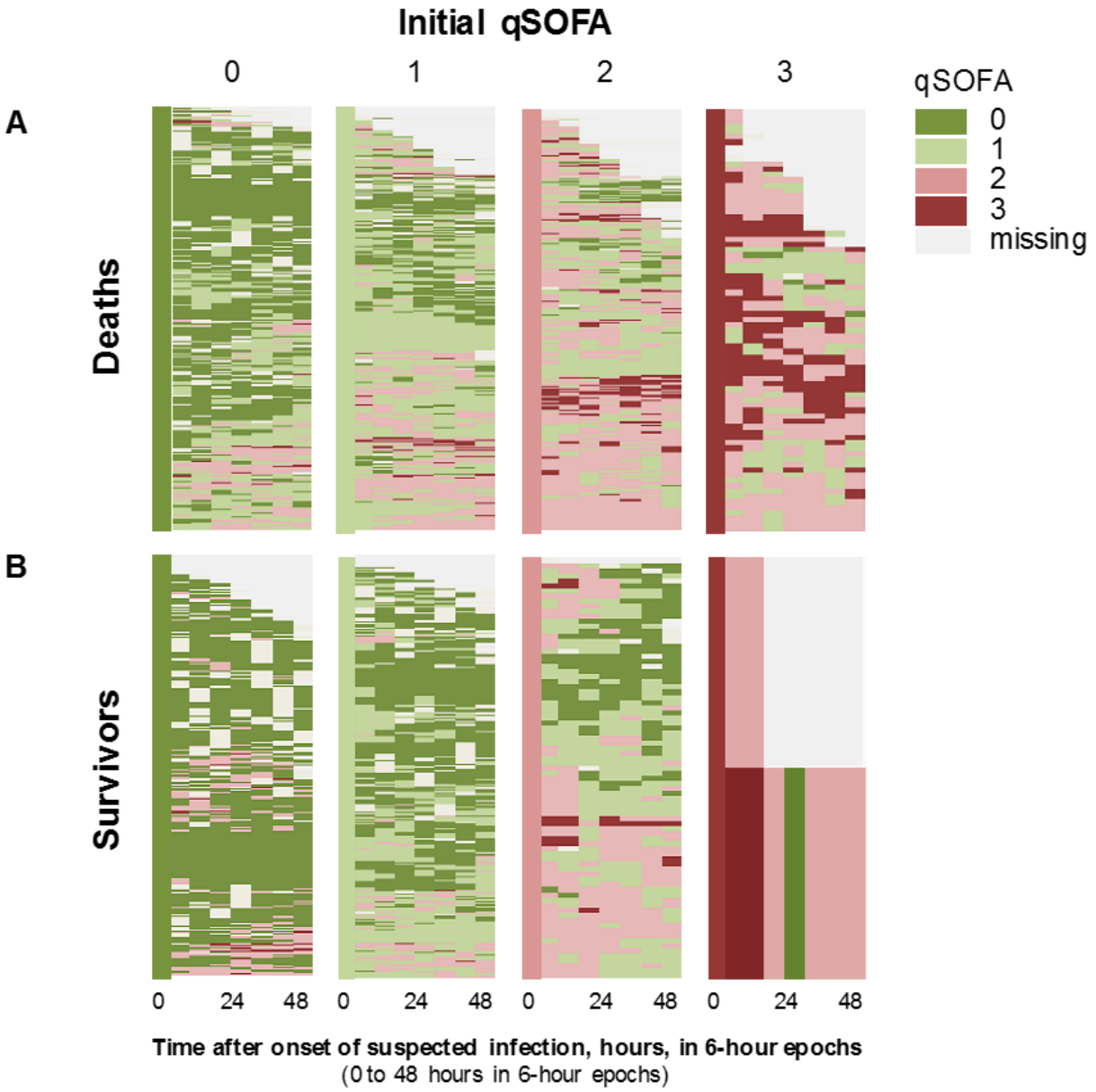


Figure 2.5: Clustered Trajectories of qSOFA scores in patients with suspected infection. Shades of red indicate the severity of qSOFA scores and gray indicates missing measurements. Patients with low baseline qSOFA scores (qSOFA=0,1, survivors) were more likely to remain low over 48 hours than nonsurvivors. Two patients with a baseline qSOFA score of 3 survived.

ation of repeated qSOFA measurements among patients with suspected infection; awaiting submission to American Journal of Respiratory and Critical Care Medicine).

A retrospective cohort study among adult hospital encounters (age /*geq* 18 years) with suspected infection in 2012 at twelve community and academic hospitals in southwestern Pennsylvania was conducted. From this data, qSOFA scores were calculated for each patient. Within the first 48 hours, data was broken into six-hour intervals and the maximum qSOFA score was calculated from each interval. The heatmap hierarchical clustering outlined in this section was used to group these trajectories together to evaluate the trends in qSOFA trajectories among differing baseline values and survivors/nonsurvivors.

The cohort consisted of 37,591 subjects with baseline qSOFA measurements. Of these subjects, 4.7% (1,769) died in the hospital. Figure 2.5 illustrates that survivors tended to have improving qSOFA trajectories. Nonsurvivors' qSOFA scores remained higher for longer. The majority of the baseline qSOFA=3 subjects belonged to the nonsurvivor group and generally did not see any decrease in their qSOFA scores over time. Among the low baseline qSOFA subjects, those that survived tended to remain low while nonsurvivors' qSOFA trajectories tended to increase over 48 hours. Furthermore, the survivors with baseline qSOFA=2 tended to improve over time while the nonsurvivors tended to deteriorate. This suggests that in clinics where regular SOFA scores are difficult to obtain, qSOFA monitoring over time provides a fast risk/mortality assessment method.

2.3.3 Summary

In summary, these findings suggest that biomarkers offer discriminatory power upon clinical presentation to segregate subjects into clusters of distinct clinical trajectories. Each cluster, dubbed an endophenotype due to their basis in observable clinical features, has distinct rates of mortality and organ failure. These clusters suggest the existence of separate endotypes in sepsis, which is currently not well understood. The results of this work motivate the following chapters under the hypothesis that research into these endophenotypes will reveal richer features and improved mechanistic understanding behind these clusters. The end goal is to classify incoming septic patients from readily available clinical and serum biomarkers into an endotype which suggests potential dysregulation mechanisms and may lead to targeted therapies.

3.0 MIXTURE MODELING APPROACH TO IDENTIFY SEPSIS ENDOTYPES

Chapter 2 identified clusters of subjects with similar clinical features and demonstrated the existence of distinct sepsis subtypes. This motivated further research into sepsis endotypes by refining these clusters and characterizing the biological mechanisms behind these distinctions. This Chapter builds upon this work by addressing two major clinical concerns associated with sepsis in critical care. First, the identified clusters in Chapter 2 were deemed endophenotypes because no mechanistic considerations were involved during the clustering process. There is currently a lack of mechanistic understanding behind sepsis and much is unknown about the dysregulated pathways that lead to sepsis [67, 68]. Mathematical models of sepsis have previously demonstrated success in characterizing the dynamics and mechanisms of inflammatory pathways in sepsis, but have not been validated against human data [3, 4, 69, 70]. This deficiency is caused by the second concern: the clinical practice of defining subject baselines at the time of clinical presentation renders characterization of sepsis difficult. From a dynamics point-of-view, analyzing data without initial conditions introduces significant ambiguity to mathematical findings. The clinical status quo is the result of a lack of knowledge of a subject's pre-hospital time and trajectory. When queried about the duration of their acute inflammatory response, subjects in the ProCESS cohort estimated time-spans of up to one year [59]. The lack of consideration of pre-hospital durations weakens the findings and translatability of current sepsis research, including the results from Chapter 2.

This section improves upon the previous method of hierarchical clustering by addressing clinical concerns via the development of a novel mathematical modeling-based approach to the identification of sepsis endotypes. Considerations for inflammatory dynamics and subject

pre-hospital times were taken into account using a mathematical model that initiated at the beginning of the inflammatory response. A mixture-modeling-based approach is introduced where each subject was parameterized in accordance with the mathematical model and then clustered based on their parameters. The resulting clusters were deemed endotypes rather than endophenotypes because of the mechanistic focus on the inflammatory response rather than the phenotype-based method in Chapter 2.

As readouts of the inflammatory response, cytokines are desirable candidates for modeling and characterizing inflammatory pathway endotypes. Although cytokines represent a minute portion of the extensive inflammatory response to a pathogen, they are highly potent mediators that are vital for initiating, sustaining, and suppressing inflammation [37, 71]. The very small concentrations of cytokines that are produced upon encountering pathogens elicit large systemic changes [10]. Interleukin-6 and Interleukin-10 (IL-6 and IL-10, respectively) are well-known cytokines with distinctive trajectories during sepsis [39]. Thus, it is reasonable to posit that endotypic differences might be reflected in cytokine expression patterns, even if such patterns do not represent a comprehensive endotypic characterization. The current clinical consensus is that both the pro-inflammatory response and the anti-inflammatory response are important in clearing an infection, allowing tissue recovery, and determining sepsis outcomes [5]. A cytokine was selected as the proxy of the pro-/anti- inflammatory response.

Interleukin-6 (IL-6) is a pro-inflammatory cytokine that is characterized by a fast rise during the inflammatory response. IL-6 is initially secreted by tissue-resident macrophages and is responsible for mobilizing and recruiting neutrophils and monocytes to the site of infection. Recruited monocytes may differentiate into macrophages of the pro-inflammatory phenotype, M1, which further produce IL-6 [14]. This positive feedback is responsible for the fast rise in IL-6. Experiments involving endotoxin challenges in healthy volunteers reveal that many cytokines follow a fast rise and fall dynamic. In the case of IL-6, it peaks at 3 hours and returns to baseline by 8 hours after endotoxin challenge [39]. However, research has shown that cytokine responses differ between endotoxin challenges and sepsis, which is likely attributed to the fast clearance of endotoxin versus the sustained presence of infection in sepsis [37]. In sepsis, IL-6 peaks later and has a long protracted decline [37, 39]. As such,

IL-6 serves as a good proxy of the pro-inflammatory response and provides information on infection severity/injury and the type of host response [68].

Similarly, IL-10 is the *de facto* anti-inflammatory cytokine. IL-10 has a fast rise, peaks slightly after IL-6, and has a long protracted decline during sepsis [15, 39]. It is secreted by monocytes and M2 macrophages and prevents the synthesis of pro-inflammatory mediators [13, 39]. The timing and trajectory of IL-10 are extremely important and, if inappropriate, can lead to uncontrolled infection or dysregulated inflammation [15].

Knowledge of sepsis endotypes, as defined by pro- and anti-inflammatory responses, can reveal the underlying pathobiological differences between different host responses to sepsis and capture some of the clinical variability observed during sepsis. Consideration of endotypes in sepsis, which dispels the notion of a single therapeutic approach, can lead to better targeted and more effective treatment [45]. It was assumed that a finite number of characteristic sepsis responses exist and that each case manifests into distinct IL-6 and IL-10 trajectories, given their importance as master regulators of inflammation.

Identification of sepsis endotypes is an exercise in unsupervised clustering and proved an interesting engineering challenge. First, there is a lack of clinical understanding regarding the mechanisms that lead to sepsis because current clinical work is heavily focused on identifying phenotypes that predict outcomes. Mathematical models of sepsis (such as mechanistic compartmental ordinary differential equation models) can supplement or improve clinical understanding, but few exist in the literature that has been validated against human sepsis data. Second, human sepsis data is inherently left-censored (because data collection can begin only after clinical presentation) and this causes problems in generating unique clusters. Figure 3.1 illustrates the problems posed by left-censorship. Suppose an arbitrary cytokine is measured from two separate subjects at 0, 6, and 24 hours after they arrive in the hospital (or are enrolled in the study, as shown in Figure 3.1A) and that these subjects belong to the same sepsis endotype (demonstrated in Figure 3.1B). By disregarding the dynamics that describe how these trajectories manifest, they would be binned into two clusters (Figure 3.1A) by popular algorithms such as K-means or 'proj-traj' [72].

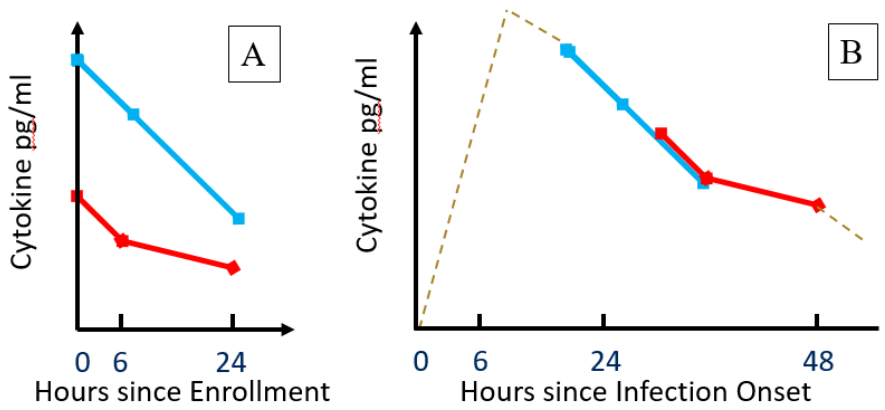


Figure 3.1: Schematic of clustering challenge associated with endotype identification. A: Diagram illustrating a sample result from traditional clustering methods based on time since enrollment. Magnitude heavily affects clusters. Blue and red time courses are binned into separate clusters. B: Diagram illustrating that the previous blue and red time sources may come from similar trajectories. Magnitude differences are attributed to different sampling time points along a single trajectory.

Due to the sparsity of temporal cytokine data from sepsis subjects and the variability in patient pre-hospital time prior to measurements, traditional clustering techniques were ineffective and necessitated a dynamics-based approach. This requires a mixture modeling approach that can simultaneously (i) recover dynamics from censored trajectories of IL-6 and IL-10 and (ii) conduct unsupervised clustering on recovered dynamics. The resulting method utilized a combination of systems analysis, hierarchical clustering, and mixture models.

3.1 METHODS

A retrospective analysis of the Protocol-Based Care for Early Septic Shock (ProCESS) trial was conducted [59]. A description of this trial is available in Chapter 2. The ProCESS trial collected rich longitudinal and high-fidelity clinical data from its subjects. A convenience cohort of 390 subjects was selected from the database of 1341 subjects in accordance to the selection criteria of IL-6 and IL-10 measurement availability at 0, 6, and 24 hours, where the zero baseline was the time of trial enrollment. This was different from the 493 subject cohort in Chapter 2, where the selection criteria were the availability of certain

Table 3.1: Comparison of convenience cohort for mixture model against the rest of ProCESS database

Characteristic	**Analyzed Cohort** n=390	**Rest of ProCESS** n=951	
Age	61.0±16.9	60.9±17.9	
Male sex – no. (%)	236 (60.5%)	512 (53.8%)	*
Race – no. (%)			
White	268 (68.7%)	648 (68.1%)	
Black	88 (22.6%)	245 (25.8%)	
Other	34 (8.7%)	58 (6.1%)	
Source of sepsis – no. (%)			
Pneumonia	133 (34.1%)	310 (32.6%)	
Urinary Tract Infection	93 (23.8%)	191 (20.1%)	
Intraabdominal infection	37 (9.5%)	140 (14.7%)	*
Infection of unknown source	46 (11.8%)	124 (13.0%)	
Skin or soft-tissue infection	25 (6.4%)	71 (7.5%)	
Catheter-related infection	15 (3.8%)	23 (2.4%)	
Central nervous system infection	3 (0.8%)	7 (0.7%)	
Endocarditis	3 (0.8%)	4 (0.4%)	
Other	29 (7.4%)	56 (5.9%)	
Determined not to have infection	6 (1.5%)	25 (2.6%)	
Positive blood culture – no. (%)	138 (35.4%)	258 (27.1%)	*
Injury Severity			
APACHE II	20.8±7.4	20.7±7.7	
APACHE III	61.4±22.8	61.1±22.5	
Charleson	2.5±2.4	2.7±2.7	
SOFA	7.3±3.6	7.1±3.6	
Physiological variables			
Systolic blood pressure – mmHG	97.9±24.1	100.7±24.4	*
Heart rate – beats/min	104.1±20.5	102.8±20.4	
Temperature – C	37.3±1.4	37.3±1.3	
Respiratory rate – breaths/min	22.7±7.2	22.8±6.6	
Total bilirubin – mg/deciliter	1.3±1.6	1.5±2.3	
Serum lactate – mmol/liter	4.5±2.9	5.0±3.4	*
Outcome			
14 Day All-cause Death – no. (%)	39 (10.0%)	200 (21.0%)	*
Mortality – days	72.0±88.2	45.6±71.6	*
Multiple Organ Failure – no (%)			
14 Days	250 (64.1%)	631 (66.4%)	
Stay in hospital – days	11.8±9.5	11.5±11.6	*
ICU	5.2±5.2	4.9±6.9	*
Serious Adverse Events – no. (%)	15 (3.8%)	67 (7.0%)	*

$*$ denotes $p \leq 0.05$ from Dunn' or Chi-Squared tests to determine differences between continuous or categorical variables, respectively.

biomarkers at 0 hours. Treatment arm was not included in the selection criteria because the ProCESS trial concluded that clinical outcomes among all three arms were effectively similar [59]. A comparison of the cohort against the rest of the ProCESS database is available in Table 3.1, which demonstrates that the convenience cohort was not entirely representative of the ProCESS cohort. Specifically, the convenience cohort contained a higher population of males, lower serum lactate, higher 14-day all-cause mortality outcomes, and fewer numbers of serious adverse events than the rest of the ProCESS cohort. This was reasonable because the selection criteria identified a group of subjects that (i) survived up to 24 hours and (ii) was clinically stable enough to have blood drawn for cytokine assays. In contrast, the convenience cohort contained a higher percentage of patients with blood infections and lower systolic blood pressures, which are correlated with negative outcomes. This suggested that the convenience cohort also responded better to sepsis despite having worse infections.

Of the 390 subjects, 165 had an additional IL-6 and IL-10 measurement at 72 hours post-trial enrollment. 72-hour cytokine measurements for the remaining 225 subjects were imputed via a nearest neighbor technique based on their measured vital signs and blood-based biomarkers (refer to Chapter 5) [41]. Specifically, the first 6 hours of measurements of respiratory rate, heart rate, temperature, urine output, systolic blood pressure, tumor necrosis factor α (TNF), IL-6, and IL-10 were used. These biomarkers were matched against those of the 165 subjects with 72-hour cytokine measurements. The one nearest neighbor within the 165 cohort provided the imputation results. Finally, any measurements less than 10 pg/ml were treated as 10 pg/ml. Due to the different assaying techniques used by the multiple centers that participated in the ProCESS trial, there were many different lower limits of cytokine detection (LLD) present in the data, with 10.0 being the highest LLD.

3.1.1 Systems Analysis and Mixture Modeling

The pro- and anti-inflammatory responses were modeled after a mass-spring-damper response to a bump (see Figure 3.2). This physical system was simplified using transfer functions and their step response. Transfer functions provide a mathematical formalism where model parameters can be directly linked to dynamics characteristics of the system response (such

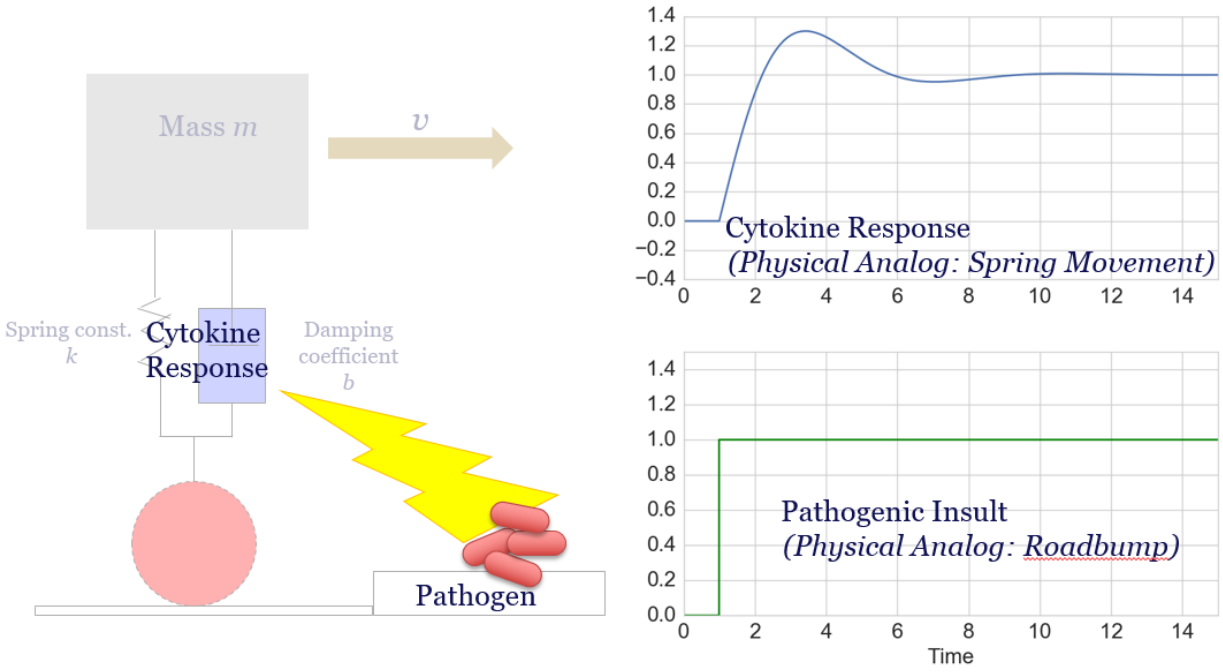


Figure 3.2: Schematic of a mass-spring-damper system serving as a physical analog to the inflammatory response. The cytokine response was likened to a mass-spring-damper physical system. Cytokine trajectories were modeled as the unit step response of a second-order transfer function to a pathogen presence (analogous to a car suspension responding to a "road bump"). The steady state gain of the response was set to unity.

as oscillation, overshoot, etc.), thereby facilitating model identification from data. Furthermore, transfer functions represent entire classes of mathematical systems exhibiting a desired behavior. The inflammatory response (output) may be characterized as the response to a rectangular wave of infection for a specified duration (input). However, the short temporal length of the cytokine data from the ProCESS trial (up to 72 hours after trial enrollment), in combination with an assumption that subjects still had ongoing infection by 72 hours, led to the simplification of this input to a step change (as shown in Figure 3.2, bottom right). A similar cytokine modeling approach (authors used a state-space realization of second-order transfer functions) was taken in a journal article by Yiu, et al [73].

Analysis of the cohort cytokine data revealed three types of step response dynamics, which are represented by the second-order transfer functions in Figure 3.3). The first transfer

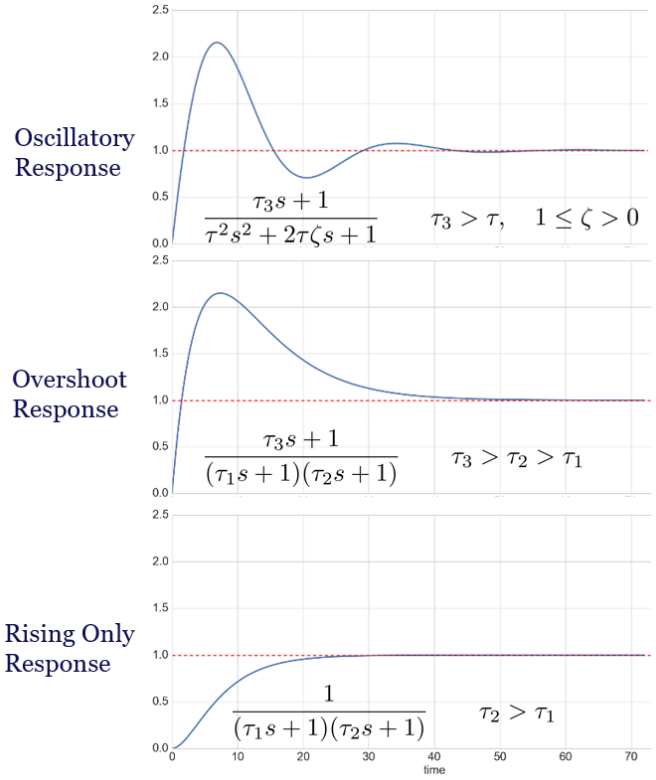


Figure 3.3: Sample unit step responses of the three transfer functions used. Each response reaches a steady state output value of 1.0 but the trajectory to reaching steady state is characteristically different for each model. The first category, oscillatory response, is a three-parameter system (τ, ζ, τ_3) characterized by a damped oscillating approach to 1.0. The second category, overshoot response, is a three-parameter system (τ_1, τ_2, τ_3) characterized by a fast rise followed by a decline to 1.0. The third category, rising only response, is a two-parameter system (τ_1, τ_2) characterized by a rising trajectory to 1.0 with no overshoot.

function enforced an oscillatory response to capture the primary rise and fall motif and a secondary lower-magnitude peak. Such peaks have previously been observed in the clinic [74]. The damping parameter of this transfer function, ζ, was of particular interest. ζ controls the oscillatory nature of the system and bounds were set to prevent it from taking low values near zero, thereby preventing the system from oscillating for many periods at a relatively high magnitude. The second transfer function captured the primary rise and fall motif only. Finally, the third transfer function enforced a rising only behavior to capture subjects whose inflammatory responses keep rising.

The data from the convenience cohort was normalized to account for extensive magnitude differences between subjects.The normalization criterion consisted of dividing each patient's IL-6 and IL-10 trajectory by their respective 72-hour measurements. The normalization procedure was assumed to normalize both the magnitude of inflammatory response and the magnitude of infection for each patient. A benefit of this normalization was that it removed the need to estimate a gain for the transfer function model. The dynamics of the resulting normalized pro- and anti-inflammatory responses were clustered.

The first step was fitting the 390 subjects' IL-6 and IL-10 responses (after taking the log10 of cytokine concentrations) to each type of transfer function response. The 0, 6, and 24 cytokine measurements were normalized by dividing by their respective IL-6 or IL-10 72-hour measurement. The normalized IL-6 and IL-10 trajectories were fit against each of the three transfer function categories in Figure 3.3. Parameter fitting was performed using the Levenburg-Marquardt algorithm provided by the Python *lmfit* package.

This approach fully addressed the problem previously illustrated by Figure 3.1. Step responses for second order transfer functions always begin at 0. This enabled the objective function to also fit a discrete hourly time shift parameter (bounded to [0,78] hours) to estimate the pre-hospital time for each patient. Each patient was then classified into one of nine categories (three possibilities each for IL-6 and IL-10, yielding 9 combinations) based on the lowest sum of squared errors.

The next step further split each of the 9 categories by identifying similar clusters of subjects in parameter space. Splitting within each category allowed for the possibility of different behaviors of the oscillatory, overshoot, and rising-only responses. Parameters for each subject were assumed to be distributed about a Gaussian mixture. Gaussian mixtures were identified via the expectation maximization algorithm and subjects were segregated in accordance to the highest probability of membership to a certain Gaussian. Expectation maximization is a well known statistical algorithm that (i) calculates a probability of membership for each subject (expectation-step), (ii) calculates all Gaussian component parameters (μ_i, σ_i) via the maximum likelihood approach in accordance to each component's membership (maximization-step), and then (iii) repeat steps (i)-(ii) iteratively. The learning step and mixture modeling were performed via the *sklearn* package in Python. The number

of components in the Gaussian mixture model for each of the 9 categories was a user input and was varied over multiple simulations until a minimum Bayesian Information Criteria (BIC) was obtained. Low membership clusters ($n < 10$) were removed and those subjects were reassigned to the next most similar cluster.

The third step generated a master inflammatory response trajectory for each remaining cluster. The parameters from each cluster were averaged and used to generate a representative IL-6 and IL-10 trajectory for the cluster. Subjects within each cluster were then shifted in time (original measurement intervals were preserved) to best align their IL-6 and IL-10 trajectories to the respective master response trajectory.

Clinical outcomes were defined by 14-day all-cause mortality and 14-day multiple organ failure rates. This was done for the same reasons outlined in the previous chapter. No distinction was made between baseline MOF and 14-day MOF because this method identifies patient time zeros, which overrides the clinical definition of baseline (time of clinical presentation).

Statistical analysis of the resulting clusters was conducted on patient demographics, outcomes, and clinical biomarkers taken within 6 hours of trial enrollment. Statistically significant differences ($p < 0.05$) were identified using the non-parametric Dunn's test for continuous variables and the Chi-squared test for categorical variables. Multiple pair-wise testing, to specifically identify which clusters were different from each other, was performed via Dunn's test and Chi-squared tests. Bonferroni corrections were applied to counteract the error effects of multiple comparisons.

3.2 RESULTS

Figure 3.4 shows the resulting five distinct inflammatory responses that were identified across the 390 subjects. Cluster 1 illustrates a fast and high magnitude pro- and anti-inflammatory response to infection. This cluster had the highest 14-day mortality and the highest incidence of 14-day multiple organ failure (MOF). This high-risk cluster also described the most subjects, with nearly 30% membership. Cluster 2 illustrates a fast and medium magnitude pro-

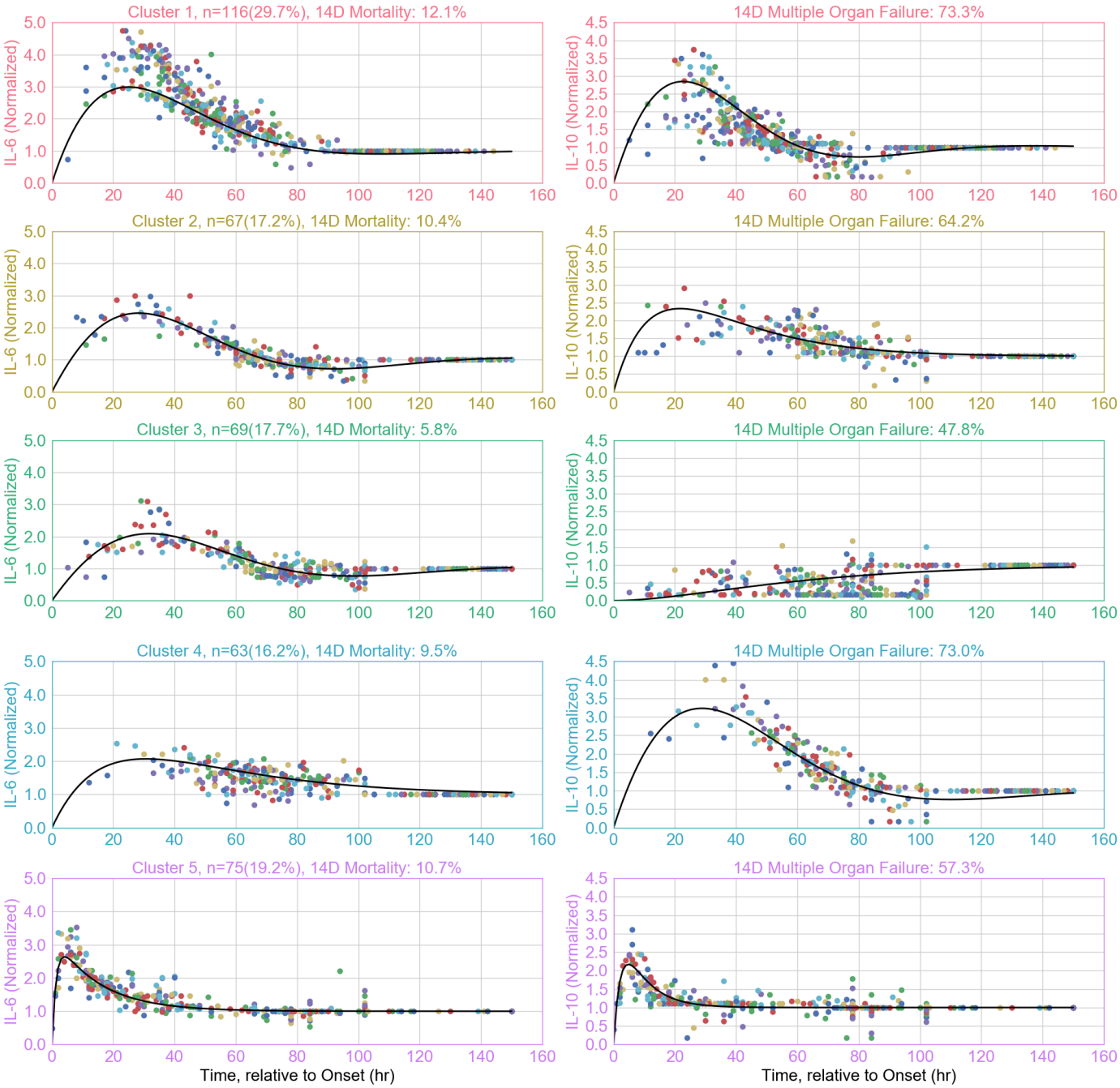


Figure 3.4: Results from cluster identification process (5 sepsis subtypes). Each row of subplots represents a distinct cluster. The left subplots show the IL-6 (pro-inflammatory) response and the right subplots show the IL-10 (anti-inflammatory) response. The Y-axis represents the normalized log10 values of the associated cytokine. The bold black curves indicate the master response for each cluster. Subject data (dots) was time-shifted to best align with these trajectories.

and anti-inflammatory response. Compared to Cluster 1, the anti-inflammatory response of Cluster 2 had a slower, protracted decline. Cluster 3 illustrates a low magnitude and slower pro-inflammatory response. The anti-inflammatory response was steadily increasing. This cluster had the lowest 14-day mortality and lowest incidence of 14-day MOF. Cluster 4 represents a fast rise and then sustained pro-inflammatory response. The anti-inflammatory response rose to a magnitude higher than that of Cluster 1 and then descended swiftly. Cluster 5 shows a medium magnitude (similar to Cluster 2) pro- and anti-inflammatory response characterized by a faster rise and decline. Cluster 5 was unique in that it represents a fast-inflammatory response while all other clusters demonstrated protracted responses. Table 3.2 summarizes the dynamics of five clusters.

Table 3.2: Summary of the Five Identified Endotypes

Endotype	Description
1	Overwhelming inflammatory response to infection (elevated cytokinemia)
2	Prolonged anti-inflammatory response, (mild immunosuppression)
3	Rising only anti-inflammatory response, (immunosuppression)
4	Prolonged pro-inflammatory response, (sustained inflammation)
5	Fast dynamic response that ends within two days (immunodeficiency)

The calculation of a model-based master response trajectory within each cluster allowed the time-axis to start at the estimated onset of infection rather than to the start of data collection. Subjects were shifted up to 78 hours (representative of pre-hospital time) to best align with cluster master trajectories. It appears that subjects in Cluster 5 were enrolled into the ProCESS trial earlier along their disease timeline due to the shorter estimated onset times. In contrast, Cluster 1 subjects were enrolled later, with the majority of subjects within 20-60 hours after infection onset. The subjects within Cluster 4 were enrolled even later, between 40-80 hours after infection onset, as evidenced by their sustained IL-6 response. Clusters 2 and 3 appeared to have a larger variety of patient onset times.

Table 3.3: Comparison of baseline characteristics and outcomes between each endotype.

	Endotype					
Characteristic	1 n=116	2 n=67	3 n=69	4 n=63	5 n=75	
Age	63.4±17.7	60.6±16.5	58.1±16.8	62.6±15.1	58.9±16.9	
Male sex – no. (%)	74 (63.8%)	38 (56.7%)	38 (55.1%)	36 (57.1%)	50 (66.7%)	
Race – no. (%)						
White	81 (69.8%)	50 (74.6%)	50 (72.5%)	43 (68.3%)	44 (58.7%)	
Black	23 (19.8%)	13 (19.4%)	13 (18.8%)	16 (25.4%)	23 (30.7%)	
Other	12 (10.3%)	4 (6.0%)	6 (8.7%)	4 (6.3%)	8 (10.7%)	
Source of sepsis – no. (%)						
Pneumonia	45 (38.8%)	24 (35.8%)	24 (34.8%)	15 (23.8%)	25 (33.3%)	
Urinary Tract Infection	24 (20.7%)	17 (25.4%)	13 (18.8%)	20 (31.7%)	19 (25.3%)	
Intraabdominal infection	17 (14.7%)	3 (4.5%)	5 (7.2%)	6 (9.5%)	6 (8.0%)	
Infection of unknown source	12 (10.3%)	8 (11.9%)	8 (11.6%)	9 (14.3%)	9 (12.0%)	
Skin or soft-tissue infection	9 (7.8%)	7 (10.4%)	5 (7.2%)	1 (1.6%)	3 (4.0%)	
Catheter-related infection	5 (4.3%)	3 (4.5%)	2 (2.9%)	2 (3.2%)	3 (4.0%)	
Central nervous system infection	0 (0.0%)	0 (0.0%)	1 (1.4%)	1 (1.6%)	1 (1.3%)	
Endocarditis	0 (0.0%)	1 (1.5%)	0 (0.0%)	0 (0.0%)	2 (2.7%)	
Other	4 (3.4%)	4 (6.0%)	7 (10.1%)	9 (14.3%)	5 (6.7%)	
Determined not to have infection	0 (0.0%)	0 (0.0%)	4 (5.8%)	0 (0.0%)	2 (2.7%)	*
Positive blood culture – no. (%)	52 (44.8%)	21 (31.3%)	13 (18.8%)	29 (46.0%)	23 (30.7%)	*
Injury Severity						
APACHE II	21.0±7.5	21.2±6.8	19.7±6.4	20.8±8.3	21.1±7.9	
APACHE III	65.1±20.8	52.3±19.1	60.7±21.5	63.0±25.2	62.8±25.4	
Charleson	2.5±2.4	2.5±2.5	2.5±2.2	1.9±2.1	2.9±2.7	
SOFA	8.1±3.7	7.0±3.1	5.9±3.2	7.7±3.8	7.0±3.8	
Physiological variables						
Systolic blood pressure – mmHG	94.9±23.4	96.8±21.4	100.2±24.8	100.0±24.6	99.9±25.6	
Heart rate – beats/min	108.9±17.9	102.0±19.5	97.7±21.6	103.5±22.9	104.8±20.1	*
Temperature – C	37.5±1.2	37.2±1.3	37.0±1.1	37.0±1.7	37.3±1.7	
Respiratory rate – breaths/min	23.9±7.0	22.9±7.2	20.4±5.8	22.4±8.2	23.1±7.4	*
Total bilirubin – mg/deciliter	1.5±1.9	1.3±1.7	0.9±0.5	1.7±2.3	1.2±0.9	
Serum lactate – mmol/liter	4.9±2.9	4.1±2.9	3.6±2.5	5.3±3.6	4.6±2.4	*
Outcome						
14 Day All-cause Death – no. (%)	14 (12.1%)	7 (10.4%)	4 (5.8%)	6 (9.5%)	8 (10.7%)	
Mortality – days	84.8±110.9	63.2±67.8	109.6±108.9	47.1±41.7	47.5±55.4	
Multiple Organ Failure – no (%)						
14 Days	85 (73.3%)	43 (64.2%)	33 (47.8%)	46 (73.0%)	43 (57.3%)	*
Stay in hospital – days	11.5±10.7	12.3±8.2	11.5±9.8	13.6±10.1	10.5±7.4	*
ICU	5.6±6.0	4.7±4.1	4.7±5.4	5.4±4.3	5.2±5.4	
Serious Adverse Events – no. (%)	7 (6.0%)	2 (3.0%)	5 (7.2%)	1 (1.6%)	0 (0.0%)	

$*$ denotes $p \leq 0.05$ from Dunn' or Chi-Squared tests to determine differences between continuous or categorical variables, respectively.

3.2.1 Statistical Analysis of Endotypes

Patient demographics, vital signs at hospital admission, baseline physiological variables, and outcomes from each cluster endotype are shown in Table 3.3. The most striking difference between endotypes was the differing levels of 14-day incidence of multiple organ failure (MOF). Clusters 1 and 4 had the largest population of subjects that tested positive for blood culture infections, which corresponded with high rates of MOF. Clinical biomarkers such as heart rate, respiratory rate, and serum lactate demonstrated significant difference as well. Overall, Table 3.3 does not reveal many statistically significant differences among the five endotypes. This lack of significant differences reveals that patient demographics and admission vital signs were unable to fully classify subjects into a cytokine response cluster.

Figure 3.5 shows the 47 clinical biomarker trajectories that were analyzed. Of these, 29 biomarkers were identified to be significantly different between clusters. This demonstrated that, despite being similar to measurements taken at hospital admission, these biomarkers diverged within six hours. Other biomarkers such as cytokines, urine output, and white blood cell adhesion molecules were significantly different between clusters. Furthermore, pairwise testing revealed that these biomarkers offer high discriminatory power among clusters. For example, all pair-wise testing of IL-6 values involving cluster 1 were significantly different from all other clusters.

A comparison of endotype memberships between the cohort described in Section 3.1 and Chapter 2 is illustrated in Table 3.4. Chi-Squared testing revealed that the endotypes from both Chapters were not independent. The subjects of high cytokinemia endotype 1 predominantly belonged to high-risk endophenotypes A from Chapter 2 and some were split into high-risk endophenotypes E and F. The anti-inflammatory dominant endotype 3 was mostly represented by endophenotype D. However, endotypes 2, 4, and 5 were spread out across endophenotypes A-F.

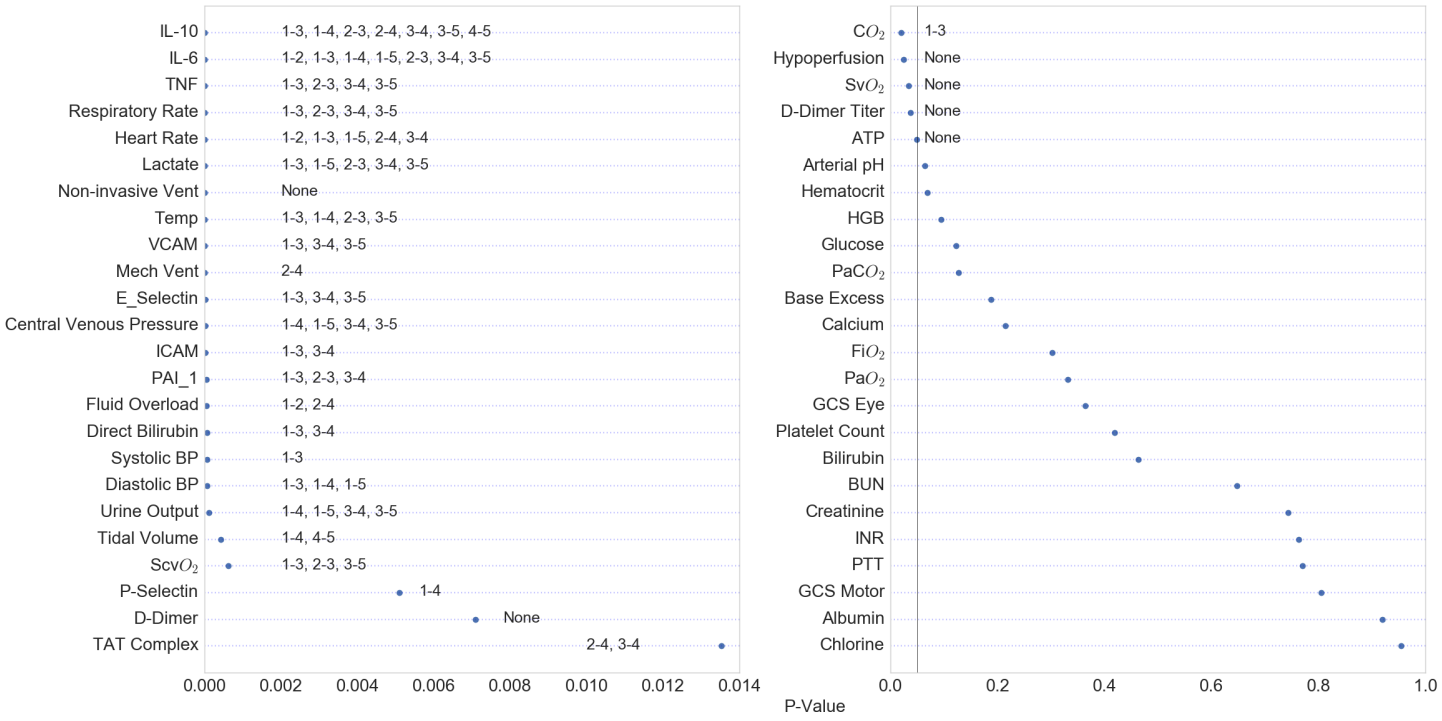


Figure 3.5: Omnibus testing (Dunn's test) on 6-hour biomarker trajectories from each endotype. A significance cutoff of p$\leq$ 0.05 was applied (shown on right panel). Paired numbers on each row reveal pair-wise differences revealed by Dunn's test pertaining to that biomarker. Note that x-axis spans of left and right plots are different. The phrase "None" indicates that single pair-wise endotype differences did not exist.

3.3 DISCUSSION

A model-based approach to identify clusters of similar cytokine responses was presented. Five distinct types of IL-6 and IL-10 responses were discovered. Table 3.3 shows that patient demographics and baseline vital signs were unable to distinguish between patients in different clusters. While seemingly similar at the beginning of the ProCESS trial, differences were clearly expressed in cytokine responses over time. This suggested that patient differences were rooted deeper than clinical phenotypic manifestations.

Endotypes 1 and 3 represent drastically different inflammatory responses and highlight the need for a more personalized approach to sepsis therapy. Intuitively, the predominately

Table 3.4: Contingency table comparing endotype memberships between Chapters 2 and 3.

		Chapter 2 Cohort 321 of 493						
		A	B	C	D	E	F	$\sum$
	1	47	8	1	4	20	16	96
Chapter 3 Cohort	2	19	12	2	1	9	13	56
321 of 390	3	2	5	8	36	6	10	61
	4	20	4	6	0	4	16	50
	5	19	13	2	4	6	14	58
	$\sum$	107	42	19	39	45	69	321

Chi-Squared test of independence between the endotype memberships between the two chapters rejected the null hypothesis with a p-value < 0.00001.

anti-inflammatory response of cluster 3 corresponded with a low MOF and mortality while the characteristically high magnitude responses of endotype 1 resulted in high MOF and mortality. Furthermore, 29% of patients in endotype 3 returned to baseline IL-6 and IL-10 values (lower limit of detection) by 72 hours while only 6% returned in endotype 1. Endotype 1 may not benefit from immune-stimulating therapies and endotype 3 may not benefit from immune-suppressing therapies. Interestingly, the other endotypes had low rates of returning to baseline ($< 5\%$).

Figure 3.6 illustrates the progression of commonly used clinical biomarkers in sepsis for each endotype. To normalize biomarkers for easy comparison, an empirical cumulative distribution function (eCDF) was calculated for each biomarker over the entire convenience cohort. This process was described in detail in Chapter 2.1. The resulting normalized dataset ranged between [-10, 10] for all biomarkers, and 0 represented median values. Each patient data column was then organized in the order they were sorted in Figure 3.4 to present a clear map of biomarker progression over each endotype's disease timeline. Biomarkers within each endotype clearly diverged over time. Endotype 3 had overall lower levels of bilirubin, lactate, TNF, IL-6, and IL-10. Endotypes 1 and 4 were defined by higher initial levels of creatinine,

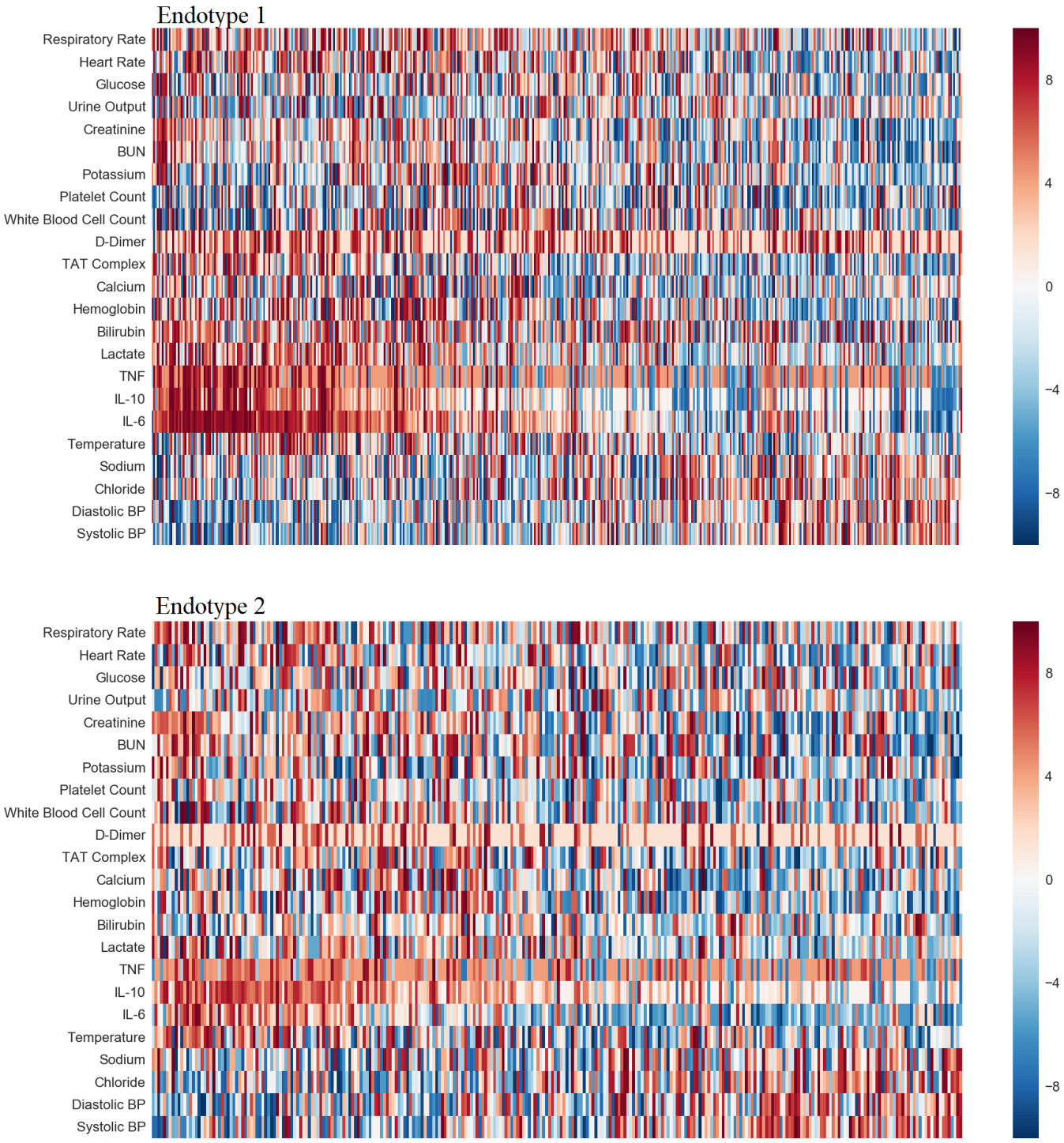

Endotype 1
Respiratory Rate
Heart Rate
Glucose
Urine Output
Creatinine
BUN
Potassium
Platelet Count
White Blood Cell Count
D-Dimer
TAT Complex
Calcium
Hemoglobin
Bilirubin
Lactate
TNF
IL-10
IL-6
Temperature
Sodium
Chloride
Diastolic BP
Systolic BP
8
4
0
−4
−8
Endotype 2
Respiratory Rate
Heart Rate
Glucose
Urine Output
Creatinine
BUN
Potassium
Platelet Count
White Blood Cell Count
D-Dimer
TAT Complex
Calcium
Hemoglobin
Bilirubin
Lactate
TNF
IL-10
IL-6
Temperature
Sodium
Chloride
Diastolic BP
Systolic BP
8
4
0
−4
−8

Figure 3.6

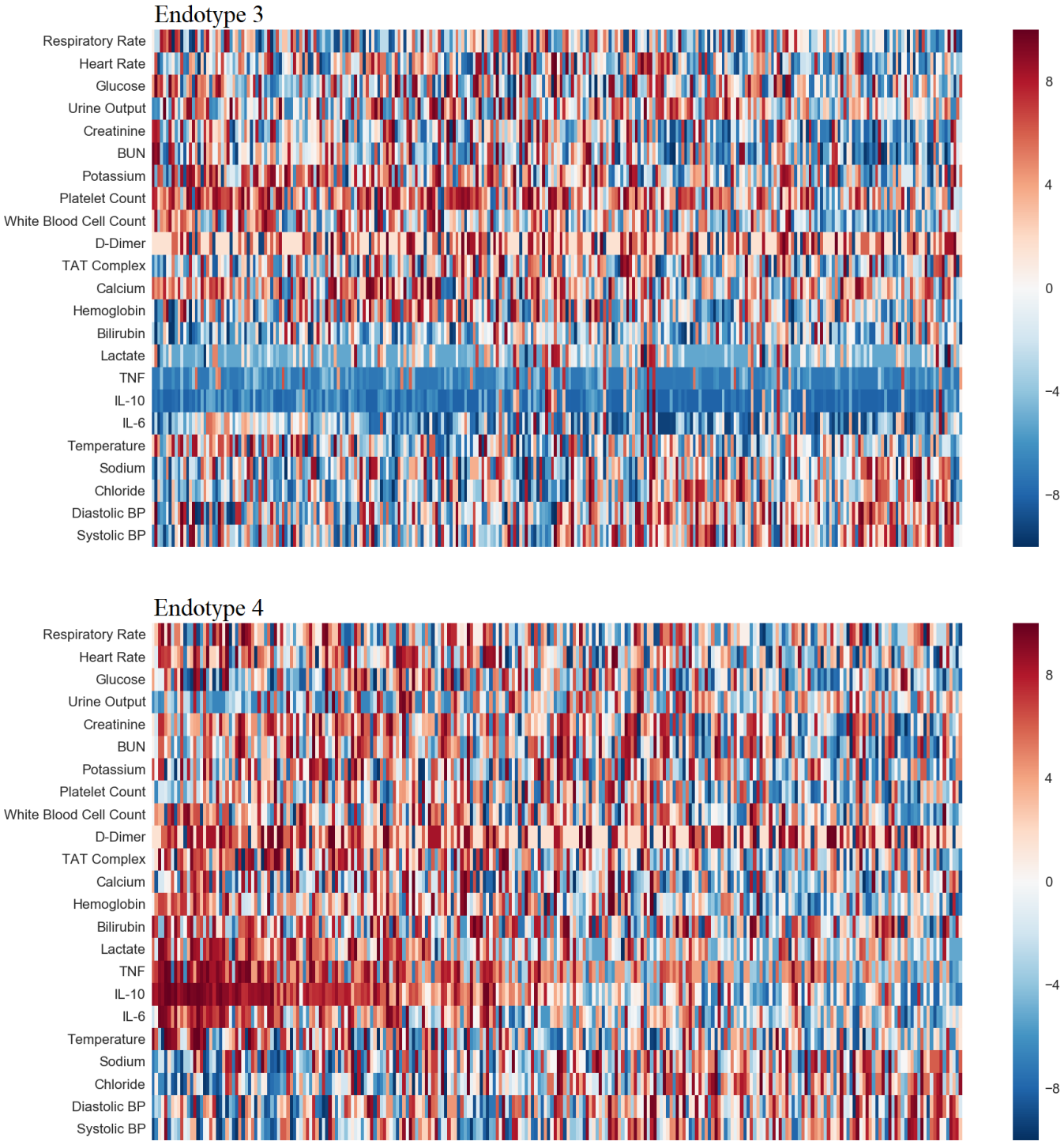

Endotype 3
Respiratory Rate
Heart Rate
Glucose
Urine Output
Creatinine
BUN
Potassium
Platelet Count
White Blood Cell Count
D-Dimer
TAT Complex
Calcium
Hemoglobin
Bilirubin
Lactate
TNF
IL-10
IL-6
Temperature
Sodium
Chloride
Diastolic BP
Systolic BP
8
4
0
−4
−8
Endotype 4
Respiratory Rate
Heart Rate
Glucose
Urine Output
Creatinine
BUN
Potassium
Platelet Count
White Blood Cell Count
D-Dimer
TAT Complex
Calcium
Hemoglobin
Bilirubin
Lactate
TNF
IL-10
IL-6
Temperature
Sodium
Chloride
Diastolic BP
Systolic BP
8
4
0
−4
−8

Figure 3.6

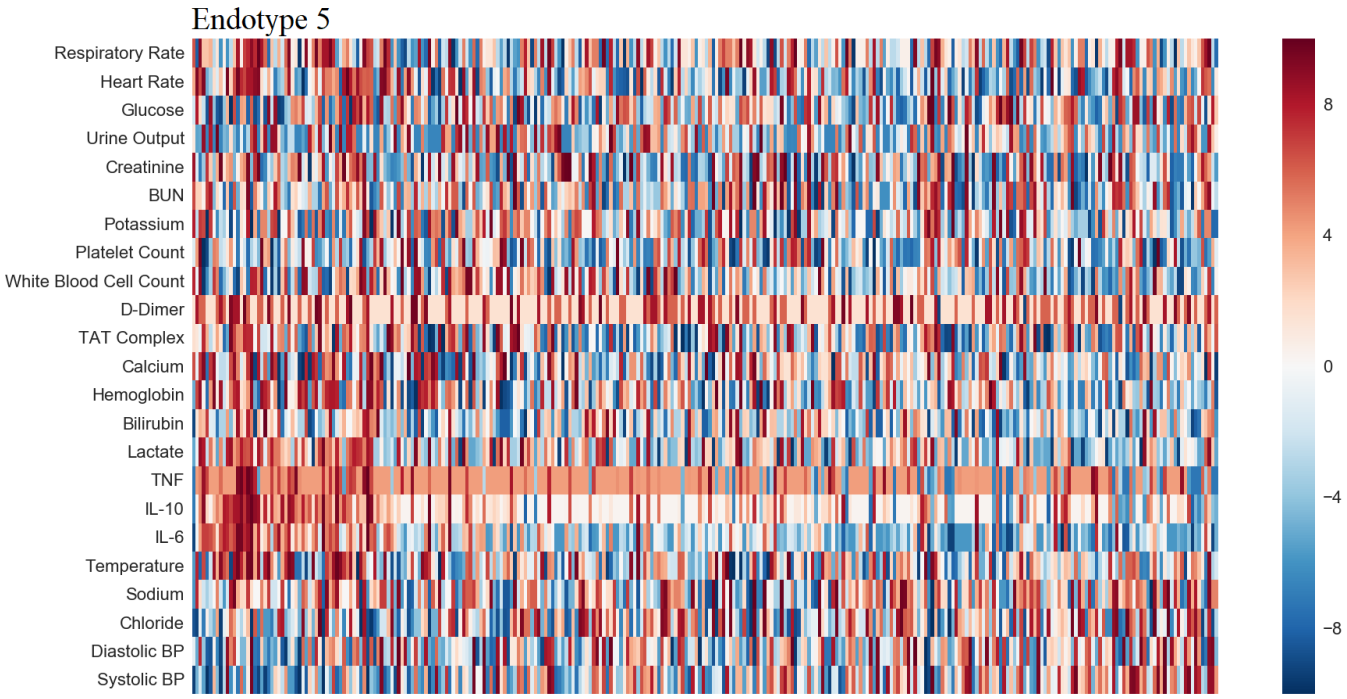


Figure 3.6: Heatmap illustrating clinical biomarker progression over each endotype. Columns (unlabeled) consist of time, relative to estimated time of infection as estimated from Figure 3.4. Individual subjects were not distinguished in heatmaps. Rows consist of commonly measured clinical biomarkers. A cumulative distribution function-based method was used for biomarker normalization to enable comparability (refer to Chapter 2.1).

TAT complex, blood urea nitrogen (BUN), and potassium. Blood pressure, sodium, chloride, platelet count, and white blood cell counts followed similar trends in both clusters. Figure 3.6 demonstrate that the identified endotypes manifest into several observable phenotypic differences once pre-hospital times are used to temporally organize the data.

3.3.1 Study Limitations and Potential Improvements

The biggest assumption with this method was the use of a 72-hour measurement for normalization. Analyzing step response behavior forced the method to assume a nonzero steady-state value. Converting each subject's IL-6 and IL-10 trajectories to end at a value of unity was convenient for the method, but required an assumption that cytokines were at pseudo-

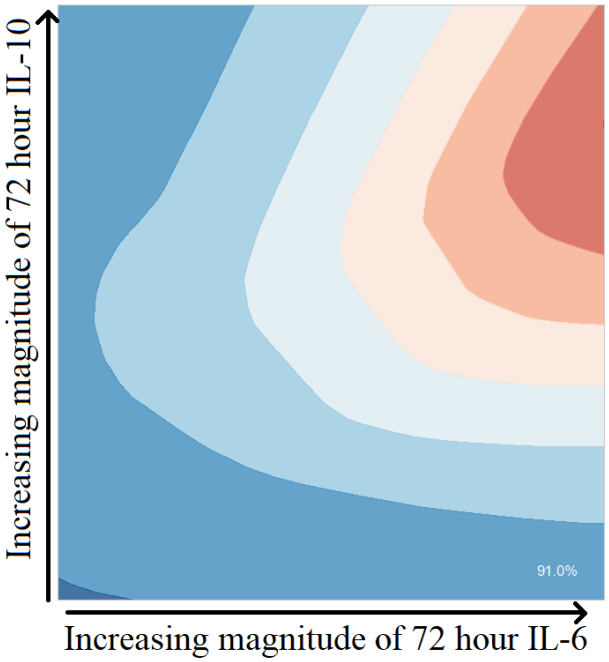


Figure 3.7: Machine Learning Derived regions of mortality based on IL-6 and IL-10 values at 72-hours. Using a 60/40 training-test split on the convenience cohort, a multi-layer perceptron classifier (neural network) was able to learn the regions of varying 14-day all-cause mortality based on 72-hour measurements of IL-6 and IL-10. The *sklearn* package in Python (*MLPClassifier* function, $\alpha = 1$ and other options set to default) was used. The algorithm was able to generate mortality predictions with 91% accuracy. Regions of uniform color represent regions of similar outcomes. The top right, associated with elevated levels of 72-hour IL-6 and IL-10, corresponds with high rates of mortality.

steady-state by hour 72. As a result, the mixture model method focused on the dynamics, rather than magnitude, of the inflammatory response up to 72 hours. The lack of magnitude consideration removed the need to estimate a steady state gain from the transfer function models but can be improved upon, considering a large body of clinical literature associating cytokine magnitudes with sepsis severity and mortality [74–77]. Furthermore, long-term daily cytokine measurements in a separate study revealed that cytokine values in severely septic patients do not stabilize by 72 hours [66]. While the ProCESS dataset is one of the richest sepsis datasets, the method can benefit from a dataset with longer and more frequently measured cytokine trajectories.

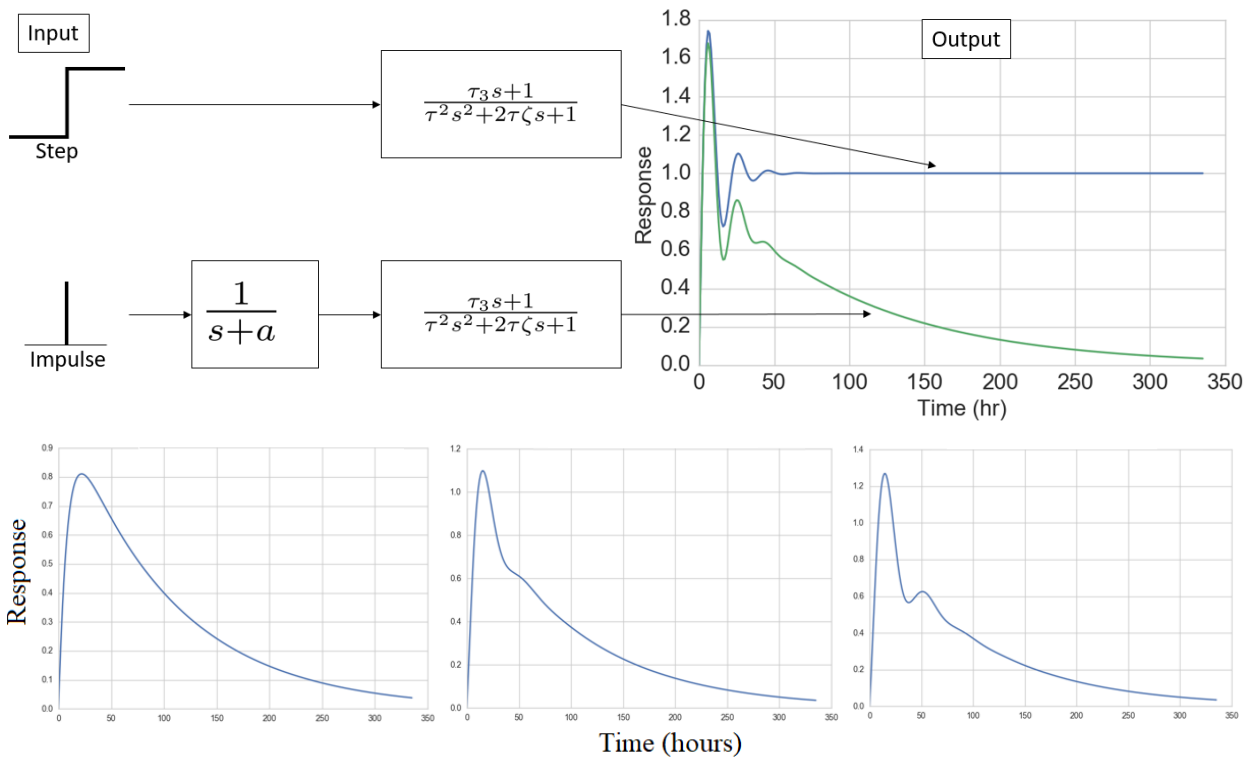


Figure 3.8: Schematic of impulse-response behaviors to explore in order to address current methodological weaknesses. Impulse-response behaviors prevent the need to assume any steady-states in cytokine trajectories. The bottom set of three graphs illustrate the types of behaviors this system is capable of depending on the value of the damping parameter, ζ.

Magnitude considerations may be integrated into the current method by using the IL-6 and IL-10 values at 72-hours to provide further stratification. For example, each cluster may be further stratified into low, medium, or high IL-6 and IL-10 72-hour magnitudes. Statistical modeling revealed distinct mortality regions in their magnitudes. These mortality regions were distinctive (see Figure 3.7) and machine learning algorithms can learn these regions and generate updated risks of 14-day all-cause moralities.

The next step was to consider transfer functions with impulse-response inputs, rather than the steps used in Section 3.1. A major benefit of this approach is that impulse-response behaviors always decay back to 0, removing the need for any steady-state assumptions in the inflammatory response. A system is proposed in Figure 3.8, which illustrates a slight

modification to the oscillatory system in the method. This system can generate a variety of responses depending on the value of the damping parameter, ζ.

Finally, a major difficulty with the use of mixture models is the tendency of such methods to collapse all subjects into one cluster. The iterative process of recovering dynamics from cytokine segments and clustering them was very sensitive to (i) initial conditions, steady-state gain, and/or magnitude of the input, (ii) the types of transfer functions and how different their responses were, and (iii) the time shift parameter to realign subjects to a master trajectory. Currently attempts with the impulse response model illustrated in Figure 3.8 have resulted in one cluster because the model was subject to all three problems. In contrast, the method in Section 3.1 removed the need to estimate the parameters that cause issue (i) and used 3 transfer function models with step-responses that allowed sufficient separation (issue (ii)). A potential solution to improve the impulse response model is to estimate pre-hospital times via a separate approach (thereby tackling issue (iii)), which may improve the identifiability of model parameters, including steady-state gain.

3.3.2 Bedside Endotype Classification Tool

The endotypes identified within this section necessitated the development of a bedside classification tool to guide clinical decisions. The requirements of this tool are twofold: (i) it makes accurate endotype predictions and (ii) it generates predictions quickly upon clinical presentation of a subject. To accommodate requirement (ii), predictors for this tool were artificially limited to the first six hours of measurements (relative to the ProCESS trial start).

First, the entire selection of machine learning classification algorithms from the Python SciKit-Learn package was tested. The 29 significant biomarkers identified in Figure 3.5 were used to train each classifier and were broken into 70%-30% training-test sets. The best performing algorithms were decision tree-based algorithms, which yielded up to 54% accuracy on the test set. Confusion matrices of algorithmic performance revealed that algorithms were able to predict endotypes 1 and 3 accurately, but missed endotypes 2, 4, and 5. This suggested that endotypes 1 and 3 were clearly distinguishable, which was important considering the differing clinical support required by either endotype.

Second, a decision-tree support vector machine (SVM) was trained on the data [78]. This approach breaks a multi-class classification problem into separate problems: use a SVM to separate one class from others, then repeat over the remaining classes to generate predictions in a decision-tree-like process. The algorithm, using radial basis functions, yielded between 75% to 89% accuracy on one-class-versus-rest problems. However, overall endotype accuracy was 30% to 45%, which is an improvement on a null result of 20% accuracy (random assignment to 5 clusters). Further analysis revealed that the SVM accurately identified subjects from endotypes 1 and 3, but were unable to fully distinguish between endotypes 2, 4, and 5.

Finally, a nearest neighbor algorithm yielded the best results [41]. The imputation method to obtain 72-hour cytokine measurements of 225 subjects utilized only biomarkers collected within 6 hours. A one nearest neighbor approach was tested to verify if 0 and 6-hour measurements of IL-6 and IL-10 can provide endotype segregation. Matching normalized (by imputed 72-hour values) IL-6 and IL-10 measurements, limited to 0 and 6 hours, to one of the 5 master response curves in Figure 3.4 yielded 73% endotype prediction accuracy. This result suggests that a 6-hour temporal trajectory of vital signs combined with 0 and 6-hour measurements of cytokines can be used to predict patient endotypes with good accuracy. This algorithm satisfied both requirements of the bedside classification tool. However, this tool warrants further research via a traditional machine learning algorithm given the promising statistical segregation suggested by Figure 3.5.

3.3.3 Clinical Implications

The 73% classification accuracy of the aforementioned classifier demonstrated that a clinician, rather than having to measure a subject's cytokines over 72 hours, may instead identify a patient's endotype with high accuracy using a mixture of cytokines and vital signs measured over six hours. This provides a clinician a potential bedside tool to gauge the endotype-associated risk of mortality and MOF, which may improve patient outcomes in the clinic. High-risk patients may receive preemptive organ support therapies before the onset

of MOF. For patients with low risk of succumbing to the initial septic shock in 14-days, clinicians can focus on preventing downstream complications.

Endotypes of sepsis have not been extensively studied, but the ability to identify them has profound implications on how sepsis research is conducted in the future. The results suggest that a one-size-fits-all approach to sepsis therapy may be inappropriate or even detrimental to certain scenarios. For example, Figure 3.6 highlighted significant biomarker differences between endotypes 1 and 3, suggesting that an endotype-specific therapeutic approach might be warranted. Specifically, an immuno-modulatory drug may benefit the high magnitude responses characteristic of endotype 1, but may be harmful to patients exhibiting the low magnitude IL-6 response in endotype 3. By identifying cohorts that may be harmed or benefit from certain treatments, endotype stratification may explain why so many clinical trials in sepsis have failed in the past [57]. Furthermore, improvements can be made in patient populations enrichment in future clinical trials. This method may allow clinicians to pre-screen candidates for endotype, and possibly pre-hospital times, to select those who may benefit from a proposed therapy.

3.3.4 Summary

This Chapter provides an extension of the work in Chapter 2 by incorporating mechanistic understandings into the clustering of the inflammatory response to sepsis. A novel methodology for endotype identification was developed by simultaneously reconstructing the dynamics of IL-6 and IL-10 trajectories and clustering on these dynamics. Reconstruction of these dynamics was performed by parameterizing each subject via mathematical models and estimating pre-hospital time. Representative master trajectories of the pro- and anti-inflammatory response were calculated for each endotype. These master trajectories revealed distinct and varied responses within a 390 subject cohort from the ProCESS trial. Statistical analysis of these endotypes revealed significant differences within 6 hours of patient enrollment, but an accurate early-detection classification tool (current accuracy is at 73%, which needs improvement) from phenotypic patterns is not yet available. The five endotypes identified in this Chapter were not fully explainable by phenotypes alone. Additional

research will further improve mechanistic understandings of these endotypes, discover novel segregation strategies, and reveal richer features, all of which will enable the translation of sepsis endotyping into the clinic.

4.0 MECHANISTIC DIFFERENCES BETWEEN ENDOTYPES

The key to improving sepsis therapies starts with understanding the mechanisms involved in the dysregulations that cause a normal inflammatory response to turn into a sepsis response. For example, Chapter 6.1 reveals that there are age-related immune pathway differences in mice which manifests into longer inflammatory durations within older mice. Such differences support the notion of mechanistic differences driving the five distinct sepsis behaviors seen in Chapter 3. Improved understanding of these pathobiological differences can provide targets for pharmaceutical intervention.

Mathematical modeling provides a quantitative framework to understand the mechanisms behind the complex sepsis syndrome. A myriad of mechanistic sepsis models exist to capture the representative inflammatory behaviors, but few have been calibrated against human data [3, 4, 52, 55, 70, 79, 80]. The models that have been calibrated against human data are based on endotoxin responses, which is mechanistically different from sepsis [37, 81–83].

A mechanistic model, tailored for the ProCESS data, was generated in response to the shortcomings of existing mathematical models. The model focused on fitting the temporal white blood cell (monocytes and neutrophils) and cytokine (TNF-α, IL-6, IL-10) measurements within the ProCESS data in a compartmental-based (blood, bone marrow, infected tissue) framework. The model characterized known phenomenological dynamics and interactions of each state to mimic the inflammatory response. Model structure was modified to include additional pathways that were deemed necessary to capture desired behaviors in the data. First, the methods of model calibration are discussed. Second, the final iteration of the model is presented. Finally, an endotype-level (from Chapter 3) parameter analysis of the subjects is conducted and discussed.

4.1 METHODS

The model initiated as a project to fit a porcine model of sepsis (unpublished data, Soheyl Bahrami, Ph.D., Ludwig Boltzmann Institute of Experimental and Clinical Traumatology). 14 pigs were subject to surgically induced peritonitis (introduction of bacteria into the abdominal organs), 7 of which were given an experimental super-oxide dismutase therapy. The 7 control pigs were used for model fits. Measurements of serum TNF-α (TNF), IL-6, IL-8 (a chemokine that attracts neutrophils from serum into tissue), IL-10, neutrophils, and monocytes were available in 2-hour increments up to hour 12 post sepsis induction. The model development phase was broken into three steps. First, as drivers of the inflammatory response, cytokines TNF, IL-6, IL-8, and IL-10 were independently fit using the step-response transfer function method outlined in Chapter 3.1. The step-response of the overshoot-type transfer function was used. The step-response was appropriate here because the nonzero lower limit of detection was treated as a steady-state value, which was modeled by a steady-state gain. A delay term ($e^{-\theta s}$) was included to account for the delay between the onset of sepsis (start of peritonitis) to the initial rise of each cytokine. Once these cytokine models were developed, the next step was to use them to develop the model structure governing white blood cell dynamics. Filter functions of TNF and IL-6 were generated to account for the delays and timing of downstream white blood cell dynamics. Three compartments, **Bone**, **Tissue**, and **Blood**, were designed to account for the relevant recruitment dynamics of white blood cells. The third step was to use the resultant white blood cell model to replace the transfer function cytokine models with a mechanistic one.

With model structure defined, calibration was performed using a human endotoxin infusion experiment [84]. Although the endotoxin inflammatory response is not entirely mechanistically similar to sepsis, the model structure tweaks in this step served as a transitional step to adapt the model for human data. Endotoxin was administered at 10, 20, and 40 ng/kg for three groups of subjects. Measurements were taken of serum TNF, IL-6, IL-8, IL-10, neutrophils, and monocytes at hours 0 (start of experiment), 1, 1.5, 2, 3, 4, 6, 8, 12, and 24.

The porcine experimental data helped develop the white blood cell and cytokine model structure. The human endotoxin experiment was used to adjust the IL-6 and IL-8 filter functions to reflect the slower dynamics in humans compared to the porcine response. Furthermore, the fast circulating neutrophil dip within 2 hours justified the addition of an LPS state to allow for IL-8 independent recruitment of neutrophils into the **Tissue** compartment.

The ProCESS data described in Chapter 2.1 was used for model fitting. Temporal trajectories of serum TNF-α, IL-6, IL-10, and neutrophils were available for model calibration. Serum monocytes were not explicitly available but given measurements of total white blood cell counts and neutrophil counts, an upper bound for serum monocytes were available. Under normal conditions, monocytes account for less than 10% of white blood cells. During inflammatory periods, monocyte composition increases, but numbers were not available in literature. An assumption was to limit monocytes to an upper limit of 40% of the non-neutrophil white blood cell counts. Dynamics began at the onset of infection and patient data were shifted in accordance with the estimated pre-hospital times in Chapter 3.2. Simulation time was 0 to 150 hours post-infection (first week of sepsis) to fit the assumptions made in Chapter 3.1 (maximum pre-hospital time was bounded at 78 hours and a patient may have up to 72 hours of cytokine measurements).

4.1.1 Model Formulation and Simulation

Parameter-fitting was performed using the *Pyomo* Python package and the IPOPT solver using the *ma86* linear solver. Pyomo is a framework of tools designed to convert a model specified in Python into the algebraic modeling programming language (AMPL) for use with a variety of numerical optimization solvers [85, 86]. IPOPT was used for optimization and is a solver that uses the interior point method. This method uses barrier functions to solve linear programming problems and a well-written explanation is provided by Robere [87]. Take, for instance, a generic LP:

$$\begin{aligned} & \min c^T x \\ s.t. \quad & Ax = b \\ & x \geq 0 \end{aligned} \tag{4.1}$$

The inequality constraint can be replaced with a barrier function and reformulated into a similar problem while preserving the Karush–Kuhn–Tucker conditions of the original problem (when $\mu = 0$).

$$\min c^T x - \mu \sum_{i=1}^{N} \ln x_i \tag{4.2}$$
$$st. \quad Ax = b$$

The benefit of solving Equation 4.2 over Equation 4.1 is that it relaxes the inequality constraints via nonzero values of μ. Interior point methods iteratively solves Equation 4.2 for decreasing values of μ, which effectively traverses the interior of the feasible region as it approaches an inequality constraint. At $\mu = 0$, the optimum to Equation 4.2 is also the optimum to Equation 4.1 as well.

The Pyomo framework comes with a differential algebraic equation (DAE) package which enables numerical optimization of problems involving ordinary/partial differential equations. The DAE package converts algebraic differential equation expressions into forward or backward finite difference approximations or orthogonal collocation on finite elements. For this Chapter, all parameter fitting was conducted on ordinary differential equations (ODEs) with the Pyomo DAE package set to the backward finite difference approximation option. The objective function was formulated as the minimization of the weighted sum of squared errors between the data and their relevant model states. Weights were applied on the TNF and white blood cell states to capture the quick responses to inflammation.

On the ProCESS database, a regularization term was included ($\lambda = 1$) to penalize nonzero values of cytokines by hour 150, which was a phenomenon that tended to occur in the absence of a data-driven reason. To prevent numerical issues pertaining to scaling, all data was scaled (divided by 10^i) to homogeneously range within [0,100]. Finally, an additional inequality constraint was posed on monocytes to enforce the aforementioned maximum monocyte limit.

Pyomo requires setting bounds on each parameter, which was a difficult task because most of the parameters have no literature sources. Initial parameter-fitting simulations were performed in APT-MCMC (see Chapter 6.2) using large, uninformative bounded uniform

distributions to obtain the posterior distribution of parameters. A posterior distribution is the distribution of a parameter after it has been conditioned upon observed data (fitted in accordance to the maximum likelihood estimator) and APT-MCMC calculates these distributions via Baye's Theorem (Equation (6.18)). The posterior distributions from the initial APT-MCMC simulations on the ProCESS data informed the Pyomo parameter bounds.

4.1.1.1 Initiation of Inflammation and the Cytokine Storm

$$\frac{dTNF_T}{dt} = (V_{mt}Macro + Rmacro)\frac{P}{K_{pmt}+P} - \mu_{t_t}TNF_T - d_tTNF_t \tag{4.3}$$

$$\frac{dIL6_T}{dt} = V_{m6}Macro\left(\frac{P}{K_{pm6}+P}\right)\left(\frac{K_{106}}{IL10_T+K_{106}}\right) - \mu_{6_t}IL6_T - d_6IL6_T \tag{4.4}$$

$$\frac{dIL8_T}{dt} = (V_{m8}Macro + Rmacro)\left(\frac{P}{K_{pmt}+P}\right)\left(\frac{K_{86}}{IL10_T+K_{86}}\right) - \mu_{8_t}IL8_T - d_8IL8_T \tag{4.5}$$

$$\frac{dIL10_T}{dt} = (V_{m10}Macro + Rmacro)\left(\frac{P}{K_{pmt}+P}\right) + V_{n10}\mu_{n_t}N_T\left(1 - \frac{TNF_T}{K_{tn}+TNF_T}\right) \tag{4.6}$$

$$- \mu_{10_t}IL10_T - d_{10}IL10_T \tag{4.7}$$

The inflammatory response begins with the presence of Pathogen (P) in the infected **Tissue** compartment. Resident macrophages that reside in the infected tissue detect the presence of pathogen via a Michaelis-Menten dynamic: $P/(K_i + P)$ and produce cytokines: TNF, IL-8, and IL-10 ($TNF_T, IL8_T, IL10_T$, respectively) [13, 55, 73]. IL-6 is also produced by macrophages (recruited and resident macrophages), but the model does not allow resident macrophages to produce IL-6 to enforce a delay between TNF and IL-6 peaks [39]. These cytokines leave the **Tissue** compartment via two pathways: natural decay/degradation (μ_{i_t} term) and diffusion into the **Blood** compartment (d_i term). Once cytokines diffuse into the blood, they each take on different inflammatory roles.

Eventually, blood monocytes are recruited towards the infected tissue. Upon arrival, recruited monocytes differentiate into macrophages ($Macro$) where they begin to produce TNF, IL-6, IL-8, and IL-10. This production is governed by a rate constant (V_{mi}) and the aforementioned Michaelis-Menten dynamic. Together, recruited and resident macrophages

produce large amounts of cytokines to emulate the cytokine storm phenomena during the early phases of sepsis [73]. While neutrophils are able to produce cytokines, monocytes are able to synthesize cytokines at a rate that is 50-100 times faster [88]. The model simplifies this dynamic by delegating the role of cytokine production to monocytes/macrophages only.

IL-10 is an anti-inflammatory cytokine and its primary mechanism for countering the pro-inflammatory response is to halt macrophage production of TNF, IL-6, and IL-8 [15]. This mechanism is governed by a non-competitive inhibitory Michaelis-Menten dynamic:

$$1 - IL10_T/(K_i + IL10_T) = K_i'/(IL10_T + K_i')$$

This inhibitory behavior blocks macrophage production of IL-6 and IL-8. IL-10 inhibition of TNF was not included in the model because TNF was used to drive the early dynamics of the inflammatory response. A fast and high magnitude peak of TNF is desired and necessary to trigger model dynamics and the inclusion of IL-10 inhibitory effects damped such behaviors. Finally, there is an additional IL-10 dynamic associated with neutrophils that will be discussed in Section 4.1.1.3.

4.1.1.2 Cytokine Diffusion and their Roles

$$\frac{dTNF_B}{dt} = d_t TNF_T - \mu_{t_b} TNF_b - K_{Tconsume} TNF_B (M_C + N_C) \tag{4.8}$$

$$\frac{dIL6_B}{dt} = d_6 IL6_T - \mu_{6_b} IL6_b - k_{6consume} IL6_B (M_C + N_C) \tag{4.9}$$

$$\frac{dIL6_{F1}}{dt} = d_6 IL6_b - d_6 IL6_{F1} \tag{4.10}$$

$$\frac{dIL6_{F2}}{dt} = d_6 IL6_{F1} - d_6 IL6_{F2} \tag{4.11}$$

$$\frac{dIL6_{F3}}{dt} = d_6 IL6_{F2} - d_6 IL6_{F3} \tag{4.12}$$

$$\frac{dIL6_{F4}}{dt} = d_6 IL6_{F3} - d_6 IL6_{F4} \tag{4.13}$$

$$\frac{dIL8_B}{dt} = d_8 IL8_T - \mu_{8_b} IL8_B - k_{8consume} IL8b \frac{N_c}{K_{nc8} + N_c} \tag{4.14}$$

$$\frac{dIL8_{F1}}{dt} = d_8 IL8_b - d_8/3 IL8_{F1} \tag{4.15}$$

$$\frac{dIL10_B}{dt} = d_{10} IL10_t - \mu_{10_b} IL10_b \tag{4.16}$$

Once cytokines diffuse into the **Blood** compartment at a rate of d_i, they undergo natural decay/degradation (μ_{i_b}). To capture the fast decline of TNF, IL-6, and IL-8 within the porcine and human LPS data, an additional sink term, with the rate constant $k_{iconsume}$, was added. Biologically, this represents the binding of these cytokines onto the receptors of circulating (in the blood) neutrophils and/or monocytes to activate/prime them. While more detailed neutrophil models exist differentiating circulating neutrophils into activated and not-activated pools, none of the neutrophil data in the available data was granular enough able to justify such distinction [70]. The $k_{iconsume}$ rate constant has no effect on neutrophil or monocyte behavior and exists merely as a **Blood** compartment cytokine sink.

IL-8 is a chemokine that is responsible for recruiting neutrophils into the tissue by way of extravasation (movement through endothelial walls) [89]. IL-6 is a pro-inflammatory cytokine that is able to recruit neutrophils and monocytes [14]. IL-6 and IL-8 both have filter functions which serve as delays in their induction of specific inflammatory pathways. These were deemed necessary during the migration from time-delayed transfer function cytokine models to mechanistic cytokine models during the porcine model calibration step. Furthermore, the filter functions captured cytokine migration towards their appropriate target (such as diffusion into the spleen to recruit marginated neutrophils into the blood) without explicitly modeling additional compartments.

The primary role of blood TNF in the model is to recruit monocytes into the infected **Tissue** compartment to sustain the inflammatory response. Monocyte recruitment is a complex process that is not well understood in humans [12]. Human monocyte recruitment is likely the result of a combination of CC-chemokine receptor 2 (CCR2) and CX_3C-chemokine receptor 1 (CX_3CR1), which was not measured in any of the calibration datasets [12]. Blood TNF took on the role of CCR2 in the model due to its fast acting pro-inflammatory profile.

Blood IL-10 has no role in the model because the majority of its role lies in the **Tissue**-level inhibition of the macrophage production of cytokines [15]. Blood IL-10 was explicitly modeled because all three data sets contained serum measurements of IL-10.

4.1.1.3 Neutrophil Dynamics

$$f_{Nexchange} = k_P(1 + 5\frac{IL8_{F1}}{K_{nr8} + IL8_{F1}})N_M - 5k_P(1 - \frac{IL8_{F1}}{K_{nr8}/5 + IL8_{F1}})N_C \tag{4.17}$$

Under homeostatic conditions, the majority of neutrophils are under intravascular margination, where they are stored in the spleen (primarily), bone marrow, and liver [90]. Despite being kept in storage, radio-labeling techniques have revealed that neutrophils are able to migrate from the marginated pool into blood within 2-10 minutes [90]. The size of the marginated neutrophil pool is currently unknown and under much debate. An assumption was made that the marginated neutrophil pool was 5 times larger than basal circulating neutrophil levels to capture the fast rise in neutrophils seen in the human endotoxin data. Furthermore, because neutrophil margination storage occurs in several compartments throughout the body, the model simplified this mechanism to a single state within the **Blood** compartment. Equation (4.17) describes the rate of exchange between the circulating and marginated neutrophil pools (N_C and N_M, respectively) and was formulated in a journal article by Ho, et al [91]. In Equation (4.17), a positive sign indicates flow into the circulating neutrophil pool.

During non-inflammatory conditions, to maintain a 1:5 (circulatory to marginated pool) ratio, neutrophils flow into the circulating rate at k_p and flow back into the marginated pool at $5k_p$. Under inflammatory conditions, a delayed IL-8 signal (due to transit time from blood flow into these marginated compartments) was used to modify k_p to greatly prefer transit into the circulating pool [92]. During inflammation, circulating neutrophils typically dip within the first 2 hours and is followed by a fast rise to above basal levels [88]. This can also be seen in the porcine and human endotoxin data sets. The marginated-circulating mechanism prevented neutrophil depletion and provided a source for the fast increase in neutrophil count.

$$\frac{Prog}{dt} = K_{trm}Prog\left(\frac{M_{c_{basal}}}{M_c}\right)^{\gamma_m} - K_{trm}Prog \tag{4.18}$$

$$+ K_{trn}Prog\left(\frac{N_{c_{basal}}}{N_c}\right)^{\gamma_n} - K_{trn}Prog \tag{4.19}$$

$$\frac{dT_{1_N}}{dt} = K_{trn}Prog - K_{trn}T_{1_N} \tag{4.20}$$

$$- K_{n6}\frac{IL6_{F4}}{K_{nr} + IL6_{F4}}T_{1_N}\frac{1}{N_M + K_{T_3N_M}}\frac{1}{K_{Tinhibit} + T_{3_N}}\frac{1}{K_{Tinhibit} + T_{2_N}} \tag{4.21}$$

$$\frac{dT_{2_N}}{dt} = K_{trn}T_{1_N} - K_{trn}T_{2_N} - K_{n6}\frac{IL6_{F4}}{K_{nr} + IL6_{F4}}T_{2_N}\frac{1}{N_M + K_{T_3N_M}}\frac{1}{K_{Tinhibit} + T_{3_N}} \tag{4.22}$$

$$\frac{dT_{3_N}}{dt} = K_{trn}T_{2_N} - K_{trn}T_{3_N} - K_{n6}\frac{IL6_{F4}}{K_{nr} + IL6_{F4}}T_{3_N}\frac{1}{N_M + K_{T_3N_M}} \tag{4.23}$$

The time span of the ProCESS data was three days and with the pre-hospital time shift, the model simulated one week of inflammatory dynamics from the infection onset. The long timespan necessitated a slower-acting source of neutrophils to supplement the dead neutrophils within the body. Neutrophil recruitment from the bone marrow is governed by G-CSF and various chemokine receptors [91, 93, 94]. Within the bone marrow, progenitor cells called hematopoietic stem cells differentiate into common myeloid progenitor (CMP) cells that can differentiate into monocytes or neutrophils [13, 89]. Under the neutrophil differentiation pathway, the CMP cells ($Prog$) differentiate into metamyelocytes (T_{1_N}), band cells (T_{2_N}), and then segmented granulocytes (T_{3_N}) [91, 95]. The results provided by Friberg, et al., and Ho, et al. both reveal that the bone marrow is capable of increasing the circulating neutrophil pool within 5-10 days, which is an appropriate time frame for the ProCESS data [91, 95]. Between $Prog$ and T_{1_N}-T_{3_N}, K_{trn} governs the natural maturation rate that it takes for cells to mature into each classification and includes the non-inflammatory migration of segmented granulocytes into the circulating pool. The progenitor cells receives a feedback signal from the **Blood** compartment: $\frac{N_{c_{basal}}}{N_c}^{\gamma_n}$, which induces progenitor cell growth if circulating neutrophils levels fall below normal levels.

In the model, the immature neutrophils within the bone marrow are available for emergency recruitment source should the marginated neutrophil pool subceed a certain level (governed by an inhibitory mechanism: $1/(N_M + K_{T_3N_M})$). The body first draws upon the segmented granulocytes (T_{3_N}) if the marginated neutrophil pool is too low. It then draws

upon the band cells (T_{2_N}) and finally the metamyelocytes (T_{1_N}).This model formulation was provided by Ho, et al [91].

$$\frac{dN_M}{dt} = -f_{Nexchange} \tag{4.24}$$

$$\frac{dN_C}{dt} = f_{Nexchange} - V_{n8}N_c\frac{IL8_T}{K_{n8}+IL8_T} + k_{trn}T_{3_N} - \mu_{n_b}N_c(1-\frac{TNF_B}{K_{tn}+TNF_B}) \tag{4.25}$$

$$+ K_{n6}\frac{IL6_{F4}}{K_{nr}+IL6_{F4}} \tag{4.26}$$

$$* \frac{1}{N_M+K_{T_3N_M}}\left(\frac{1}{K_{Tinhibit}+T_{3_N}}\left(T_{1_N}\frac{1}{K_{Tinhibit}+T_{2_N}}+T_{2_N}\right)+T_{3_N}\right) \tag{4.27}$$

$$- K_{ln}N_cLPS \tag{4.28}$$

The model uses filtered IL-6 to drive the recruitment from bone marrow into blood. The fourth IL-6 filter was used as a way of approximating the pro-inflammatory signaling cascade (chemokines and chemokine receptors) that interact with the relevant bone marrow cells [96]. This interaction was modeled with Michaelis-Menten kinetics. The recruited neutrophils migrate into the circulating neutrophil pool where they can be stored in the marginal pool, undergo natural degradation, or migrate into infected tissue.

IL-8 forms a chemotactic concentration gradient which attracts circulating neutrophils [89]. Specifically, circulating neutrophils migrate towards regions of increasing IL-8 concentrations via endothelial wall rolling, adhesion, and then intraluminal crawling [96]. This process, called extravasation, is governed by many chemokines and receptors, but the model simplifies these dynamics with a single Michaelis-Menten term on **Tissue** IL-8: $V_{n8}IL8_T/(K_{n8}+IL8_T)$. The attempts involving a difference term to represent the chemotactic gradient $(IL8_B - IL8_T)$ caused unexpected model behavior because the term often changed signs.

Neutrophils are capable of also following chemotactic gradients produced by endotoxins and lipopolysaccharide (LPS), which are molecular patterns indicative of a pathogen [88]. This mechanism provides quick neutrophil response into infected tissue prior to IL-8 driven recruitment and explains why the inflammatory response is typically faster in endotoxin experiments than in sepsis [37]. This term is represented by a simple kinetic rate constant, K_{ln}. LPS is modeled separately from pathogen to represent the release of endotoxin/LPS

into nearby tissue, thus forming a separate chemotactic gradient that can independently recruit neutrophils.

$$\frac{dN_T}{dt} = V_{n8} N_C \frac{IL8_T}{K_{n8} + IL8_T} - \mu_{n_t} N_T \left(1 - \frac{TNF_T}{K_{tn} + TNF_T}\right) + K_{ln} N_c LPS \tag{4.29}$$

Once neutrophils have entered the infected **Tissue** compartment, its only sink is death by degradation. Natural neutrophil degradation occurs at rates $\mu_{n_b} N_C$ and $\mu_{n_t} N_T$ in the blood and tissue, respectively. However, TNF is known to prolong neutrophil lifespans [11, 97]. Within the model, the natural degradation term is modulated by $1 - TNF_j/(K_{tn} + TNF_j)$ to represent a prolonged lifespan under inflammatory conditions.

The death of neutrophils contributes to the resolution of inflammation [97]. Neutrophils may die via one of two pathways: necrosis or apoptosis. Necrosis contributes to further inflammation because the cellular walls of neutrophils break down and release toxic, pathogen killing, molecules into the tissue. Apoptosis represents a "healthy" neutrophil death which involves signaling nearby macrophages for safe clearance of the dying neutrophil and its toxic contents [97–99]. Neutrophil apoptosis promotes the synthesis of IL-10 and other anti-inflammatory signals [97]. The model uses a rate parameter, V_{n10}, to govern the induction of IL-10 production due to apoptotic neutrophil death:

$$V_{n10} \mu_{n_t} N_T \left(1 - \frac{TNF_T}{K_{tn} + TNF_T}\right)$$

The model does not currently separate neutrophil death into apoptotic or necrotic stages due to the lack of monocyte measurements in the ProCESS database.

4.1.1.4 Pathogen Dynamics

$$\frac{dP}{dt} = K_{p1}P\left(1 - \frac{P}{P_\infty}\right)\left(\frac{P-\epsilon}{P+\epsilon}\right) - k_{NP}N_T\frac{P}{K_{p2}+P} \tag{4.30}$$

$$\frac{dLPS}{dt} = k_{pp}P - k_{pp2}LPS \tag{4.31}$$

Pathogen is located in the infected **Tissue** compartment and is introduced via a nonzero initial condition. Pathogen grows at a rate of K_{p1} but is limited by the P_∞ term to represent growth limitations due to the conditions within the **Tissue** compartment (nutrition availability, unfavorable environmental conditions, etc). A similar approach was taken in a journal article published by Reynolds, et al [4]. The ϵ term was introduced to prevent oscillations at low values of pathogen close to zero. **Tissue** neutrophils kill pathogen with a Michaelis-Menten kinetic, $k_{NP}P/(K_{p2}+P)$. Finally, pathogens release LPS (LPS) into the tissue, which forms a chemotactic mechanism that can drive **Tissue** neutrophil recruitment.

4.1.1.5 Macrophage Dynamics

$$f_{Mexchange} = k_{pm}(1 + \frac{IL6_{F1}}{K_{m6}+IL6_{F1}})M_M - 3.5k_{PM}(1 - \frac{IL6_{F1}}{K_{m6}+IL6_{F1}})M_C \tag{4.32}$$

Similar to neutrophils, monocytes circulate the blood and are sequestered in marginated compartments during non-inflammatory conditions at an estimated 1:3.5 circulating to marginated count ratio in human studies [12, 13, 100].Marginated monocytes, M_M, was represented as a single state within the **Blood** compartment. The dynamic sequestration exchange of M_M with circulating monocytes, M_C, is represented by Equation (4.32). Current knowledge monocyte recruitment from marginated compartments are not well known. The model assumes that a pro-inflammatory signal, represented by a filtered IL-6 ($IL6_{F1}$), biases Equation (4.32) towards M_C to increase circulating monocyte levels.

$$\frac{dT_{1_M}}{dt} = K_{trm}Prog - K_{trm}T_{1_M} \tag{4.33}$$
$$\frac{dT_{2_M}}{dt} = K_{trm}T_{1_M} - K_{trm}T_{2_M} \tag{4.34}$$
$$\frac{dT_{3_M}}{dt} = K_{trm}T_{2_M} - K_{trm}T_{3_M} - K_r\frac{IL6_{F4}}{K_{t6}+IL6_{F4}}T_{3_M}\frac{1}{M_M + K_{T_3M_M}} \tag{4.35}$$

Monocytes are derived from the same progenitor cells (CMP, represented by $Prog$) that differentiate into neutrophils [12]. Decreases in basal monocyte levels result in a growth signal to the CMP cells: $\frac{M_{c_{basal}}}{m_c}^{\gamma_M}$. Monocyte development can also be staged into 3 phases: progenitor granulocyte-macrophage progenitor (T_{1_M}), macrophage/dendritic cell progenitor (T_{2_M}), and common monocyte progenitor (T_{3_M}). While little is known about immature monocyte recruitment during inflammatory conditions, recent literature suggested that this mechanism occurs in mice [101]. A conservative assumption was made to use only the common monocyte progenitor state to replenish monocytes under inflammatory conditions. The model utilizes $IL6_{F4}$-driven recruitment from the T_{3_M} immature state when the marginal monocyte pool becomes low.

$$\frac{dM_M}{dt} = -f_{Mexchange} \tag{4.36}$$
$$\frac{dM_C}{dt} = f_{Mexchange} \tag{4.37}$$
$$- k_{mmt}M_c * \frac{TNF_B}{K_{tm}+TNF_B} \tag{4.38}$$
$$+ k_{trm}T_{3_M} - \mu_{m_b}M_c \tag{4.39}$$
$$+ K_r\frac{IL6_{F4}}{K_{t6}+IL6_{F4}}T_{3_M}\frac{1}{M_M + K_{T_3M_M}} \tag{4.40}$$

Circulating monocytes are affected by the exchange dynamic in Equation (4.32) and replenished from the T_{3_M} state (k_{trm} represents normal replenishment and K_r represents inflammatory driven replenishment). M_C decays at a rate of μ_{mb} in the blood. Unlike neutrophils, there are no lifespan prolonging mechanisms for monocytes. Monocyte recruitment

into tissue is driven by a complex chemokine signaling pathway that is approximated by a Michaelis-Menten interaction involving TNF_B.

$$\frac{dMacro}{dt} = k_{mmt} M_C \frac{TNF_B}{K_{tm} + TNF_B} - \mu_{m_t}(1 + Macro\frac{P}{k_{mtp} + P}) \tag{4.41}$$

Monocyte, upon entering the **Tissue** compartment, differentiate into macrophages ($Macro$). While macrophages are able to eliminate pathogens directly via phagocytosis, this interaction occurs on the orders of minutes compared to seconds for neutrophil-related elimination [99]. Macrophage elimination of pathogens is not modeled due to this rate difference. The primary role of macrophages is to synthesize cytokines. The secondary role of macrophages is phagocytosis and efferocytosis. Efferocytosis refers to the removal of apoptotic cells (mostly neutrophils) [98, 99]. Macrophages that undergo phagocytosis have been shown to exhibit elevated rates of death [102, 103]. The model includes macrophage death at a constant basal rate of μ_{mt} due to the longevity of macrophages. Under inflammatory conditions, this rate can be accelerated due to phagocytosis via a simplified mechanism involving pathogen levels: $P/(k_{mtp} + P)$.

4.1.2 Model Post-hoc Analysis

Figure 4.1 illustrates the aforementioned model structure. The model has 58 parameters and 28 states. Of the 28 states, the initial conditions for circulating monocytes, circulating neutrophils, and initial pathogen concentration were fitted (as part of the 58 total parameters). Initial conditions for marginated monocytes and neutrophils were set to 3.5x and 5.0x their respective circulating counterpart's initial conditions. The initial conditions for the progenitor cells and the immature white blood cell states within the **Bone** were set to 1000 $10^5/mL$. All other initial conditions were set to zero.

Individual fits for every subject within the 390 cohort in Chapter 3 was conducted using Pyomo and IPOPT. Cytokines TNF, IL-6, and IL-10 were scaled to fall within a range of [0,100] and their units were $10^1 pg/mL$, $10^2 pg/mL$, or $10^3 pg/mL$. Circulating neutrophils and monocytes were scaled to $10^5/mL$ and $10^4/mL$, respectively.

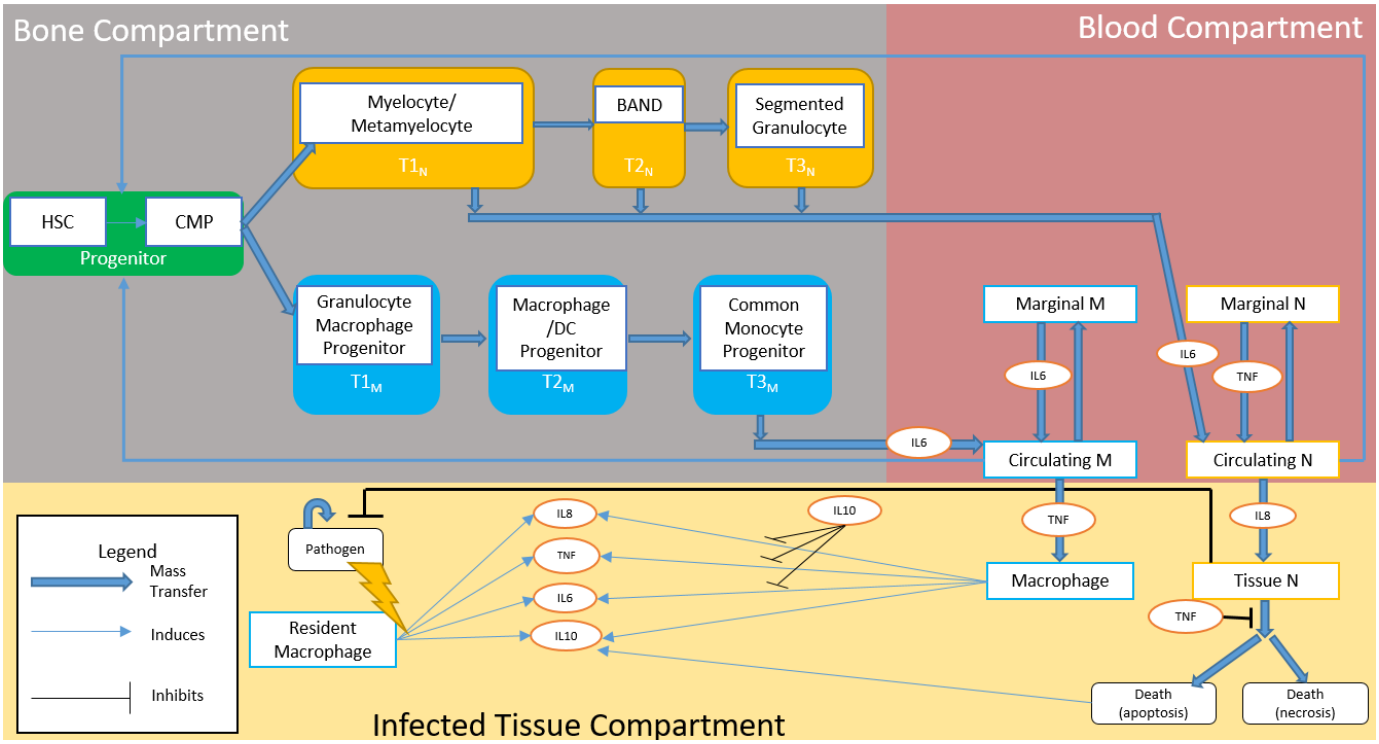


Figure 4.1: Diagram illustrating the sepsis ODE model. The tissue compartment is in yellow, blood in red, and bone in gray.

From each endotype, the 10 subjects with the lowest objective function values were considered to be "representative" patients of that endotype and selected for post-hoc analysis. Two tests were used to identify endotype-specific parameter differences. First, the Anderson-Darling omnibus test was used to identify differences between parameters within each endotype. Pair-wise differences (1 vs 4 and 2 vs 3) were tested with the Anderson-Darling test and a Bonferroni correction. Second, a nonlinear fixed-effect model using the covariates age, gender, and endotype was used. This model explicitly included age and gender, two well-known covariates that confound parameter differences between groups.

$$K = \theta \exp\left(\eta_0 + \sum_i \eta_i X_i\right) \tag{4.42}$$

$$\log K = \log\theta + \eta_0 + \sum_i \eta_i X_i \tag{4.43}$$

Equation (4.42) represents the nonlinear fixed-effects of X={Age, Gender, Endotype} on an arbitrary parameter, K. Taking the log of both sides yields Equation (4.43), which results in a linear model. The package *statsmodels* in Python was used to conduct linear regression for 57 parameters (the parameter for the initial condition of pathogen was excluded). The F-statistic of overall significance was calculated to reveal the significance of the η_i coefficients. A p-value$\leq$ 0.05 cutoff was used to determine if a covariate was significantly better than an intercept-only model.

4.2 RESULTS

Figures 4.2-4.6 shows sample fits from each one of the five endotypes. The filtered states for IL-6 and IL-8 are not shown. For the subjects that had neutrophil data, monocyte upper bounds were calculated and represented in green dashed lines. The bottom row of each figure demonstrate the cytokines. For TNF and IL-6, there were often scale differences between the blood (red trajectories) and tissue (blue trajectories) concentrations. As a result, their units varied to ensure similar scaling. Blood and tissue y-axes were the left and right, respectively.

Figure 4.2 illustrates a subject from endotype 1, which was representative of a high magnitude pro-inflammatory and anti-inflammatory response followed by a fast decline. Large amounts of pathogen were estimated for this subject. LPS immediately recruited circulating neutrophils, which accounted for the fast dip during the first day. Cytokines rose slowly and peaked on the second day, which was the estimated time of clinical presentation. IL-6 and IL-10 declined sharply during this day, which was a characteristic feature of endoype 1. The fourth filtered signal of IL-6 reached the bone marrow during day 3, which recruited immature neutrophils and monocytes from the **Bone** compartment into the **Tissue**, causing the hour 80 peaks. At 5 days, IL-8 was no longer present in the body and enabled the accumulation of circulating neutrophils.

Figure 4.3 illustrates a subject from endotype 2, which was representative of a protracted anti-inflammatory response and a fast-acting pro-inflammatory response. This subject's neutrophils and monocytes were quickly depleted following an aggressive and fast pro-

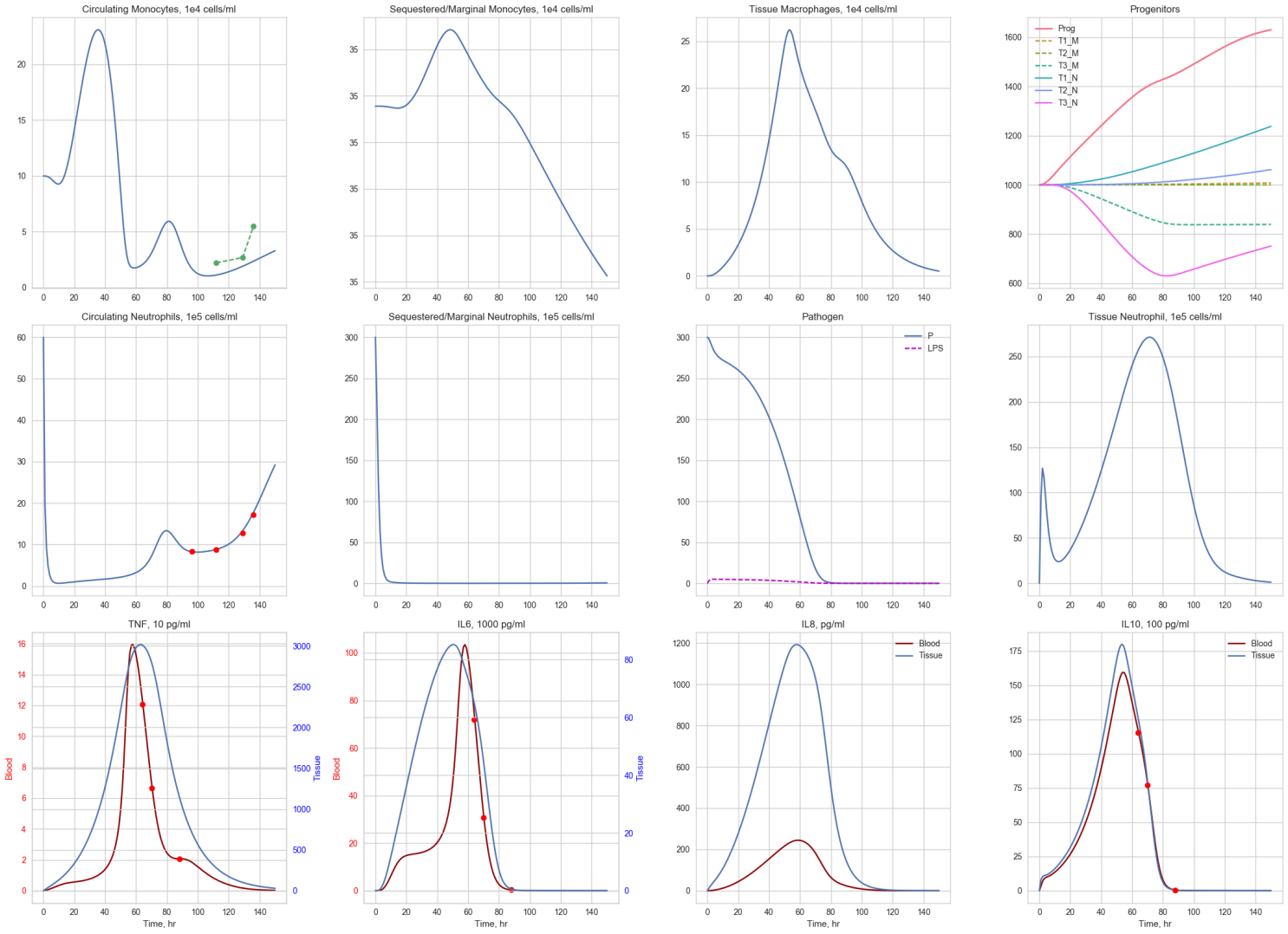


Figure 4.2: Sample fit for endotype 1, which was characterized by overwhelming inflammation. Subject 200039 is shown.

inflammatory response on the first day of infection. There was a sustained IL-10 response during the entire week, which indicated an immuno-suppressed state. This was further evidenced by a nonzero pathogen level by the end of the simulation.

Figure 4.4 illustrates a subject from endotype 3, which was representative of a rising-predominant anti-inflammatory response. This subject eliminated all pathogen by the 5th day, but the pro-inflammatory cytokines TNF, IL-6, and IL-8 were elevated during the entire

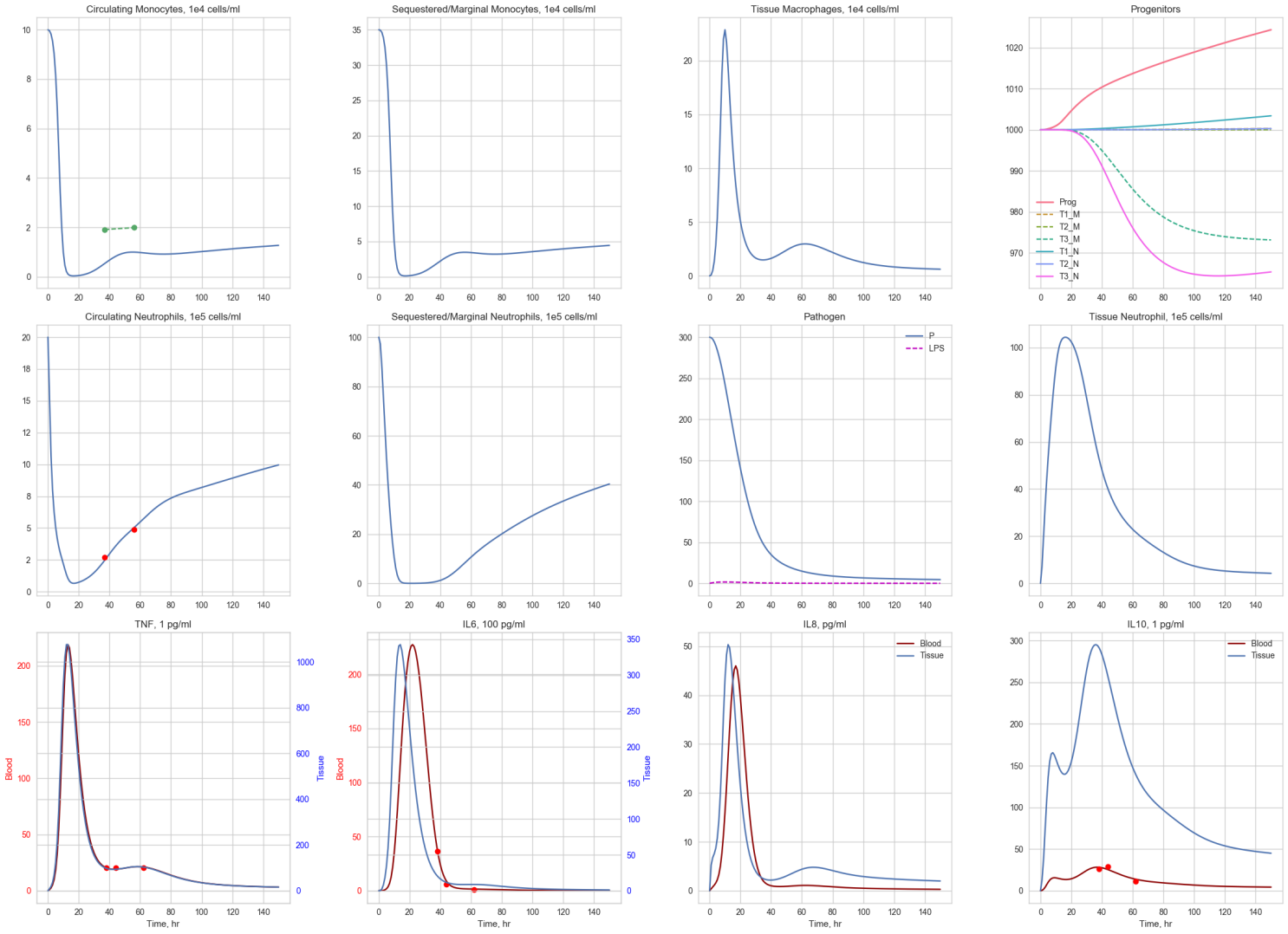


Figure 4.3: Sample fit for endotype 2, which was characterized by a protracted anti-inflammatory response. Subject 560054 is shown.

duration. IL-10 was sustained during this duration and caused the shark fin-like shape in the pro-inflammatory cytokines. The secondary IL-10 peak was caused by neutrophil apoptosis following the elimination of pathogens at hour 100. Finally, the lag between the steep decline of IL-6 at hour 100 (on day four) and the filter functions of IL-6 caused the circulating neutrophil dip around 110 hours, which ended immature neutrophil recruitment.

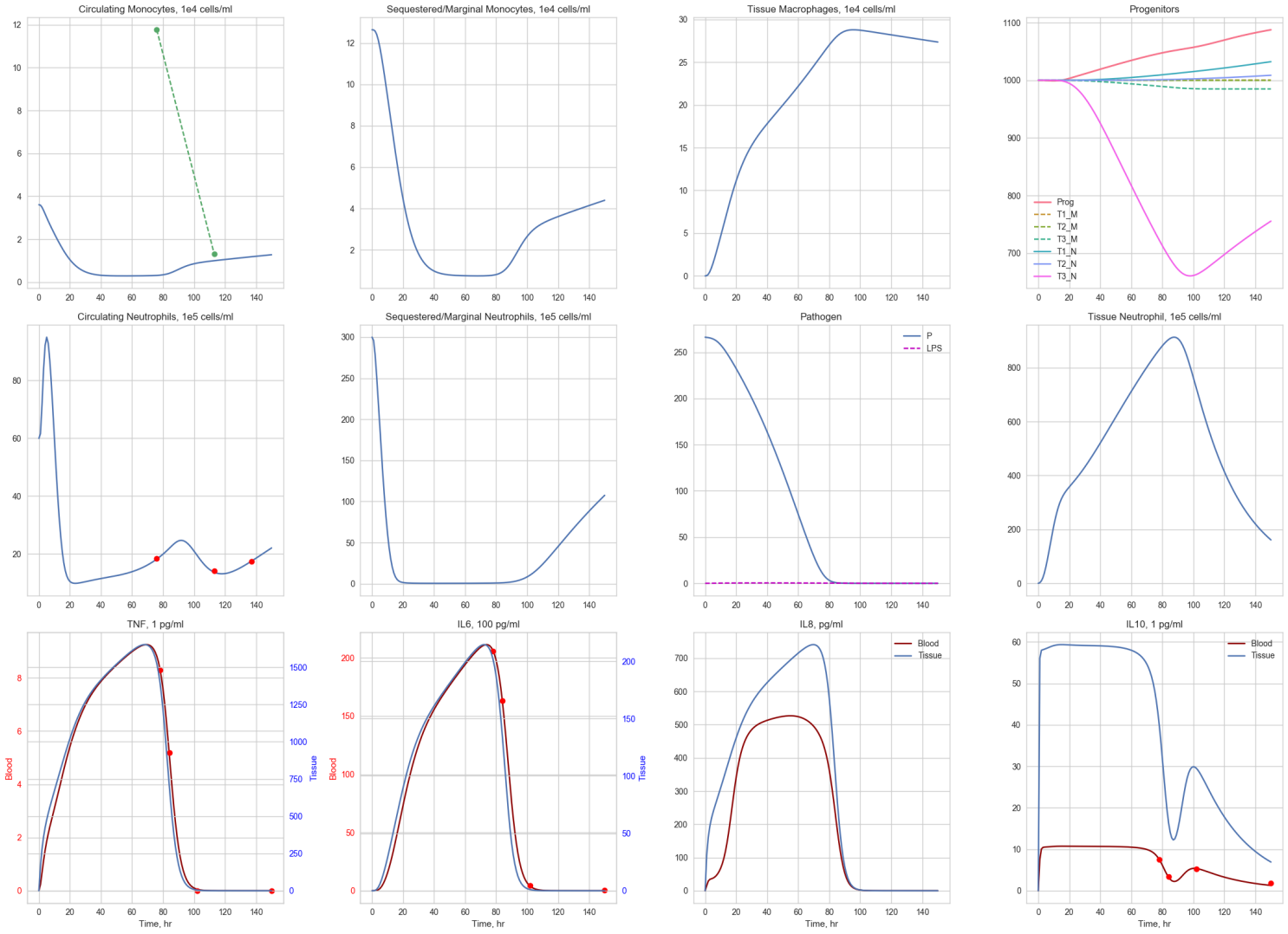


Figure 4.4: Sample fit for endotype 3, which was characterized by immunosuppression. Subject 210085 is shown.

Figure 4.5 illustrates a subject from endotype 4, which was representative of a sustained inflammatory response and an ineffective anti-inflammatory response. The TNF and IL-10 response in the **Tissue** rises quickly and remains elevated until pathogen elimination at 100 hours. However, the IL-6 presence remained elevated for days after pathogen elimination. The sustained inflammatory response caused rapid depletion of neutrophils and monocytes in both the circulating and marginal compartments. The unique aspect of this subject was

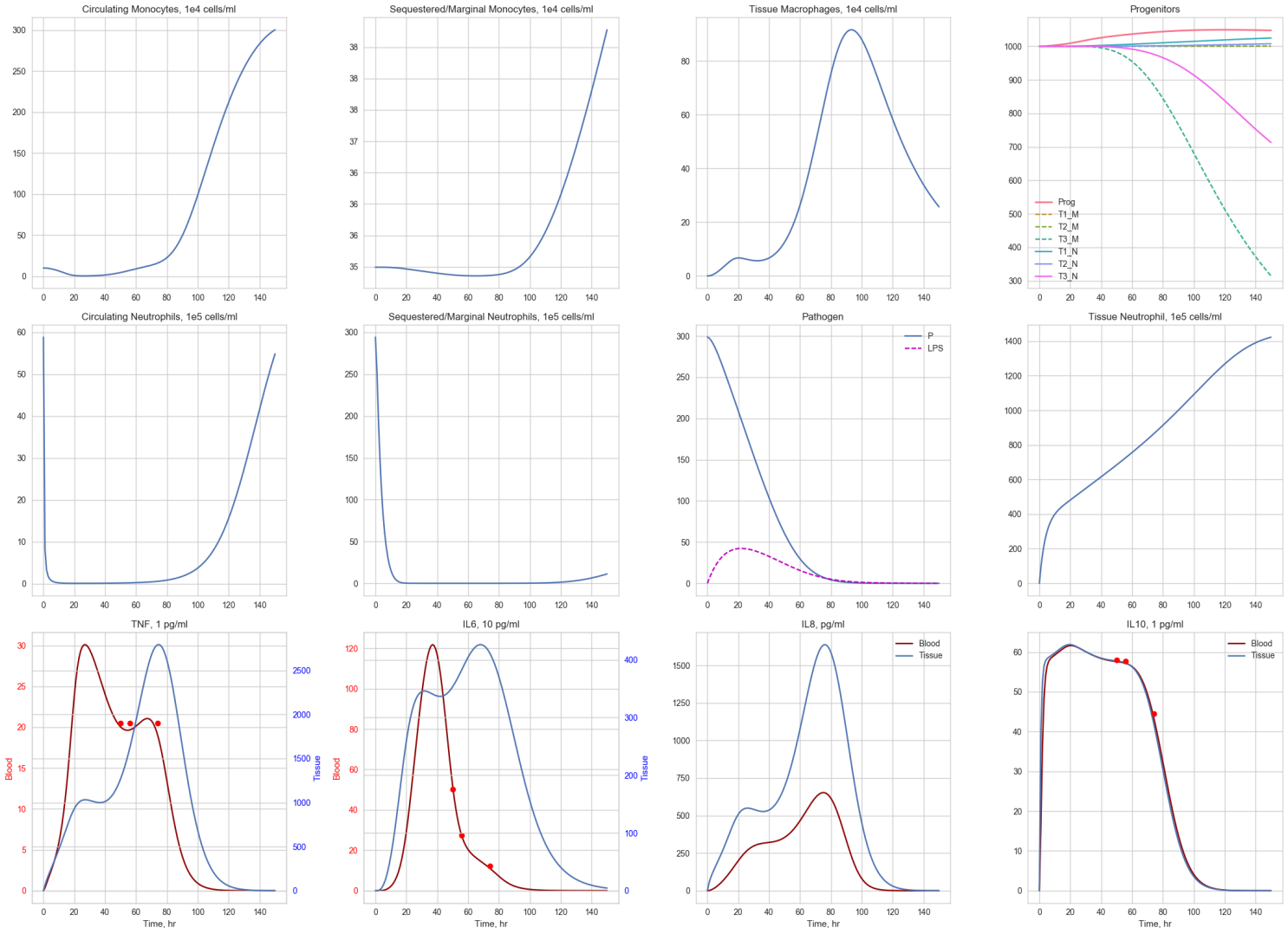


Figure 4.5: Sample fit for endotype 4, which was characterized by sustained inflammation. Subject 360025 is shown.

that the pathogen generated a large number of LPS, which also contributed to neutrophil depletion. Within one week, the values of T_{3_M} and T_{3_N} significantly declined. Due to the sustained pro-inflammatory response, neutrophils were recruited into tissue after pathogen elimination.

Figure 4.6 illustrates a subject from endotype 5, which was representative of a quick rise and fall of both pro- and anti-inflammatory responses. The TNF, IL-6, and IL-8 responses

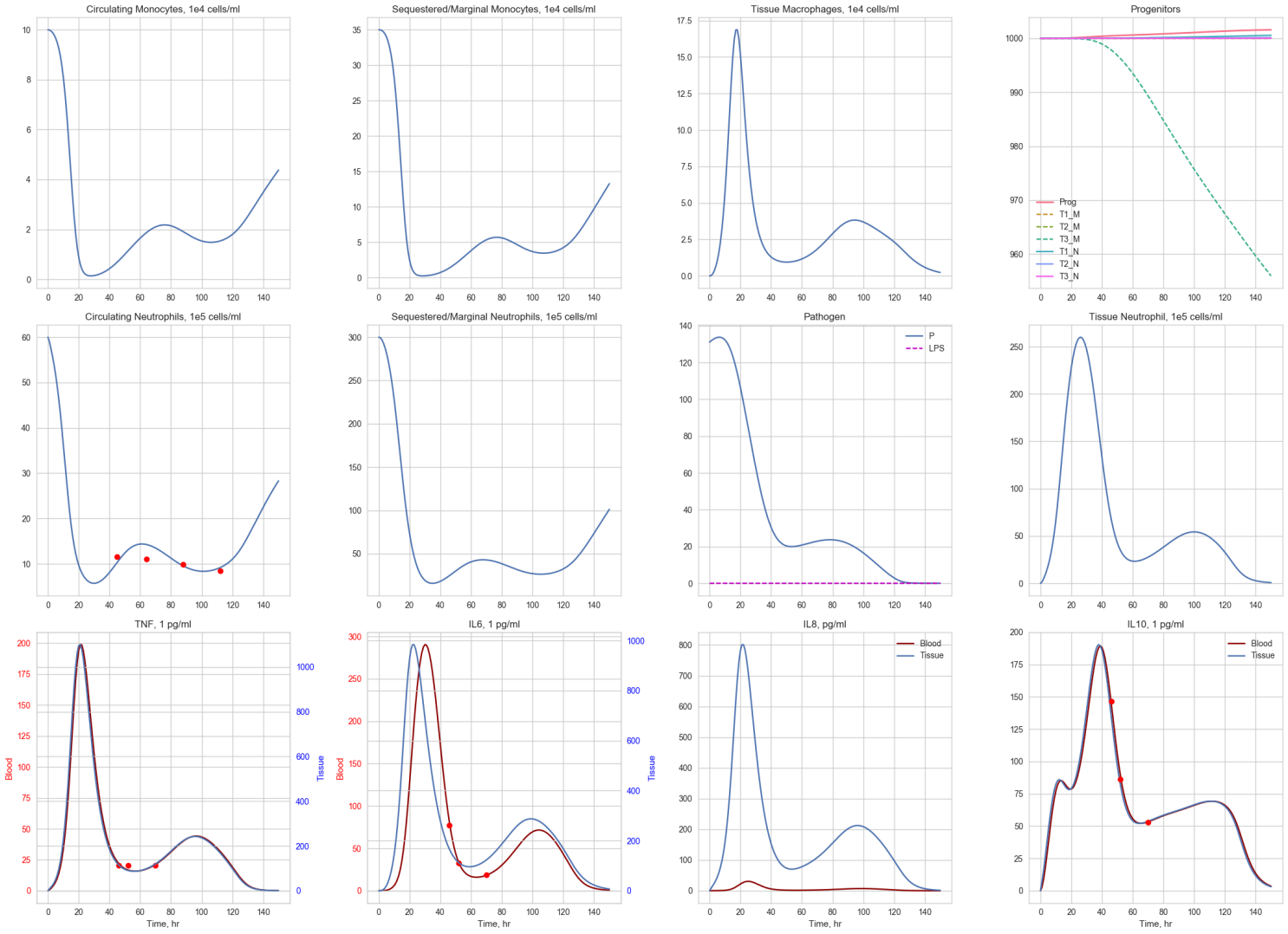


Figure 4.6: Sample fit for endotype 5, which was characterized by a quick and short inflammatory response. Subject 420021 is shown.

for this subject comprised of a fast up-down response and reach a nadir by hour 60 (on day 2). The body detected a rise in pathogen around this time and a secondary response occurred. As a result, full pathogen elimination occurred late (on day 5) compared to the other endotypes (with the exception of Figure 4.3 representing a subject from endotype 2). IL-10 dipped around hour 20 due to the decrease in pathogens. The subsequent IL-10 rise was caused by the neutrophil apoptosis mechanism (due to elevated levels of tissue neutrophils).

For this subject, IL-10 appeared to be heavily influenced by the apoptosis mechanism, which indicated a preference towards immuno-suppression.

The Anderson-Darling statistical test revealed two parameters to be different between endotypes. The first parameter was μ_{m_t}, which governed macrophage cell death rate (which could be accelerated by phagocytosis). Pair-wise testing revealed that μ_{m_t} was significantly different between endotypes 4 and 5 versus 1, 2, and 3. Furthermore, it was also revealed that μ_{m_t} within endotype 5 was different from 1, 2, 3, and 4. The second parameter was μ_{10_b}, which governed the decay of IL-10 in the blood. Pair-wise testing revealed that this value was significantly different between endotypes 1 and 2 versus endotypes 3, 4, and 5.

The covariate model revealed several parameters of interest. First, endotypes had a significant effect on the parameters k_{mtp}, μ_{10_b}, and μ_{t_b}. The parameter k_{mtp} is the half-max term of the Michaelis-Menten term describing accelerated macrophage death due to phagocytosis. Specifically, the relevant term is $1 + MacroP/(k_{mtp} + P)$. The parameter μ_{t_b} describes the natural decay rate of TNF within the **Blood** compartment. Second, age had a significant effect on the parameters K_{nr} and $K_{Tconsume}$. The parameter K_{nr} is the half-max term of the Michaelis-Menten kinetic that governs immature neutrophil recruitment from the bone compartment. The parameter $K_{Tconsume}$ governs the mechanism of neutrophil activation by blood TNF. Finally, age had a significant effect on the parameters d_t, μ_{t_b}, and $K_{Tconsume}$.

4.3 DISCUSSION

Figure 4.7 shows the estimated distributions for the four parameters that were found to be significantly different among endotypes. Parameter k_{mtp} demonstrated bimodal peaks. Endotypes 1 and 2 were predominantly distributed at low values for k_{mtp}, which resulted in a fast rate of macrophage death. These two endotypes were characteristic of high cytokinemia. One possibility was that overproduction of cytokines caused macrophage death. Alternatively, the overproduction of cytokines may have been a response to macrophage death: more TNF was needed to replace the dying macrophages to sustain the inflammatory effort. Endotypes

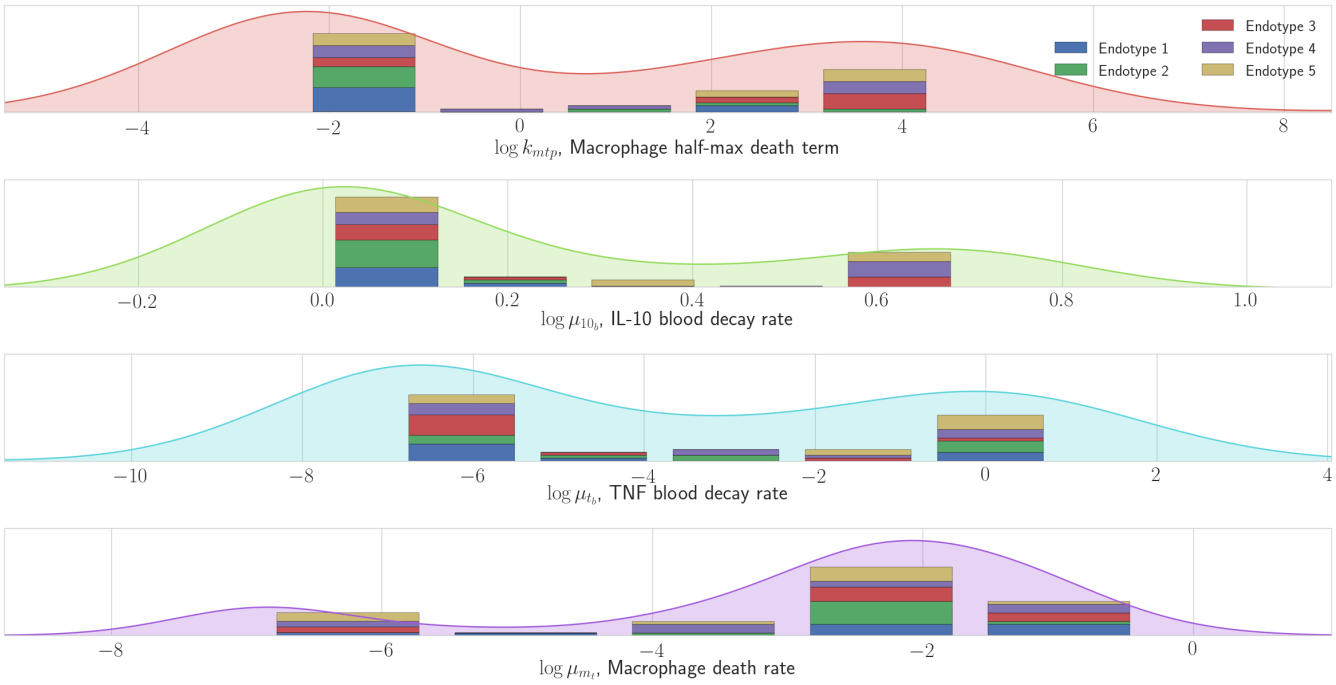


Figure 4.7: Distribution of significant parameters among endotypes. Stacked histograms, color coded by endotype, are shown. A Gaussian kernel density estimation (shaded trajectory) within each panel is shown. X-axis represents the log value of the parameter. Y-axis is not shown because the Gaussian densities are not normalized and scaled to encompass the histograms.

4, and 5 were distributed between high and low values of k_{mtp}. Most subjects within endotype 3 consisted of high values of k_{mtp}, which resulted in a slower macrophage death rate. Slower macrophage death resulted in a smaller TNF signal because monocyte recruitment was unnecessary given the large presence of macrophages in tissue. The presence of large numbers of macrophages may have also contributed to the resolution of inflammation (and therefore a predominant anti-inflammatory state) via the neutrophil apoptosis mechanism.

Parameter μ_{10_b} demonstrated a similar bimodal behavior: endotypes 1 and 2 preferred low values and endotypes 3, 4, and 5, were distributed between high and low values. Slow IL-10 degradation may partially explain the high cytokinemia response in endotypes 1 and 2.

Parameter μ_{t_b} had bimodal peaks as well. Endotype 3 preferred low values of the TNF blood decay parameter while the other endotypes were fairly spread out. It is possible that

TNF signals are prolonged (due to longer half-life) within endotype 3 and the high anti-inflammatory response is caused by a compensatory mechanism to offset the TNF response. Subjects from endotype 1 also dominated low values of μ_{t_b}. TNF may have contributed to the high cytokinemia response that was characteristic of this endotype. A negative consequence of high cytokinemia and longer cytokine lifespans is the improper recruitment of neutrophils, misdirecting them into healthy tissue, which may explain the high organ failure rates associated with endotype 1 [9, 91].

Parameter μ_{m_t} was primarily distributed within the [-4,0] log values. However, several subjects from endotypes 3, 4, and 5 formed a smaller peak at lower values of μ_{m_t}, which indicated longer overall macrophage lifespans.

Two parameters involving macrophage death dynamics were identified to be significantly different among endotypes and were subject to further analysis. Figure 4.8 illustrates a joint density plot for log parameter values of k_{mtp} and μ_{m_t}. The majority of subjects had low values of k_{mtp}, which the lower bound set in Pyomo for that parameter during optimization. Two subjects from endotype 1 exhibited both low k_{mtp} and low μ_{m_t}, a combination that resulted in a decreased natural macrophage death rate, but increased macrophage death rate during inflammatory conditions. There was a cluster of subjects, mostly consisting of endotype 5, that exhibited high values of k_{mtp} and low values of μ_{m_t}. This behavior was associated with decreased natural macrophage death rate and decreased inflammatory-driven death. The long lifespan macrophages precluded the need for a sustained inflammatory signal and may explain the fast decline of cytokines in endotype 5. For high values of μ_{m_t}, there were two regions. The region with high k_{mtp} and high μ_{m_t} consisted of subjects from every endotype. This parameter space described a high macrophages natural death rate and a low inflammatory death rate. Endotype 3 had a greater presence in this area. The region with low k_{mtp} and high μ_{m_t} also consisted of subjects from every endotype. However, endotypes 1 and 2 was almost located within this region. This parameter space described fast dying macrophages: both the natural death rate and the inflammatory death rate were high.

Figure 4.9 illustrates the joint density plot for log parameter values of μ_{10_b} and μ_{t_b}. This plot describes the parameter space concerning cytokine decay for blood IL-10 and TNF. Endotypes 1 and 2 were located exclusively in regions of low μ_{10_b} values, which may be

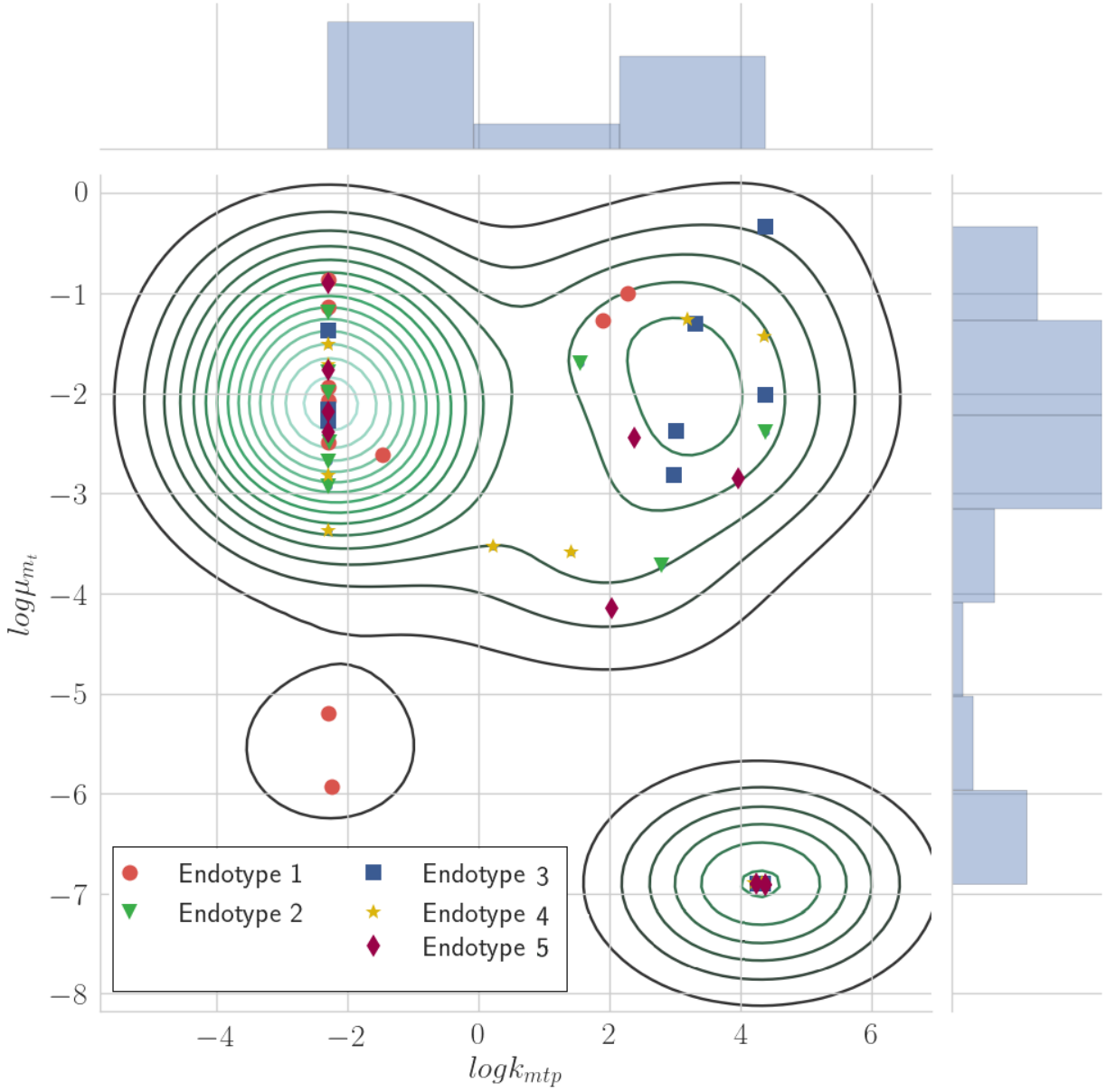


Figure 4.8: Joint distribution plot of significant macrophage parameters. X-axis is the log of k_{mtp}, the half-max parameter in Michaelis-Menten term that accelerates macrophage death. Y-axis is the log of μ_{m_t}, the non-inflammatory death rate of macrophages. Histograms of the parameters are shown at the top and right edge of the figure. Contour lines represent Gaussian kernel estimates. Scatter plots of parameter locations for each color-coded endotypes are shown.

responsible for the high cytokinemia in those groups. Endotype 3 was centered around low μ_{10_b} and low μ_{t_b}, a parameter region that resulted in the slow decay of IL-10 and TNF.

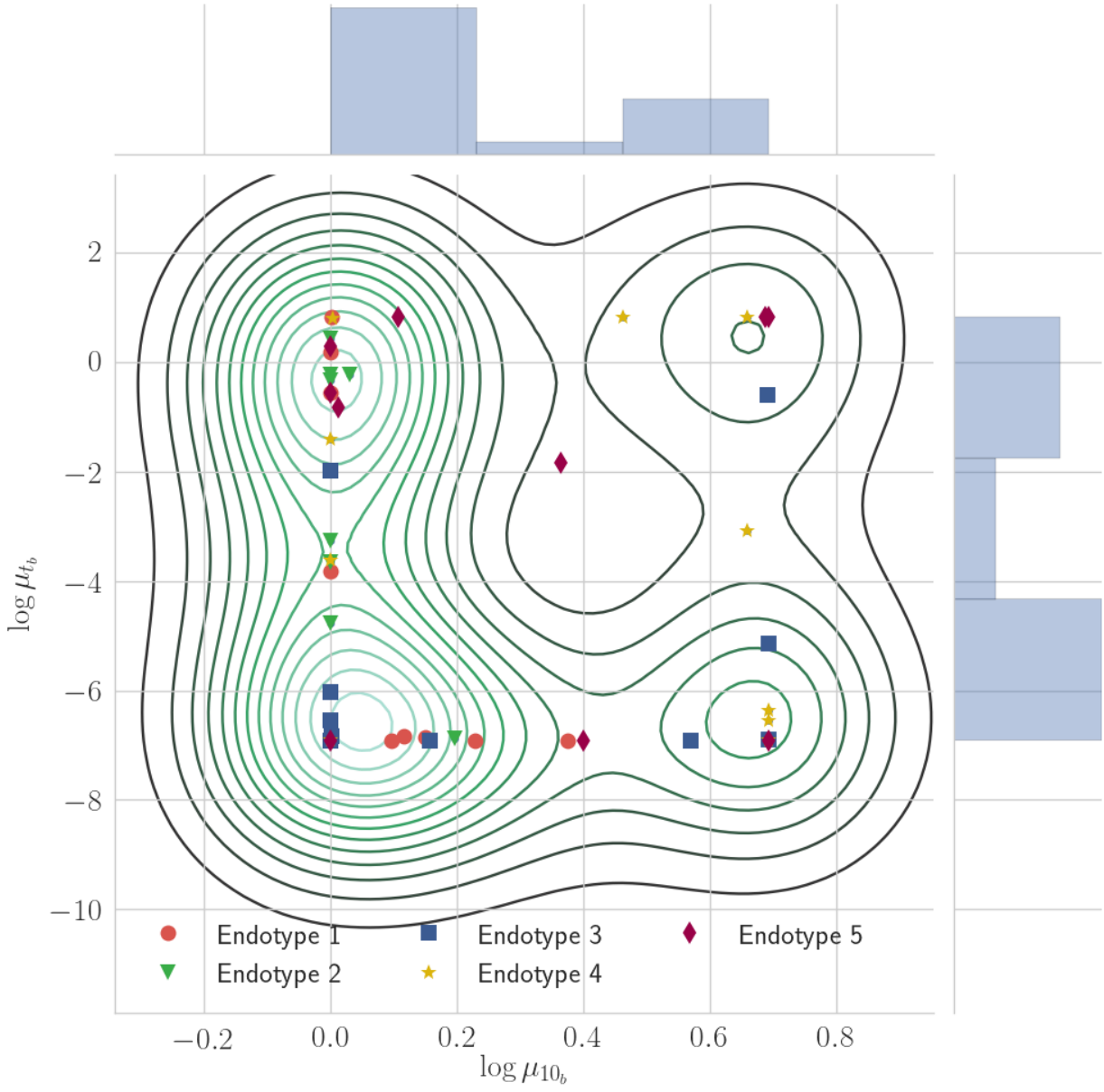


Figure 4.9: Joint distribution plot of significant cytokine decay parameters. X-axis is the log of k_{10_b}, the decay of blood IL-10. Y-axis is the log of μ_{t_b}, the decay of blood TNF. Histograms of the parameters are shown at the top and right edge of the figure. Contour lines represent Gaussian kernel estimates. Scatter plots of parameter locations for each color-coded endotypes are shown.

Endotype 4 was centered in regions of high μ_{10_b} (fast IL-10 decay), which supported the hypothesis in Chapter 3 that it represented a sustained inflammatory response. Endotype 5 was spread out across the entire distribution and did not appear to have a central location.

4.3.1 Study Limitations

The 390 subjects from the ProCESS trial offered many challenges in parameter fitting: fitting initial conditions, unknown pre-hospital time, and the lack of dense temporal data. White blood cell measurements were rarely collected in the 390 patients. Furthermore, IL-8, a crucial driver of neutrophil dynamics, was not measured. These deficiencies limit the translatability of these results into the clinic.

The lack of dense temporal data resulted in high parameter confidence intervals. Profile likelihood, a popular method in systems biology to determine parameter identifiability and confidence intervals, could not be applied to the existing model [104]. A profile likelihood simulation was run but did not finish by the end of two months. The combination of the lack of dense fitting data and the complexity of the model caused non-identifiability problems and extremely large confidence intervals, rendering profile likelihood, as a method, infeasible for this application.

Mathematical insights from sepsis ODE models cannot be obtained without dense temporal data. There are numerous difficulties associated with obtaining human sepsis data, but the results in this Chapter suggests that data collection may be worth the resources and effort.

4.4 SUMMARY

A mathematical model of sepsis was developed and fitted to human sepsis data from the ProCESS trial. The 390 subjects from Chapter 3 were individually fit to the model and an omnibus post-hoc analysis was conducted to identify population-level differences in parameters between endotypes. Macrophage cell death was an important distinguishing fac-

tor between endotypes because two parameters affecting the death rates of macrophages were deemed statistically significant. Furthermore, cytokines IL-10 and TNF had endotype-specific rates of decay, which may explain why some septic patients have high cytokinemia while others have low cytokinemia. These mechanistic insights begin to shed light on the innate differences between sepsis endotypes and the causes of variability in the syndrome. Such insights may also provide targets for pharmaceutical therapy to address the underlying cause of inflammatory dysregulation in sepsis rather than treat the symptoms of it.

5.0 IDENTIFICATION OF TRUE SEPSIS TIME-ZERO AND QUANTIFYING SEPSIS-INDUCED DAMAGE

5.1 PRE-HOSPITAL TIME IDENTIFICATION

Chapter 3 lacked to the ability to identify pre-hospital times for sepsis subjects in the clinic. Such an ability can improve clinical outcomes and significantly change how sepsis therapies are developed. Because of the dynamic nature of the physiologic response to infection, both the nature and the timings of potential interventions are likely to be determinant factors in influencing outcome [3, 105]. For example, an abrogation of the early TNF response increased mortality in some animal models, while most pre-clinical treatment models showed benefit [106, 107]. This may explain, at least in part, why many sepsis therapies that showed promise from animal models, where timing is known, failed in human studies.

In addition to clinical benefits of a pre-hospital time estimation tool, mathematical models of sepsis can significantly improve and reveal valuable mechanistic insight. Several mathematical and statistical models have been posed to elucidate the fast-acting dynamics of sepsis and to offer predictions regarding the potential effects of interventions and their timing [3, 4, 51, 69, 70, 108]. However, training and fitting parameters of these models for individual patients is challenging, in part due to the aforementioned timing issues with human data collection. For the purposes of such models, a population mean is typically computed from data pooled at time points relative to the time of enrollment. Naïve pooling becomes a problem because these data points are located at various points along temporal trajectories of individual patients. Methodological obstacles in developing robust models, such as inter-individual timing and variability and response, combined with a lack of familiarity of the

research community with such computational tools have delayed the introduction of such advanced tools as core to the design of clinical trials of sepsis.

Identifying the time of onset of infection offers two advantages in human sepsis research. First, it potentially enables the revisiting of previously failed trials with the purpose of analyzing, a posteriori, possible relationships between the elapsed time from onset and effectiveness. Second, identifying onset time enables patient biomarker data to be shifted relative to the time of infection, therefore, allowing more effective translation between animal results and human expectations. Further, mathematical models can be properly trained and provide more accurate predictions if mechanistically-based, biomarker-driven interventions are contemplated.

There are currently no proposed methods to identify the time of onset of an inflammatory challenge for sepsis or non-sepsis data. This motivated the development of a tool to estimate pre-hospital times with the requirements of (i) estimation generation within a few hours of hospital admission and (ii) the use of commonly measured clinical features.

5.1.1 Methods

A baboon sepsis dataset was retrospectively analyzed for this work [109]. The original experimental design was to investigate the therapeutic effects of a nitric oxide synthase (NOS) inhibitor of septic shock. Thirty-three baboons of the species *Papio ursinus* were sedated and $2 * 10^9$ colony forming units/kg of *Escherichia coli* (*E. Coli*) were infused intravenously into each subject over two hours. Fluid resuscitation and antibiotic therapy were provided to all subjects throughout the experiment. The proposed NOS inhibitor treatment began after hour 12 on sixteen subjects. Animals were treated in accordance with National Institutes of Health guidelines. The experimental protocol was reviewed and approved by the Institutional Animal Care and Use Committee of Biocon Research Laboratories, Pretoria, South Africa. 73 biomarkers were obtained as time series for each baboon including vital signs, arterial blood gases and lactate, hemodynamic parameters, complete blood counts and differential, and biochemistry. Baseline measurements were taken 30 minutes prior to *E. Coli* infusion. Additional measurements were collected at specified times throughout the

experiment. Subjects that survived the experiment had final measurements taken prior to sacrifice. Biomarkers that were intermittently measured throughout the experiment were eliminated from the analysis. This reduced the number of biomarkers to 29, where measurements were available for all baboons at hours -0.5, 0, 0.5, 1, 2, 3, 4, 5, 6, 11, and 12, where hour 0 marked the beginning of the *E. Coli* infusion. This time point was considered the true time of onset of infection. Time points past 12 hours post-infection were not utilized in this analysis. Baboons in the sham and treatment groups were combined for analysis. The biomarkers evaluated herein are listed in Table 5.1.

For validation of the method, two pig datasets were used. The first dataset contained 14 pigs that were subjected to surgically induced peritonitis, of which 7 subjects received a super-oxide dismutase treatment. Measurements were collected at 2, 4, 6, 8, 10, 12 hours after abdominal closure. Both groups were combined during analysis because trajectories between groups were not significantly different. The second dataset included 22 pigs subjected to one hour infusions: 12 subjects received 1 μg/kg/h and 10 received 10 μg/kg/h. Half of the animals received biliverdin. Measurements were collected at 0, 1, 2, 3, 4, 5, 7 hours after the start of infusion. Despite significant differences in trajectories between *E. Coli* doses, all groups were combined during analysis to test the method's performance on a non-homogeneous database. In both porcine studies, baboon-comparable biomarkers were utilized.

5.1.1.1 One-Nearest-Neighbor

Figure 5.1 provides an illustration of the process to estimate the time of onset of infection for a given subject. The temporal trajectories of biomarkers for this subject were left censored at all possible time points to simulate prospective monitoring at any given moment along the subject's sepsis trajectory. Trajectories were right censored to 1,2,3, or 4 consecutive points to simulate a time period of monitoring. The one-nearest-neighbor method compared the subject's sub-trajectory against a database of equal length from the remaining 32 subjects, identified the most similar sub-trajectory, and assigned the time of onset as the known time of onset for the most similar sub-trajectory.

Table 5.1: Dictionary of biomarker acronyms. Acronyms used for physiological measurements obtained longitudinally in the animal experiments, the means through which these measurements were obtained, and whether these means are considered invasive, minimally invasive, or noninvasive. Column three indicates how the biomarker was measured within the baboon data. Column four indicates the feasibility of measuring the biomarker in patients within a clinical setting.

Acronym	**Meaning**	**Method of Measurement**	
WBC	White blood Cell count	Arterial and mixed venous blood sample	M
HR	Heart Rate	Straightforward	N
HCO3A	Bicarbonate	Arterial blood gas analysis	M
SVR	Systemic vascular resistance	Arterial catheter	I
HB	Hemoglobin	Arterial and mixed venous blood sample	M
CO	Cardiac output	Thermal-dilution technique with Swan-Ganz catheter	I[a]
CI	Cardiac index	Calculated from CO and body surface area	N
Cc_{O_2}	Capillary oxygen content	Calculated from AP_{O_2}	I
ABEA	Arterial base excess	Calculated from HCO3A and pHA	M
MAP	Systemic arterial pressure	Arterial catheter	N[b]
PVR	Pulmonary vascular resistance	Arterial catheter	I
Pa_{O_2}	Arterial oxygen tension	Arterial blood gas analysis	M
RBC	Red blood cell count	Arterial and mixed venous blood sample	M
Ca_{O_2}	Arterial oxygen content	Calculated from arterial and mixed venous blood sample	I
PLT	Platelet count	Arterial and mixed venous blood sample	M
AP_{O_2}	Alveolar oxygen tension	Arterial catheter	M
HCT	Hematocrit	Arterial and mixed venous blood sample	M
PWP	Pulmonary wedge pressure	Pulmonary artery catheter	I
RAP	Central venous pressure	Arterial catheter	I
Pa_{CO_2}	Arterial carbon dioxide tension	Arterial blood gas analysis	M
TEMP	central blood temperature	Swan-Ganz catheter	N[b]
O_2DEL	Oxygen delivery	Calculated from respirometry and AaD_{O_2}	I
$SATA_{O_2}$	Arterial oxygen saturation	Calculated from arterial blood sample	N[b]
MPAP	Mean pulmonary artery pressure	Arterial catheter	I
RR	Respiratory rate	Straightforward	N
HOROW	Horowitz index	Calculated from Pa_{O_2} and fraction of inspired oxygen	M
QUOTIENT	Respiratory quotient	Respirometry	N
PHA	Arterial pH	Arterial blood gas analysis	M
AaD_{O_2}	Alveolar-arterial oxygen difference	Calculated from Pa_{O_2} and AP_{O_2}	I

N=noninvasive-to-measure, M=minimally invasive-to-measure, I=very invasive-to-measure.
[a]Cardiac output can be estimated via an ultrasound technique, but this technique is not widely adopted.
[b]Widely accepted noninvasive methods exist to obtain or closely estimate this value.

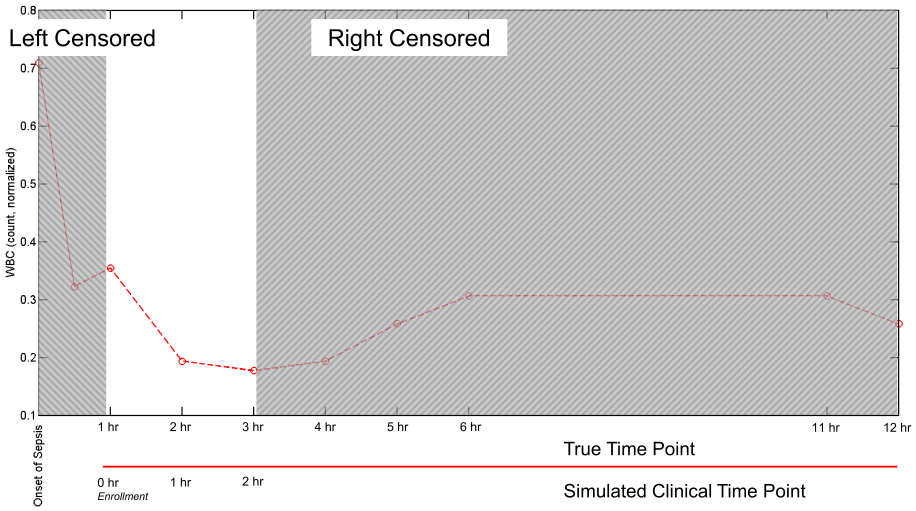


Figure 5.1: Schematic demonstrating data preparation process for testing of the pre-hospital time estimation tool. An entire white blood cell count trajectory (normalized to healthy baseline) is shown for a baboon and is censored prior to testing the nearest-neighbor method. Left censoring (left shaded area) was performed to simulate the passage of time between the onset of sepsis and the first measurement taken at simulated hospital enrollment. All subsequent time points were renumbered to simulate clinical time points where data is relative to hospital enrollment time. Right censoring (right shaded area) was performed to emulate sparseness of human data. In this case, measurements ended at 2 hours post enrollment for a total of 3 subsequent measurements (right censor level 3).

Suppose a three hour WBC data segment was available for a study subject. A Euclidean distance was calculated between this segment and all possible three consecutive WBC points in the database. Figure 5.2 conceptualizes this methodology. This process was repeated for each additional biomarker and their distances were summed. The three-hour length sub-trajectory in the database with the shortest distance to the study subject's was identified as the nearest-neighbor. The first time point of this sub-trajectory was used to estimate

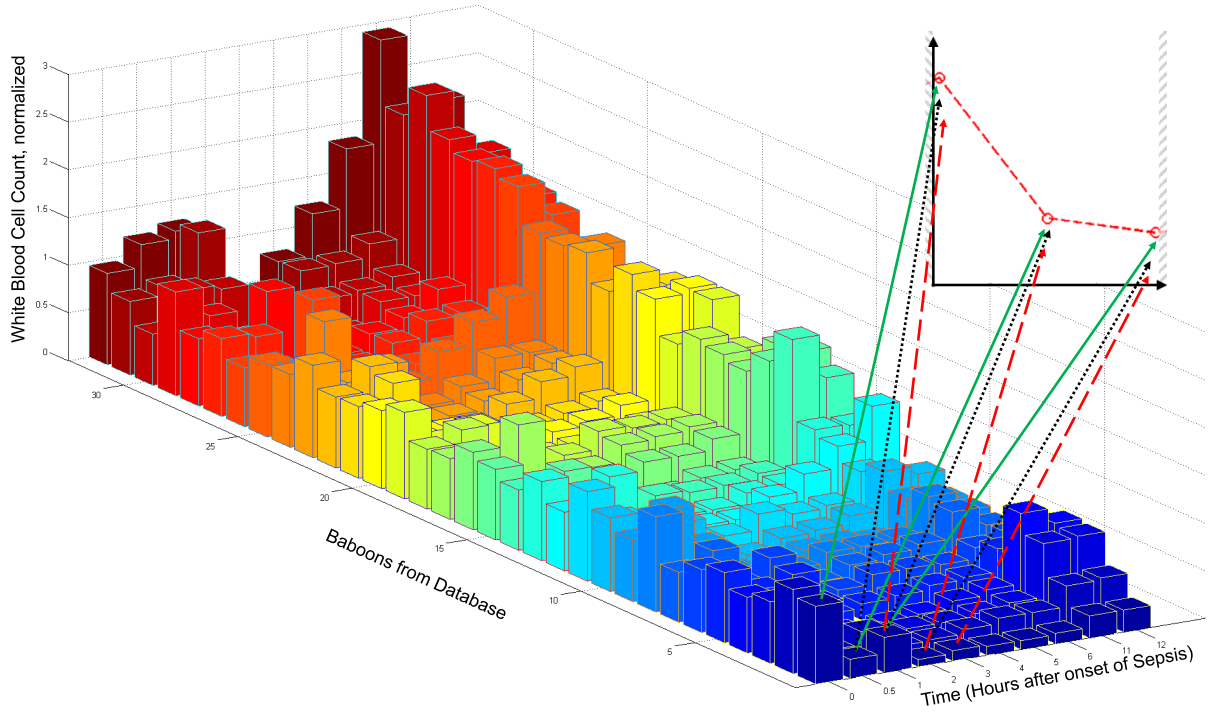


Figure 5.2: Schematic demonstrating the method of the pre-hospital time estimation tool. The normalized and censored white blood cell count (WBC) trajectory from Figure 1 was compared against the WBC database of the remaining 32 baboons (normalized to respective healthy baseline). The 3-point WBC trajectory was compared against every group of three sequential points in the database by calculating a Euclidean distance. Each type of line visualized represents one such comparison. If additional biomarkers were included for analysis, the Euclidean distances from all biomarkers were summed.

the study subject's elapsed time since infection. Using more than one nearest-neighbor was tested, but it did not improve the accuracy of results.

$$X_{baseline_i} = \beta_0 + \sum_{j=1,2,..N;j\neq i}^{N} \beta_j X_{baseline_j}) \tag{5.1}$$

To account for scaling differences across biomarkers and for inter-baboon variability, data were normalized on a per-baboon basis. Each subject's biomarker trajectories were normalized to its respective baseline (value at t=-0.5hr prior to infusion). However, the baseline values for the study subject were unknown due to left censoring. A linear regression model for estimating baseline was created for each biomarker using all time points in the database.

This is represented in Equation 5.1, where the baseline value for biomarker X_i estimated using available baseline measurements. The intercept, β_0, represents a population average baseline and the slope parameter, β_1, represents a subject-specific shift to this baseline. Once all missing baseline values were imputed, each study subject's biomarker trajectory was normalized by their specific baseline measurements.

To test estimation accuracy, we use leave-one-out validation. For each of the 33 subjects, every possible left censoring within the interval [0.5hr, 12hr] was tested to emulate a maximum of 9 possible "arrival" times. Estimations were generated for each of these cases and accuracy was determined by dividing the number of correct estimations by the total number of estimations generated. An estimated infection time within a tolerance of pm one-time point from the actual infection time was deemed correct.

5.1.1.2 In Silico Experiments and Validations A combinatorial search was performed in order to find the set of biomarkers that yielded the highest accuracy in estimating time of infection while minimizing the number of invasive clinical measurements required. The majority of the 29 biomarkers within the dataset were the result of invasive measurements and some of these are difficult to collect from human patients. To improve the clinical feasibility of the method, the search was performed involving single point measurements of minimally invasive biomarkers, i.e., blood samples and vital signs.

The first accuracy experiment tested the individual estimation capacity for the time of onset of each of the 29 biomarkers across various levels of right censoring. The best biomarkers were identified by calculating the mean accuracy across the four right censoring durations and selecting the top ten. Additionally, the null hypothesis was tested by making estimations based on randomly generated trajectories sampled from a zero mean lognormal distribution. The second accuracy experiment exhaustively searched all possible combinations of the previously identified top ten biomarkers. A combinatorial ${}_{10}C_n$ search, where $n \in 2, 3, ..., 10$, was performed to identify the best n combinations of biomarkers that identify infection time. This search was performed for each of the right censoring options.

The first feasibility experiment tested the ability of a mixture of time series biomarkers and single point measure biomarkers to estimate infection time. Vitals heart rate (HR),

mean arterial blood pressure (MAP), temperature (TEMP), oxygen saturation ($SATA_{O_2}$), and respiratory rate (RR) are noninvasive-to-measure and were made available in time series. Minimally invasive biomarkers white blood cell count (WBC), Pa_{O_2}, and platelet count (PLT) were chosen due to their diagnostic abilities as listed in the Surviving Sepsis Guidelines (1). Additionally, the top three minimally invasive biomarkers from the first accuracy experiment were included as well. All combinations of these vitals and minimally invasive biomarkers were tested for their infection time estimation accuracy. This experiment was repeated for right censoring at 2, 3, and 4 hours. The second feasibility experiment further explored this combination of time series biomarkers and single-point biomarkers by comparing two diagnostic panels that can be realistically performed at the time of patient enrollment. Two minimally invasive diagnostic panels were chosen: arterial blood gas test (yielding: arterial base excess [ABEA], Pa_{O_2}, arterial bicarbonate [HCO3A]) and blood analysis (yielding: WBC, PLT, HB). Similar to before, vitals were available in time series and this panel of biomarkers was available at the time of simulated enrollment for a given emulated patient. All combinations of the aforementioned vitals (except for RR) were tested with data from either or both diagnostic panels. RR was excluded because it was not among the top performers in the previous experiment. This experiment was repeated for right censoring at two, three, and four hours.

For validation, the first accuracy experiment (single biomarker search) and the first feasibility experiment (vitals+1 search) were repeated on each of the pig datasets. The biomarkers used in those experiments were selected to be comparable to those of the baboons'.

5.1.2 Results

Table 5.2 shows the infection time estimation accuracy of individual biomarkers. Only the top 15 best-performing biomarkers are shown in addition to the null hypothesis test. The entire table displaying the results for all tested biomarkers are available in Appendix A1. The first right censor duration tested only a single time measurement and yielded low accuracies throughout the table. Accuracy increased as more of the study subject's temporal points were included in the search. Biomarkers were sorted by decreasing mean accuracy across

Table 5.2: Time-of-infection estimation accuracy over varying right censor values (temporal durations).

	Baboon Accuracy					Pig Peritonitis Accuracy				Pig LPS Accuracy			
Biomarker	R 1	R 2	R 3	R 4	Mean	R 1	R 2	R 3	Mean	R 1	R 2	R 3	Mean
WBC	51.8%	67.0%	83.0%	93.1%	73.7%	60.7%	58.6%	76.8%	65.4%	46.5%	64.4%	73.4%	61.4%
HR	42.1%	58.2%	74.6%	84.8%	65.0%	64.6%	83.8%	87.0%	78.5%	47.2%	61.0%	69.1%	59.1%
HCO3A	48.5%	56.2%	67.8%	75.3%	62.0%	-	-	-	-	-	-	-	-
SVR	32.7%	59.3%	72.0%	79.7%	60.9%	50.0%	57.4%	63.0%	56.8%	50.7%	63.6%	66.0%	60.1%
HB	34.8%	58.2%	70.8%	79.2%	60.8%	53.6%	67.1%	83.9%	68.2%	52.8%	73.7%	84.0%	70.2%
CO	33.6%	54.5%	68.9%	84.8%	60.5%	46.3%	51.5%	64.8%	54.2%	45.8%	55.1%	63.8%	54.9%
CI	32.1%	51.9%	72.0%	84.4%	60.1%	-	-	-	-	-	-	-	-
Cc_{O_2}	33.9%	55.9%	70.8%	79.7%	60.1%	-	-	-	-	-	-	-	-
ABEA	44.8%	51.5%	65.9%	74.5%	59.2%	59.8%	73.5%	77.8%	70.4%	50.4%	63.9%	69.6%	61.3%
MAP	33.0%	50.8%	65.9%	82.7%	58.1%	43.9%	50.0%	61.1%	51.7%	53.5%	65.3%	77.7%	65.5%
PVR	38.5%	52.5%	66.7%	74.5%	58.0%	-	-	-	-	-	-	-	-
Pa_{O_2}	45.2%	49.5%	63.6%	70.6%	57.2%	53.7%	51.5%	64.8%	56.6%	47.2%	49.2%	61.7%	52.7%
RBC	29.7%	49.8%	67.4%	78.8%	56.4%	52.4%	74.3%	87.5%	71.4%	54.2%	73.7%	83.0%	70.3%
CA_{O_2}	30.9%	51.2%	64.0%	77.1%	55.8%	51.2%	50.0%	57.4%	52.9%	47.9%	56.8%	57.4%	54.0%
PLT	50.9%	49.2%	56.1%	60.6%	54.2%	61.9%	70.0%	82.1%	71.3%	41.5%	50.0%	61.7%	51.1%
Mean Null	26.6%	29.0%	32.3%	36.7%	31.2%	38.1%	44.7%	55.0%	45.9%	42.7%	49.1%	57.7%	49.8%

Results were generated with available biomarkers from the baboon data and then sorted by mean accuracy. The top 15 performing biomarkers are shown in addition to the null hypothesis (Full Table is available in Appendix A1). The null hypothesis, tested with randomized biomarkers, performed worse than all of the tested biomarkers. This method was repeated on the two porcine data sets for validation. Only 3 right censor durations were tested due to the sampling rate and duration limitations of those experiments.

the four right censor durations. WBC was the top biomarker in all columns and yielded a maximum estimation accuracy of 93.1% when using four sequential hourly points. Jackknife resampling of the data revealed that the standard deviation of accuracies across biomarkers and right censor durations had a mean of 0.7%. No entry within Table 5.2 fell below their respective right censor level null accuracy. The mean null hypothesis accuracy increased with higher right censor levels due to the shrinking of estimation possibilities.

All possible combinations of the top ten biomarkers from Table 5.2 were tested to maximize the accuracy at each level of right censorship. The top result for each right censor level is shown in Table 5.3. Remarkably, single time point measurements of six biomarkers yielded roughly 70% estimation accuracy. WBC was selected in all cases. There was a strong preference given to arterial blood gas measurements, hemoglobin, and cardiovascular measurements.

5.1.2.1 Feasibility Experiments

The goal of the feasibility experiments was to identify a parsimonious set of biomarkers, in

Table 5.3: Multiple biomarker prediction accuracy. All possible n biomarker combinations of the top 10 (baboon) biomarkers from Table 5.2 were tested for their predictive accuracy. The best biomarker set is shown for each right censor duration under baboon accuracy. Validation of these biomarkers were performed on the porcine datasets. Arterial bicarbonate, cardiac index, and capillary oxygen content were unavailable for the porcine datasets and were respectively substituted by pH, cardiac output, and Pa_{O_2}.

		Accuracy		
Right Censor Length	Biomarkers	Baboon	Pig Peritonitis	Pig LPS
1 time points	WBC HCO3A HR SVR CI ABEA	72.70%	71.95%	58.87%
2 time points	WBC HCO3A MAP HB CO Cc_{O_2}	85.90%	70.59%	72.41%
3 time points	WBC HR HCO3A MAP SVR HB	93.20%	87.04%	93.48%
4 time points	WBC HR MAP CO Cc_{O_2} HCO3A HB CI	97.80%	-	-

LPS = lipopolysaccharide, HCO3A = arterial bicarbonate, HR = heart rate, SVR = systemic vascular resistance, CI = cardiac index, MAP = mean arterial blood pressure, HB = hemoglobin, CO = cardiac output, Cc_{O_2} = capillary oxygen content.
Dashes indicate measurements not available in porcine experiments.

both time series and single-point measurements, which were minimally invasive to measure in humans and yielded a good accuracy in their ability to estimate the time of infection onset. HCO3A, hemoglobin (HB), and cardiac output (CO) were the top performing minimally invasive biomarkers in Table 5.2 and were included in the experiment. Table 5.4 shows the best results from each measurement collection (one minimally invasive biomarker plus a combination of vitals in time series), organized by right censor duration. For comparison, accuracies were calculated for each entry with vitals alone and with all biomarkers in time series. The inclusion of a point measurement had the greatest impact on accuracy for the lower right censor durations. A single value of WBC combined with two hours of HR data improved accuracy from 58.2% to 71.4%. HCO3A and WBC were consistently selected throughout the table. Point measures of biomarkers helped many entries achieve over 90% estimation accuracy.

Time series of invasive measurements did not improve accuracy, thus not warranting the extra probing of subjects. For example, two measurements of HR and WBC yielded an estimation accuracy of 81.1%. The same accuracy was achieved by using two measurements of MAP and HR along with a single measurement of HCO3A. Furthermore, the use of MAP, $SATA_{O_2}$, and HR + 1x HCO3A outperformed the majority of entries in each right censor category.

The second feasibility experiment tested the accuracy of using multiple single point biomarkers with time series vitals. The top three results from the combinatorial search are shown in Table 5.5 with results sorted based on accuracies from the "Both Panels" column. The information gained by administering both diagnostic panels aided in almost all entries of the 2 and 3-time point right censor levels, achieving maximum accuracies of 84.2% and 90.2%, respectively. Regardless, despite the added data from the diagnostic panels, many entries from Table 5.4 using a point biomarker yielded equal or higher accuracies.

5.1.2.2 Validation on Porcine Data

Table 5.2 shows single biomarker accuracies for both pig experiments. Accuracies were, in general, equivalent or higher than those of the baboons. Time of onset estimation differed slightly among the models. For example, HB, TEMP, and HCT performed better in both

Table 5.4: Prediction accuracy of longitudinal vital signs With a single blood biomarker.

		Accuracy			
Right Censor	Time Series Biomarkers	Point Biomarker	No Point	With Point	Time Series Point
2 time points	HR	WBC	58.2%	71.4%	81.1%
	HR	HCO3A	58.2%	70.0%	71.7%
	MAP	HCO3A	50.8%	69.4%	70.0%
	MAP HR	HCO3A	70.0%	81.1%	81.1%
	MAP HR	WBC	70.0%	77.8%	81.1%
	$SATA_{O_2}$ HR	WBC	60.3%	76.1%	81.8%
	MAP $SATA_{O_2}$ HR	HCO3A	76.8%	81.8%	81.5%
	TEMP MAP HR	HCO3A	73.4%	81.1%	81.1%
	TEMP MAP HR	PA_{O_2}	73.4%	79.1%	81.1%
	TEMP MAP $SATA_{O_2}$ HR	HCO3A	75.4%	81.8%	81.8%
	TEMP MAP $SATA_{O_2}$ HR	WBC	75.4%	80.1%	82.2%
	TEMP MAP $SATA_{O_2}$ HR	PA_{O_2}	75.4%	80.1%	81.8%
3 time points	HR	WBC	74.6%	76.9%	86.4%
	MAP	HCO3A	65.9%	75.8%	81.8%
	HR	HCO3A	74.6%	75.4%	78.4%
	MAP HR	HCO3A	87.5%	89.8%	92.0%
	MAP HR	HB	87.5%	89.0%	90.5%
	MAP HR	PA_{O_2}	87.5%	88.3%	92.8%
	MAP $SATA_{O_2}$ HR	HCO3A	88.3%	92.0%	92.8%
	MAP $SATA_{O_2}$ HR	PA_{O_2}	88.3%	91.3%	93.9%
	MAP $SATA_{O_2}$ HR	HB	88.3%	89.4%	91.7%
	TEMP MAP $SATA_{O_2}$ HR	HCO3A	87.5%	92.0%	92.8%
	TEMP MAP $SATA_{O_2}$ HR	PA_{O_2}	87.5%	90.5%	93.9%
	TEMP MAP $SATA_{O_2}$ HR	HB	87.5%	89.4%	91.3%
4 time points	MAP	HCO3A	82.7%	88.3%	89.2%
	MAP	WBC	82.7%	86.1%	94.4%
	HR	HCO3A	84.8%	85.7%	88.7%
	MAP HR	WBC	94.4%	94.8%	94.4%
	MAP HR	HCO3A	94.4%	94.8%	95.2%
	MAP HR	HB	94.4%	93.9%	94.4%
	MAP $SATA_{O_2}$ HR	HCO3A	95.2%	96.1%	96.5%
	TEMP MAP HR	WBC	93.9%	95.2%	93.9%
	MAP $SATA_{O_2}$ HR	WBC	95.2%	95.2%	94.8%
	TEMP MAP $SATA_{O_2}$ HR	WBC	95.7%	95.2%	94.8%
	TEMP MAP $SATA_{O_2}$ HR	HCO3A	95.7%	95.2%	96.5%
	TEMP MAP $SATA_{O_2}$ HR	PA_{O_2}	95.7%	95.2%	97.0%

HR = heart rate, HCO3A = arterial bicarbonate, MAP = mean arterial blood pressure, $SATA_{O_2}$ = arterial oxygen saturation, TEMP = temperature, HB = hemoglobin.
Results from the first feasibility experiment where vital signs in time series were combined with single-point measurements (at the time of simulated enrollment) of minimally invasive biomarkers to estimate infection time. Vitals temperature, mean arterial blood pressure, arterial oxygen saturation, heart rate, respiratory rate, or in any combination thereof, was tested in conjunction with a single blood biomarker for their prediction accuracy. The top three biomarker sets for each combination and right censor duration are shown, with their accuracies listed in the "with point" column. "No point" shows the estimation accuracy for the vitals without the point biomarker. "Time series point" shows the estimation accuracy when all biomarkers (vitals and point) were available in time series.

Table 5.5: Prediction accuracy of longitudinal vitals With different diagnostic blood panels.

		Accuracy		
Right Censor	Time Series Biomarkers	Both Panel	Blood Panel	ABG Panel
2 time points	MAP	81.8%	74.7%	69.0%
	HR	79.5%	71.0%	70.4%
	TEMP	77.4%	65.0%	58.3%
	MAP HR	82.2%	78.8%	78.8%
	MAP $SATA_{O_2}$	77.8%	76.1%	70.0%
	TEMP MAP	77.4%	76.1%	70.0%
	TEMP MAP HR	82.8%	79.8%	78.5%
	MAP $SATA_{O_2}$ HR	82.2%	80.5%	78.5%
	TEMP MAP $SATA_{O_2}$	77.8%	77.1%	69.7%
	TEMP MAP $SATA_{O_2}$ HR	84.2%	81.1%	77.8%
3 time points	MAP	86.7%	77.7%	75.8%
	HR	85.6%	70.5%	75.4%
	TEMP	81.1%	62.2%	54.9%
	MAP HR	83.7%	84.5%	84.5%
	TEMP MAP	79.9%	78.4%	75.8%
	MAP $SATA_{O_2}$	79.6%	80.7%	76.9%
	TEMP MAP HR	83.7%	85.2%	84.8%
	MAP $SATA_{O_2}$ HR	83.7%	86.0%	85.2%
	TEMP MAP $SATA_{O_2}$	80.7%	80.3%	77.3%
	TEMP MAP $SATA_{O_2}$ HR	90.2%	86.7%	85.2%
4 time points	MAP	90.5%	84.8%	81.8%
	HR	89.2%	80.5%	78.8%
	$SATA_{O_2}$	87.5%	68.4%	64.5%
	MAP HR	89.6%	95.2%	88.7%
	MAP $SATA_{O_2}$	87.0%	87.0%	83.5%
	TEMP MAP	86.2%	85.3%	82.7%
	TEMP MAP HR	90.0%	94.8%	89.6%
	MAP $SATA_{O_2}$ HR	89.6%	95.7%	90.9%
	TEMP MAP $SATA_{O_2}$	87.5%	87.0%	83.5%
	TEMP MAP $SATA_{O_2}$ HR	93.5%	95.2%	90.9%

ABG = arterial blood gas, MAP = mean arterial blood pressure, HR = heart rate, TEMP = temperature, $SATA_{O_2}$ = arterial oxygen saturation. Results from the comparison of two types of diagnostic panels: arterial blood gas (arterial bicarbonate, Pa_{O_2}, and arterial base excess) and blood analysis (WBC, platelet count and hemoglobin). Vitals temperature, mean arterial blood pressure, arterial oxygen saturation, heart rate, or in any combination thereof were tested in conjunction with either or both diagnostic panels for their estimation accuracy. The top performing combinations are shown.

pig data sets. Alternatively, PLT performed similarly between the baboon and the pig LPS data but was more informative in the pig peritonitis data. This may be the result of the differences in sepsis induction protocols across the three models and suggestive towards the existence of sepsis endotypes characterized by distinctive biomarker trajectories.

The vitals+1 biomarker search on the porcine data revealed MAP, TEMP, and SATAO2 to provide highly accurate estimates (LPS: 70 to 90+%, peritonitis: 80 to 90+%) when used in conjunction with a single WBC or PHA measurement. Detailed porcine results from this search are available in Appendix A3-A2 .

5.1.3 Discussion

A one-nearest-neighbor approach was selected to tackle the problem of identifying infection time from left and right censored data. Serial measurements of non-invasive vital signs, when combined with a single minimally invasive, yet routinely done, blood work, yielded good accuracy to identify the time of onset of infection in an experimental baboon model of sepsis. The method was further confirmed in two additional animal models. The one-nearest-neighbor method was developed based on the hypothesis that cohorts of septic subjects exist with similar characteristic biological responses. The baboon study chosen for analysis represented a homogenous cohort that all exhibited leukopenia following the *E. Coli* infusion. Pig validation sets did not share this feature. Given a censored and normalized trajectory for a study baboon, one-nearest-neighbor identified the most similar trajectory within the database. The temporal information of this trajectory provided the time-of-infection estimate for the study baboon. Nearest-neighbor is a popular non-parametric approach in data mining and was selected here to find similarities among biomarker trajectories.

The exhaustive biomarker search revealed that certain biomarkers provide highly accurate estimations of time of onset of infection. WBC, HR, HB, MAP, and HCO3A were the top performers in Table 5.2 and had many appearances in Table 5.3. In contrast to CO or SVR, both of which requires the insertion of an arterial catheter, WBC, HR, HB, MAP, and HCO3A are relatively easy to measure. The modest accuracy increases from the inclusion of CO or SVR is of doubtful clinical significance and did not seem to justify the

invasive measurement. This suggested that more invasive to measure biomarkers might be unnecessary and that this time of infection onset can be estimated with easily-measured, patient-friendly biomarkers. It is also interesting to note the consistent appearance of many biomarkers across animal models.

Most clinicians do not perform consecutive hourly blood sample tests, and only noninvasive biomarkers are likely to be acquired in time series. Many biomarkers in the baboon data cannot be obtained hourly (or ever) in human patients due to practical reasons (no available commercial assay, slow turnovers) or to unjustifiable expenses. Patients with suspected sepsis upon enrollment sometimes have a panel of diagnostic tests administered, including arterial blood gas tests and a WBC measurement. These measurements are typically only taken once at the time of enrollment. In contrast, vital signs such as HR, MAP, TEMP, and oxygen saturation are continuously monitored and universally measured in time series. The feasibility experiments addressed these issues by using clinically-obtainable data from humans and yielded interesting results. Specifically, they showed that: (i) the addition of a minimally invasive point measurement generally improved time-of-infection accuracy over the use of time series vitals alone, (ii) taking minimally invasive measurements in time series may be unnecessary for estimating time-of-infection, and (iii) there existed a data saturation limit where one-nearest-neighbor did not benefit from additional data.

5.1.3.1 Study Limitations

The main limitation associated with this method was the small sample size of all three datasets. This method worked well for the heterogeneous pig datasets at least in part because of the uniformity of the insult within datasets, thus improving the chances of a relevant "closest looking" animal in the cohort. If the database does not contain sufficiently similar subjects to an incoming subject for which an estimation is to be made, the method might not be able to generate a meaningful estimate.

Three main challenges currently prevent this method from being directly translated to human patients. First, developing a human model requires access to a temporally rich collection of physiologic markers at baseline and following time-of-infection. The second challenge is human variability, demanding that the dataset is of sufficient size to be representative of

most endotypes. Third, human sepsis often happens in conjunction with other inflammatory stressors, such as surgery or trauma. To address these challenges, a human dataset of sufficient breadth could potentially be assembled in a cohort of patients developing sepsis in association with an invasive procedure, therefore bounding time of onset to within a few hours. The impact of the procedure itself on biomarkers time series would constitute useful additional information.

5.1.4 Translation into GLUE Grant Human Trauma Subjects

The Inflammation and the Host Response to Injury "Glue Grant" database was the result of a multi-center initiative to collect rich longitudinal data for trauma patients undergoing hemorrhagic shock (www.gluegrant.com). This database consists of 2007 adult subjects with known pre-hospital times (time of trauma was known, in minutes). While trauma is different from sepsis, the inflammatory dynamics occurring between the two are similar.

The nearest-neighbor method was validated on the Glue Grant database. The parallels between the syndromes of hemorrhagic shock in trauma and sepsis enable using trauma data as a human validation cohort. Given the aforementioned problems associated with collecting human sepsis data, trauma serves as a decent proxy. Figure 5.3 illustrates the distribution of pre-hospital times within the Glue Grant database. Correlation analysis yielded few clinical features with linear relationships to pre-hospital time. Subsequently, patients were binned into 10 discrete intervals of pre-hospital times and one-way non-parametric ANOVA (via Kruskal-Wallis tests) was conducted on all measured clinical features to identify nonlinear relationships with pre-hospital time.

Several biomarkers were identified to have significant differences between various pre-hospital interval groups ($p \leq 0.05$). Fluid resuscitation data was discovered to be significantly different between each group, but these measurements were discarded due to potential causal relationships with pre-hospital time. The remaining clinical features were systolic blood pressure, hypotension, diastolic blood pressure, heart rate, Glasgow coma score, lactate, respiratory rate, hemoglobin, and the INR (international normalization ratio, a measure of coagulopathy).

The one-nearest-neighbor method was tested on this dataset. Unfortunately, hourly measurements were not available in the Glue Grant. As a result, temporal trajectories of pre-hospital measurements (in-ambulance measurements) and emergency room measurements were used for analysis, if possible. Otherwise, biomarkers were used as point measurements. After removing missing values, a cohort of 946 subjects. A 60-40 training-test split revealed that the method estimated pre-hospital times correctly, to within 60 minutes, with a test set accuracy of 68.1%. Various modifications of the nearest-neighbor methods were attempted, ranging from filtering the nearest-neighbor database by injury severity score to modifying the list of predictive biomarkers. Accuracy to 75-82%, but at the cost of reduced cohorts sizes.

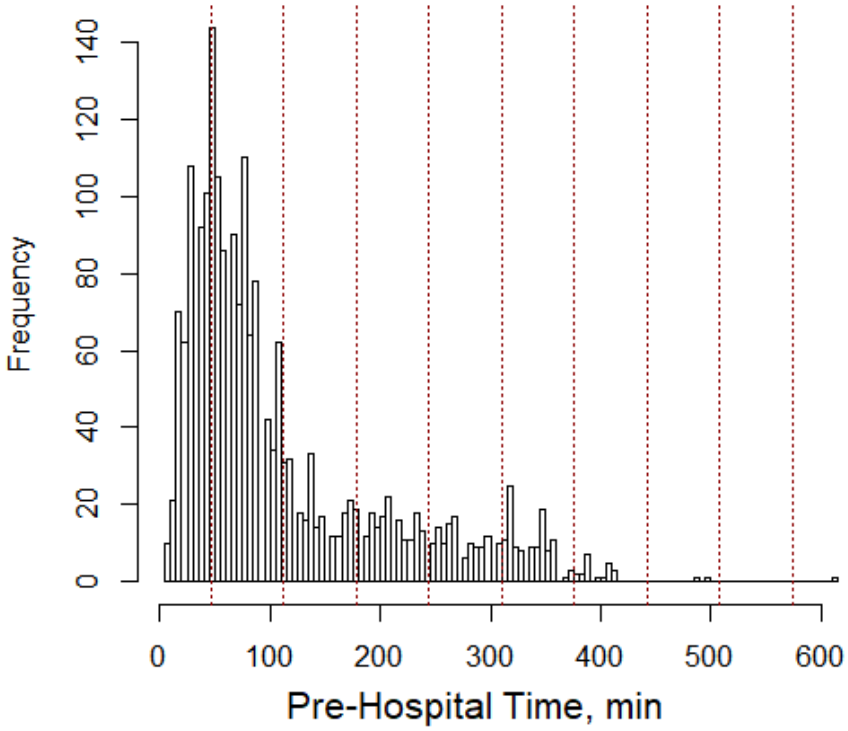


Figure 5.3: Histogram of pre-hospital times in the Glue Grant database. Vertical dashed lines represent each hour.

5.1.5 Summary

A novel application of the nearest-neighbors application was developed to estimate the onset time of infection (specifically, a patient pre-hospital time relative to onset) within artificially censored baboon and porcine sepsis data. High accuracies were achieved with varying sets of biomarkers, but some biomarkers are difficult to measure clinically in humans. A compromise was made between minimizing the invasiveness of measurement and maximizing the estimation accuracy. A feasible set of clinically-measurable biomarkers was identified; combinations of vital signs in time series and minimally invasive point measurements yielded similar but slightly lower accuracies.

The clinical tool presented in section 3.2.1 provided suggested an accurate early classification of a subject's endotype was possible, but such a tool lacked estimations of pre-hospital time. However, promising results from a human trauma database suggested that translation of the nearest neighbor approach may be viable for sepsis data, but such a temporally-rich sepsis database is not yet available.

There are no other approaches to estimating time-of-infection in sepsis, and the method developed here may have profound, paradigm-shifting implications if successfully extended for use with human data. Septic patients would be able to be grouped into early, middle, and late infection times and treated differently, which may play a clinically meaningful role in patient outcome. Clinical trials may be revisited or new therapeutic targets may emerge because treatment timings can be more effectively controlled with respect to the temporal span of infection. Finally, computational models of sepsis and immunomodulatory interventions used in the design of such trials may immediately benefit from the knowledge of infection times.

5.2 SYSTEMIC DAMAGE MODELS

Chapter 3 presented endotypes with distinct rates of multiple organ failure (MOF). The ProCESS data revealed that subjects on any type of organ support therapy (e.g dialysis or

respiratory support) fared much worse than those without. MOF is a serious clinical incident and a clinical tool to detect MOF early on can be beneficial. Clinicians, with early knowledge of imminent MOF, can apply preventative therapies and respond accordingly. The clinical need for quantifiable sepsis-induced damage motivates this section.

Mathematical models of sepsis may similarly benefit from quantifiable damage. Figure 1.1 outlined the basic acute inflammatory response. Currently, well-studied proxies for the inflammatory and anti-inflammatory responses exist. The damage compartment is a key player in inflammatory dynamics, but the lack of quantifiable damage during sepsis prevents this compartment from being accurately modeled. As a result, current mathematical models insufficiently characterizes the feedback loops that may lead to a sustained inflammatory response. Mathematical models of sepsis in literature either ignore this dynamic or exist within the realm of theory [3, 4]. Malkin, *et al* published a mathematical model of neutrophils in sepsis, and modeled kidney damaged based on creatinine clearance data [80]. However, creatinine clearance is a measure of kidney health and only proxies kidney damage rather than other organs which can fail during sepsis such as the respiratory system or the liver.

Quantifying the amount of sepsis-induced damage to the body can be done by characterizing the level of systemic tissue damage. Acute Physiology and Chronic Health Evaluation (APACHE) and Sequential Organ Failure Assessment (SOFA) are commonly used methods to assess damage in the clinic. Furthermore, multiple organ failure is based on the SOFA score (≥ 2 failed organ systems). However, these assessment scores are granular assessments and do not provide clear trajectories of how patients evolve temporally. A novel and data-driven damage assessment tool is presented to augment APACHE and SOFA scores. The goals of this tool are twofold: (i) provide damage trajectories using easy-to-measure biomarkers to reduce measurement burden and (ii) provide robustness to missing measurements to accommodate hectic conditions within the clinic.

5.2.1 Methods

The aforementioned baboon sepsis dataset was retrospectively analyzed for this work due to the temporally rich data and because pre-hospital times are a non-issue with animal data

[109]. Some baboons had occasional missed measurements, but many biomarkers were measured at $[-0.5, 0, 0.5, 1, 2, 3, 4, 5, 6, 1112, 23, 24, 35, 36, 47, 48, 72, 73, 144]$ hours post *E. Coli* infusion. Some baboons had a 672-hour measurement if they belonged to the long-term survival cohort where they had a final measurement taken prior to sacrifice at the 28-day mark.

The risk of death was associated with global tissue damage. Logistic regression (LR) mathematically characterized the risk of death for each baboon via biomarker odds ratios. Odds ratios multiplied by baboon-specific biomarkers characterized the baboon's risk of death and level of tissue injury. A discrete time equation was formulated to generate damage based on the odds ratios obtained from LR.

Biomarker selection was performed statistically with no *a priori* bias applied. Initial preprocessing reduced the set of biomarkers to 63 by eliminating predictors that were not fully measured for each of the 33 baboons. Collinearity was checked between the 63 biomarkers. Each biomarker, across all time points, was compared to each other using the Spearman rank correlation coefficient. A cutoff of Spearman's rho ($|\rho| \leq 0.75$) removed collinear biomarkers. Finally, a univariate analysis was performed on the biomarkers to identify the time points at which they were significant with respect to mortality outcome. Student's t-test was used to test the remaining biomarkers for significance with respect to mortality at each temporal measurement ($p \leq 0.10$ to provide a larger pool of biomarkers). Any biomarkers that were deemed insignificant across all time points were subsequently removed. The remaining biomarkers were further filtered based on their ease-of-measurement in order to improve the translatability of this approach into the clinic. The significant time point measurements of the easy-to-measure biomarkers served as inputs into LR models.

LR models were created for each time point from 0.5 hours to 144 hours post infusion, with biomarkers as the input and mortality probability as the output. Mortality output was defined as the interval $(0, 1)$, with 0 being survival and 1 being non-survival. For each time point, all significant biomarkers from the 0.5 hour time point up to the current time point were included as inputs. This strategy was chosen in order to capture changes between biomarkers at different time points. All LR models were calibrated against the experimental mortality outcomes of the baboons. Baboons that were sacrificed at 144 hours post-infusion

were assumed to be survivors. Each LR model was trained with one hundred replicates of three-fold cross-validation. The odds ratios across replicate models were averaged to obtain mean odds ratios for each biomarker.

Elastic net regularization was applied to each LR model in order to reduce the number of inputs in the higher time point models. Elastic net regularization is a mixture of lasso regularization and ridge regularization and provides the benefits of predictor selection when the number of predictors exceeds the number of observations and predictor coefficient shrinkage [110].

$$\min \beta \frac{1}{2N} \sum i = 1^{N} (y_i - x_i^T \beta)^2 + \lambda \left[(1-\alpha) \frac{1}{2} ||\beta||2 + \alpha ||\beta||1 \right] \tag{5.2}$$

A parameter, α, determines this regularization mixture, with $\alpha = 0$ being fully lasso regularization and $\alpha = 1$ being fully ridge regularization. $\alpha = 0.25$ was used to prioritize biomarker selection. Elastic net employs a single regularization parameter, λ, to control both ridge and lasso regularization. In order to select the regularization penalty, λ, a vector of evenly spaced λ in natural log space was generated and LR models were fit for each λ. The model with the best area under the receiver operating curve (AUROC) was compared. $\lambda = e^{-6}$ yielded consistently high AUROC (near maximum) for the majority of the time point LR models. For consistency, elastic net regularization with this value of λ was applied to LR at all time points. The ensemble of beta coefficients associated with this λ (3000 sets of coefficients, due to 3 folds and 1000 replicates) were averaged to produce a set of time-varying odds ratios for each clinical predictor.

A discrete time damage equation was implemented due to the irregular sampling rate of biomarkers. Time was discretized into intervals of 0.5 hours to accommodate the fastest sampling time in the data. The formulation, shown in Equation 5.3, consists of a damage elimination term and a generation term. The negative term including k represents the body's healing rate. A linear mechanism was proposed as a best-case scenario; in reality, high levels of injury typically result in cascading organ failures that impede a patient's healing ability. The value of k was defined as 0.1768, which corresponded to a damage half-life of about 4 hours. Humans may take months to recover from endotoxemia challenges while murine models demonstrate remarkably fast recovery in the order of hours [8]. The value of k was set to characterize a similarly fast damage recovery rate in baboons. Damage was generated

by the logit (the mean odds ratio multiplied by the value of the associated biomarkers at that time) obtained from the time relevant LR model.

$$DMG(\ell+1) = DMG(\ell) - k * DMG(\ell) + exp(\underline{\beta_\ell}\,\underline{\underline{X_\ell}}) \tag{5.3}$$

Two cases were defined for missing data. In the case of a missing measurement, the logit was calculated by multiplying the most recently measured biomarker by the current odds ratio, β_ℓ. For example, a baboon has one missing biomarker measurement at hour 47. That specific term was calculated with $\beta_{47}X_{36}$. In the case of an ℓ at a time point that was not sampled, the logit was calculated by multiplying the most recently measured biomarker by the most recently calculated odds ratio. For example, the next measurement after 6 hours post infusion was 11 hours. To calculate damage at hours 6.5 through 10.5, the logit at hour 6 was used: $exp(\underline{\beta_{6hr}}\,\underline{\underline{X_{6hr}}})$.

5.2.2 Results

5.2.2.1 Damage Trajectories

Pre-processing reduced the pool of biomarkers from 73 to 27. 10 biomarkers were removed because they were unavailable among all 33 baboons. 30 biomarkers were removed due to identified collinearities in the data. Student's t-test on the remaining 33 biomarkers revealed 6 biomarkers that were not significant at any time point. However, certain time points such as hour 23 revealed seventeen different significant biomarkers. To satisfy the requirements of low measurement burden (assessing 17 measurements to evaluate damage would be an unreasonable task) and ease of measurements (quick, commonly taken measurements in the clinic), 20 additional biomarkers were excluded, leaving 13 biomarkers. These biomarkers are listed in Table 5.6.

A total of 18 LR models were created: one for each time point. Each LR model was performed with three-fold cross-validation and 1000 replicates. The predictor selection frequency resulting from regularization was tabulated and normalized between 0 and 1. A specific predictor can appear in an LR model a maximum of 3000 times (appears in each of the three cross-validation models across all 1000 replicates). Additionally, a predictor at

Table 5.6: Acronym dictionary of biomarkers for damage results.

Acronym	Meaning	Measurement Method
HR	Heart Rate	Straightforward
MAP	Mean Arterial Pressure	Calculated
TEMP	Temperature	Straightforward
HB	Hemoglobin	Vein Blood Sample
PHA	Arterial PH	Arterial Blood Sample
ABEA	Acid Base Excess	Arterial Blood Sample
RR	Respiratory Rate	Straightforward
SHUNT	Amount non oxygenated blood returning to heart	Calculated
PAO2	Partial arterial pressure of Oxygen	Arterial Blood Sample
PLT	Platelet Count	Vein Blood Sample
LACTATE	Lactate	Vein Blood Sample
GLUCOSE		Finger Stick
RAP	Right Atrial Pressure	Need Central Line*

*While central line measurements are hard to obtain due to the risks associated with their insertion, severely septic patients typically have one inserted.

time 0.5 hour can appear in all 18 LR models while a predictor at time 144 hour can only appear once. Normalization accounted for both of these factors when scaling down the selection frequency. Figure 5.4 presents the predictor selection frequency for all 18 LR models. For example, the ensemble of LR models selected 30-minute measurements of right atrial pressure (RAP) quite often and arterial pH (PHA) roughly 50% of the time.

The odds ratios from all 18 LR models were used to calculate the logit in Equation 5.3 and a damage trajectory was generated for each of the 33 baboons. Figure 5.6 shows all of the trajectories. Some baboons' trajectories were incomplete due to right-censoring via sacrifice in accordance with experimental protocol. Trajectories with straight line segments were the result of the zero order hold assumption. Survivor and nonsurvivor trajectories were distinct from one another and complete separation occurs at 7-8 hours post infusion. Intuitively, nonsurviving baboons with high damage levels die at earlier time points while nonsurviving baboons with lower damage levels die at later time points. Interestingly, certain nonsurvivor trajectories exhibited rapid decreases of damage. However, these baboons eventually succumbed to the septic shock. Surviving baboon trajectories were characterized

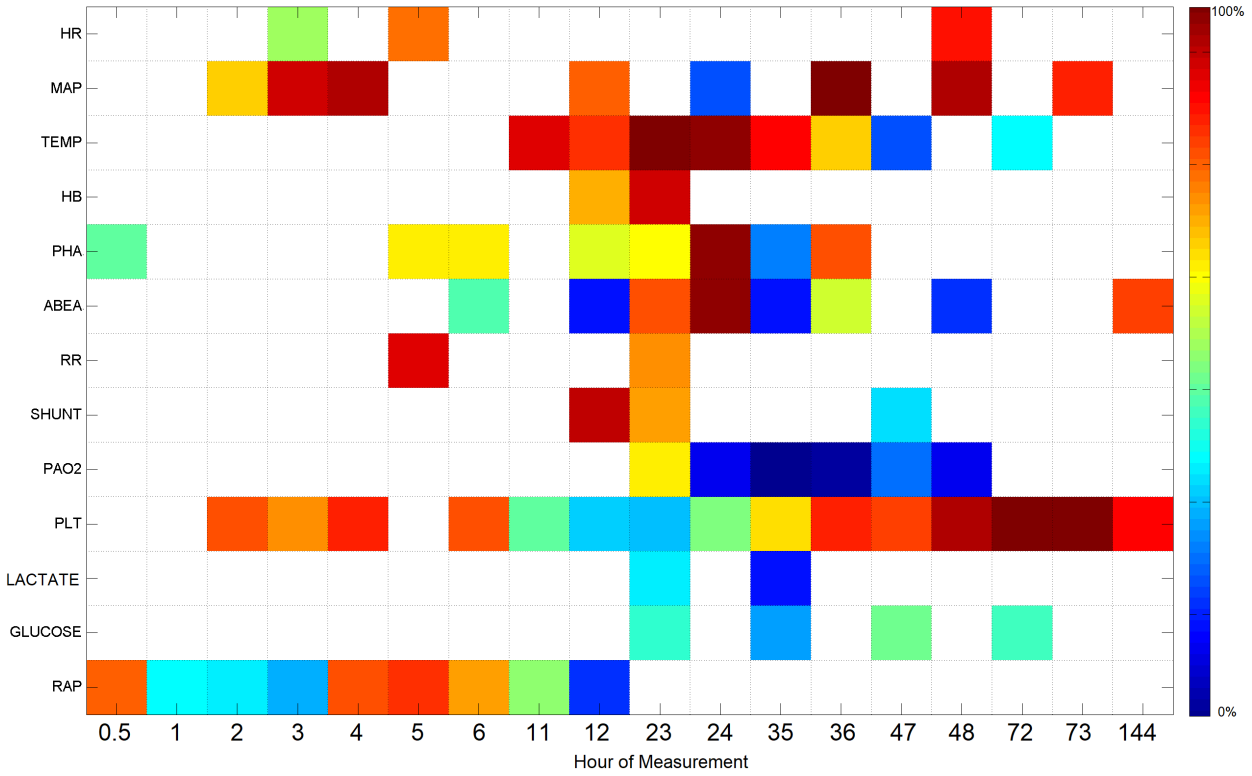


Figure 5.4: Predictors (and their frequency) used to generate damage curves. Y-axis specifies the predictors and X-axis specifies the time point of measurement at which the predictors were used. White cells indicate that the corresponding predictor at that time point was not used due to a lack of significance from Student's t-test. Colored cells indicate that the corresponding predictor was selected by the regularized LR model with a frequency corresponding to that color: dark red means frequently (a maximum of 3000 times) and dark blue means very few (a minimum of 0 times) and everything in between.

by an initial damage spike followed by a steady decline to steady states within $10^{-}1 - 10^{-}3$. The rate of damage was assumed to be constant across all baboons, which indicated that the survivor trajectory declivities may be attributed to the resolution of infection and inflammation. Furthermore, damage levels for survivors stayed below a magnitude of 10, which suggested a threshold of no return in terms of cascading organ damage.

To further test the robustness of this algorithm to missing measurements, entire time points of measurements were removed from those shown in Figure 5.4 to emulate missed measurements and/or sampling limitations in the clinic. Figure 5.5 illustrated the remain-

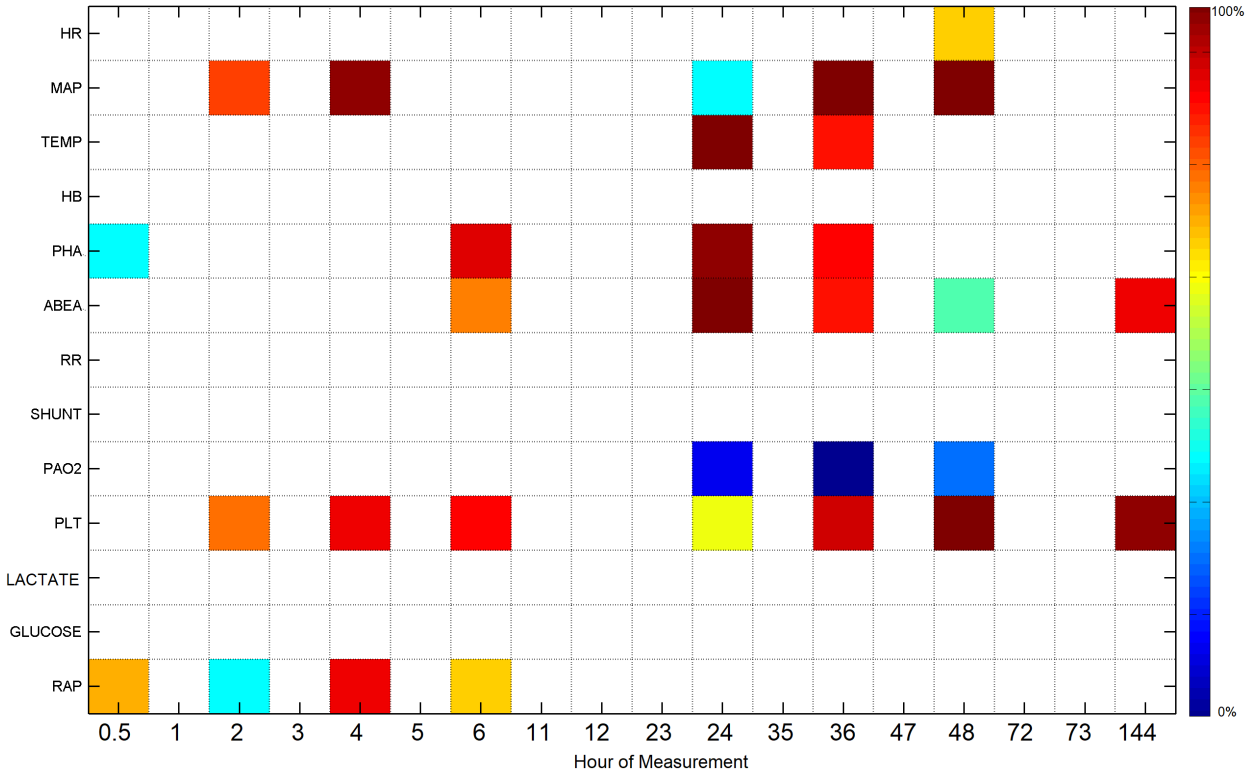


Figure 5.5: Predictors (and their frequency) used to generate damage curves in a missed measurement scenario. This heatmap is organized in the same fashion as in Figure 5.4. However, a sparse measurement scenario was enacted, causing many of the measurements in Figure 5.4 to be missed. Of the remaining predictor measurements (non-white cells), selection frequencies by regularized LR are shown.

ing predictors and time points. The algorithm was re-applied to this set of predictors and the updated color cells represented new LR selection frequencies. Once fully trained, these LR models were used to recalculate the logits in Equation 5.3. The damage trajectories that were recreated are shown in Figure 5.7. The missing measurements at early time points $(1, 3, 5, 11, 12, 23$ hours post infusion) contributed to the significantly slower separation. Separation occurred at 25-26 hours, which indicate that once new measurements were available (the last set of measurements were at hour 6), damage levels updated to reflect new information. Additionally, overall nonsurvivor damage magnitudes decreased as compared to those of Figure 5.6. However, the damage threshold of 10 still appeared to hold in this sparse data scenario.

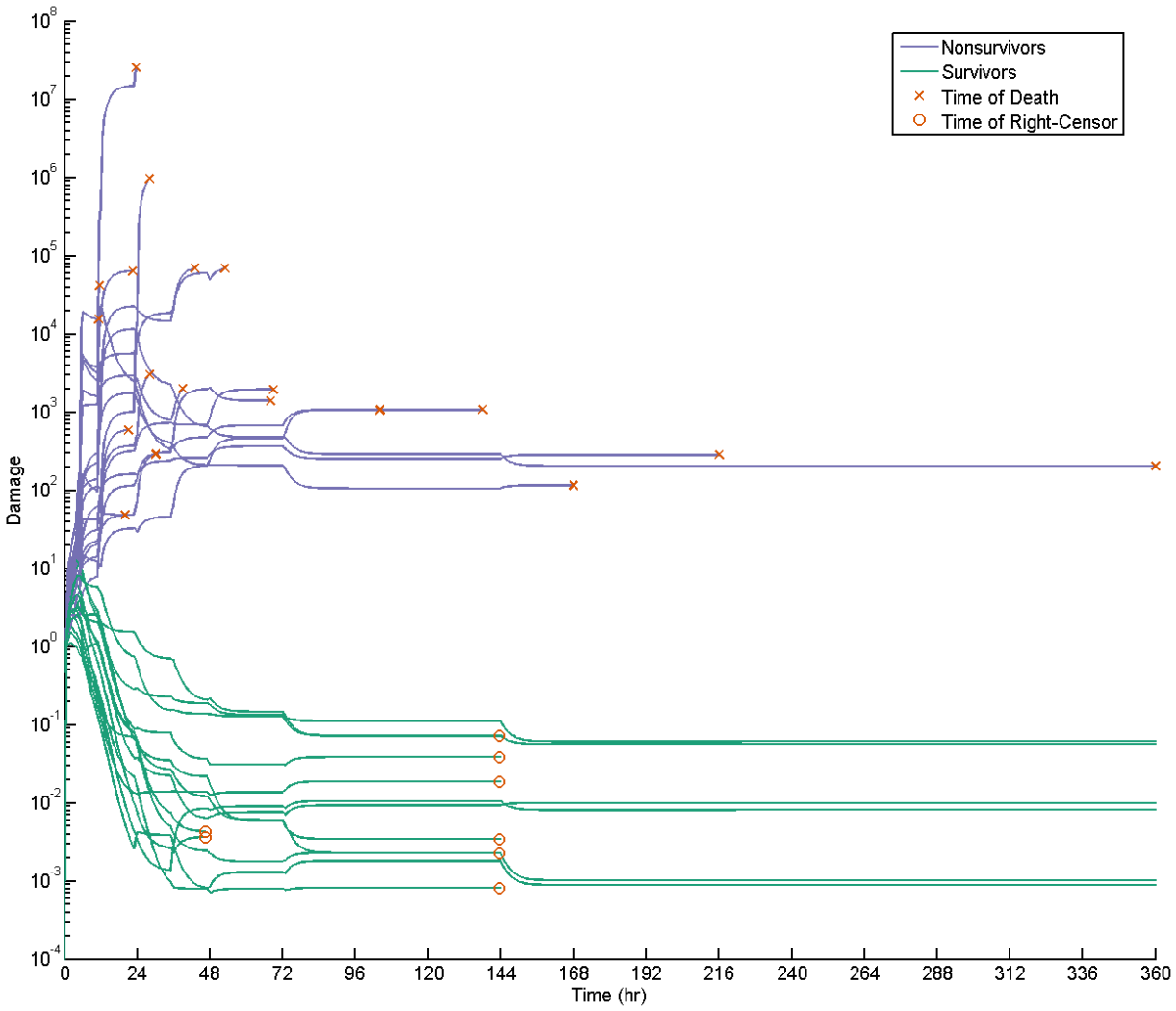


Figure 5.6: Systemic tissue damage trajectories for baboons using easy to measure biomarkers. Green trajectories represent baboons that ultimately survived sepsis while the purple trajectories represent nonsurvivors. Within the baboon survivors, many were sacrificed at the end of 6 days as per experimental protocol, which is represented by the red circles capping those trajectories. Nonsurvivors' times of death are represented by a red x capping the end of their trajectories. There was a clear separation of survivors versus nonsurvivors by hour 11, where nonsurvivors exceeded a damage of 10^1 and proceeded to increase sharply. While some survivors may have reached this level of damage, their bodies recovered.

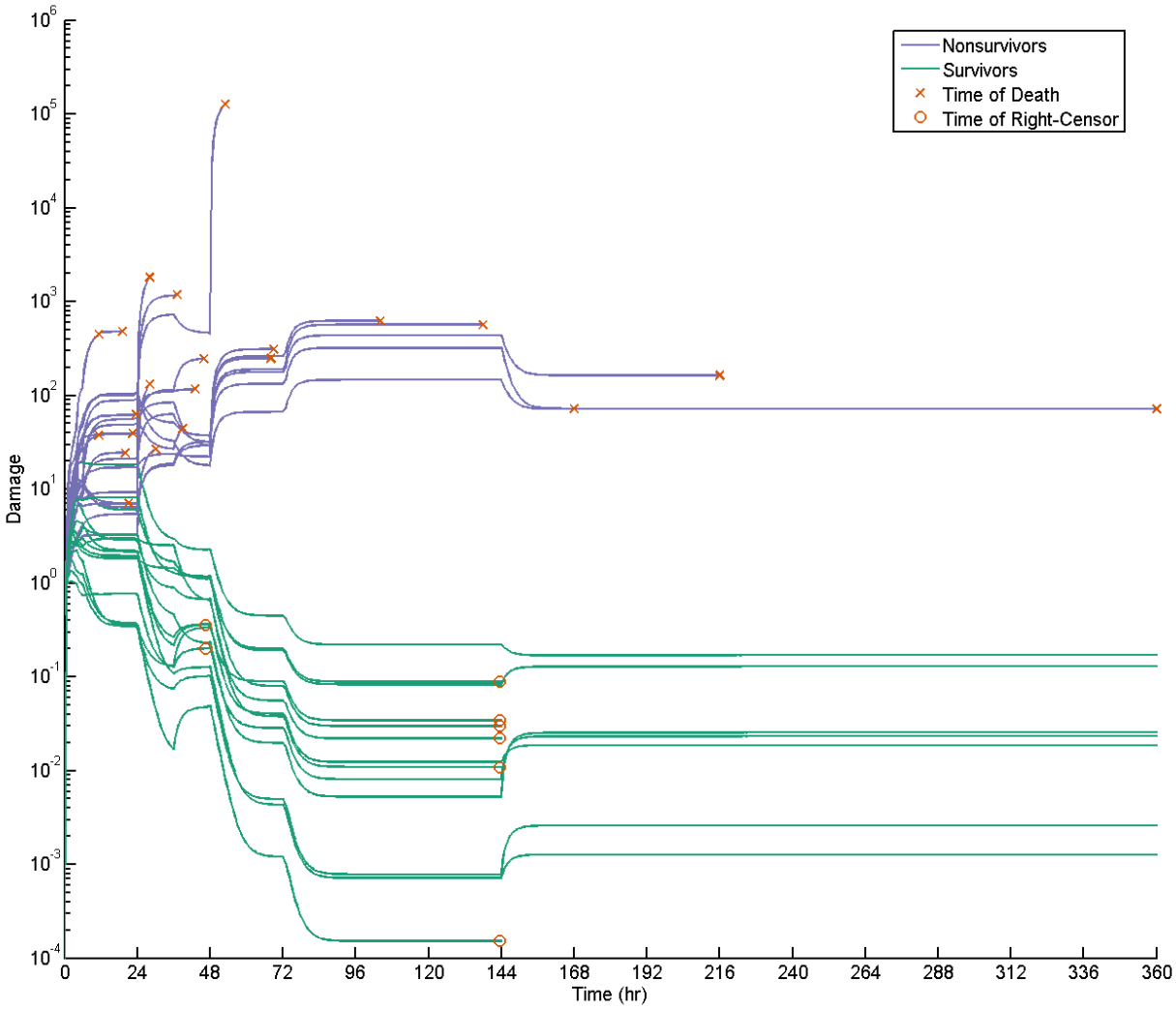


Figure 5.7: Systemic tissue damage trajectories for baboons using sparse biomarkers. Damage trajectories were calculated from significantly missing data (see Figure5.5. Separation was achieved after a longer period of time (24 hours) due to missing measurements and the zero-order-hold rule. However, whenever new data became available, the damage trajectories moved appropriately in the survivor/nonsurvivor directions. 10^1 appeared to remain a valid damage cutoff in determining mortality.

5.2.3 Discussion

Predictor selection frequency via LR identified several key biomarkers that were currently used to calculate disease severity scores such as SOFA and APACHE II. For example, the

method identified and favored biomarkers MAP (cardiovascular health), WBC (immune capability),Pa_{O_2} (respiratory health), and PLT (coagulation) across many points in time. Initial tests, prior to selecting only for biomarkers that were easily measured, revealed other biomarkers used to calculate SOFA or APACHE II including bilirubin (liver health) and creatinine (kidney health). These biomarkers were identified purely out of statistical methods and *a priori* information from existing damage scoring systems was not taken into account.

Figure 5.6 suggests that early organ support is critical to shift subjects towards a surviving trajectory. For the first 12 hours, nonsurviving subjects had a characteristic rising-only damage trajectory, indicating that their organ systems were worsening by the hour. Alternatively, survivor trajectories began declining between 6 to 12 hours post infection. This indicated that organ function began returning in the survivors because the healing term in Equation 5.3 dominated the dynamics.

While multiple organ failure is not clearly defined within baboons, Figures 5.6-5.7 clearly demonstrate that there were levels of catastrophic systemic damage after which surviving was not possible. These damage trajectories, combined with the fact that many of the selected predictors chose known organ damage biomarkers, suggest that this approach can be translated into the clinic.

The algorithm proposed here serves as a proof of concept of a damage assessment method that can both be feasible and practical in a clinical setting. The biomarkers used for predictions in the sparse measurement scenario demonstrated the robustness of this algorithm with respect to extreme missing measurements. Furthermore, the easy-to-measure set of identified biomarkers included vital signs and venous and arterial blood samples, all of which are easy and fast to obtain. A large and temporally rich sepsis database, combined with knowledge of pre-hospital times (estimated or otherwise) would be suitable to validate the algorithm.

This method may be possible to translate to the ProCESS trial, which is one of the largest and richest sepsis databases in existence. Preliminary testing revealed that complete cases of the easy-to-measure biomarkers yielded a large enough convenience cohort to apply this algorithm. Pre-hospital times (such as those estimated from Chapter 3) would need to be accounted for because the algorithm is heavily dependent on calibrating the logit term in Equation 5.3 via comparable population data. One potential fallacy that the current

algorithm does not account for would be the issue of therapy. The baboons within the experiment only received fluid resuscitation while the subjects in the ProCESS trial received interventions. The current algorithm is currently unable to process interventions as an input into the ensemble of LR models. As a result, clear mortality separation is currently difficult to achieve because a potentially life-threatening state, which would otherwise be predicted by the LR to die, may be modulated by some intervention and survive the initial septic shock. As a result, future work is warranted to explore the effects of treatment and therapies on damage.

5.2.4 Summary

A data-driven algorithm was introduced to quantify sepsis-induced systemic tissue damage. The premise of this method was to identify subjects who perished due to septic shock and calibrate the model with the state of these subjects' various organ systems. Biomarkers deemed to be significant with respect to mortality via Student's t-test served as proxies of the health of organ systems and tissue that were significantly affected by septic shock. At each time point, an ensemble of LR models was trained on the data available. Organ damage was calculated as odds ratio and characterized by the logit portion in Equation 5.3. This approach was able to characterize the initial inflammatory-induced damage to sepsis in all subjects sepsis and demonstrated that nonsurvivors exceeded a damage threshold of no return. These results are applicable in the clinic because subjects remained in a state of uncertainty for several hours before setting upon a trajectory that sealed their fates. This demonstrates both the existence of a clinical therapeutic window and the possibility of the ability to modulate a trajectory from a "death" state to a surviving one. Combined with the ability to estimate pre-hospital time, this systemic damage algorithm may provide clinicians the power to apply and time preventative therapies to improve patient outcomes. This algorithm is ready for translation and testing with human subjects.

6.0 RELEVANT MODELING TOOLS

Several tools are introduced here to assist with the challenging task of mathematical modeling of sepsis dynamics. First, a network optimizer is introduced to assist with the establishment of a model framework structure. This optimizer analyzes time series data that has been converted into Boolean variables (two states: on/off) and uses mixed integer linear programming to optimize the best network that produces the data. Second, a Markov Chain Monte Carlo-based algorithm is introduced that focuses on parameter fitting. This algorithm is written with user-friendliness in mind and focuses on algorithmic speed, efficiency, and ease-of-use. These tools are intended to be used in harmony; the former produces latent relationships within the data inform and refine model structure and the latter tackles the inverse-problem of parameter fitting to said model.

6.1 BOOLEAN-LP, A NETWORK PATHWAY OPTIMIZER

A major problem associated with mathematical modeling of sepsis is the sheer number of potential biomarkers and the myriad of interactions between them [39]. Compounded with the sparsity of human sepsis data, it is often difficult to pinpoint why few mathematical models of sepsis have been validated against literature: model structure problems, failure of the inverse-problem algorithms, or a combination of both.

Boolean Linear-Programming (Boolean-LP) is introduced as a tool to assist with the initial steps of model creation. Boolean-LP was created in an effort to identify the underlying inflammatory pathways between elderly and young mice (Mochan, Zhang, *et al*, Discrete

dynamical modeling of influenza infection suggests age-dependent differences in immunity; submitted to the Journal of Virology).

Model structure development is one of the initial phases of mathematical modeling, where the modeler combines literature knowledge and intuition to propose a mathematical characterization of certain dynamical behaviors. This begins an iterative process of fitting data to this proposed model and then making structural model changes to accommodate missing behaviors or other issues. Boolean-LP speeds up this approach by taking a set of proposed network pathways between model states and identifying the network pathway that best describes the given data.

Boolean-LP first requires the user to first discretize the data into Boolean variables. Second, the user needs to propose a set of model pathways via Boolean rules. For example, the rule $C(t+1) = A(t)$and not $B(t)$ may be proposed, describing the production of state C by state A but this process is inhibited by state B. A set of rules describing the relationships between states A, B, and C and data are provided to Boolean-LP and it returns the rule that best describes the data. In the case of a solution pool (multiple rules yield the same objective function), Boolean-LP will return all of the possibilities.

To find the optimal set of Boolean rules that best described the data, the rule discovery problem may be formulated as a mixed integer linear programming (MILP) problem. A similar formulation has been previously reported [111]. We define our objective function as the minimum difference between the model and the measured data, given by equation (6.1).

$$\min \sum_{s \in S} \sum_{t \in T_s} |D_{t,s} - M_{t,s}| \tag{6.1}$$

This was later reformulated as a linear objective function using dummy variables (see Section 6.1.2, specifically equation 6.17) in order to conver the problem into a mixed-integer linear programming (MILP). $D_{t,s}$ and $M_{t,s}$ represents the measured data and model, respectively, for state s at time t.

All potential Boolean rules were expressed as a series of logical equivalences (if and only if statements: $clause_1 \leftrightarrow clause_2$). These rules were expanded into their equivalent conjunctive normal form, which is the conjunction (AND) of several inclusive OR clauses.

This form allows us to represent each of the OR clauses with a single linear inequality [112]. As long as each of these linear inequalities are satisfied, the overall conjunction expression is satisfied. The result is:

$$\begin{aligned} & Q_1 \wedge Q_2 \wedge \cdots \wedge Q_n \\ & Q_1 = P_{1_1} \vee P_{1_2} \vee \cdots \Longrightarrow y_{1_1} + y_{1_2} + \cdots \geq DV \\ & Q_2 = P_{2_1} \vee P_{2_2} \vee \cdots \Longrightarrow y_{2_1} + y_{2_2} + \cdots \geq DV \end{aligned} \tag{6.2}$$

where y_{i_j} represents the Boolean value of expression P_{i_j} and DV is a Boolean decision variable. The decision variable allows the optimizer to apply this constraint ($DV = 1$) or turn it off ($DV = 0$).

6.1.1 MILP Formulation of Network Optimizer Problem

Specifically, all Boolean Rules can be expressed in the conjunctive normal form, which comprises of a series of overarching AND clauses consisting of OR operators:

$$\text{Rule}R = Q_1 \wedge Q_2 \wedge \cdots \wedge Q_n \tag{6.3}$$

$$Q_i = P_{i1} \vee P_{i2} \cdots \vee P_{ir} \tag{6.4}$$

$$P_{i1} \in 0, 1 \tag{6.5}$$

where Q_i is a series of inclusive OR operators. Let y_i represent the Boolean value of clause P_i. Each of the Q_i logical OR constraints can be expressed as

$$P_{i1} \vee P_{i2} \cdots \vee P_{ir} \Rightarrow y_1 + y_2 + \cdots + y_r \geq 1 \tag{6.6}$$

The AND constraint, R, does not need to be explicitly constrained because (6.6) ensures that each of its sub-clauses, Q_i, are true. NOT clauses, $\neg P_1$, can be expressed as:

$$1 - y_1 \tag{6.7}$$

Implications, e.g. $P_1 \Rightarrow P_2$, can be expressed as $\neg P_1 \vee P_2$, which is an OR constraint:

$$1 - y_1 + y_2 \geq 1 \tag{6.8}$$

Using this framework, we can formulate this as a MILP. Potential rules are always of the form: $S_{i,t+1} \leftarrow S_t$, where S_t represents a series of logical operations acting upon the states at the current time t. This logical clause will generate an update to the ith state, S_i, at time $t+1$. If this rule were true, then $S_{i,t+1} \Leftrightarrow S_t$ for all time, t. Applying the above equivalences, we obtain:

$$\begin{aligned} &\neg S_{i,t+1} \vee S_t \\ &\neg S_t \vee S_{i,t+1} \end{aligned} \tag{6.9}$$

which we expand into the conjunctive normal form and apply the appropriate linear constraints.

Finally, in order to perform rule optimization, Boolean decision variables, DV_{ij}, are initialized for every proposed rule j in each state i. The k OR constraints generated from the conjunctive normal form of Rule ij is now represented as:

$$\begin{aligned} &\text{Constraint 1: } y_1 + y_2 + \cdots + y_r \geq DV_{ij} \\ &\text{Constraint 2: } y_1 + y_2 + \cdots + y_r \geq DV_{ij} \\ &\qquad\qquad \cdots \\ &\text{Constraint k: } y_1 + y_2 + \cdots + y_r \geq DV_{ij} \end{aligned} \tag{6.10}$$

which represents a slight modification from equation (6.6) in order to allow the optimizer to turn a constraint on or off. If DV_{ij} is 0, the values of y_i are unconstrained and potential rule ij does not apply. If DV_{ij} is 1, the rule applies. A final constraint is set:

$$\sum_{j=1}^{J} DV_{ij} == 1 \tag{6.11}$$

such that each state i may only have 1 rule selected.

6.1.2 Detailed Formulation Example

As a detailed example, take a rule governing Interleukin-6, activated macrophages (ActiveM), and Interleukin-8 (IL-8, a neutrophil attractant): $ActiveM(t+1) \leftarrow IL8(t) \wedge IL6(t)$ Working out one implication at a time, this is equivalent to constraints (6.13),(6.14), (6.16).

$$\begin{aligned} \text{ActiveM}(t+1) \Rightarrow \big[\text{IL8}(t) \wedge \text{IL6}(t)\big] &= \neg\text{ActiveM}(t+1) \vee \big[\text{IL8}(t) \wedge \text{IL6}(t)\big] \\ &= \big[\neg\text{ActiveM}(t+1) \vee \text{IL8}(t)\big] \wedge \big[\neg\text{ActiveM}(t+1) \vee \text{IL6}(t)\big] \end{aligned} \tag{6.12}$$

$$1 - \text{ActiveM}(t+1) + \text{IL8}(t) \geq DV \tag{6.13}$$

$$1 - \text{ActiveM}(t+1) + \text{IL6}(t) \geq DV \tag{6.14}$$

$$\begin{aligned} \big[\text{IL8}(t) \wedge \text{IL6}(t)\big] \Rightarrow \text{ActiveM}(t+1) &= \neg\big[\text{IL8}(t) \wedge \text{IL6}(t)\big] \vee \text{ActiveM}(t+1) \\ &= \text{ActiveM}(t+1) \vee \neg\text{IL8}(t) \vee \neg\text{IL6}(t) \end{aligned} \tag{6.15}$$

$$ActiveM(t+1) + 1 - IL8(t) + 1 - IL6(t) \geq DV \tag{6.16}$$

To further simplify the problem, the nonlinear objective function from equation (6.1) may be linearized via the introduction of dummy variables $A_{t,s}$:

$$\begin{aligned} \text{minimize} \quad & \sum_{s \in S} \sum_{t \in T_s} A_{t,s} \\ \text{subject to} \quad & D_{t,s} - M_{t,s} \leq A_{t,s} \\ & D_{t,s} - M_{t,s} \geq -A_{t,s} \end{aligned} \tag{6.17}$$

where $D_{t,s}$ and $M_{t,s}$ represents the measured data and model, respectively, for state s at time t.

A Python (version 3.5) package was written to accept Boolean data and a list of potential rules for each state. This package reformulates the inputs into a MILP problem for use with the Python Optimization Modeling Objects package (Pyomo) [85, 86]. Pyomo then converts this script into a solver-friendly file, which was then solved by the IBM ILOG CPLEX optimization studio. CPLEX was set to populate all optimum solutions via its solution pool

feature. Finally, our Python package parses through this solution pool and returns all valid rules that yield the minimum objective function value. The GLPK solver was successfully tested, but it was significantly slower then CPLEX and lacked the ability to populate a solution pool in the case of multiple solutions. Boolean LP is available as an open source project on GitLab.

6.1.3 Applications in Age-related Immune Pathways

The BooleanLP algorithm was created to address optimization problems in a project quantifying the immunosenescence (age-related) differences in immune pathways between adult and elderly mice. Adult (12-16 weeks) and elderly (72-76 weeks) BALB/c mice were subject to an innocula of the Influenza virus [113]. Near daily sacrifices (in triplicate) were taken from each group for measurements of inflammatory markers (chemokines and cytokines) and white blood cells, including day zero baseline measurements. An ANOVA analysis was conducted ($p \leq 0.05$) to convert the entire dataset into zero (off) and one (on). A library of possible immune pathway rules were generated and BooleanLP was applied to identify the optimum inflammatory network that best described the data. BooleanLP indicated no age-dependent changes in macrophage recruitment between the elderly and adult mice, but macrophage cytokine expressions were different. Cytokine and chemokine pathways differed vastly between the age groups and was responsible for the two day delay in the immune response within elderly mice. The manuscript of this work is currently being updated after the first round of reviewers.

6.2 APT-MCMC, A C++/PYTHON IMPLEMENTATION OF MARKOV CHAIN MONTE CARLO FOR PARAMETER IDENTIFICATION

While Boolean-LP provided a method to quickly identify the optimal model structure from a potential list of possibilities, the next daunting step is the inverse problem. The inverse problems that mathematical modelers tackle are often very challenging due to parameter

dimensionality, parameter correlation, and multiple optima. A topical and popular set of inverse problems is fitting parameters to a system of ordinary differential equations (ODEs). Most deterministic optimization algorithms can be broken down into two types: line search and trust region [114]. During nonlinear fitting, these algorithms often get trapped within local minima because they are highly dependent on the provided starting point. Additionally, many advanced solvers calculate or estimate the Jacobian or Hessian of the objective function, a process which may not converge for complex ODE parameter-fitting problems. This results in heuristic implementations that employ multi-start optimization to characterize the optimal parameter space, and parameter estimates that are likely only locally optimal and for which confidence intervals cannot necessarily be easily constructed. These problems are very topical in the case of modeling the dynamics of sepsis due to the sparsity of human data and the tightly regulated interactions that persist throughout the inflammatory response.

Markov Chain Monte Carlo (MCMC) is a stochastic sampling technique typically used to gain information about a probability distribution that lacks a closed form. It has been described as a "bad method" for parameter estimation to be used when all alternatives are worse [115]. MCMC has been used in a variety of fields, such as cryptography, statistical mechanics, and astrophysics [116, 117]. Long runtimes, simulation stochasticity, and the lack of a robust convergence and stationarity criteria contribute to the stigma against MCMC. However, MCMC is immediately relevant to ill-posed ODE inverse problems. Unlike deterministic algorithms, MCMC is not dependent on the starting point and does not require the computation of Jacobians or Hessians. Furthermore, an MCMC simulation provides parameter probability distributions, which can help elucidate parameter correlations and identify multiple optima (see Figure 6.1).

MCMC is formulated as a stationary Markov Chain with a transition probability derived from the parameter probability distribution. Bayesian formalism is used to reformulate a parameter fitting problem as a search for probability distribution. That distribution can be sampled using MCMC. Equation (6.18) describes Bayes theorem, where θ is a parameter and D is the data.

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \tag{6.18}$$

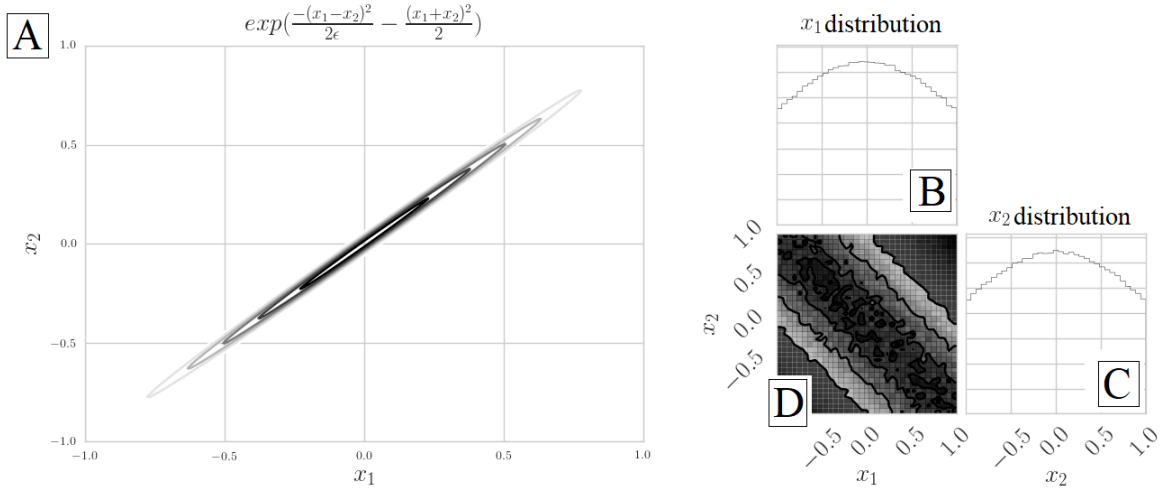


Figure 6.1: MCMC results from exploring a highly correlated probability function. **A**: Contour plot for an anisotropic distribution. Epsilon was set to 0.001. **B** and **C**: MCMC results yields marginal probability distributions for x_1 and x_2. **D**: These distributions can be combined into a joint distribution plot to identify parameter correlations.

The $P(\theta|D)$ term is the Bayesian posterior probability, which represents the probability of the parameter set, θ, given the data. The first term in the numerator describes the likelihood of obtaining the data, D, given the parameters, θ. The objective function for a parameter fitting problem, reformulated as a likelihood, is used here. The second term in the numerator describes the prior distribution of the parameters. This can vary depending on the problem, but the simplest case is a bounded uniform distribution to provide upper and lower bounds on θ. The denominator is a normalization constant that represents the probability of the data, typically an unknown, but constant, term. This term is not necessary for MCMC simulation, as shown in Equation (6.18).

To recover the Bayesian posterior distribution from a parameter fitting problem, MCMC needs to generate many samples from the parameter space in accordance to Equation (6.18). The MCMC sampler stochastically proposes the next parameter set, $\theta_{proposed}$, based on the location of the current parameter set, $\theta_{current}$ [118, 119] . The Metropolis-Hastings criterion is typically used to accept or reject $\theta_{proposed}$ by calculating the posterior probability, $P(\theta_{proposed}|D)$. If this value is higher than the posterior probability of the current parameter

set, $P(\theta_{current}|D)$, then $\theta_{proposed}$ will be accepted with 100% probability. If the probability is lower, then $\theta_{proposed}$ may be accepted with the nonzero probability calculated by Equation (6.21) (with T=1). This step is repeated many times to allow the samplers to fully explore the parameter space.

Over the course of an MCMC simulation, all the values of θ (derived from each MCMC sample via the calculation of Equation (6.18)) are preserved for post-hoc analysis. The posterior distribution can be recovered via histogram analysis or kernel density estimations from the chain of θ values. θ may be broken down into its individual parameters to obtain a marginal posterior probability distribution for each fitted parameter. These marginal distributions can yield generate parameter confidence intervals, identify inter-parameter correlations via joint probability distributions, and to identify the locations of multiple minima in the parameter space based on multi-modal marginal parameter distributions. Furthermore, irreducibility and aperiodicity properties of Markov Chains guarantees convergence of the posterior distribution [115]. The major benefit of these properties is that an MCMC simulation will yield the correct Bayesian posterior based on the data and model structure in finite time.

Despite the guarantee of convergence, it is impossible to test if a MCMC sampler has converged towards its equilibrium distribution or a temporary meta-stable distribution [115]. As a result, MCMC simulations typically sample a very large number of points. The initial samples that are taken prior to reaching the equilibrium distribution may bias the posterior distribution. The two ways possible ways to deal with this issue are: (i) discard some initial percentage of simulation samples or (ii) run a long simulation to minimize the bias. In either scenario, simulations take a long time because a large number of samples are needed for convergence.

Many variations have been proposed over the years in order to improve the efficiency of the sampling. One issue is that MCMC samplers may get trapped in local minima for a long time before escaping. An escape consists of several consecutive low probability parameter moves because it requires movement towards increasing deviation from the data. At any point, the MCMC sampler may change direction and move back into the local minima. Parallel tempering addresses this by employing multiple MCMC interacting simulations in

parallel [52, 120]. Each simulation is assigned a temperature and the Boltzmann distribution, Equation (6.21), is used to accept a parameter proposal. The Bayesian posterior probability for the proposed and current parameter vectors are calculated as normal (Equation (6.18)), but are treated as energies (Equations (6.19)-(6.20)) in order to fit the formalism of the Boltzmann distribution. These energies, E_1 and E_0, are scaled by the simulation temperature before determining the proposal acceptance probability. In effect, the objective function is divided by temperatures so that samplers have higher probabilities of escaping local minima. Figure 6.2 demonstrates how the energy "landscape" of a probability distribution can change given a high enough temperature. The higher temperature landscape at $T = 3$ or $T = 10$ demonstrates the relative ease at which an MCMC sampler can explore the entire parameter space compared to the lower temperature landscape where the sampler may be prone to getting trapped in local minima.

$$E_1 = \log P(\theta_{proposed}|D) \tag{6.19}$$

$$E_0 = \log P(\theta_{current}|D) \tag{6.20}$$

$$\text{Parameter Acceptance} = \min(1, \exp\left(\frac{E_1 - E_0}{T}\right)) \tag{6.21}$$

$$\text{Swap Acceptance} = \min\left[1, \exp\left(-(E_i - E_j)(\frac{1}{T_i} - \frac{1}{T_j})\right)\right] \tag{6.22}$$

These energies, E_1 and E_0, are scaled by the simulation temperature before determining the proposal acceptance probability. In effect, the objective function is divided by temperatures so that samplers have higher probabilities of escaping local minima. Figure 6.2 demonstrates how the energy "landscape" of a probability distribution can change given a high enough temperature. The higher temperature landscape at $T = 3$ or $T = 10$ demonstrates the relative ease at which an MCMC sampler can explore the entire parameter space compared to the lower temperature landscape where the sampler may be prone to getting trapped in local minima.

A unique feature of parallel tempering algorithms is that the tempered Monte Carlo chains are allowed to evolve independent of each other for some time. After a user-set hyperparameter, n_{steps} (a "swap interval") has passed, an inter-chain swap of parameters is

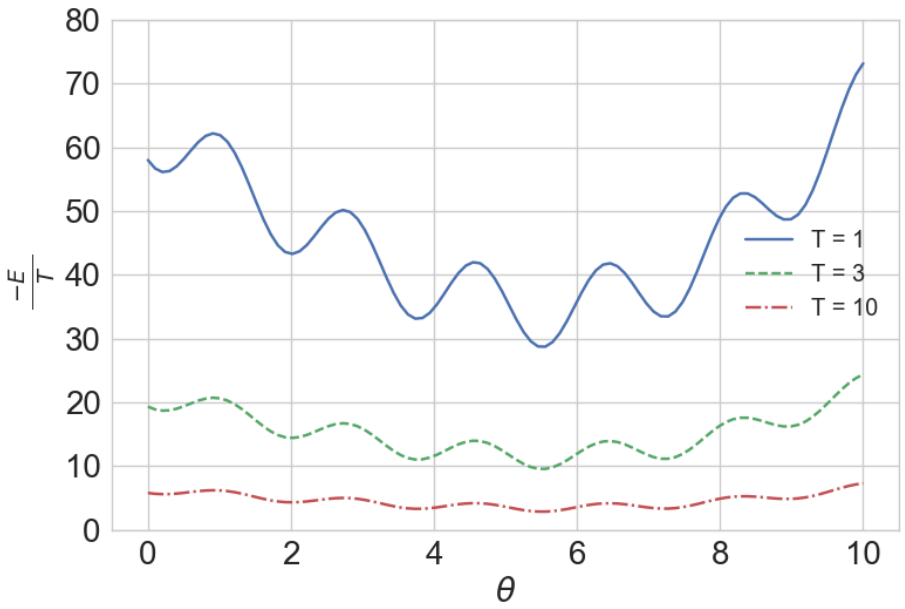


Figure 6.2: Effects of varying temperatures on the Boltzmann distribution. Shown is a single parameter, θ, and the system's energy at that parameter value. The steep local minima are difficult to escape because the escape probability as defined by Metropolis-Hastings criterion becomes exponentially smaller. Higher temperatures modulate the energy landscape such that local minima become shallow and easier to escape.

attempted to pass the information about potential energy minima (higher probability parameter locations) from higher temperature samplers to the lower temperature samplers. This swap probability is calculated with Equation (6.22) by comparing the energy and temperature of chain i against those of chain j [121]. If a higher temperature sampler is in a parameter location with a larger energy value (higher Bayesian posterior and more favorable location), then that information will pass down to a lower temperature sampler with 100% probability. The opposite scenario, where a swap is proposed to provide a low temperature sampler with a lower energy parameter location (less favorable), is possible, but less probable (small, nonzero swap acceptance). The net effect is that high temperatures perform a broad search over the parameter landscape and pass locations of local minima to the lower temperatures, which conducts a depth search within these minima.

An issue that often arises from inverse problems that estimate ODE parameters is parameter correlation. Correlations cause parameter distributions to be highly anisotropic. Figure 6.1A illustrates a highly anisotropic system that would be very difficult for MCMC to efficiently sample. The significant correlation between x_1 and x_2 can be clearly seen in the joint probability distribution in Figure 6.1D. The majority of MCMC samples will lie outside regions of interest and provide no information that can be used to recover the probability distribution. As a result, to uncover the marginal distributions of x_1 and x_2, as illustrated in Figures 6.1B and C, a lengthy simulation would be required. To improve the sampling efficiency of anisotropic distributions, Goodman and Weare proposed an MCMC variant that utilizes an ensemble of samplers with affine invariance [122]. The affine invariance property of this algorithm means that sampling efficiency is unaffected by the anisotropic nature of a parameter topology. This means that the ensemble of samplers can sample from the distribution in Figure 6.1A just as efficiently as it can sample from an isotropic distribution (e.g. a perfectly circular density plot).

ODE parameter fitting can greatly benefit from MCMC techniques, but several barriers of entries prevent it from being commonly adopted. Most existing packages suffer from shortcomings related to one of two categories. First, many packages exist in "prototyping language" such as Python, R, or MATLAB in order to facilitate the coding process. Prototyping an MCMC simulation is more convenient, but the user pays for it greatly with the lengthy simulation time. The second category of packages are those written in "fast languages" such as C(++) or FORTRAN. MCMC simulations compiled from these languages are considerably faster, but suffer from the opposite problem: difficulty of programming. A large barrier to entry needs to be overcome before research code can pass through the layers of abstraction into a "fast language." APT-MCMC is a C++ implementation of MCMC with the aforementioned parallel tempering and affine-invariant ensemble of samplers [52, 122]. It aims to bridge the gap of programming accessibility by allowing the user to define their simulation solely within a Python package, which provides code conversion to compilable C++ code.

The rest of this section is structured as follows. First, a description of the APT-MCMC software is provided, including a brief user guide. Second, results of a standard series of

optimization benchmark are presented. APT-MCMC is shown to converge to the optimal parameter set and are shown to be comparable to a popular Python MCMC package, *emcee*. Finally, generalizable MCMC hyperparameters are explored and tested against these optimization benchmarks; hyperparameter tuning heuristics are provided.

6.2.1 Methods

6.2.1.1 APT-MCMC Package

APT-MCMC (**A**ffine-invariant ensemble of samplers with **P**arallel **T**empering **M**arkov **C**hain **M**onte **C**arlo) began as a C++ implementation of MCMC merging PTempEst, a parallel tempering MCMC algorithm with Boltzmann distribution acceptance in MATLAB, and *emcee* (`http://dan.iel.fm/emcee/`), a (parallel tempering optional) affine invariant ensemble of samplers MCMC algorithm in Python [52, 117]. APT-MCMC initializes with n_{chains} concurrent ensembles at temperatures arranged in a geometric progression. The geometric step was taken from the source code of *emcee* [117]. Within each ensemble, $n_{ensemble}$ samplers simultaneously generate the next proposed parameter step according to the "stretch-move" technique from Goodman and Weare [122]. In short, the stretch-move utilizes the location of another parameter set within the same ensemble to determine a movement direction. The proposed parameter is the result of a random movement along this direction and is determined by a step size hyperparameter. This value, z, allows movement between $[\frac{1}{z}, z]$ times the distance between the two parameters. Proposed parameter acceptance is governed by Equation (6.21). Samplers within each temperature ensemble are allowed to move independently of other temperature ensembles for n_{steps}, after which, a parameter location swap is attempted between another sampler in a higher temperature ensemble. The probability of this swap is governed by Equation (6.22). Table 6.1 summarizes the nomenclature used in this work.

Despite the impossibility of determining convergence in an MCMC simulation, APT-MCMC computes two statistical tests to provide evidence for convergence: integrated autocorrelation and the potential scale reduction factor (PSRF) of the simulation [115, 123]. The PSRF is a value that compares the parameter distributions between and within ensembles.

Table 6.1: Overview of the Markov Chain Monte Carlo nomenclature used.

Nomenclature		Meaning
n_{steps}	Hyperparameter	Number of MCMC sampler steps to take before a parameter location swap is attempted between tempered chains
$n_{ensemble}$	Hyperparameter	Size of MCMC ensemble of samplers within each tempered chain
n_{chains}	Hyperparameter	Number of tempered chains to simultaneously explore the problem
Swap Length	Hyperparameter	Number of proposed parameter location swaps among tempered chains; determines simulation length
τ	Statistical Measure	Integrated autocorrelation, measure the how well an MCMC simulation explored the parameter space (lower is better)
PSRF	Statistical Measure	Potential scale reduction factor, a measurement of simulation variance between tempered chains and within-chain ensembles (lower is better)

Values close to 1.0 are desired because this statistic provides evidence for convergence by indicating that the inter- and intra-ensemble distributions are identical [123]. APT-MCMC calculates the PSRF value using the Welford's variance algorithm. Integrated autocorrelation time, τ, measures the correlation of a signal at every possible time lag and is calculated according to Equations (6.23)-(6.24), where θ_t refers to the state of the Markov Chain (a parameter vector) at a point in time and M represents the length of the MCMC simulation. A property of Markov Chains is that each state depends solely on the one before it. As a result, a high autocorrelation indicates that the Markov Chain was correlated to multiple previous states, suggesting that the simulation was stuck in a certain parameter region. In contrast, a low autocorrelation indicates that the simulation sampled the parameter efficiently and provides evidence that the simulation has reached the true distribution of the system [115]. APT-MCMC uses Goodman's Acor C code in order to calculate the integrated autocorrelation value of the simulation [122].

$$C(t) = \frac{1}{M-t} \lim_{t' \to \infty} \left[\theta_{t+t'} - \langle\theta\rangle\right]\left[\theta_{t'} - \langle\theta\rangle\right] \tag{6.23}$$

$$\tau = 1 + 2\sum_{t=1}^{\infty} \frac{C(t)}{C(0)} \tag{6.24}$$

A property of Markov Chains is that each state depends solely on the one before it. As a result, a high autocorrelation indicates that the Markov Chain was correlated to multiple previous states, suggesting that the simulation was stuck in a certain parameter region. In contrast, a low autocorrelation indicates that the simulation sampled the parameter efficiently and provides evidence that the simulation has reached the true distribution of the system [115]. APT-MCMC uses Goodman's Acor C code in order to calculate the integrated autocorrelation value of the simulation [122].

ODE solving is handled by Sundials CVODE v2.6.2 [124]. Each temperature ensemble can be run in parallel and is handled by OpenMP. It is recommended to have at least one processor thread per temperature ensemble. The stretch-move step is nested-parallelizable if additional threads are available (threads $\geq$ 2x number of temperature ensembles). Initialization code to set the sequence of temperatures for a simulation was ported from Python package *emcee* [117]. Random number generation is handled by TRNG, a package designed for use with parallel Monte Carlo simulations [125]. TRNG provides a leapfrog technique, which ensures statistically-independent pseudo-random number streams for each parallel thread. Initial seeds for random number generators are provided by the current time plus the thread number.

As depicted in Figure 6.3, the user is responsible for providing the parameters and their prior distributions, the likelihood/objective function (the $P(D|\theta)$ portion from equation 6.18), fitting data, and MCMC options/hyperparameters to setup a simulation. This can be performed in C++ or by using the Python auto-generator package tailored for ODE inverse problems. APT-MCMC has been extensively tested on a Linux-64 platform, but it compiles and runs under OSx with the appropriate developer tools and under Windows 10 x64 under Cygwin and the Windows 10 Ubuntu Subsystem.

For user convenience, APT-MCMC comes with a Python 3 package to automatically generate the necessary C++ files to run a simulation. Equation 6.25 shows the type of

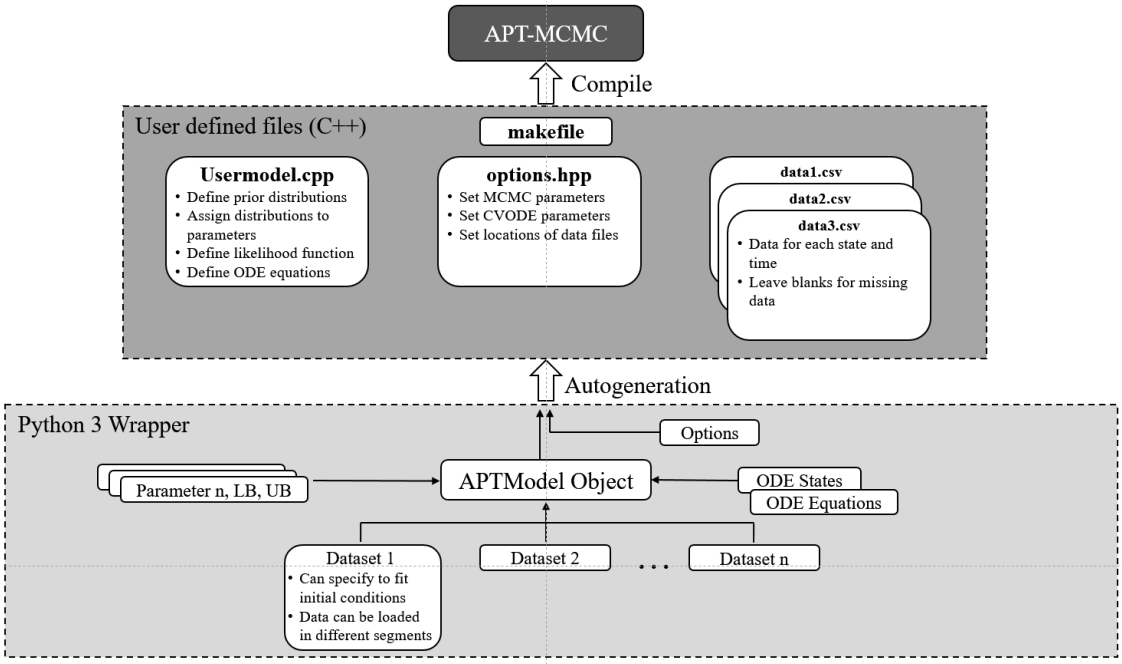


Figure 6.3: Schematic of generating an APT-MCMC simulations. Users need to specify their parameter-fitting problem using Usermodel.cpp and set MCMC options in options.hpp. The Python 3 package can autogenerate these files by providing the ODE states and their equations, parameter lower and upper bounds, and the dataset(s) for fitting.

problems that the package generates.

$$\begin{aligned} \underset{p}{\text{minimize}} \quad & \sum_t \left[\log(\text{Data}(t)) - \log(X(t))\right]^2 \\ \text{subject to} \quad & \dot{X}(t) = f(X, \theta, t, u(t)) \\ & X(0) = X_0 \end{aligned} \tag{6.25}$$

X represents states, p represents a parameter set, t represents time, u represents an input (optional). This encompasses most parameter-fitting usage cases. The user defines the parameters, ODE equations, fitting data, and MCMC options within an "aptmodel" object. Data handling is done by defining experiment(s), where each experiment represents a separate dataset to simultaneously fit. This Python object will then generate readable and compilable C++ code. Advanced features such as implementation of ODE inputs, fitting unknown initial conditions, and adjusting CVODE integration options are supported within

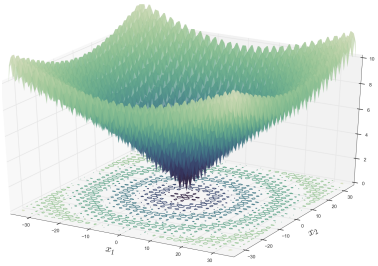
the Python interface. Within the objective function, the sum of squared error term is calculated after taking the natural log of the data and states to prevent numerical overflow. This addresses magnitude differences between data and model states, but may be toggled off within the Python code.

At runtime, APT-MCMC provides information on the status of each temperature ensemble, current energies, as well as the location of the best parameter set found since starting the run. Best parameter sets are identified to the θ that minimizes Equation (6.25). Once the simulation is completed, all relevant simulation results and variables are saved in a binary file with a ".mcmc" extension. Python and MATLAB code is provided to read this file into a results object for post-processing and analysis.

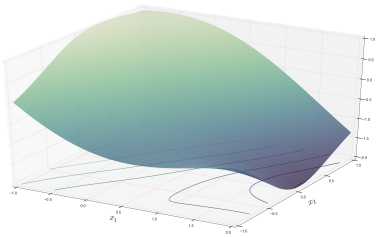
6.2.1.2 APT-MCMC Performance Evaluation

To evaluate the performance of APT-MCMC, several optimization benchmark functions were tested. They were selected from the list provided by Jamil, et al [126]. APT-MCMC was tested for its ability to recover the optimal parameter set for each function and for its runtime and memory usage. The functions tested were Ackley 1, Adjiman, Alpine 1, Bard, Beale, Bird, Bohachevskey 3, Booth, Bukin 6, Corana, Damavandi, Devilliers-Glasser, Eggholder, and Griewank (most functions are visualized in Figure 6.4). They were chosen to represent a large variety of functional landscapes: basins, valleys, multiple optima, dimensionality, and differentiability. Function equations may be found in Appendix B. Simulations were performed for each benchmark function in order to test if APT-MCMC reaches the optimum parameter set. All simulations were performed with the hyperparameter values noted in Table 6.2. Parameter histograms for each benchmark are available in the Appendix B. A bounded uniform prior was used to provide upper and lower bounds on the parameters specified by each benchmark function. Simulation autocorrelation and PSRF was calculated for each benchmark.

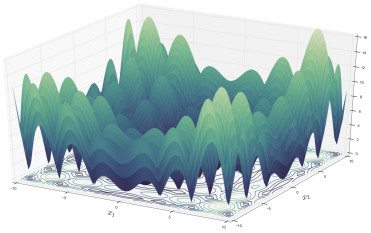
Speed and memory for the benchmarks were compared against the popular Python MCMC package *emcee*, which also features an MCMC algorithm using parallel tempering and sampler ensembles. However, several differences exist between the two packages. First, *em-*



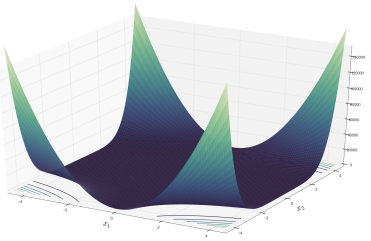
(a) Ackley 1



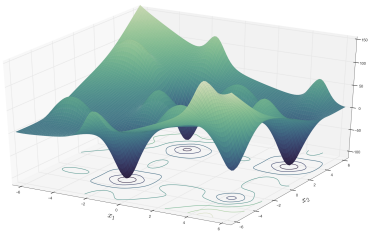
(b) Adjiman



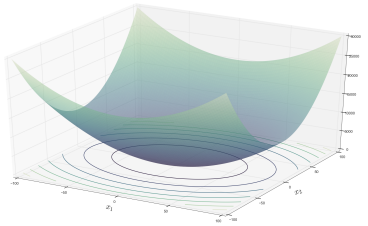
(c) Alpine 1



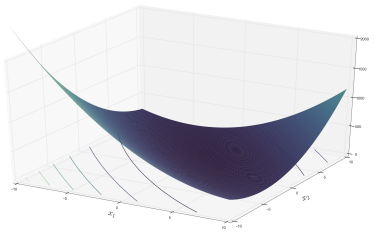
(d) Beale



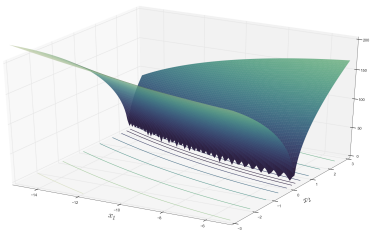
(e) Bird



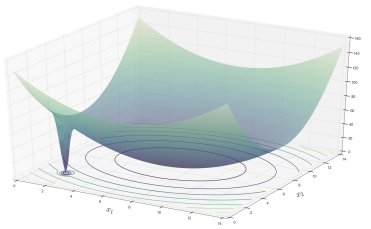
(f) Bohachevsky



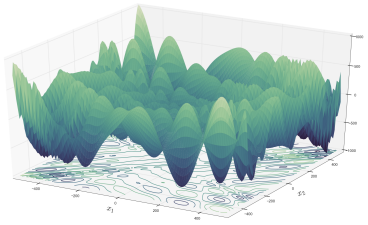
(g) Booth



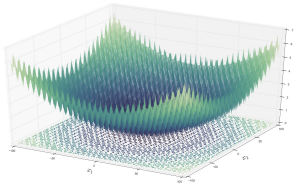
(h) Bukin 6



(i) Davamandi



(j) Eggholder



(k) Griewank

Figure 6.4: Visualizations for the benchmarks tested. Bard, Corana, Devilliers-Glasser are not shown due to high parameter dimensionality.

cee attempts a temperature swap after every attempted stretch-move whereas APT-MCMC will make a user-specified number of "stretch-moves" before attempting a swap. Second, the

Table 6.2: Simulation settings for benchmark tests.

Hyperparameter	Setting
n_{chains}	4
$n_{ensemble}$	100
n_{steps}	25
Burn-In Phase	1000
Swap Length	10^{4}*

*: Devilliers Glasser swap length set to 10^{6}

"burn-in" phase in *emcee* is performed by running the MCMC simulation and initializing at the last location. Within APT-MCMC, this is handled by sampling from the parameter priors for a user-specified number of times and then initializing at the locations of the best samples. To ensure comparable testing, benchmark simulations were repeated with different settings from that of Table 6.2. n_{steps} was set to 1 in APT-MCMC to attempt a temperature swap after every single stretch-move. "Burn-in", while not entirely similar between the two algorithms, was set to 1 in both *emcee* and APT-MCMC, which effectively initializes the MCMC simulation with a single sample from the prior. Simulation settings were set to four temperatures, eight processor threads, an ensemble size of 100, and a total swap length of 1e4.

At the end of the tests, this swap length was determined to be sufficient for both algorithms to reach convergence for each benchmark. All of the aforementioned benchmarks were tested in triplicate for each algorithm. All tests were performed on the same machine with the following specifications: 2x Intel Xeon E5-2670 v3, 256GB DDR4 memory, and Ubuntu 14.04. APT-MCMC and *emcee* memory benchmarks were performed by requesting the resident memory usage from the Linux kernel via the *readproc* command.

The hyperparameters required by APT-MCMC were tested to provide heuristic guidelines. Specifically, APT-MCMC introduces certain hyperparameters that the user must set: $n_{ensemble}$, n_{chains}, step size, and n_{steps}. Each hyperparameter was tested over a range of values

and each value was tested with 100 replicate simulations. For comparability, the random number generator for each replicate simulation was seeded from a predetermined value for each replicate. All other hyperparameters were kept equal according to Table 6.2 (unless otherwise noted). To quantify the effects of each hyperparameter, two statistics were used: simulation parameter percent error (or deviation if optimum parameter is 0) and maximum parameter autocorrelation.

The ensemble size parameter was tested as a function of problem dimensionality. The Ackley benchmark was chosen for its dimensional scalability. A 5, 10, 20, and 30 dimensional problem was tested and ensemble size was set at 1x, 2x, 5x, 10x, and 20x the number of parameters for each problem. In order to prevent ensemble sizes from biasing the number of simulation samples, swap length was adjusted each time to ensure that a total of 600,000 samples were taken per simulation. The step size parameter controls how far the MCMC samplers traverse the parameter space with each iteration. It was tested against a parameter fitting benchmark, the Bard function. The parameter dictating the number of moves per swap attempt controls the swap attempt frequency. This adjusts how long the lower temperature simulation is able to search a parameter location of interest before it potentially jumps to another optimum. This technique can greatly aid in highly multimodal problems and therefore the 10 dimension Ackley and the Griewank benchmarks were tested. The number of temperatures may be beneficial in highly multimodal problems due to the higher temperature ensemble's ability to identify optima. It was tested against the 10 dimension Ackley and the Bukin benchmarks.

$$\begin{aligned} A &\xrightarrow[k_{AB}]{} B \xrightarrow[k_{BC}]{} C \\ 2A &\xrightarrow[k_{AD}]{} D \end{aligned} \tag{6.26}$$

APT-MCMC was tested against a sample ODE fitting problem using the classic Van de Vusse reaction scheme (see Equation 6.26). It was assumed that this reaction takes place within an isothermal continuously stirred tank reactor and the feed is pure A with the following properties: $\frac{F}{V} = 4/7\text{min}^{-1}$ and $C_{AI} = 10$ mol/L. The resulting ODE system is presented in Equation 6.27. Unsteady startup data for this system was generated for

states $A(t)$, $B(t)$, $C(t)$, and $D(t)$ at 0.5 minute increments from 0 to 10 minutes and subject to a sampling noise of $\mathcal{N}(\mu = 0, \sigma^2 = 0.1)$. APT-MCMC fit the three kinetic parameters k_{AB}, k_{BC}, k_{AD} to concentration data of species. Simulation settings were set the those in Table 6.2 with the exception of swap length set at 10^3. Each parameter was bounded within $[0, 10]$.

$$\begin{aligned}
\frac{dC_A(t)}{dt} &= \frac{F}{V}(C_{AI} - C_A(t)) - k_{AB}C_A(t) - k_{AD}C_A(t)^2 \\
\frac{dC_B(t)}{dt} &= -\frac{F}{V}C_B(t) - k_{BC}C_B(t) + k_{AB}C_A(t) \\
\frac{dC_C(t)}{dt} &= -\frac{F}{V}C_C(t) + k_{BC}C_B(t) \\
\frac{dC_D(t)}{dt} &= -\frac{F}{V}C_D(t) + \frac{1}{2}k_{AD}C_A(t)^2 \\
k_{AB} &= 0.833, k_{BC} = 1.667, k_{AD} = 0.167
\end{aligned} \tag{6.27}$$

APT-MCMC also tested a bioreactor system describing glucose to ethanol fermentation by *Saccharomyces cerevisiae* in a batch tank reactor [127]. The equations describing this reactor are shown in Equation 6.28. The goal was to fit six parameters: $\mu_{max}, K_S, k_d, Y_{S/C}, m, Y_{P/C}$ against time-series data for each state, $C_C(t)$ (yeast concentration), $C_S(t)$ (glucose concentration), $C_P(T)$ (ethanol concentration), at 0.5 hour increments up to 14 hours. Noise sampled from $\mathcal{N}(\mu = 0, \sigma^2 = 2)$ was added to the data. The initial yeast, glucose, and product concentrations were set to 1.0 g/dm^3, 250 g/dm^3, and 0.0 g/dm^3. Each parameter prior was set to a bounded uniform prior about $[0, 20]$ to test APT-MCMC's ability to recover the parameters given a large uninformative prior.

$$
\begin{aligned}
r_g(t) &= \mu_{max}\left(1-\frac{C_P(t)}{C_P^*}\right)^{0.52}\frac{C_C(t)C_S(t)}{K_S+C_S(t)} \\
\frac{dC_C(t)}{dt} &= r_g - k_d C_C(t) \\
\frac{dC_S(t)}{dt} &= -Y_{S/C}r_g - mC_S(t) \\
\frac{dC_P(t)}{dt} &= Y_{P/C}r_g \\
\mu_{max} &= 0.33h^{-1}, K_S = 10.7g/dm^3, k_d = 0.01h^{-1}, \\
Y_{S/C} &= 12.5g/g, m = 0.03g/(\text{g cells h}), Y_{P/C} = 0.45g/g, C_P^* = 93g/dm^3
\end{aligned}
\tag{6.28}
$$

6.2.2 Results

6.2.2.1 Benchmarks

Table 6.3: Overview of APT-MCMC Performance on several benchmark functions.

Benchmark	τ	PSRF	Param 1	Param 2	Param 3	Param 4	Param 5
Ackley	1.16	1.00	$5.58*10^{-3}$ (0)	-1.53 $*10^{-2}$ (0)	9.78 $*10^{-1}$ (0)		
Adjiman	1.85	1.03	2.00 (2)	$9.60*10^{-2}$ (0.106)			
Alpine	2.34	1.00	$6.80*10^{-2}$ (0)	0.143 (0)	$3.19*10^{-2}$ (0)	$1.52*10^{-2}$ (0)	$2.67*10^{-2}$ (0)
Bard	1.85	1.20	$8.11*10^{-2}$ ($8.24*10^{-2}$)	1.09 (1.13)	2.38 (2.34)		
Beale	1.95	1.06	3.00 (3)	0.500 (0.5)			
Bird	4.05	1.02	-1.58 (-1.58)	-3.13 (-3.13)			
Bohachevsky	1.18	1.00	$1.86*10^{-3}$ (0)	$1.33*10^{-3}$ (0)			
Booth	1.85	1.02	1.00 (1)	3.00 (3)			
Bukin	1.85	1.05	-10.56 (-10)	0.753(1)			
Corana	1.93	1.00	3.86 $*10^{-2}$ (0)	$8.42*10^{-3}$ (0)	$3.29*10^{-2}$ (0)	$4.76*10^{-2}$ (0)	
Damavandi	1.85	1.00	2.00 (2)	2.00 (2)			
Devilliers	1.95	1.18	53.9 (53.81)	1.27 (1.27)	2.95 (3.01)	65.0 (2.13)	.507 (.507)
Eggholder	1.85	1.35	512 (512)	404 (404.23)			
Griewank	1.10	1.00	$-4.05*10^{-2}$ (0)	$3.58*10^{-2}$ (0)			

Benchmark functions are provided in Appendix B. For each benchmark, the integrated autocorrelation time(τ), potential scale reduction factor (PSRF), the best parameter set, and the true optimum in parenthesis are reported. Each simulation was performed using the hyperparameters in Table 6.2.

Table 6.3 shows the maximum likelihood parameter for each benchmark function. The swap length of simulations were arbitrarily chosen to be 10,000; however, the low values of the autocorrelation and PSRF statistics in Table 6.3 support that simulations converged over this run length. APT-MCMC returned accurate parameter estimates for most benchmarks. For example, the Damavandi benchmark contains a very large local minimum well near $(7, 7)$(Figure 6.4i), but APT-MCMC was able to locate the global minimum at $(2, 2)$. The Bukin benchmark (Figure 6.4h) contains a sharp ridge with multiple minima and APT-MCMC was able to get close to this point.

The notable exception is with the DeVilliers-Glasser function, represented in equation 6.29. This benchmark is a sum of squares parameter fitting problem and, despite not being previously reported, appears to contain multiple global minima. The highly multimodal histogram (see Appendix B) for parameter x_4 and the bimodality of x_1 suggest that this problem has multiple solutions. The hyperbolic tangent is an odd function and its sym-

metricity explains the bimodal distribution at x_1 ± 54. The cyclical nature of $\sin(x)$ explains the multimodal nature of x_4.

$$\begin{aligned}
&\sum_{i=1}^{24} \left[x_1 x_2^{t_i} \tanh[x_3 t_i + \sin(x_4 t_i)] \cos(t_i e^{x_5}) - y_i\right] \\
&t_i = 0.1(i-1) \\
&y_i = (53.81)(1.27^{t_i}) \tanh[3.012 t_i + \sin(2.13 t_i)] \cos(t_i e^{0.507}) \\
&x_i \in [-500, 500] \quad i = 1, 2, ..., 5
\end{aligned} \tag{6.29}$$

The head-to-head comparison tests between APT-MCMC and *emcee* is summarized in Table 6.4. For each benchmark tested, APT-MCMC was many times faster than the Python counterpart. This speed came at the cost of using 14% more memory during simulations. The reported times for APT-MCMC do not include generation time (time required for Python to generate C++ code) nor do they include C++ compilation time, but the inclusion of those times (roughly an extra 5-7 seconds) will not significantly change Table 6.4. These results are expected due to Python overhead as an interpreted language. While useful for prototyping, the slower execution speed is less desirable MCMC simulations. As a result, APT-MCMC may lack some of the flexibility of *emcee* but succeeds in being the faster tool at runtime. Lastly, the symmetric mean absolute percentage error was computed for each benchmark simulation instead of percentage error to deal with division by zero issues. The error calculations demonstrate that APT-MCMC is able to achieve similar a computational performance with *emcee*.

6.2.2.2 Hyperparameters

The $n_{ensemble}$ hyperparameter dictates how many simultaneous samplers exist for each temperature ensemble. These samplers utilize each other‘s locations to generate search directions for the next state. Test results for $n_{ensemble}$ are shown in Figure 6.5. Autocorrelation increases with problem dimensionality due to the increasing amounts of local minima that are introduced. Parameter error was unaffected by ensemble size. In general, the higher the ensemble sizes (as a function of problem dimensionality) resulted in lower integrated

Table 6.4: Comparison of computing resources used by APT-MCMC and Python package *emcee*.

	emcee			APT-MCMC		
Benchmark	Time (s)	Memory (MiB)	Error	Time (s)	Memory (MiB)	Error
Ackley	118	189	$9.63 * 10^{-01}$	2	193	$9.63 * 10^{-01}$
Adjiman	125	159	$9.73 * 10^{-03}$	2	194	$9.50 * 10^{-03}$
Alpine	110	251	$9.63 * 10^{-01}$	3	255	$9.63 * 10^{-01}$
Bard	118	190	$7.05 * 10^{-01}$	2	193	$1.86 * 10^{-02}$
Beale	128	160	$2.58 * 10^{-04}$	2	193	$5.34 * 10^{-04}$
Bird	124	157	$1.14 * 10^{-05}$	2	193	$6.67 * 10^{-01}$
Bohach	116	158	$9.63 * 10^{-01}$	2	193	$9.63 * 10^{-01}$
Booth	116	158	$2.06 * 10^{-04}$	2	193	$1.98 * 10^{-04}$
Bukin	115	158	$4.89 * 10^{-01}$	2	193	$1.01 * 10^{-01}$
Corana	118	220	$9.63 * 10^{-01}$	3	255	$9.63 * 10^{-01}$
Damavandi	116	157	$5.35 * 10^{-01}$	2	193	$2.86 * 10^{-04}$
deVilliersGlasser	187	252	$8.90 * 10^{-01}$	9	255	$2.88 * 10^{-01}$
Eggholder	114	157	$6.48 * 10^{-06}$	2	193	$1.40 * 10^{-04}$
Griewank	116	158	$9.63 * 10^{-01}$	2	193	$9.63 * 10^{-01}$
Average	123	180	$5.32 * 10^{-01}$	2	206	$4.21 * 10^{-01}$

Results shown are averaged from $n = 3$ test simulations. Memory was the physical resident memory in mebibytes, as reported by the Linux kernel. Each benchmark function simulation was set according to the hyperparameters in Table 6.2 with the exception of $n_{steps} = 1$ (for comparability between APT-MCMC and *emcee*). Burn-in for both algorithms were set to 1 iteration. APT-MCMC was faster and used less memory for each case. The error columns represent the calculated symmetric mean absolute percentage error averaged over all parameters and across replicates.

autocorrelation values. Larger ensemble sizes reduced autocorrelation while utilizing the same amount of computational resources (simulation times were similar among each problem dimensionality). The reduced autocorrelation is because larger ensembles offer more varied and robust search directions. For smaller problems such as the 5 parameter case, there were no distinguishable effects due to the ease of convergence. For large problems, preference should be given towards increasing ensemble size rather than swap length. As a reference, the source code behind *emcee* enforces the minimum ensemble size at 2x the problem parameter dimensionality [117].

The n_{chains} hyperparameter in a simulation helps the system explore highly modal parameter spaces by running many tempered simulations simultaneously. Samplers in high-temperature ensembles can locate alternative energetically favorable regions (potential lo-

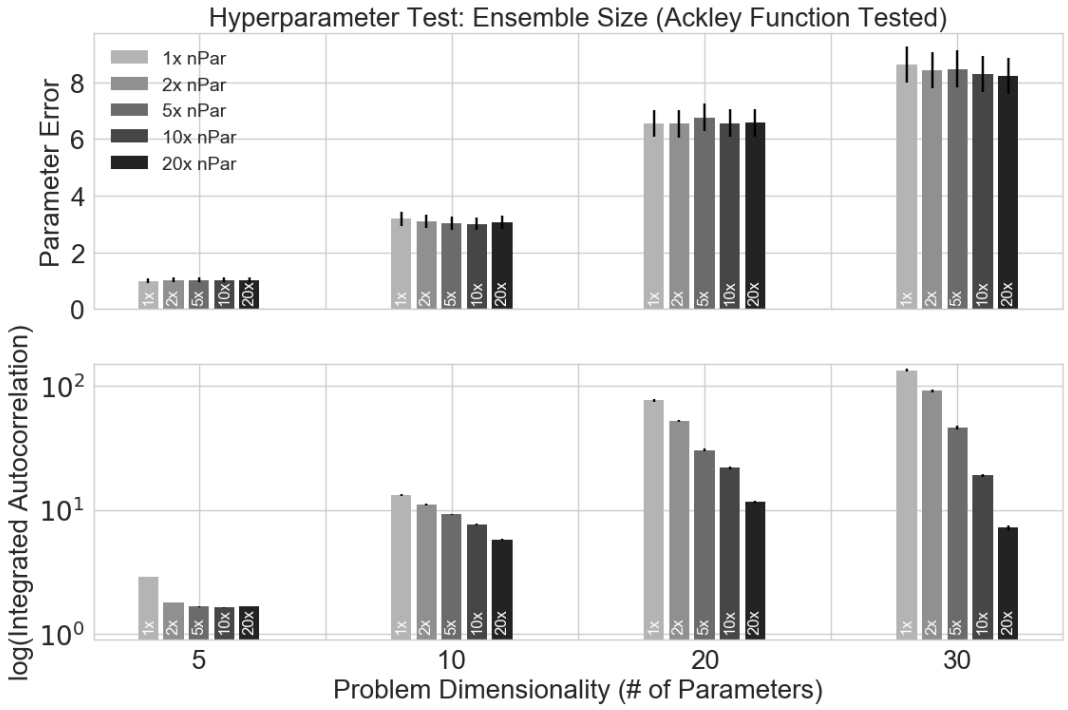
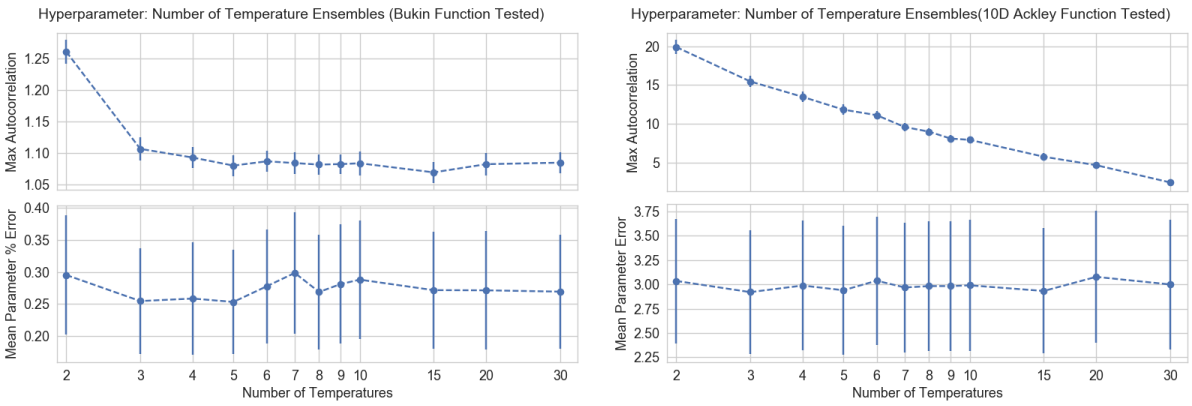


Figure 6.5: 100 repeated simulations were used to test the effects of the $n_{ensemble}$ hyperparameter by using the dimension-scalable Ackley benchmark function. The value of $n_{ensemble}$ was set as a function of the dimensionality of the Ackley function and denoted by #x nPar (multiplier x # of parameters). Mean ± SEM of best parameter error and integrated autocorrelation shown. Raw parameter error is reported because the optimum parameters are located at 0. Problems benefit from larger ensemble sizes, but this comes at the cost of increased computation time.

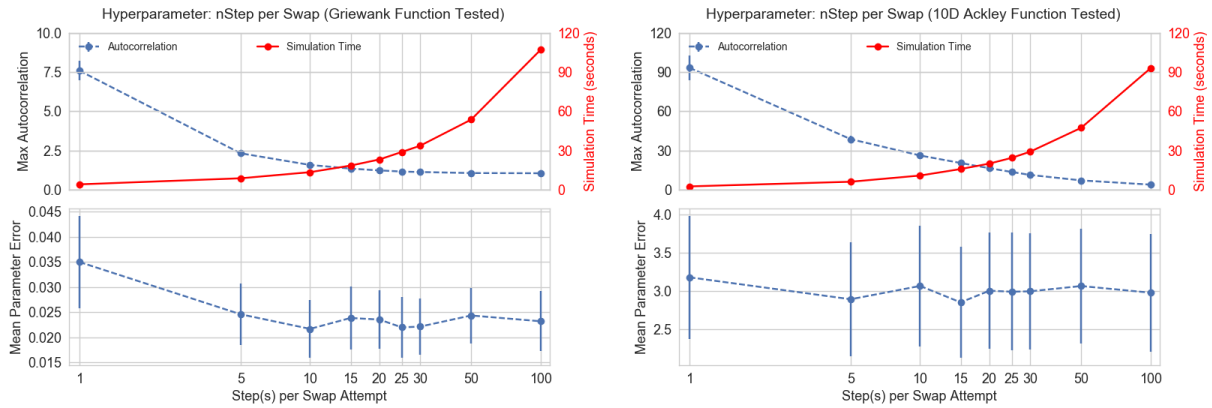
cations of local minima) with higher probabilities and pass this information to the lower temperature samplers for minima identification. Figure 6.6a shows the results of the Bukin benchmark. There were diminishing decreases of autocorrelation at a higher number of chains, which indicate that this benchmark function had local minima that were relatively easy to escape from. In contrast, Figure 6.6b shows the results for the 10-dimension Ackley benchmark and demonstrates a linear relationship between autocorrelation and the number of chains. The 10-dimension Ackley benchmark contains more local minima than the Bukin benchmark and benefits more from the additional chains. n_{chains} cannot be set lower than 2 because parameter swapping would not be possible with a single temperature Markov



(a) 2-dimension Bukin benchmark function (b) 10-dimension Ackley benchmark function

Figure 6.6: 100 repeated simulations were used to test the effects of the n_{chains} hyperparameter by using benchmark functions Bukin and Ackley. Mean ± SEM of integrated autocorrelation shown and parameter error shown. Additional temperature ensembles decreased simulation autocorrelation, although there were diminishing returns after 4. A higher number of temperatures results in improved parameter mixing and simulation efficiency. There was no significant effect on the accuracy of the maximum likelihood parameter vector.

Chain. Increasing n_{chains} increases the magnitude of the temperature sequence and higher magnitudes benefit from the aforementioned favorable region exchange. This information exchange leads to a faster exploration of the parameter space and therefore lower autocorrelation values. While the number of tempered chains increases computational cost, they can in parallel and scale well via OpenMP. As a result, this hyperparameter can be set to equal the number of CPU threads available on the machine running the simulation. APT-MCMC is able to utilize up to the number of detected CPU threads (logical cores) on the machine it runs on. OpenMP synchronization locks in the APT-MCMC OpenMP code ensure that all chain movements are completed prior to a swap attempt. Setting n_{chains} to a value that exceeds the number of CPU threads will greatly increase simulation time because some threads have to calculate multiple tempered chains while others are forced to idle. Computational resources are better spent on longer simulations with an n_{chains} equivalent to the number of cores. It is recommended to set n_{chains} to as many available CPU cores as possible because simulation efficiency improves and parallelization offsets any computational trade-offs.



(a) 2-dimension Griewank benchmark function (b) 10-dimension Ackley benchmark function

Figure 6.7: 100 repeated simulations were used to test the effects of the n_{steps} hyperparameter by using benchmark functions. Mean ± SEM of best parameter error and integrated autocorrelation shown and parameter error shown. Increasing the steps per swap decreased maximum autocorrelation but increased the simulation time. Parameter error decreased only in the Griewank case.

The n_{steps} hyperparameter describe how much a sampler ensemble explores a local region of the parameter space before it is potentially swapped to a different parameter location. The effect of this hyperparameter is shown in Figure 6.7a for the Griewank benchmark. Both autocorrelation and parameter error decreases with increasing n_{steps}. In contrast, simulation time increased. Figure 6.7b shows the results for the 10-dimension Ackley function, which demonstrate similar trends. Larger n_{steps} allow for the ensemble of samplers to explore a region of interest for longer before they are potentially transferred to a different region. Very small values may prevent the samplers from exploring optima sufficiently. As n_{steps} became larger, autocorrelation was unaffected, but simulation time increased. For both benchmark functions, there were diminishing returns in autocorrelation decreases compared to the increases in simulation time. Values within 20-30 offered trade-offs between autocorrelation, simulation time, and parameter accuracy.

The step size hyperparameter determines how far a proposed parameter can move in accordance with the stretch-move scheme. Small step sizes prevent proper parameter space exploration. High step sizes allow potential movements farther away from the rest of the en-

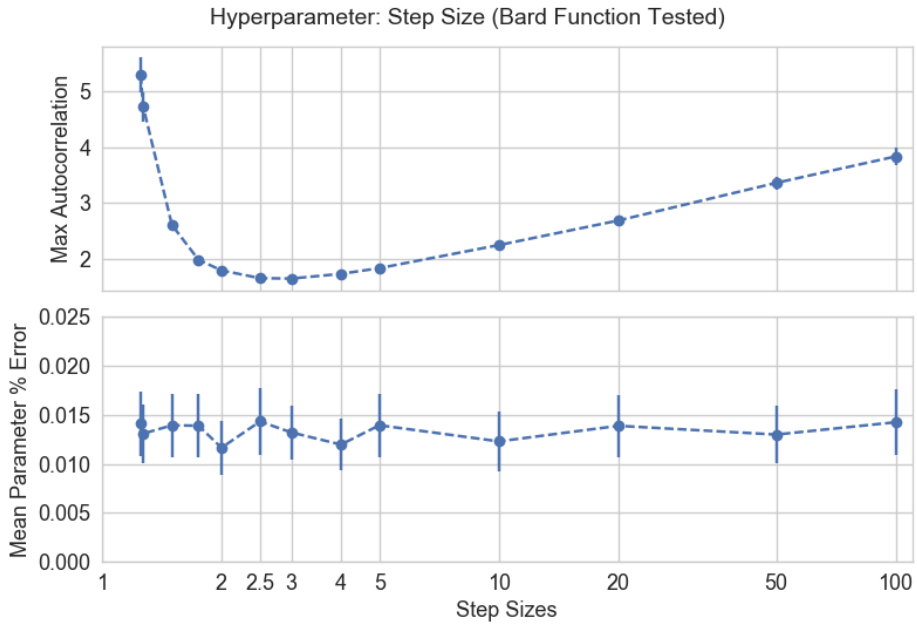


Figure 6.8: 100 repeated simulations were used to test the effects of the step size hyperparameter by using the 2-dimension Bard benchmark function. Mean ± SEM of best parameter percent error and integrated autocorrelation shown. The minimum autocorrelation was achieved at a clear optimum at 2.5-3. There was no effect on the accuracy of the maximum likelihood parameter vector.

semble samplers. Figure 6.8 shows the results for the Bard benchmark. The suggested value for this hyperparameter is between 2-3 where there is a clear minimum in autocorrelation. As a reference, Goodman and Weare suggested a value of 2 for this hyperparameter [122].

6.2.2.3 Van de Vusse Reaction Scheme

Figure 6.9 illustrates the Bayesian posterior results from an ODE parameter fitting problem involving the Van de Vusse reaction scheme. Parameter histograms are shown along with a fitted kernel density estimation. The peaks of each parameter's probability density correspond with the true parameters used to generate the data for the Van de Vusse reactions. Furthermore, the best parameter vector from the simulation (associated with the highest likelihood) demonstrated 1%-3% difference from the true parameter vector. The fit from the best parameter set is shown in Figure 6.10.

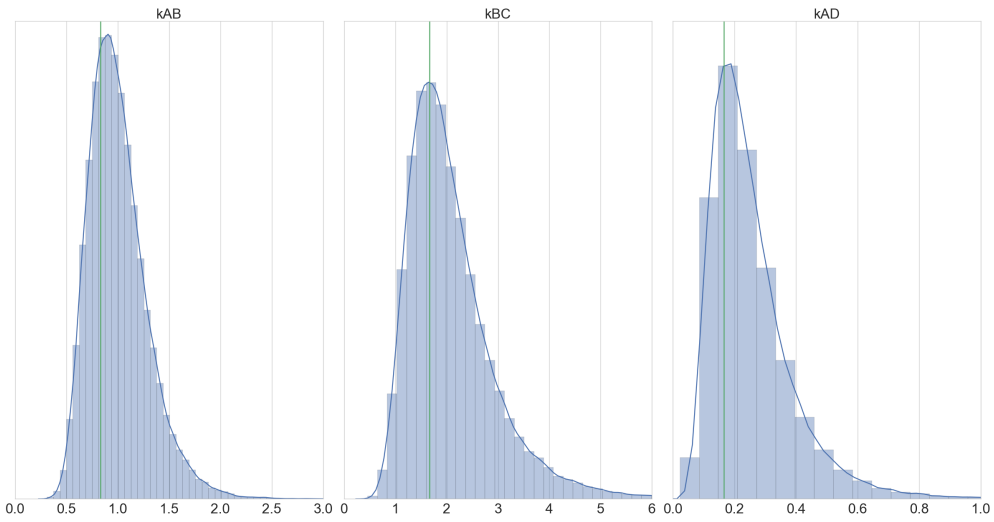


Figure 6.9: Parameter histogram from simulation to fit 3 parameters to Van de Vusse reaction data. The vertical line in each subplot represents the true parameter values used to generate the data. The smooth line in each subplot is a kernel density estimation of the histogram.

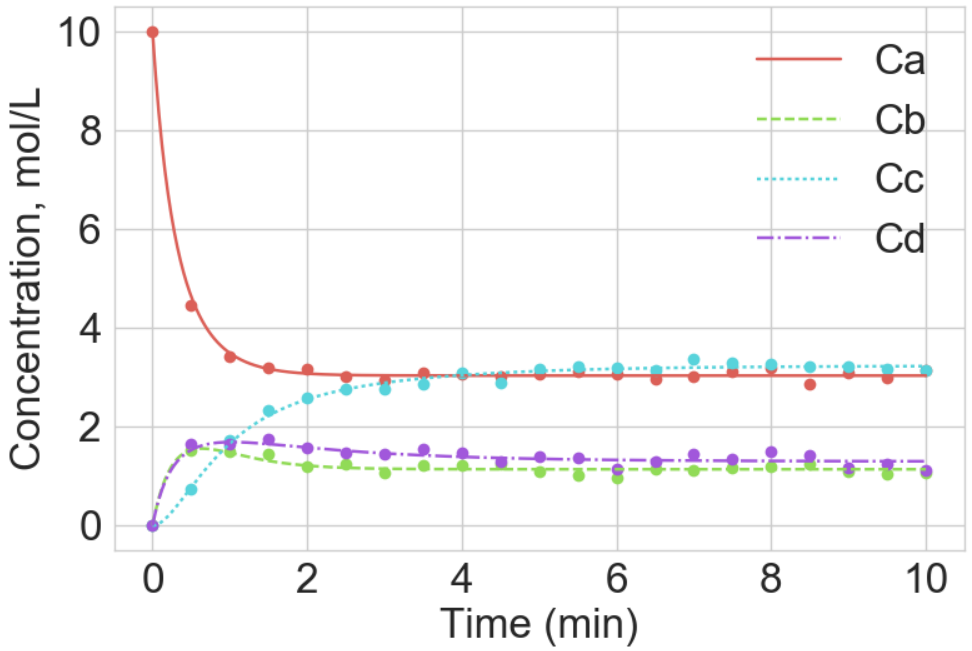


Figure 6.10: Data and best parameter fit for the Van de Vusse reaction scheme system. Simulated state data (with noise) are shown as dots in either graph. Concentration trajectories from the best parameter fit are shown as lines for each state.

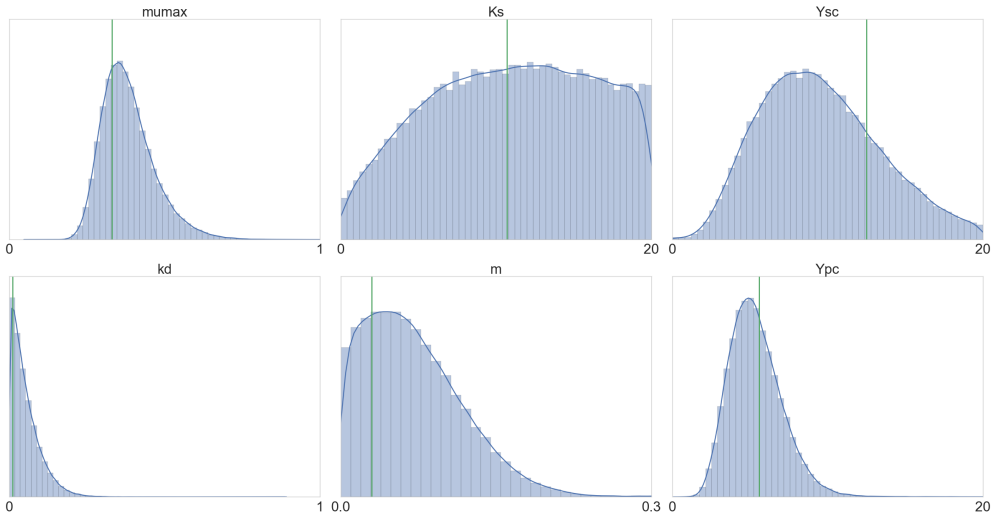


Figure 6.11: Parameter histogram from simulation to fit 6 parameters in a batch bioreactor. The vertical line in each subplot represents the true parameter values used to generate the data. The smooth line in each subplot is a kernel density estimation of the histogram.

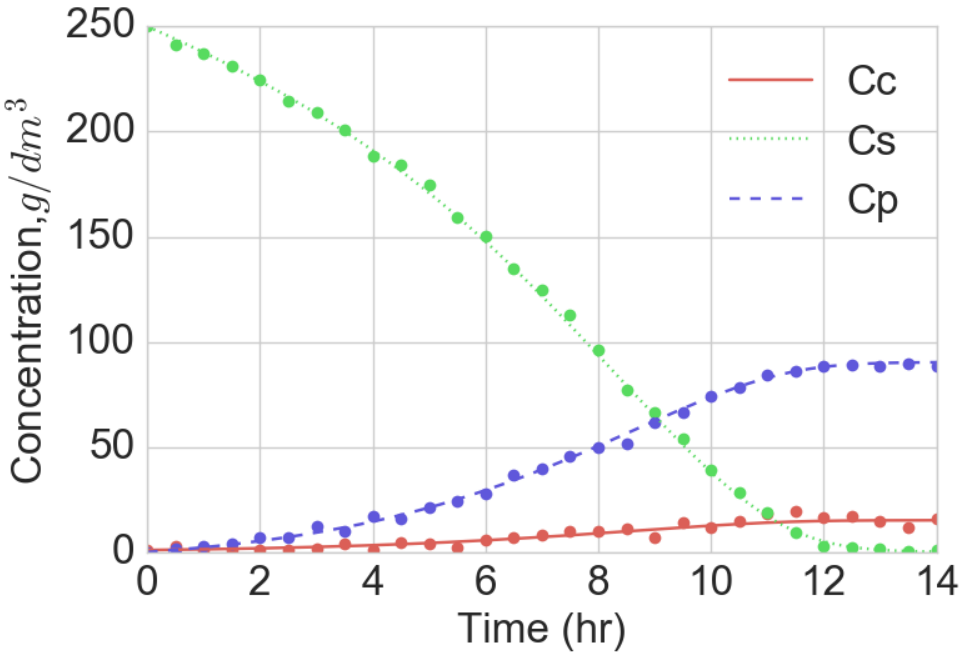


Figure 6.12: Data and best parameter fit for the bioreactor system. Simulated state data (with noise) are shown as dots in either graph. Concentration trajectories from the best parameter fit are shown as lines for each state.

6.2.2.4 Bioreactor

Figure 6.11 illustrates the Bayesian posterior results from the ODE fitting of a bioreactor. Histogram peaks for most of the parameters matched the true values. The histogram peaks for the $Y_{S/C}$ and $Y_{P/C}$ were under-predicted. This was likely due to the difficulty the algorithm had in fitting K_S, a Michaelis-Menten half-max concentration term, as indicated by the high variance within the K_S distribution. As a result, the biomass growth term, r_g, differed from the true growth rate and the two yield parameters were used to compensate. Michaelis-Menten half-maximum constants are difficult to fit without carefully characterized reaction rate data that are often experimentally challenging to collect. If such data were available, the parameter estimates would be significantly improved based on the tightening of the bounds on the K_S parameter. Regardless, this example highlights a major benefit of APT-MCMC. Compared with traditional nonlinear least squares fitting algorithms such as the trust region reflective or Levenburg-Marquardt, APT-MCMC does not require a good set of initial conditions to achieve a promising model fit; instead, the use of broad parameter bounds is sufficient to estimate parameter values and distributions, which is especially useful when some parameters may prove difficult to identify from the available data. The fit from the best parameter set is shown in Figure 6.12.

6.2.2.5 Limitations

One current limitation is associated with the Python auto-generation package. It assumes uniform priors on all parameters and a sum of squared error likelihood function. However, the C++ code can be easily modified to overcome these limitations. Should the user require additional features for advanced fitting procedures, the user can auto-generate the files with Python and then modify the appropriate C++ functions to their needs. Any problem that can be posed as a maximum likelihood can be programmed into APT-MCMC, but C++ programming of Usermodel.cpp is required.

Another limitation is the lack of a termination criterion, which is inherent to MCMC techniques. While APT-MCMC seeks to alleviate this issue by calculating the autocorrelation and PSRF values, MCMC statistics are necessary but not sufficient tests of convergence.

Practically, the user should set the simulation to a very long swap length. While the simulation proceeds, the user should analyze the partial APT-MCMC results as they become available and make a stopping decision based on the statistical tests and in determining if the distributions of the parameter posterior histograms are still changing. MCMC simulations may also get stuck in a meta-stable convergence state (which is not testable either), but a popular solution is to run multiple MCMC simulations from distinct initial conditions [115]. The second simulation can be run with tighter bounds on each parameter based on the posteriors from the first simulation and one can verify if simulations converge in the same parameter location.

6.2.3 Summary

APT-MCMC is presented as a fast MCMC platform tailored for use in solving parameter-fitting problems. User-convenience was prioritized during the creation of APT-MCMC. Traditionally, MCMC packages are provided in either easy-to-prototype, but slow, or difficult-to-learn, but fast, languages. The Python auto-generation package serves to provide users with an easy way to prototype and set up simulations while retaining the speed associated with a compiled and static language (C++). Additionally, MCMC hyperparameters are provided based on tests using common optimization benchmarks. APT-MCMC is ready for use and exists as an open source project on GitLab.

7.0 SUMMARY AND FUTURE WORK

The focus of this dissertation was the identification of sepsis endotypes. To achieve this, a variety of mathematical models from were developed to provide insight into the time course and the mechanisms of the syndrome. Hierarchical clustering revealed the existence of several distinct clusters primarily described by the level of cytokinemia (high, medium, low) and clinical features pertaining to the sequential organ failure assessment score. Membership within these clusters was confounded by the issue of variable pre-hospital time: sepsis patients arrive in the clinic at variable times along their inflammatory trajectory. In response, a systems analysis approach was developed to simultaneously model each subject (along with their pre-hospital times) and cluster the responses. This method revealed five sepsis endotypes with distinct behaviors of the pro- and anti-inflammatory response to infection. Endotypes are the different manifestations of disease due to mechanistic differences between patients. These five endotypes were explored with applied statistics, machine learning techniques, and ODE mathematical models. Applied statistics revealed that the traditional "baseline" approach (defined by when the subject arrives in the clinic) cannot segregate patients with different endotypes. Machine learning techniques, trained on six hours of temporal clinical features, also demonstrated difficulty in segregating endotypes. A mathematical ODE model revealed that macrophage cell death and cytokine degradation may contribute to the underlying pathomechanistic differences between endotypes.

In conjunction with the endotyping work, other mathematical models were explored to tackle other clinical challenges in sepsis: specifically, unknown pre-hospital time and unknown systemic damage for a septic subject. Knowledge of these two features can improve clinical decision-making, as well as improve the aforementioned mechanistic model. First, a nearest-neighbor model was constructed. Three sets of animal data and one set of human

trauma data demonstrated the feasibility of applying a clinical tool to estimate pre-hospital time. Second, an ensemble of regularized logistic regression models that were trained on baboon sepsis data yielded an exploratory approach to quantify the systemic inflammatory damage suffered during sepsis as a function of time.

Finally, to support this body of work, two mathematical tools were created. In exploring age-related immune pathway differences between old and young mice, a tool was created that automatically parses a list of potential immune pathways, formulates a mixed integer linear programming problem, and solves it via Pyomo and CPLEX. APT-MCMC was born out of the need for a faster MCMC ODE parameter-fitting algorithm during the development and calibration of the aforementioned mathematical ODE model of sepsis.

7.1 CONTRIBUTIONS

7.1.1 Endotype Identification and Prediction in the Clinic

The hierarchical heatmap clustering approach in Chapter 2 revealed that cytokinemia and organ dysfunction biomarkers in the ProCESS data contribute to differing 14-day all-cause mortality and multiple organ failure rates. Heatmaps provided a way to aggregate and visualize all of the factors a clinician takes into account when developing treatment plans for septic subjects. The distinctive patterns among clusters paved the way towards a novel clustering approach that (i) accounted for individual patient pre-hospital times, (ii) modeled patients using a mathematical model of the inflammatory response, (iii) recovered patient cluster dynamics from censored trajectories (Chapter 3). The result was five endotypes, each with their own characteristic master inflammatory-responses model.

The endotypes encompassed a variety of responses: overwhelming inflammation in endotype 1, sustained anti-inflammation in endotype 2, immunosuppression in endotype 3, sustained inflammation in endotype 4, and immunodeficiency in endotype 5 (refer to summary Table 3.2). The difference in response behaviors of these endotypes suggests that certain immuno-modulatory therapies may be inappropriate for certain endotypes (for ex-

ample, anti-inflammatory therapy is not suitable for an immunosuppressed subject). This insight may explain why many sepsis clinical trials fail to demonstrate positive mortality benefits. While some groups of subjects may have benefited from a certain therapy, others may have been adversely affected.

7.1.2 Mechanistic Ordinary Differential Equation Model of Sepsis

The proposed ODE model in Chapter 4 is among the first models of sepsis to be calibrated on human sepsis data. The model captures many of the complex interactions between white blood cells (neutrophils and monocytes), cytokines, and pathogens. The model, combined with the estimated pre-hospital times from Chapter 3, demonstrated an ability to fit many of the ProCESS patients well. White blood cell recruitment drew from marginal sources first during the first days of infection. After 4-5 days, immature white blood cell recruitment from the bone marrow occurred to replenish depleting marginal reserves. TNF was able to trigger monocyte recruitment into the tissue. Resident macrophages and the recruited monocytes (now macrophages) were able to sustain production of cytokines to emulate the cytokine storm. The model demonstrated an ability to capture the fast changes (both rise and fall within 6 hours) of both cytokines and white blood cell data. A population-level analysis of parameters revealed that macrophage death dynamics were important contributors to endotype-specific behavior. Furthermore, within some endotypes, cytokines seemed to have prolonged lifespans in the blood. Macrophage death dynamics and cytokine lifespans may begin to explain the causes of cytokinemia and may provide suitable pharmaceutical targets for future sepsis therapy.

7.1.3 Quantifying Pre-hospital Time and Systemic Damage

Chapter 5 describes two mathematical models to tackle other challenges that clinicians face with sepsis patients. A statistical model, based on nearest-neighbors, was introduced to address the issue of variable, unknown pre-hospital times in septic subjects. This approach was developed independently from the transfer function model in Chapter 3 to serve as a stand-alone method that precluded the need for serial cytokine measurements over a 72 hour

period. One-nearest-neighbors achieved high accuracy in estimating pre-hospital times in one baboon model of sepsis and two porcine models of sepsis. The measurement burden was demonstrably low: using temporal vital signs and a single, blood-based, point measurement was sufficient to achieve an accuracy between $70-95\%$, depending on the length of temporal data. Furthermore, this method worked well on human trauma data. Ambulatory and emergency room measurements of vital signs and organ-health biomarkers yielded up to 82% accuracy in estimating the original time of trauma. The translatability of this method, low measurement burden and temporal density requirements, and high accuracy properties of the nearest-neighbors approach indicate a promising way to provide clinicians with valuable pre-hospital/infection onset information about their sepsis patients.

The second statistical model related the risks of early mortality with important clinical features using an unbiased, feature selection algorithm. Time-varying odds ratios of mortality were calculated using ensembles of elastic-net logistic regression models. Elastic-net offered improved feature selection over traditional methods (ridge or lasso) and identified several key clinical features, such as lactate, glucose, hemoglobin, and platelet count, that contribute towards early sepsis deaths with varying odds ratios. Systemic organ damage, as a function of time, were calculated from odds ratios and revealed clear mortality separation within the first 12 hours of infection. Nonsurviving subjects demonstrated increasing-only damage trajectories, indicating that large amounts of early sepsis-induced organ damage may lead to a mortality bifurcation. Surviving subjects demonstrated one or more organ systems that were able to recover, indicating that early organ support may be a key factor in shifting sepsis patients towards better outcomes. Although the method required measurements of high temporal density, it was designed using a set of easy-to-measure biomarkers to reduce the overall measurement burden and is ready for translation with a suitable human dataset.

7.1.4 Efficient MCMC Sampling Software for Parameter-Fitting

Chapter 6 introduced APT-MCMC, a software that can perform MCMC sampling faster than its Python counterpart without sacrificing user-friendliness. APT-MCMC was tailored for parameter-fitting problems and integrated MCMC techniques that improved the search

process. Parameter landscapes for inverse problems, especially for complex models such as the one introduced in Chapter 4, are highly multi-modal and anisotropic. APT-MCMC solved benchmark functions with a variety of properties: basins, valleys, non-differentiability, and multiple global optima. APT-MCMC addresses these issues by design and offers the usability of Python to automate the task of designing simulations in C++. The utility of APT-MCMC extends beyond fast and effective simulations: convergence statistics are automatically provided at the end of simulations and visualization software is provided in the Python package to aid in the post-hoc analysis of posterior distributions. APT-MCMC is open-source to benefit the community and can be an important utility in a modeler's toolbox.

7.2 FUTURE WORK

7.2.1 Predicting Organ Failure

The endophenotypes identified within heatmaps can be improved by adding clinical outcomes to the heatmap; a breakdown of organ failures for each main organ system (liver, coagulation, respiratory, nervous, renal, and cardiovascular) will provide correlations between the health states of specific organ systems and baseline clinical features such as lactate or cytokinemia. Such visualizations may reveal facets in the ProCESS data that can be used to develop statistical models to identify patients at risk of certain organ failures. A potential clinical tool that can be derived from this work is a hazards model. Gray's hazard model has been demonstrated to effectively model the time-varying covariates of sepsis mortality [62, 128]. Another application of this model is to use patient comorbidities and time-varying clinical features to reveal the contributing factors to sepsis-related organ failure.

7.2.2 Early Endotype Classifier

The next step in the mixture-model endotype work is to transition to impulse-response transfer function models that decay to zero, which is a better biological representation of

the inflammatory response. Preliminary work, described in Chapter 3.3.1, demonstrated that this step is a difficult one since the mixture-models collapsed into one single cluster. Additional exploration is warranted to refine this approach. If this step is successful, the machine learning classification tool should be updated to reflect the new endotypes. Finally, the ProCESS data should be further evaluated for new endotype-predictors to improve the accuracy of this classification tool.

7.2.3 Improving Neutrophil Dynamics with Damage

The mechanistic ODE model lacked several key neutrophil dynamics that were ignored due to the ProCESS data granularity. Once a suitable damage model has been developed (see Chapter 5.2), a damage state can be added to the model and fitted against the calculated time-varying damage values. Neutrophils would be the main contributor to this state. Biologically, neutrophils are misdirected during the cytokine storm phase of sepsis and recruited into healthy organs, where they cause damage to healthy tissue [9, 91]. This dynamic may be characterized by:

$$f_{Nrecruitment} = V_{n8} N_c \frac{IL8_T}{K_{n8} + IL8_T} \tag{7.1}$$

$$\frac{dN_T}{dt} = ... + \left(\frac{IL8_B}{K_{directed} + IL8_B} \right) f_{Nrecruitment} \tag{7.2}$$

$$\frac{dN_{misdirected}}{dt} = \left(1 - \frac{IL8_B}{K_{directed} + IL8_B} \right) f_{Nrecruitment} - \mu_{n_t} N_{misdirected} \tag{7.3}$$

$$\tag{7.4}$$

where $f_{Nrecruitment}$ represents the total IL-8 directed neutrophil recruitment into tissue. A percentage of these neutrophils, dictated by a Michaelis-Menten kinetic, are appropriately recruited into the site of infected tissue (Equation (7.2)) and the rest are misdirected (Equation (7.3)). Furthermore, the current ODE model uses neutrophil apoptosis to drive the anti-inflammatory state but does not model neutrophil necrosis, which also increases damage. A

proposed formulation is:

$$f_{death} = \mu_{n_t} N_T \left(1 - \frac{TNF_T}{K_{tn} + TNF_T}\right) \tag{7.5}$$

$$\frac{dIL10_T}{dt} = ... + \frac{Macro}{K_{m10} + Macro} f_{death} \tag{7.6}$$

$$\frac{dDamage}{dt} = ... + (1 - \frac{Macro}{K_{m10} + Macro}) f_{death} \tag{7.7}$$

where Equation (7.5) represents f_{death}, the death rate of **Tissue** compartment neutrophils. f_{death} may be split into anti-inflammatory apoptosis (given the presence of nearby macrophages, Equation (7.6)) or damage-inducing necrosis (Equation (7.7)). Finally, the proposed damage compartment may be structured as follows:

$$\frac{dDamage}{dt} = V_{DN} \left(N_{misdirect} + \mu_{n_t} N_{misdirected}\right) + (1 - \frac{Macro}{K_{m10} + Macro}) f_{death} - \mu_d Damage \tag{7.8}$$

where a systemic damage state is increased by the presence of misdirected neutrophils (damaging healthy tissue), the eventual necrosis of misdirected neutrophils, and the necrosis of properly directed neutrophils (term from Equation (7.7)). Damage heals itself at a rate μ_d, which is an optimistic scenario given the notion of cascading systemic failures. This damage state would then be used to increase the production of pro-inflammatory cytokines TNF, IL-6 and IL-8.

7.2.4 Estimator of Pre-hospital Time

Pre-hospital estimation accuracy can be further via a deeper exploration of the human trauma GLUE grant data. After demonstrating the translatability of the nearest-neighbor approach on human data, the next step would be to test the algorithm on nosocomial (in-hospital infection) sepsis data. If the technique is successful, it can be applied in the clinic almost immediately.

7.2.5 Optimization of Therapy and Damage

The addition of damage to the mechanistic ODE model enables personalized treatment optimization. Thang Ho's dissertation described a numerical optimization-based approach to formulate chemotherapy treatments that minimize cancerous tumor growth without reaching dangerously low levels of white blood cells [129]. A reduced ODE model highlighting damage, neutrophil, and pathogen interactions would be a suitable candidate for an optimization approach. Supposing a subject has been monitored sufficiently to parameterize the ODE model, the following problem may be posed:

$$\underset{x}{\text{minimize}} \quad \sum_{t} f_{pathogen}(x,\theta,t) + f_{IL6}(x,\theta,t_end) \tag{7.9}$$

$$\text{subject to} \quad \frac{df_i(x,\theta,t)}{dt} = RHS_i(x,\theta,t),\ i = \text{Model States}, \forall t \in [t_{begin}, t_{end}] \tag{7.10}$$

$$f_{damage}(x,\theta,t) \leq 10\ , \forall t \in [t_{begin}, t_{end}] \tag{7.11}$$

where x represents a combination of immunomodulatory drugs and/or antibiotics, θ represents the ODE model parameters, $f_i(x,\theta,t)$ represents the model states, and $f_{damage}(x,\theta,t)$ is constrained to never go above 10, which was a damage threshold discovered in Chapter 5.2 that set subjects on a non-surviving trajectory. Neutrophils and antibiotics both have deleterious effects on the body and the effect is quantified into $f_{damage}(x,\theta,t)$. The goal of this optimization is to enable the elimination of pathogen without allowing the inflammatory collateral damage to exceed a threshold. Finally, the second term in the objective function penalizes nonzero values of the inflammatory mediator IL-6 to prevent a sustained inflammatory state at the end of the simulation.

Within the ProCESS data, the majority of subjects did not survive the year after sepsis. This may be attributed to complications downstream of sepsis, which indicate that even after patients survive the initial onset of sepsis (survive the first 14 days), they are still in danger. Optimized personalized treatment has the potential to reduce mortality rates in sepsis in both the short and long term.

7.2.6 Updates to APT-MCMC

APT-MCMC will be semi-actively maintained. Major bugs will be addressed, but active development will cease for the time being. APT-MCMC will be open-source to encourage community forks for the purposes of code peer-review and feature additions.

A short-term feature is planned to modify the Python package of APT-MCMC to facilitate the easy use of non-uniform priors. In its current form, the Python package assumes bounded uniform distributions. Many priors are currently supported, but require the user to make the appropriate changes in C++ code. The selection of priors can affect the posterior distributions and may be dependent on the type of data used (such as population-level versus individual data or the density of data available) [130].

A long-term feature upgrade is eventually planned to explore the option of using Hamiltonian-MCMC techniques [115]. Intuitively, a Hamiltonian sampler views a parameter landscape as a ball on a hill and may "accelerate" down a hill towards the bottom. Mathematically, the sampler calculates a Jacobian at its current location and moves toward the optimal objective function direction with momentum. This type of movement may or may not be accomplished with an affine ensemble of samplers. Jacobian calculations are computationally expensive and Broyden's method may be occasionally used to alleviate the computational requirements. Such a sampling technique involves directed search rather than random walks and can greatly reduce the autocorrelation and improve simulation efficiency.

7.3 IMPROVING SEPSIS CLINICAL OUTCOMES WITH MATHEMATICAL MODELS

The disconnect between animal and human sepsis clinical trials may be attributed to three aspects. First, the level of variability within humans is great while animal clinical trials are controlled to remove variability. Second, therapy timing is currently impossible for humans while in animals, therapies are tested at exact times relative to the onset of sepsis. Third,

experimental controls for animal trials ensure consistent levels of sepsis-related damage while there is currently no way to assess this level of damage in humans.

The work presented in this dissertation uses a variety of mathematical models to address all three of these aspects. While the applicability of these models to the clinic is in the nascent stages, such models may improve the translatability of sepsis research from animal models to clinical trials.

The long-term extension of this dissertation is to generate a single, bedside tool that takes a series of clinical measurements and cytokines over a small time frame and provides the clinician with a subject's endotype, pre-hospital time/how long they have been septic for, and current level of systemic damage. Armed with this panel of information, a clinician can design and tailor a treatment unique to this subject. Furthermore, access to this information will enable new types of sepsis therapies such an endotype-specific treatments, timing-dependent therapies, and providing early organ support. The existence of such a decision support system rejects the status quo of a "one-size-fits-all" approach to sepsis and may usher in a new paradigm of personalized medicine and improved clinical outcomes in sepsis.

APPENDIX A

DETAILED RESULTS FOR NEAREST-NEIGHBOR PRE-HOSPITAL TIME ESTIMATION TOOL

This section contains additional results from the analysis involved with the nearest-neighbor-based estimation tool for subject pre-hospital times.

Table A1: Time-of-infection estimation accuracy over varying right censor values (temporal durations) for all biomarkers tested.

	Baboon Accuracy					Pig Peritonitis Accuracy				Pig LPS Accuracy			
Biomarker	R 1	R 2	R 3	R 4	Mean	R 1	R 2	R 3	Mean	R 1	R 2	R 3	Mean
WBC	51.8%	67.0%	83.0%	93.1%	73.7%	60.7%	58.6%	76.8%	65.4%	46.5%	64.4%	73.4%	61.4%
HR	42.1%	58.2%	74.6%	84.8%	65.0%	64.6%	83.8%	87.0%	78.5%	47.2%	61.0%	69.1%	59.1%
HCO3A	48.5%	56.2%	67.8%	75.3%	62.0%	-	-	-	-	-	-	-	-
SVR	32.7%	59.3%	72.0%	79.7%	60.9%	50.0%	57.4%	63.0%	56.8%	50.7%	63.6%	66.0%	60.1%
HB	34.8%	58.2%	70.8%	79.2%	60.8%	53.6%	67.1%	83.9%	68.2%	52.8%	73.7%	84.0%	70.2%
CO	33.6%	54.5%	68.9%	84.8%	60.5%	46.3%	51.5%	64.8%	54.2%	45.8%	55.1%	63.8%	54.9%
CI	32.1%	51.9%	72.0%	84.4%	60.1%	-	-	-	-	-	-	-	-
Cc_{O_2}	33.9%	55.9%	70.8%	79.7%	60.1%	-	-	-	-	-	-	-	-
ABEA	44.8%	51.5%	65.9%	74.5%	59.2%	59.8%	73.5%	77.8%	70.4%	50.4%	63.9%	69.6%	61.3%
MAP	33.0%	50.8%	65.9%	82.7%	58.1%	43.9%	50.0%	61.1%	51.7%	53.5%	65.3%	77.7%	65.5%
PVR	38.5%	52.5%	66.7%	74.5%	58.0%	-	-	-	-	-	-	-	-
Pa_{O_2}	45.2%	49.5%	63.6%	70.6%	57.2%	53.7%	51.5%	64.8%	56.6%	47.2%	49.2%	61.7%	52.7%
RBC	29.7%	49.8%	67.4%	78.8%	56.4%	52.4%	74.3%	87.5%	71.4%	54.2%	73.7%	83.0%	70.3%
CA_{O_2}	30.9%	51.2%	64.0%	77.1%	55.8%	51.2%	50.0%	57.4%	52.9%	47.9%	56.8%	57.4%	54.0%
PLT	50.9%	49.2%	56.1%	60.6%	54.2%	61.9%	70.0%	82.1%	71.3%	41.5%	50.0%	61.7%	51.1%
aP_{O_2}	40.3%	49.5%	56.4%	62.8%	52.3%	-	-	-	-	-	-	-	-
HCT	28.8%	45.5%	62.9%	70.1%	51.8%	53.6%	72.9%	91.1%	72.5%	50.0%	65.3%	81.9%	65.7%
PWP	34.5%	53.2%	54.2%	63.6%	51.4%	45.1%	64.7%	81.5%	63.8%	38.0%	46.6%	66.0%	50.2%
RAP	28.8%	46.5%	58.7%	70.6%	51.1%	43.9%	63.2%	70.4%	59.2%	41.5%	51.7%	64.9%	52.7%
Pa_{CO_2}	39.4%	52.2%	51.5%	61.0%	51.0%	51.2%	50.0%	57.4%	52.9%	47.9%	56.8%	57.4%	54.0%
TEMP	32.7%	49.8%	56.1%	62.8%	50.3%	54.9%	67.6%	83.3%	68.6%	59.9%	66.1%	79.8%	68.6%
O2DEL	31.8%	41.4%	54.5%	73.6%	50.3%	-	-	-	-	-	-	-	-
$SATA_{O_2}$	34.5%	45.8%	53.8%	62.8%	49.2%	62.2%	58.8%	70.4%	63.8%	55.6%	55.9%	73.4%	61.7%
MPAP	37.0%	43.8%	48.9%	58.0%	46.9%	59.8%	52.9%	61.1%	57.9%	47.2%	57.6%	85.1%	63.3%
RR	29.1%	35.4%	45.5%	59.3%	42.3%	50.0%	73.5%	77.8%	67.1%	36.6%	44.1%	53.2%	44.6%
HOROW	28.2%	38.0%	47.3%	54.1%	41.9%	52.4%	61.7%	72.2%	62.1%	-	-	-	-
QUOTIENT	28.2%	36.0%	43.6%	53.7%	40.4%	-	-	-	-	-	-	-	-
PHA	27.9%	38.7%	44.3%	48.1%	39.7%	42.7%	57.4%	75.9%	58.7%	55.6%	65.3%	76.6%	65.8%
AAD_{O_2}	27.6%	32.3%	45.1%	51.5%	39.1%	-	-	-	-	-	-	-	-
Mean Null	26.6%	29.0%	32.3%	36.7%	31.2%	38.1%	44.7%	55.0%	45.9%	42.7%	49.1%	57.7%	49.8%

Results were generated with available biomarkers from the baboon data and then sorted by mean accuracy. All tested biomarkers are shown in addition to the null hypothesis. The null hypothesis, tested with randomized biomarkers, performed worse than all of the tested biomarkers. This method was repeated on the two porcine data sets for validation. Only 3 right censor durations were tested due to the sampling rate and duration limitations of those experiments.

Table A2: Prediction accuracy of longitudinal vital signs with a single blood biomarker on porcine peritonitis experiment.

						Accuracy		
Right Censor	Vitals				Point	No Point	W/ Point	Time Series
	HR				HGB	83.6%	84.9%	80.8%
	HR				pH	83.6%	82.2%	82.2%
	HR				PLT	83.6%	80.8%	78.1%
	Temp	HR			HGB	87.7%	86.3%	82.2%
	$SATA_{O_2}$	HR			HGB	80.8%	86.3%	84.9%
2	$SATA_{O_2}$	HR			PHA	80.8%	83.6%	83.6%
	Temp	$SATA_{O_2}$	HR		HGB	84.9%	86.3%	87.7%
	Temp	$SATA_{O_2}$	HR		PHA	84.9%	83.6%	83.6%
	Temp	$SATA_{O_2}$	HR		PLT	84.9%	82.2%	82.2%
	Temp	MAP	$SATA_{O_2}$	HR	SBE	72.6%	79.5%	78.1%
	Temp	MAP	$SATA_{O_2}$	HR	HGB	72.6%	79.5%	75.3%
	Temp	MAP	$SATA_{O_2}$	HR	PaO2	72.6%	76.7%	78.1%
	HR				HGB	89.7%	93.1%	89.7%
	HR				PHA	89.7%	89.7%	89.7%
	Temp				HGB	89.7%	86.2%	89.7%
	Temp	HR			HGB	89.7%	91.4%	91.4%
	$SATA_{O_2}$	HR			HGB	89.7%	91.4%	89.7%
3	$SATA_{O_2}$	HR			PHA	89.7%	91.4%	89.7%
	Temp	$SATA_{O_2}$	HR		PLT	87.9%	91.4%	86.2%
	Temp	$SATA_{O_2}$	HR		PHA	87.9%	89.7%	87.9%
	Temp	MAP	HR		HGB	82.8%	87.9%	87.9%
	Temp	MAP	$SATA_{O_2}$	HR	HGB	82.8%	86.2%	86.2%
	Temp	MAP	$SATA_{O_2}$	HR	WBC	82.8%	82.8%	86.2%
	Temp	MAP	$SATA_{O_2}$	HR	SBE	82.8%	82.8%	82.8%

The combinatorial vitals+1 biomarker search was repeated for pig peritonitis experiment. This search analyzed combinations of vitals in time-series and a single point measurement of a blood-derived biomarker. The top results of this search are shown.

Table A3: Prediction accuracy of longitudinal vital signs with a single blood biomarker on porcine LPS experiment.

						Accuracy								
						Combined Groups (n=22)			1 μg/kg/h Dose Group			10 μg/kg/h Dose Group		
Right Censor	Vitals				Point	No Point	W/ Point	Time Series	No Point	W/ Point	Time Series	No Point	W/ Point	Time Series
	Temp				PaO2	66.1%	74.6%	63.6%	78.3%	78.3%	70.0%	64.6%	62.5%	54.2%
	Temp				PHA	66.1%	73.7%	72.9%	78.3%	76.7%	80.0%	64.6%	62.5%	66.7%
	Temp				HGB	66.1%	71.2%	74.6%	78.3%	71.7%	83.3%	64.6%	52.1%	60.4%
	Temp	SATAO2			PHA	74.6%	77.1%	79.7%	88.3%	88.3%	90.0%	62.5%	66.7%	66.7%
	Temp	HR			PHA	76.3%	76.3%	77.1%	71.7%	70.0%	70.0%	58.3%	62.5%	56.3%
2	MAP	SATAO2			PHA	71.2%	76.3%	75.4%	81.7%	81.7%	81.7%	66.7%	64.6%	64.6%
	Temp	MAP	SATAO2		WBC	75.4%	78.0%	78.8%	83.3%	81.7%	78.3%	62.5%	62.5%	85.4%
	Temp	MAP	SATAO2		PHA	75.4%	78.0%	78.8%	83.3%	83.3%	83.3%	62.5%	62.5%	62.5%
	Temp	MAP	HR		WBC	69.5%	76.3%	75.4%	80.0%	71.7%	71.7%	58.3%	60.4%	70.8%
	Temp	MAP	SATAO2	HR	WBC	72.9%	77.1%	77.1%	86.7%	75.0%	73.3%	56.3%	66.7%	72.9%
	Temp	MAP	SATAO2	HR	PLT	72.9%	77.1%	72.9%	86.7%	71.7%	75.0%	56.3%	68.8%	70.8%
	Temp	MAP	SATAO2	HR	PaO2	72.9%	75.4%	68.6%	86.7%	83.3%	76.7%	56.3%	58.3%	56.3%
	MAP				PaO2	77.7%	84.0%	73.4%	93.8%	89.6%	95.8%	68.4%	76.3%	73.7%
	Temp				PHA	79.8%	83.0%	80.9%	85.4%	89.6%	91.7%	63.2%	73.7%	73.7%
	MAP				WBC	77.7%	81.9%	85.1%	93.8%	93.8%	91.7%	68.4%	71.1%	89.5%
	Temp	SATAO2			PHA	87.2%	87.2%	86.2%	91.7%	91.7%	91.7%	78.9%	78.9%	76.3%
	Temp	MAP			WBC	73.4%	86.2%	85.1%	93.8%	91.7%	89.6%	65.8%	71.1%	89.5%
3	MAP	SATAO2			WBC	84.0%	85.1%	83.0%	93.8%	87.5%	91.7%	71.1%	78.9%	89.5%
	Temp	SATAO2	HR		HGB	78.7%	85.1%	84.0%	72.9%	75.0%	85.4%	73.7%	81.6%	71.1%
	Temp	MAP	SATAO2		WBC	81.9%	84.0%	84.0%	95.8%	89.6%	89.6%	71.1%	76.3%	86.8%
	MAP	SATAO2	HR		WBC	76.6%	84.0%	84.0%	91.7%	83.3%	81.3%	73.7%	78.9%	84.2%
	Temp	MAP	SATAO2	HR	WBC	80.9%	85.1%	85.1%	93.8%	83.3%	79.2%	71.1%	78.9%	81.6%
	Temp	MAP	SATAO2	HR	HGB	80.9%	81.9%	84.0%	93.8%	93.8%	95.8%	71.1%	73.7%	68.4%
	Temp	MAP	SATAO2	HR	PHA	80.9%	80.9%	80.9%	93.8%	93.8%	93.8%	71.1%	71.1%	71.1%

The combinatorial vitals+1 search was repeated for pig LPS. This search analyzed combinations of vitals in time-series and a single point measurement of a blood-derived biomarker. The top results of this search are shown. Comparable results for each of the LPS dosing groups are shown because ANOVA revealed significant differences

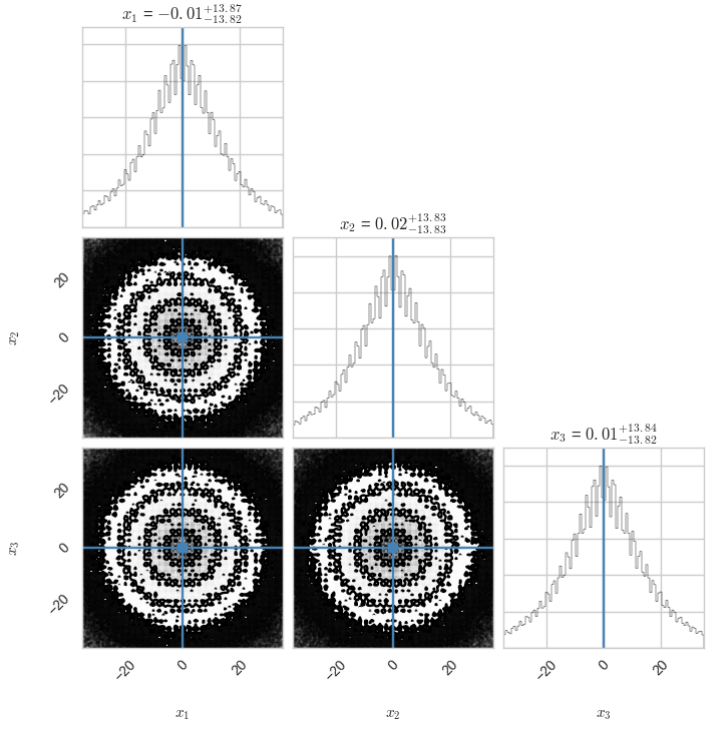
APPENDIX B

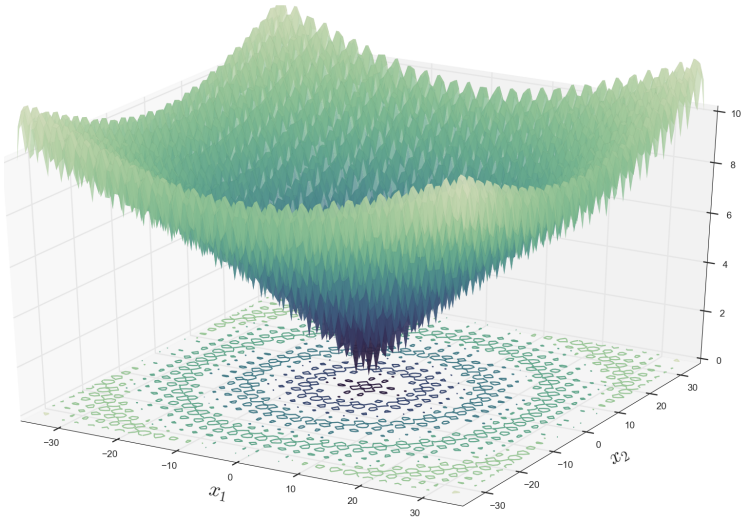
DETAILED BENCHMARK FUNCTIONS USED TO VALIDATE APT-MCMC

This section lists the benchmark functions used as well as the APT-MCMC distributions for their parameters. Functions were taken from Jamil, et al [126].

B.0.1 Ackley

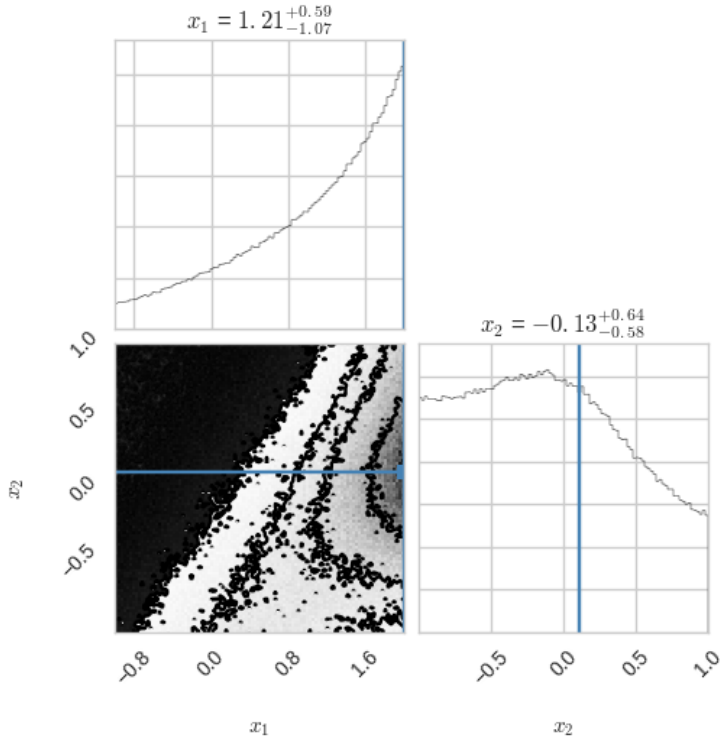
$$f_{\text{ackley}}(\underline{x}) = -20e^{-0.02\sqrt{\sum_{i=1}^{D} x_i^2/D}} - e^{\sum_{i=1}^{D} \cos(2\pi x_i)/D} + 20 + e$$

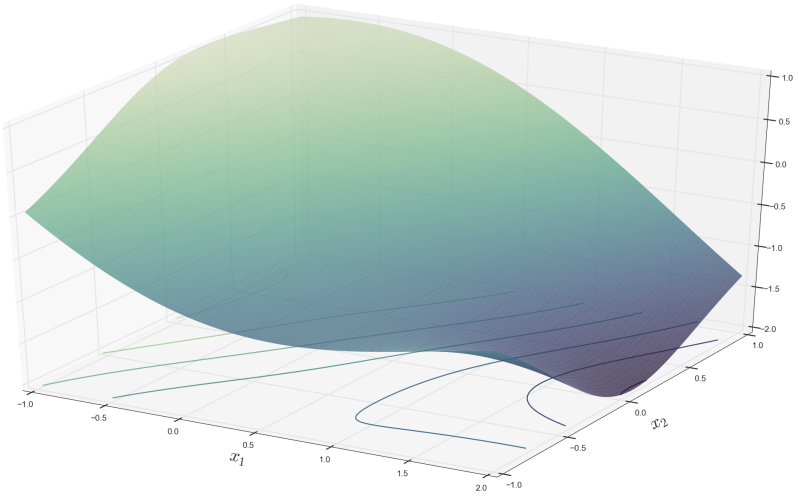



10
8
6
4
2
0
−30
−20
−10
0
10
20
30
x_1
x_2

B.0.2 Adjiman

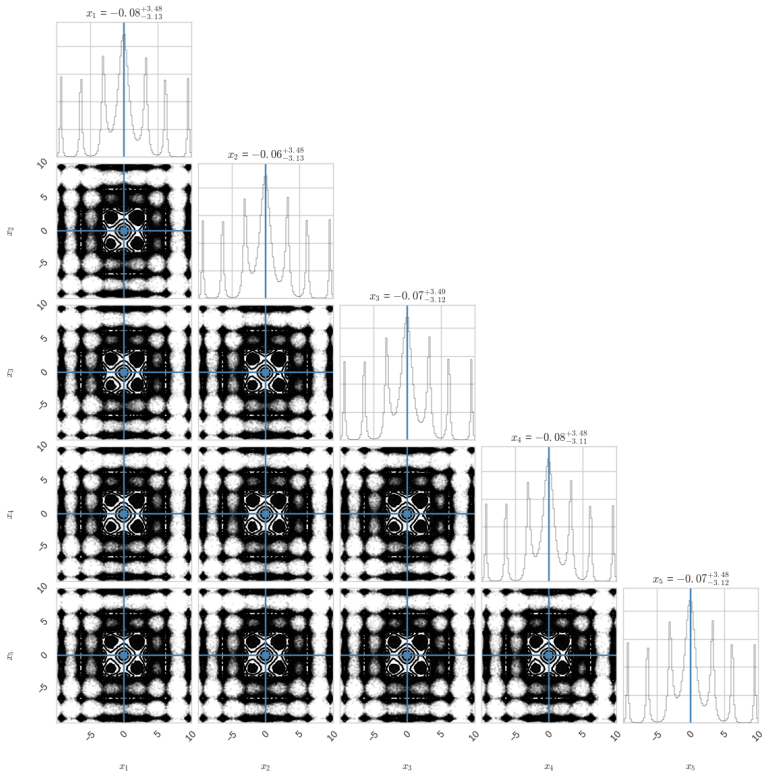
$$f_{\text{adjiman}}(\underline{x}) = \cos(x_1)\sin(x_2) - \frac{x_1}{x_2^2 + 1}$$

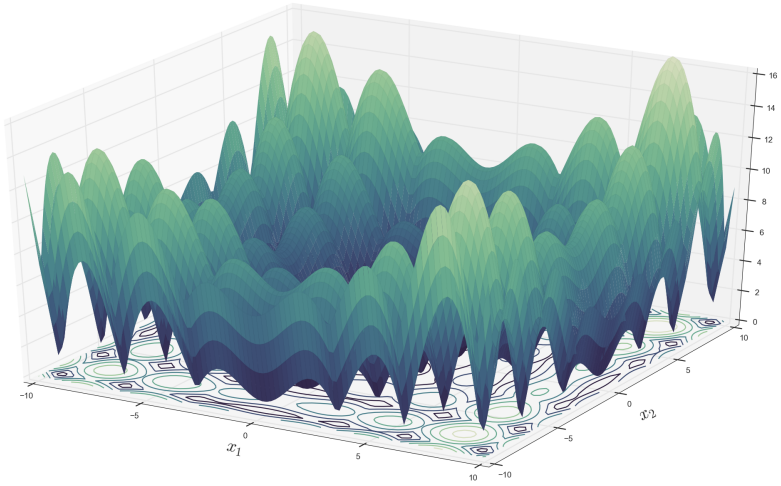




B.0.3 Alpine

$$f_{\text{alpine}}(\underline{x}) = \sum_{i=1}^{D} \left| x_i \sin(x_i) + 0.1 x_i \right|$$





B.0.4 Bard

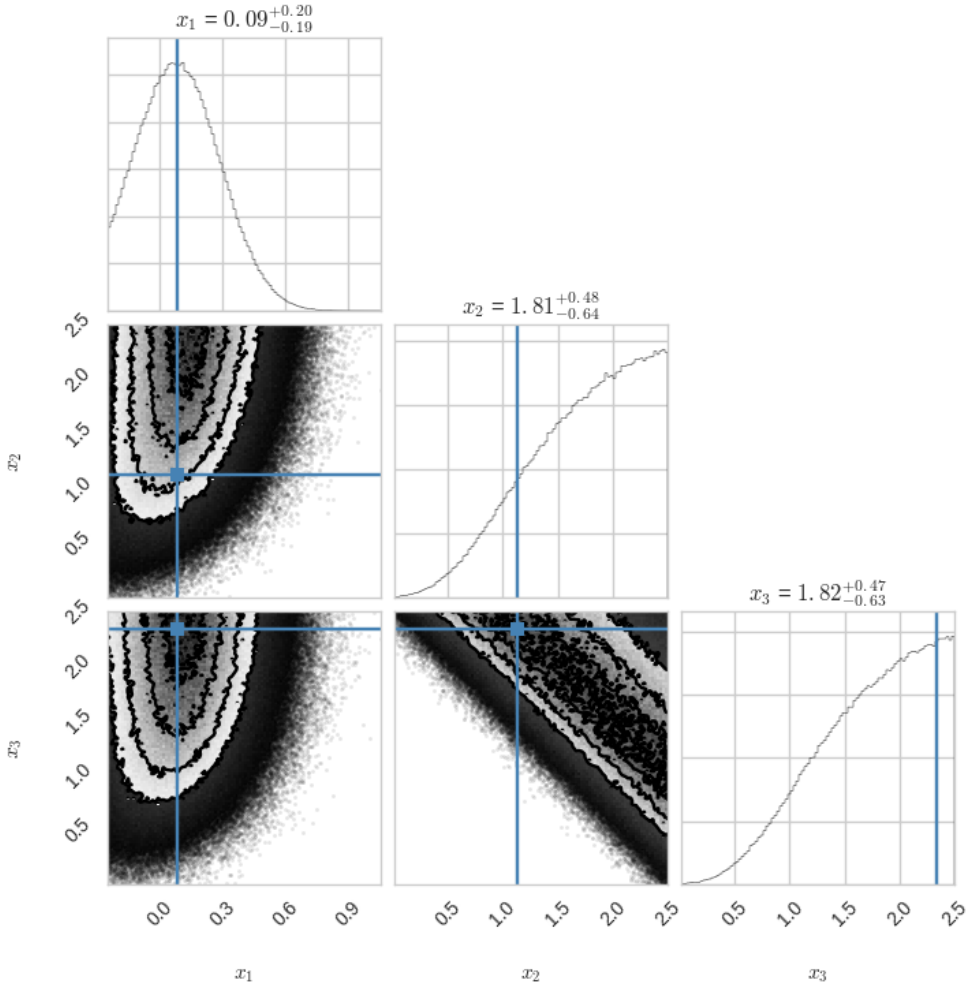
$$f_{\text{bard}}(\underline{x}) = \sum_{i=1}^{1} 5\left[\frac{y_i - x_1 - u_i}{v_i x_2 + w_i x_3}\right]^2$$

$$y_i = [0.14, 0.18, 0.22, 0.25, 0.29, 0.32, 0.35, 0.39, 0.37, 0.58, 0.73, 0.96, 1.34, 2.10, 4.39]$$

$$u_i = i$$

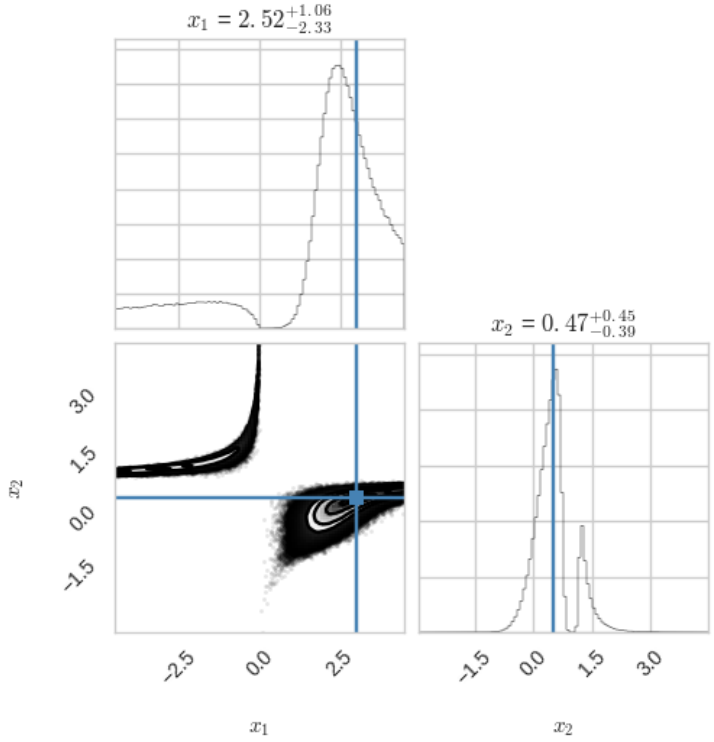
$$v_i = 16 - i$$

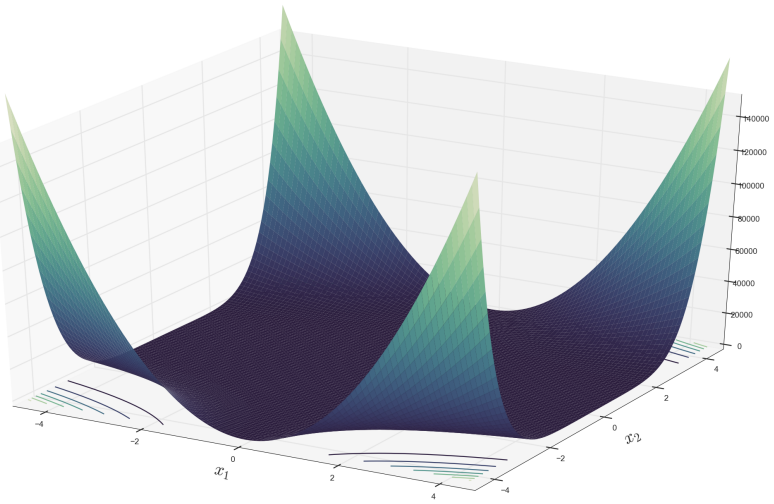
$$w_i = \min(u_i, v_i)$$



B.0.5 Beale

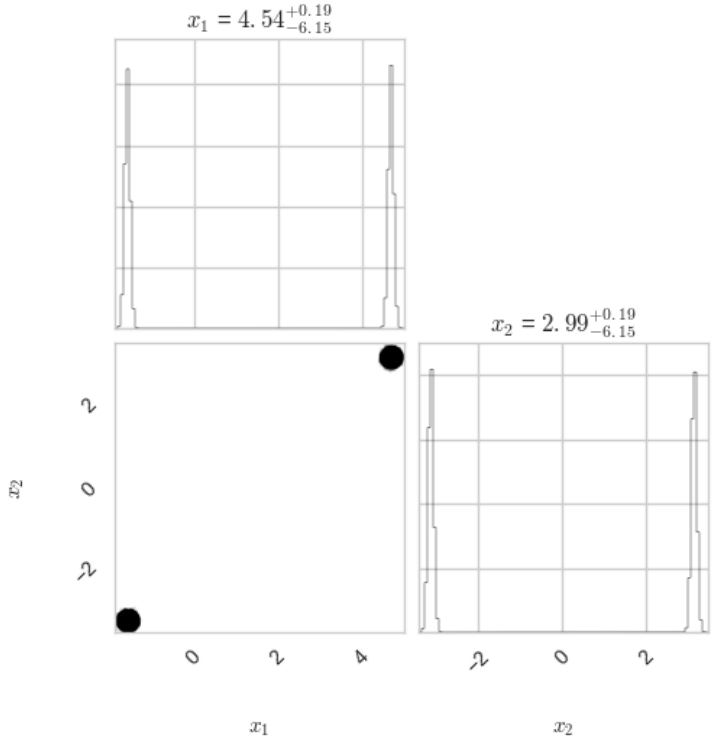
$$f_{\text{beale}}(\underline{x}) = (1.5 - x_1 + x_1 x_2)^2 + (2.25 - x_1 + x_1 x_2^2)^2 + (2.625 - x_1 + x_1 + x_2^3)^2$$

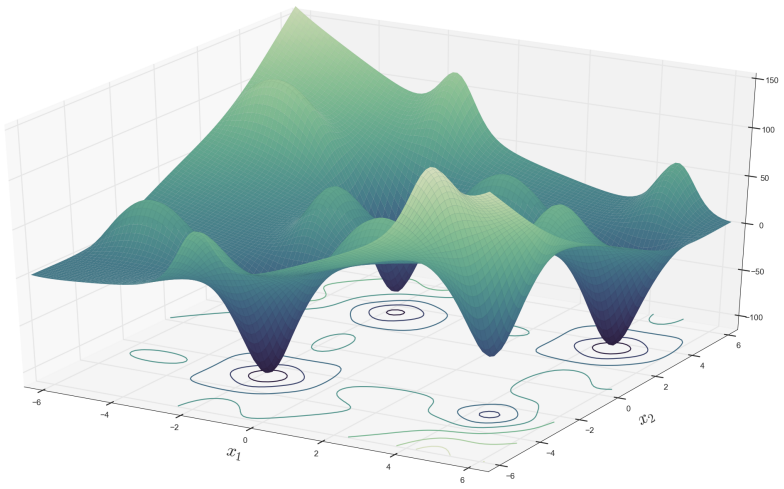




B.0.6 Bird

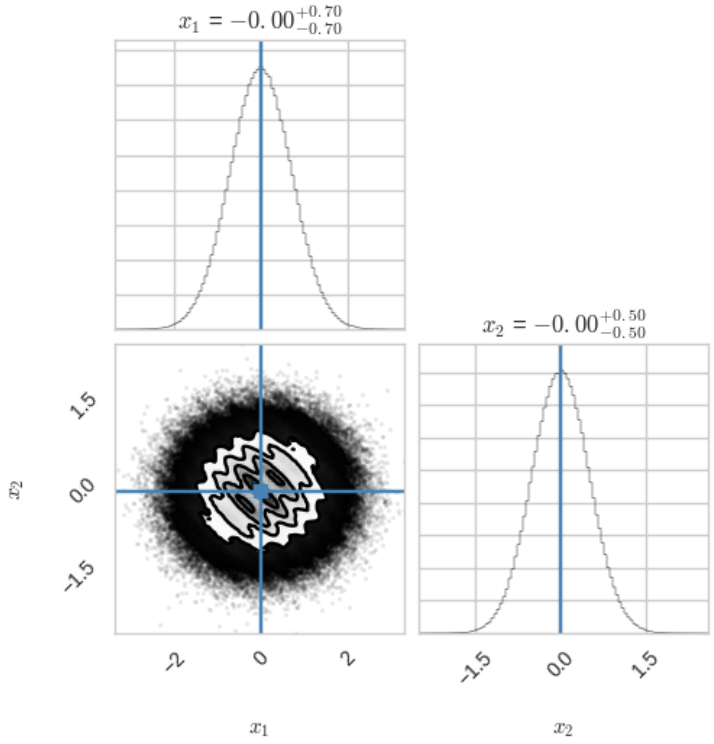
$$f_{\text{bird}}(\underline{x}) = \sin(x_1)e^{(1-\cos(x_2))^2} + \cos(x_2)e^{(1-\sin(x_1))^2} + (x_1 - x_2)^2$$

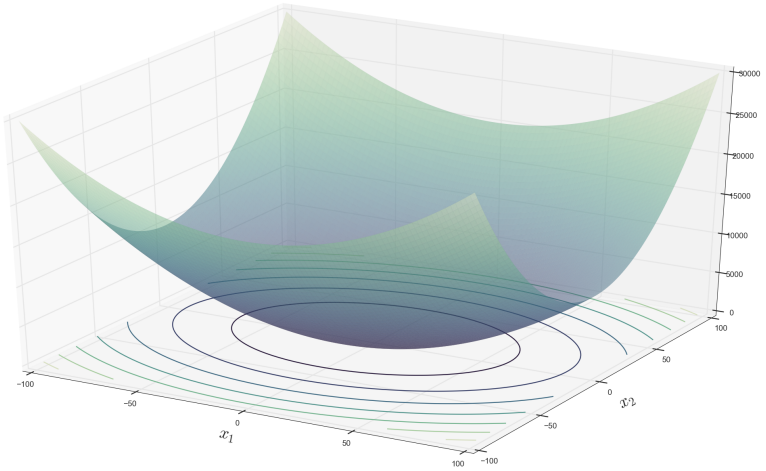




B.0.7 Bohachevsky

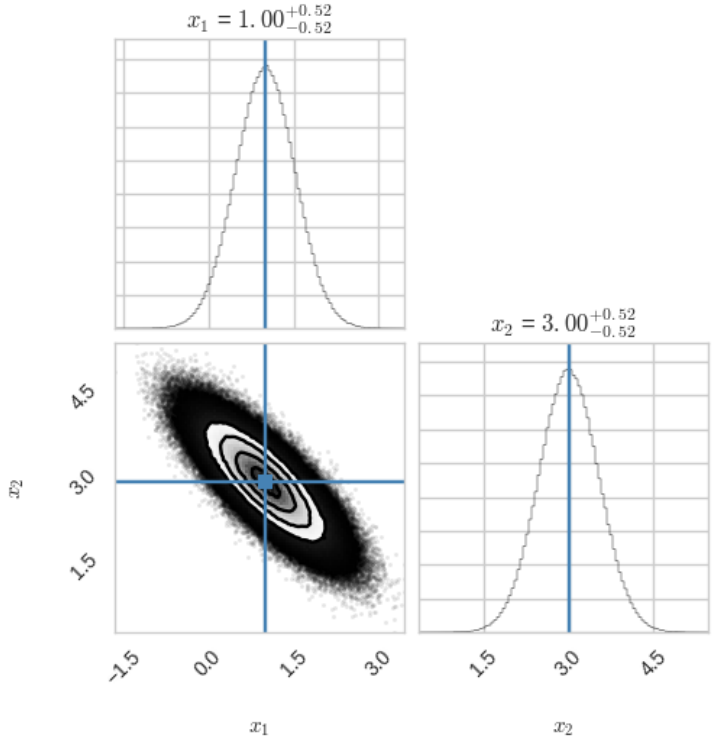
$$f_{\text{bohachevsky}}(\underline{x}) = x_1^2 + 2x_2^2 - 0.3\cos(3\pi x_1 + 4\pi x_2) + 0.3$$

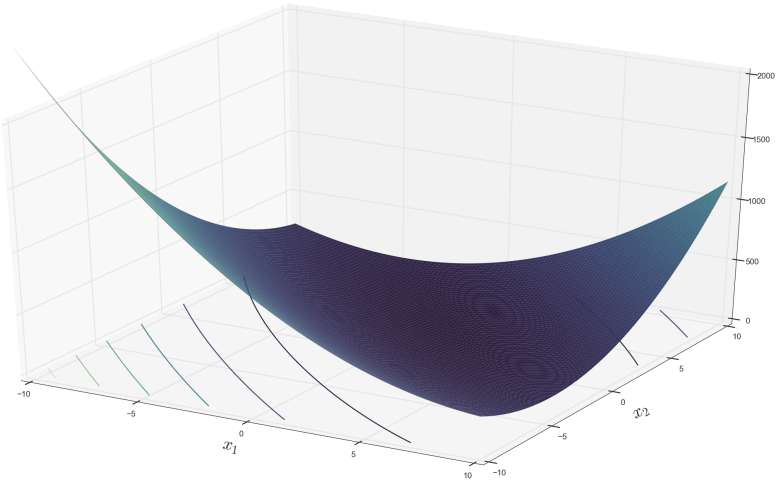




B.0.8 Booth

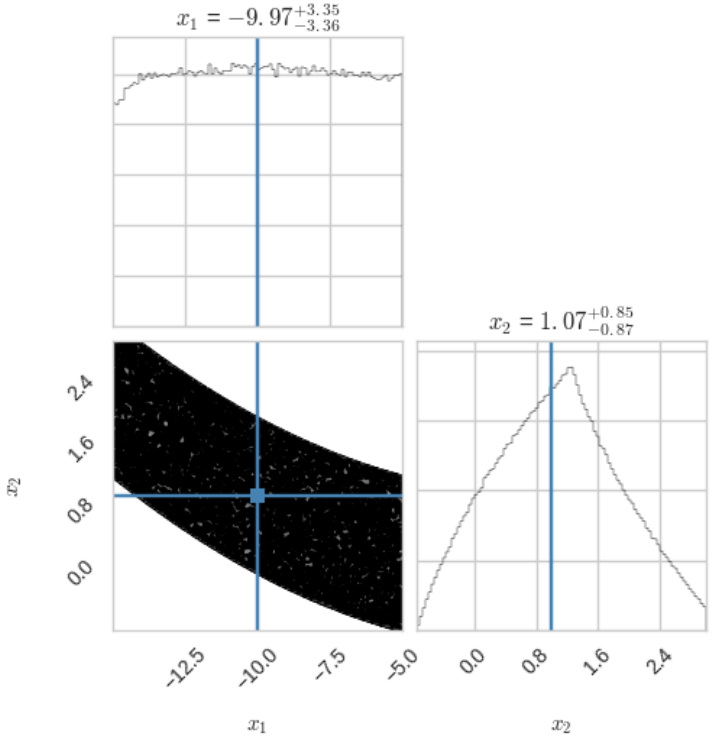
$$f_{\text{booth}}(\underline{x}) = (x_1 + 2x_2 - 7)^2 + (2x_1 + x_2 - 5)^2$$

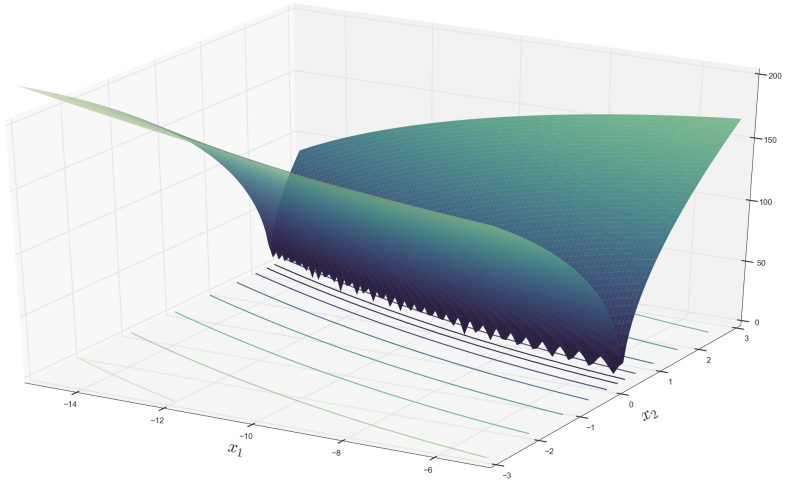




B.0.9 Bukin

$$f_{\text{bukin}}(\underline{x}) = 100\sqrt{||x_2 = 0.01x_1^2||} + 0.01||x_1 + 10||$$



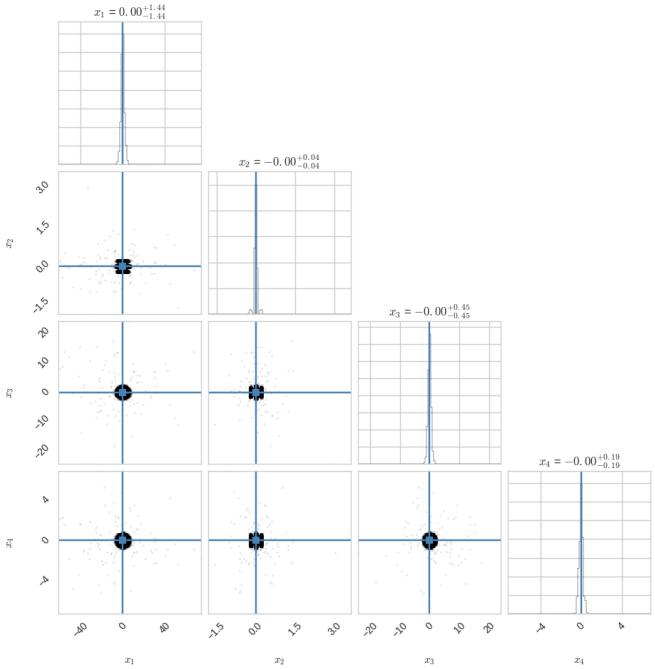


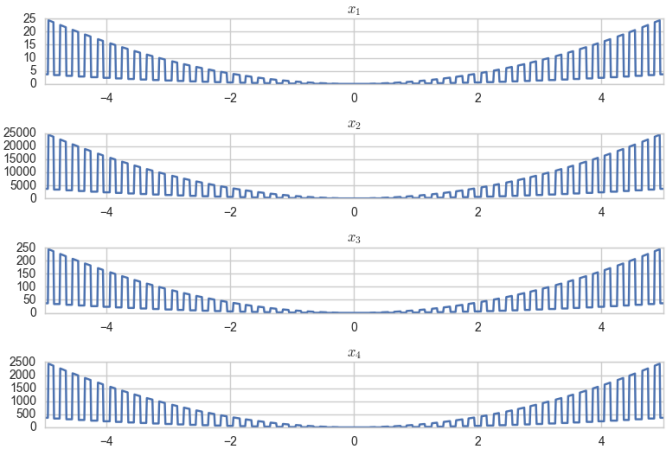
B.0.10 Corana

$$f_{\text{corana}}(\underline{x}) = \begin{cases} \sum_{i=1}^{4} 0.15(z_i - 0.5\text{sign}(z_i))^2 & , |x_i - z_i| < 0.05 \\ d_i x_i^2 & , \text{otherwise} \end{cases}$$

$$z_i = 0.2\lfloor \left| \frac{x_i}{0.2} \right| + 0.49999 \rfloor \text{sign}(x_i)$$

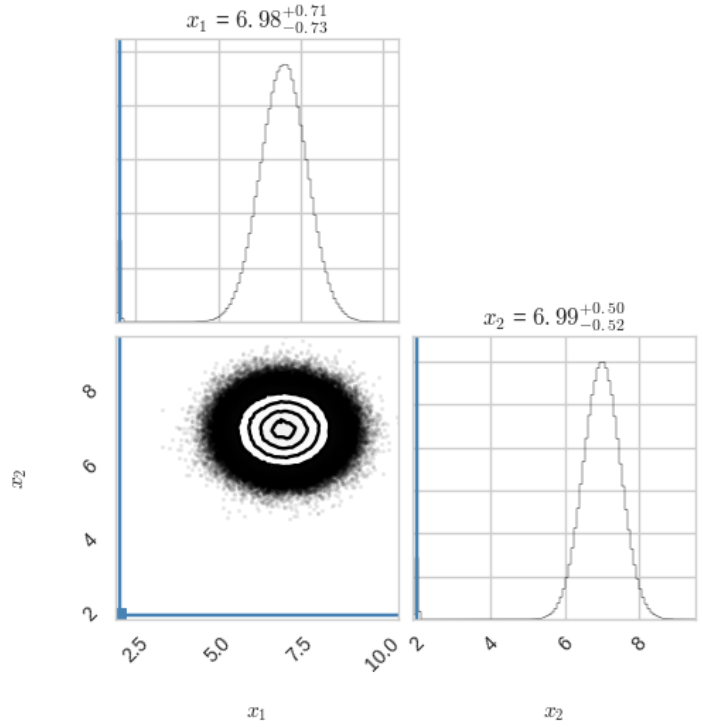
$$d_i = [1, 1000, 10, 100]$$

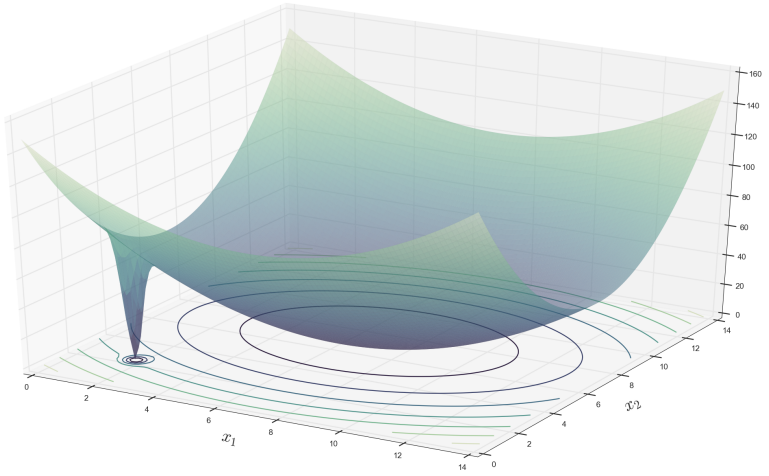




B.0.11 Damavandi

$$f_{\text{damavandi}}(\underline{x}) = \left[1 - \left|\frac{\sin[\pi(x_1 - 2)]\sin[\pi(x_2 - 2)]}{\pi^2(x_1 - 2)(x_2 - 2)}\right|^5\right]\left[2 + (x_1 - 7)^2 + 2(x_2 - 7)^2\right]$$



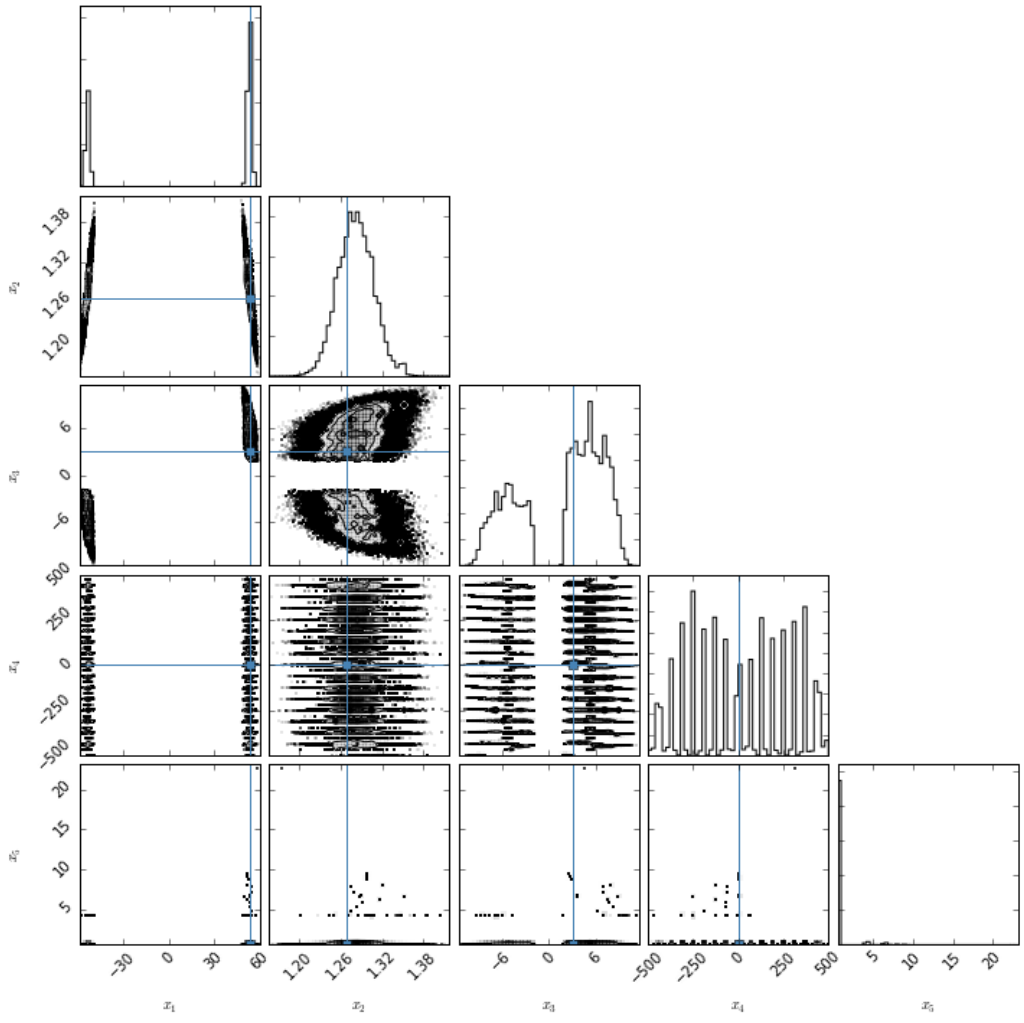


B.0.12 Devilliers-Glasser

$$f_{\text{devilliers}}(\underline{x}) = \sum_{i=1}^{1} 6 \left\{ x_1 x_2^{t_i} \tanh\left[x_3 t_i + \sin(x_4 t_i)\right] \cos(t_i e^{x5} - y_i) \right\}^2$$

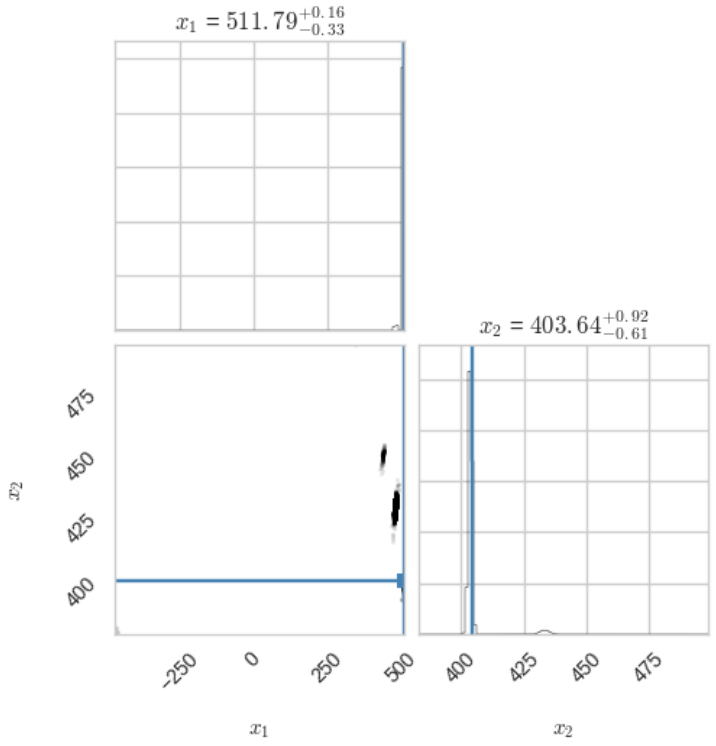
$$t_i = 0.1(i-1)$$

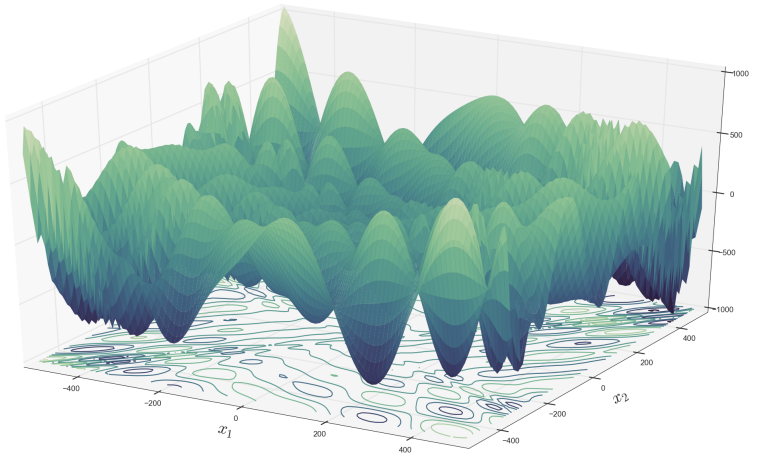
$$y_i = 53.811.27^{t_i} \tanh\left[3.012 t_i + \sin(2.13 t_i)\right] \cos(t_i e^{0.507})$$



B.0.13 Eggholder

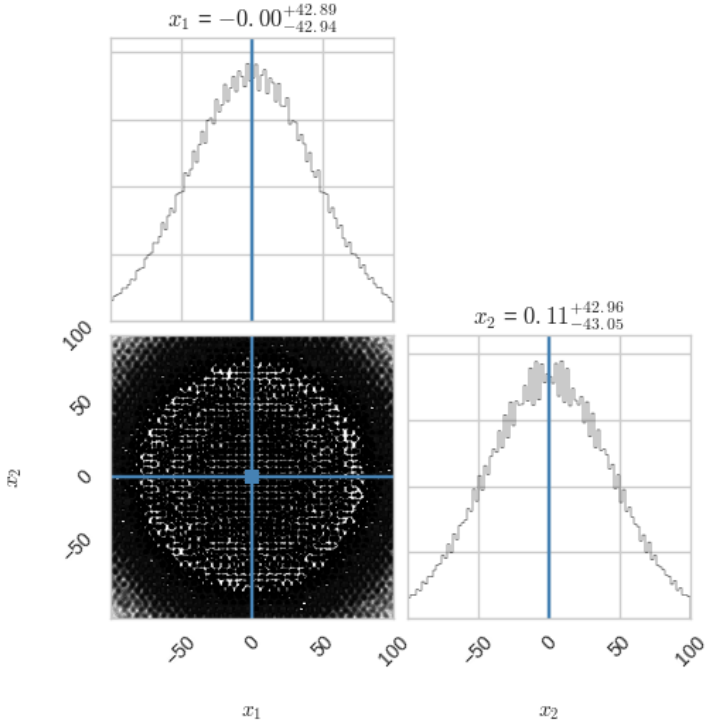
$$f_{\text{eggholder}}(\underline{x}) = \sum_{i=1}^{m-1} \left[-(x_{i+1}+47)\sin\sqrt{|x_{i+1}+x_i/2+47|} - x_i \sin\sqrt{|x_i-(x_{i+1}+47)|} \right]$$

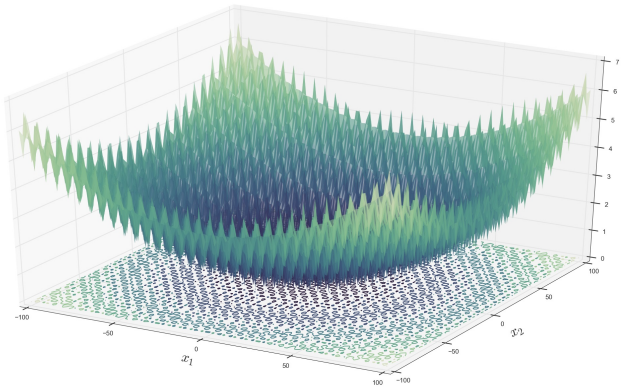




B.0.14 Griewank

$$f_{\text{griewank}}(\underline{x}) = \sum_{i=1}^{n} \frac{x_i^2}{4000} - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}} + 1\right)$$





BIBLIOGRAPHY

[1] R Phillip Dellinger, Mitchell M Levy, Andrew Rhodes, Djillali Annane, Herwig Gerlach, Steven M Opal, Jonathan E Sevransky, Charles L Sprung, Ivor S Douglas, Roman Jaeschke, Tiffany M Osborn, Mark E Nunnally, Sean R Townsend, Konrad Reinhart, Ruth M Kleinpell, Derek C Angus, Clifford S Deutschman, Flavia R Machado, Gordon D Rubenfeld, Steven a Webb, Richard J Beale, Jean-Louis Vincent, and Rui Moreno. Surviving sepsis campaign: international guidelines for management of severe sepsis and septic shock: 2012. *Critical care medicine*, 41(2): 580–637, feb 2013. ISSN 1530-0293. doi: 10.1097/CCM.0b013e31827e83af. URL http://www.ncbi.nlm.nih.gov/pubmed/23353941.

[2] D C Angus, W T Linde-Zwirble, J Lidicker, G Clermont, J Carcillo, and M R Pinsky. Epidemiology of severe sepsis in the United States: analysis of incidence, outcome, and associated costs of care. *Critical care medicine*, 29(7):1303–10, jul 2001. ISSN 0090-3493. URL http://www.ncbi.nlm.nih.gov/pubmed/11445675.

[3] Rukmini Kumar, Gilles Clermont, Yoram Vodovotz, and Carson C Chow. The dynamics of acute inflammation. *Journal of theoretical biology*, 230(2):145–55, sep 2004. ISSN 0022-5193. doi: 10.1016/j.jtbi.2004.04.044. URL http://www.ncbi.nlm.nih.gov/pubmed/15321710.

[4] Angela Reynolds, Jonathan Rubin, Gilles Clermont, Judy Day, Yoram Vodovotz, and G Bard 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):220–36, sep 2006. ISSN 0022-5193. doi: 10.1016/j.jtbi.2006.02.016. URL http://www.ncbi.nlm.nih.gov/pubmed/16584750.

[5] Derek C. Angus and Tom van der Poll. Severe sepsis and septic shock. *The New England journal of medicine*, 369(9):840–51, aug 2013. ISSN 1533-4406. doi: 10.1056/NEJMra1208623.

[6] Christopher W Seymour, Vincent X Liu, Theodore J Iwashyna, Frank M Brunkhorst, Thomas D Rea, André Scherag, Gordon Rubenfeld, Jeremy M Kahn, Manu Shankar-Hari, Mervyn Singer, Clifford S Deutschman, Gabriel J Escobar, and Derek C Angus. Assessment of Clinical Criteria for Sepsis: For the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA*, 315(8):762–74,

feb 2016. ISSN 1538-3598. doi: 10.1001/jama.2016.0288. URL http://dx.doi.org/10.1001/jama.2016.0288http://www.ncbi.nlm.nih.gov/pubmed/26903335http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5433435.

[7] Mervyn Singer, Clifford S Deutschman, Christopher Warren Seymour, Manu Shankar-Hari, Djillali Annane, Michael Bauer, Rinaldo Bellomo, Gordon R Bernard, Jean-Daniel Chiche, Craig M Coopersmith, Richard S Hotchkiss, Mitchell M Levy, John C Marshall, Greg S Martin, Steven M Opal, Gordon D Rubenfeld, Tom van der Poll, Jean-louis Vincent, and Derek C Angus. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *Jama*, 315(8):801–10, 2016. ISSN 1538-3598. doi: 10.1001/jama.2016.0287. URL http://www.ncbi.nlm.nih.gov/pubmed/26903338.

[8] Mitchell P Fink. Animal models of sepsis. *Virulence*, 5(1):143–53, jan 2014. ISSN 2150-5608. doi: 10.4161/viru.26083. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3916368{&}tool=pmcentrez{&}rendertype=abstract.

[9] Fabiane Sônego, Fernanda Vargas e Silva Castanheira, Raphael Gomes Ferreira, Alexandre Kanashiro, Caio Abner Vitorino Gonçalves Leite, Daniele Carvalho Nascimento, David Fernando Colón, Vanessa de Fátima Borges, José Carlos Alves-Filho, and Fernando Queiróz Cunha. Paradoxical roles of the neutrophil in sepsis: Protective and deleterious. *Frontiers in Immunology*, 7(APR):1–7, 2016. ISSN 16643224. doi: 10.3389/fimmu.2016.00155.

[10] Ruslan Medzhitov. Origin and physiological roles of inflammation. *Nature*, 454(July): 428–435, 2008. ISSN 0028-0836. doi: 10.1038/nature07201.

[11] Helen L Wright, Robert J Moots, Roger C Bucknall, and Steven W Edwards. Neutrophil function in inflammation and inflammatory diseases. *Rheumatology (Oxford, England)*, 49(9):1618–31, sep 2010. ISSN 1462-0332. doi: 10.1093/rheumatology/keq045. URL http://www.ncbi.nlm.nih.gov/pubmed/20338884.

[12] Chao Shi and Eric G. Pamer. Monocyte recruitment during infection and inflammation. *Nature Reviews Immunology*, 11(11):762–774, 2011. ISSN 1474-1733. doi: 10.1038/nri3070. URL http://dx.doi.org/10.1038/nri3070.

[13] Paola Italiani and Diana Boraschi. From Monocytes to M1/M2 Macrophages: Phenotypical vs. Functional Differentiation. *Frontiers in immunology*, 5 (October):514, 2014. ISSN 1664-3224. doi: 10.3389/fimmu.2014.00514. URL http://www.ncbi.nlm.nih.gov/pubmed/25368618http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4201108.

[14] 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):25–9, jan 2003. ISSN 1471-4906. URL http://www.ncbi.nlm.nih.gov/pubmed/12495721.

[15] 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, may 2008. ISSN 0022-1767. doi: 10.4049/jimmunol.180.9.5771. URL http://www.jimmunol.org/lookup/doi/10.4049/jimmunol.180.9.5771.

[16] H. R. Alexander, G. M. Doherty, C. M. Buresh, D. J. Venzon, and J. A. Norton. A recombinant human receptor antagonist to interleukin 1 improves survival after lethal endotoxemia in mice. *Journal of Experimental Medicine*, 173(4):1029–1032, apr 1991. doi: 10.1084/jem.173.4.1029. URL www.jem.org/cgi/doi/10.1084/jem.20142182.

[17] Eva Fischer, Michael A. Marano, Kimberly J. Van Zee, Craig S. Rock, Arthur S. Hawes, W A Thompson, Laura DeForge, John S. Kenney, Daniel G. Remick, and Duane C. Bloedow. Interleukin-1 receptor blockade improves survival and hemodynamic performance in Escherichia coli septic shock, but fails to alter host responses to sublethal endotoxemia. *Journal of Clinical Investigation*, 89(5):1551–1557, may 1992. ISSN 0021-9738. doi: 10.1172/JCI115748. URL http://www.jci.org/articles/view/115748.

[18] Steven M. Opal, C J Fisher, Jean-Francois F Dhainaut, J L Vincent, Rainer Brase, Stephen F. Lowry, Jerald C. Sadoff, Gus J. Slotman, Howard Levy, Robert A. Balk, Maire P. Shelly, John P. Pribble, John F. LaBrecque, Janice Lookabaugh, Hugh Donovan, Howard Dubin, Robert Baughman, James Norman, Eric DeMaria, Klaus Matzel, Edward Abraham, and Michael Seneff. Confirmatory interleukin-1 receptor antagonist trial in severe sepsis: a phase III, randomized, double-blind, placebo-controlled, multicenter trial. The Interleukin-1 Receptor Antagonist Sepsis Investigator Group. *Critical care medicine*, 25(7):1115–24, jul 1997. ISSN 0090-3493. URL http://ovidsp.tx.ovid.com/sp-3.20.0b/ovidweb.cgi?QS2=434f4e1a73d37e8c54b7bccef7797d0c9235f73fb8b1d19f41877a54327fa7ccd3d2d6cd1c44f266ca7

[19] L B Hinshaw, B K Beller-Todd, L T Archer, B Benjamin, D J Flournoy, R Passey, and M F Wilson. Effectiveness of steroid/antibiotic treatment in primates administered LD100 Escherichia coli. *Annals of surgery*, 194(1):51–6, 1981. ISSN 0003-4932. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1345194{&}tool=pmcentrez{&}rendertype=abstract.

[20] Veterans Administration Systemic Sepsis Cooperative Study Group. Effect of High-Dose Glucocorticoid Therapy on Mortality in Patients with Clinical Signs of Systemic Sepsis. *New England Journal of Medicine*, 317(11):659–665, sep 1987. ISSN 0028-4793. doi: 10.1056/NEJM198709103171102. URL http://www.ncbi.nlm.nih.gov/pubmed/2888017http://www.nejm.org/doi/abs/10.1056/NEJM198709103171102.

[21] K J Van Zee, L L Moldawer, H S Oldenburg, W A Thompson, S A Stackpole, W J Montegut, M A Rogy, C Meschter, H Gallati, C D Schiller, W F Richter, H Loetscher, A Ashkenazi, S M Chamow, F Wurm, S E Calvano, S F Lowry, and W Lesslauer. Protection against lethal Escherichia coli bacteremia in baboons (Papio anubis) by pretreatment with a 55-kDa TNF receptor (CD120a)-Ig fusion protein, Ro 45-2081.

Journal of immunology (Baltimore, Md. : 1950), 156(6):2221–30, mar 1996. ISSN 0022-1767. URL http://www.ncbi.nlm.nih.gov/pubmed/8690912.

[22] Edward Abraham, P F Laterre, Jorge Garbino, Susan Pingleton, Thomas Butler, Thierry Dugernier, Benjamin Margolis, Kenneth Kudsk, Werner Zimmerli, Paula Anderson, Marc Reynaert, Daniel Lew, Werner Lesslauer, Sharon Passe, Philip Cooper, Alex Burdeska, Marlene Modi, Anton Leighton, Miklos Salgo, Philippe Van der Auwera, and Lenercept Study Group. Lenercept (p55 tumor necrosis factor receptor fusion protein) in severe sepsis and early septic shock: a randomized, double-blind, placebo-controlled, multicenter phase III trial with 1,342 patients. *Critical care medicine*, 29 (3):503–10, mar 2001. ISSN 0090-3493. doi: 10.1097/00003246-200103000-00006. URL http://www.ncbi.nlm.nih.gov/pubmed/11373411.

[23] H Redl, G Schlag, E Paul, S Bahrami, W A Buurman, R M Strieter, S L Kunkel, J Davies, and R Foulkes. Endogenous modulators of TNF and IL-1 response are under partial control of TNF in baboon bacteremia. *The American journal of physiology*, 271 (5 Pt 2):R1193–8, nov 1996. ISSN 0002-9513. URL http://www.ncbi.nlm.nih.gov/pubmed/8945953.

[24] J Cohen and J Carlet. INTERSEPT: an international, multicenter, placebo-controlled trial of monoclonal antibody to human tumor necrosis factor-alpha in patients with sepsis. International Sepsis Trial Study Group. *Critical care medicine*, 24(9):1431–40, sep 1996. ISSN 0090-3493. URL http://www.ncbi.nlm.nih.gov/pubmed/8797612.

[25] A K Myers, J W Robey, and R M Price. Relationships between tumour necrosis factor, eicosanoids and platelet-activating factor as mediators of endotoxin-induced shock in mice. *British journal of pharmacology*, 99(3):499–502, 1990. ISSN 0007-1188. URL http://www.ncbi.nlm.nih.gov/pubmed/2110016{%}5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1917357.

[26] D M Albrecht, K van Ackern, H-J Bender, H Hof, W Kox, N Victor, P Funk, M Kieser, S Köhler, D Krausch, I Marzi, T Menges, and H Schmidt. Efficacy and Safety of the Platelet-Activating Factor Receptor Antagonist BN 52021 (Ginkgolide B) in Patients with Severe Sepsis : A Randomised, Double-Blind, Placebo-Controlled, Multicentre Trial. *Clinical drug investigation*, 24(3):137–47, 2004. ISSN 1173-2563. URL http://www.ncbi.nlm.nih.gov/pubmed/17516700.

[27] M Giral, D Balsa, R Ferrando, M Merlos, J Garcia-Rafanell, and J Forn. Effects of UR-12633, a new antagonist of platelet-activating factor, in rodent models of endotoxic shock. *British journal of pharmacology*, 118(5):1223–31, jul 1996. ISSN 0007-1188. URL http://www.ncbi.nlm.nih.gov/pubmed/8818347http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1909590.

[28] J L Vincent, H Spapen, J Bakker, N R Webster, and L Curtis. Phase II multicenter clinical study of the platelet-activating factor receptor antagonist BB-882 in the treat-

ment of sepsis. *Critical care medicine*, 28(3):638–42, mar 2000. ISSN 0090-3493. URL http://www.ncbi.nlm.nih.gov/pubmed/10752807.

[29] M. Ogata, T. Matsumoto, K. Koga, I. Takenaka, M. Kamochi, T. Sata, S Yoshida, and A. Shigematsu. An antagonist of platelet-activating factor suppresses endotoxin-induced tumor necrosis factor and mortality in mice pretreated with carrageenan. *Infection and immunity*, 61(2):699–704, feb 1993. ISSN 0019-9567. URL http://www.ncbi.nlm.nih.gov/pubmed/8423096http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC302782.

[30] A M Froon, J W Greve, W A Buurman, C J van der Linden, H J Langemeijer, C Ulrich, and M Bourgeois. Treatment with the platelet-activating factor antagonist TCV-309 in patients with severe systemic inflammatory response syndrome: a prospective, multicenter, double-blind, randomized phase II trial. *Shock (Augusta, Ga.)*, 5(5):313–9, may 1996. ISSN 1073-2322. URL http://www.ncbi.nlm.nih.gov/pubmed/9156785.

[31] Takukyu Sha, Mie Sunamoto, Tomoyuki Kitazaki, Jun Sato, Masayuki Ii, and Yuji Iizawa. Therapeutic effects of TAK-242, a novel selective Toll-like receptor 4 signal transduction inhibitor, in mouse endotoxin shock model. *European Journal of Pharmacology*, 571(2-3):231–239, 2007. ISSN 00142999. doi: 10.1016/j.ejphar.2007.06.027.

[32] Todd W. Rice, Arthur P. Wheeler, Gordon R. Bernard, Jean-Louis Vincent, Derek C. Angus, Naoki Aikawa, Ignace Demeyer, Stephen Sainati, Nicholas Amlot, Charlie Cao, Masayuki Ii, Hideyasu Matsuda, Kouji Mouri, and Jon Cohen. A randomized, double-blind, placebo-controlled trial of TAK-242 for the treatment of severe sepsis*. *Critical Care Medicine*, 38(8):1685–1694, 2010. ISSN 0090-3493. doi: 10.1097/CCM.0b013e3181e7c5c9. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage{&}an=00003246-201008000-00009.

[33] S M Opal, J E Palardy, N A Parejo, and A A Creasey. The activity of tissue factor pathway inhibitor in experimental models of superantigen-induced shock and polymicrobial intra-abdominal sepsis. *Crit Care Med.*, 29(0090-3493 SB - AIM SB - IM):13–17, 2001. ISSN 0090-3493.

[34] A A Creasey, A C Chang, L Feigen, T C Wün, F B Taylor, and L B Hinshaw. Tissue factor pathway inhibitor reduces mortality from Escherichia coli septic shock. *The Journal of Clinical Investigation*, 91(6):2850–2860, 1993. ISSN 0021-9738. doi: 10.1172/JCI116529. URL http://www.ncbi.nlm.nih.gov/pubmed/8514893.

[35] Edward Abraham, Konrad Reinhart, Steven Opal, Ignace Demeyer, Christopher Doig, Angel López Rodriguez, Richard Beale, Petr Svoboda, Pierre Francois Laterre, Stuart Simon, Bruce Light, Herbert Spapen, Judy Stone, Allan Seibert, Claus Peckelsen, Cathy De Deyne, Russell Postier, Ville Pettilä, Charles L. Sprung, Antonio Artigas, Sandra R Percell, Vincent Shu, Christian Zwingelstein, Jeffrey Tobias, Lona Poole, James C. Stolzenbach, and Abla A Creasey. Efficacy and Safety of Tifacogin (Recombinant Tissue Factor Pathway Inhibitor) in Severe Sepsis. *JAMA*,

290(2):238, jul 2003. ISSN 0098-7484. doi: 10.1001/jama.290.2.238. URL http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.290.2.238.

[36] Ganesh Suntharalingam, Meghan R Perry, Stephen Ward, Stephen J Brett, Andrew Castello-Cortes, Michael D Brunner, and Nicki Panoskaltsis. Cytokine storm in a phase 1 trial of the anti-CD28 monoclonal antibody TGN1412. *The New England journal of medicine*, 355(10):1018–28, sep 2006. ISSN 1533-4406. doi: 10.1056/NEJMoa063842. URL http://www.nejm.org/doi/full/10.1056/NEJMoa063842http://www.ncbi.nlm.nih.gov/pubmed/16908486.

[37] Daniel Rittirsch, L Marco Hoesel, and Peter a Ward. The disconnect between animal models of sepsis and human sepsis. *Journal of leukocyte biology*, 81(1):137–43, jan 2007. ISSN 0741-5400. doi: 10.1189/jlb.0806542. URL http://www.ncbi.nlm.nih.gov/pubmed/17020929.

[38] Niels C Riedemann, Ren-Feng Guo, and Peter a Ward. Novel strategies for the treatment of sepsis. *Nature medicine*, 9(5):517–524, 2003. ISSN 10788956. doi: 10.1038/nm0503-517.

[39] 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 (Augusta, Ga.)*, 39(2):127–37, feb 2013. ISSN 1540-0514. doi: 10.1097/SHK.0b013e31827dafa7. URL http://www.ncbi.nlm.nih.gov/pubmed/23324881.

[40] Michael A Puskarich, Stephen Trzeciak, Nathan I Shapiro, Ryan C Arnold, James M Horton, Jonathan R Studnek, Jeffrey a Kline, Alan E Jones, and Emergency Medicine Shock Research Network (EMSHOCKNET). Association between timing of antibiotic administration and mortality from septic shock in patients treated with a quantitative resuscitation protocol. *Critical care medicine*, 39 (9):2066–71, sep 2011. ISSN 1530-0293. doi: 10.1097/CCM.0b013e31821e87ab. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3158284{&}tool=pmcentrez{&}rendertype=abstracthttp://www.ncbi.nlm.nih.gov/pubmed/21572327http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3158284.

[41] Li Ang Zhang, Robert S Parker, David Swigon, Ipsita Banerjee, Soheyl Bahrami, Heinz Redl, and Gilles Clermont. A One-Nearest-Neighbor Approach to Identify the Original Time of Infection Using Censored Baboon Sepsis Data. *Critical Care Medicine*, 44(6):e432–e442, jun 2016. ISSN 0090-3493. doi: 10.1097/CCM.0000000000001623. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage{&}an=00003246-201606000-00057.

[42] Peter Q. Eichacker, Chantal Parent, Andre Kalil, Claire Esposito, Xizhong Cui, Steven M. Banks, Eric P. Gerstenberger, Yvonne Fitz, Robert L. Danner, and Charles

Natanson. Risk and the efficacy of antiinflammatory agents: Retrospective and confirmatory studies of sepsis. *American Journal of Respiratory and Critical Care Medicine*, 166(9):1197–1205, 2002. ISSN 1073449X. doi: 10.1164/rccm.200204-302OC.

[43] Claudio Franceschi, Massimiliano Bonafè, and Silvana Valensin. Human immunosenescence: The prevailing of innate immunity, the failing of clonotypic immunity, and the filling of immunological space. *Vaccine*, 18(16):1717–1720, 2000. ISSN 0264410X. doi: 10.1016/S0264-410X(99)00513-7.

[44] Benjamin M P Tang, Anthony S. McLean, Ian W. Dawes, Stephen J. Huang, and Ruby C Y Lin. The use of gene-expression profiling to identify candidate genes in human sepsis. *American Journal of Respiratory and Critical Care Medicine*, 176(7):676–684, 2007. ISSN 1073449X. doi: 10.1164/rccm.200612-1819OC.

[45] Hernando Gomez. *The Sepsis Syndrome and the "One Size Fits All" Construct: The Emperor has No Clothes!!* Master's thesis, University of Pittsburgh, 2014.

[46] P Arndt and E Abraham. Immunological therapy of sepsis: experimental therapies. *Intensive care medicine*, 27 Suppl 1:S104–15, jan 2001. ISSN 0342-4642. URL http://www.ncbi.nlm.nih.gov/pubmed/11307366.

[47] Jan Lotvall, Cezmi A. Akdis, Leonard B. Bacharier, Leif Bjermer, Thomas B. Casale, Adnan Custovic, Robert F. Lemanske, 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. ISSN 00916749. doi: 10.1016/j.jaci.2010.11.037.

[48] P Campo, F Rodríguez, S Sánchez-García, P Barranco, S Quirce, C Pérez-Francés, E Gómez-Torrijos, R Cárdenas, J M Olaguibel, J Delgado, Severe Asthma Workgroup, and SEAIC Asthma Committee. Phenotypes and endotypes of uncontrolled severe asthma: new treatments. *Journal of investigational allergology & clinical immunology*, 23(2):76–88, 2013. ISSN 1018-9068. URL http://www.ncbi.nlm.nih.gov/pubmed/23654073.

[49] Christopher D. Fjell, Simone Thair, Joseph L. Hsu, Keith R. Walley, James a. Russell, and John Boyd. Cytokines and signaling molecules predict clinical outcomes in sepsis. *PloS one*, 8(11):e79207, 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0079207. URL http://www.ncbi.nlm.nih.gov/pubmed/24244449http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3828333.

[50] Daniel B. Knox, Michael J. Lanspa, Kathryn G. Kuttler, Simon C. Brewer, and Samuel M. Brown. Phenotypic clusters within sepsis-associated multiple organ dysfunction syndrome. *Intensive Care Medicine*, 41(5):814–822, 2015. ISSN 14321238. doi: 10.1007/s00134-015-3764-7. URL http://dx.doi.org/10.1007/s00134-015-3764-7.

[51] 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):237–56, sep 2006. ISSN 0022-5193. doi: 10.1016/j.jtbi.2006.02.015. URL http://www.ncbi.nlm.nih.gov/pubmed/16616206http://www.sciencedirect.com/science/article/pii/S0022519306000804.

[52] JS Hogg. *Advances in Rule-based Modeling: Compartments, Energy, and Hybrid Simulation, with Application to Sepsis and Cell Signaling*. PhD thesis, University of Pittsburgh, 2013. URL http://d-scholarship.pitt.edu/19621/.

[53] Mark A. Musen, Blackford Middleton, and Robert A. Greenes. Clinical Decision-Support Systems. In Edward H. Shortliffe and James J. Cimino, editors, *Biomedical Informatics: Computer Applications in Health Care and Biomedicine*, pages 643–674. Springer London, London, 2014. ISBN 978-1-4471-4474-8. doi: 10.1007/978-1-4471-4474-8_22. URL http://link.springer.com/10.1007/978-1-4471-4474-8{_}22.

[54] Shibin Mathew, John Bartels, Ipsita Banerjee, and Yoram Vodovotz. Global sensitivity analysis of a mathematical model of acute inflammation identifies nonlinear dependence of cumulative tissue damage on host interleukin-6 responses. *Journal of theoretical biology*, 358:132–48, oct 2014. ISSN 1095-8541. doi: 10.1016/j.jtbi.2014.05.036. URL http://www.ncbi.nlm.nih.gov/pubmed/24909493.

[55] Zhenzhen Shi, Chih-hang J Wu, David Ben-Arieh, and Steven Q Simpson. Mathematical Model of Innate and Adaptive Immunity of Sepsis: A Modeling and Simulation Study of Infectious Disease. *BioMed Research International*, 2015:1–31, 2015. ISSN 2314-6133. doi: 10.1155/2015/504259. URL http://www.hindawi.com/journals/bmri/2015/504259/.

[56] Bone RC and Roger C. Bone. Why sepsis trials fail. *JAMA*, 276(7):565–566, aug 1996. ISSN 0098-7484. doi: 10.1001/jama.1996.03540070061032. URL http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.1996.03540070061032http://dx.doi.org/10.1001/jama.1996.03540070061032.

[57] John C. Marshall. Why have clinical trials in sepsis failed? *Trends in Molecular Medicine*, 20(4):195–203, 2014. ISSN 1471499X. doi: 10.1016/j.molmed.2014.01.007. URL http://dx.doi.org/10.1016/j.molmed.2014.01.007.

[58] J L Vincent, R Moreno, J Takala, S Willatts, A De Mendonça, H Bruining, C K Reinhart, P M Suter, and L G Thijs. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. *Intensive Care Medicine*, 22(7):707–710, jul 1996. ISSN 0342-4642. doi: 10.1007/BF01709751. URL http://www.ncbi.nlm.nih.gov/pubmed/8844239http://link.springer.com/10.1007/BF01709751.

[59] Donald M Yealy, John a Kellum, David T Huang, Amber E Barnato, Lisa a Weissfeld, Francis Pike, Thomas Terndrup, Henry E Wang, Peter C Hou, Frank LoVecchio, Michael R Filbin, Nathan I Shapiro, and Derek C Angus. A randomized trial of protocol-based care for early septic shock. *The New England journal of medicine*,

370(18):1683–93, may 2014. ISSN 1533-4406. doi: 10.1056/NEJMoa1401602. URL http://www.ncbi.nlm.nih.gov/pubmed/24635773.

[60] Emanuel Rivers, Bryant Nguyen, Suzanne Havstad, Julie Ressler, Alexandria Muzzin, Bernhard Knoblich, Edward Peterson, and Michael Tomlanovich. Early Goal-Directed Therapy in the Treatment of Severe Sepsis and Septic Shock. *New England Journal of Medicine*, 345(19):1368–1377, 2001. ISSN 0028-4793. doi: 10.1056/NEJMoa010307. URL http://www.nejm.org/doi/abs/10.1056/NEJMoa010307.

[61] D. C. Angus, A. E. Barnato, D. Bell, R. Bellomo, C. R. Chong, T. J. Coats, A. Davies, A. Delaney, D. A. Harrison, A. Holdgate, B. Howe, D. T. Huang, T. Iwashyna, J. A. Kellum, S. L. Peake, F. Pike, M. C. Reade, K. M. Rowan, M. Singer, S. A.R. Webb, L. A. Weissfeld, D. M. Yealy, and J. D. Young. A systematic review and meta-analysis of early goal-directed therapy for septic shock: the ARISE, ProCESS and ProMISe Investigators. *Intensive Care Medicine*, 41(9):1549–1560, 2015. ISSN 14321238. doi: 10.1007/s00134-015-3822-1.

[62] Jan Kasal, Zorana Jovanovic, Gilles Clermont, Lisa A Weissfeld, Vladimir Kaplan, R Scott Watson, and Derek C Angus. Comparison of Cox and Gray's survival models in severe sepsis. *Critical care medicine*, 32(3):700–7, mar 2004. ISSN 0090-3493. doi: 00003246-200403000-00013[pii]. URL http://www.mdconsult.com/das/article/body/234387631-2/jorg=journal{%}26source={%}26sp=14432357{%}26sid=0/N/404723/szrz00304000700.pdf?issn=0090-3493http://www.ncbi.nlm.nih.gov/pubmed/15090950.

[63] Gábor J. Székely and Maria L. Rizzo. Energy statistics: A class of statistics based on distances. *Journal of Statistical Planning and Inference*, 143(8):1249–1272, 2013. ISSN 03783758. doi: 10.1016/j.jspi.2013.03.018. URL http://dx.doi.org/10.1016/j.jspi.2013.03.018.

[64] W. A. Knaus, D P Wagner, E A Draper, J E Zimmerman, M Bergner, P G Bastos, C A Sirio, D J Murphy, T Lotring, and A Damiano. The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest*, 100 (6):1619–1636, 1991. ISSN 0012-3692. doi: 10.1378/chest.100.6.1619.

[65] M E Charlson, R E Charlson, J C Peterson, S S Marinopoulos, W M Briggs, and J P Hollenberg. The Charlson comorbidity index is adapted to predict costs of chronic disease in primary care patients. *J Clin Epidemiol.*, 61(1878-5921 (Electronic)):1234–1240, dec 2008.

[66] John A. Kellum, L Kong, M P Fink, L A Weissfeld, D M Yealy, M R Pinsky, J Fine, A Krichevsky, R L Delude, and D C Angus. Understanding the Inflammatory Cytokine Response in Pneumonia and Sepsis. *Archives of Internal Medicine*, 167(15):1655, aug 2007. ISSN 0003-9926. doi: 10.1001/archinte.167.15.1655. URL http://archinte.jamanetwork.com/article.aspx?doi=10.1001/archinte.167.15.1655.

[67] Jeremy D. Scheff, Panteleimon D. Mavroudis, Panagiota T. Foteinou, Gary An, Steve E. Calvano, John Doyle, Thomas E. Dick, Stephen F. Lowry, Yoram Vodovotz, and Ioannis P. Androulakis. A multiscale modeling approach to inflammation: A case study in human endotoxemia. *Journal of Computational Physics*, 244:279–289, 2013. ISSN 00219991. doi: 10.1016/j.jcp.2012.09.024. URL http://dx.doi.org/10.1016/j.jcp.2012.09.024.

[68] Daniel G Remick. Pathophysiology of sepsis. *The American journal of pathology*, 170(5):1435–44, may 2007. ISSN 0002-9440. doi: 10.2353/ajpath.2007.060872. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1854939{&}tool=pmcentrez{&}rendertype=abstract.

[69] Yoram Vodovotz, Gilles Clermont, Carson Chow, and Gary An. Mathematical models of the acute inflammatory response. *Current Opinion in Critical Care*, 10(5):383–390, oct 2004. ISSN 1070-5295. doi: 10.1097/01.ccx.0000139360.30327.69. URL http://journals.lww.com/co-criticalcare/Abstract/2004/10000/Mathematical{_}models{_}of{_}the{_}acute{_}inflammatory.14.aspxhttp://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage{&}an=00075198-200410000-00014.

[70] Sang O K Song, Justin Hogg, Zhi-Yong Peng, Robert Parker, John a Kellum, and Gilles Clermont. Ensemble Models of Neutrophil Trafficking in Severe Sepsis. *PLoS Computational Biology*, 8(3):e1002422, mar 2012. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1002422. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3297568{&}tool=pmcentrez{&}rendertype=abstracthttp://dx.plos.org/10.1371/journal.pcbi.1002422.

[71] Jean-marc Cavaillon, Minou Adib-conquy, Catherine Fitting, Christophe Adrie, and Didier Payen. Cytokine Cascade in Sepsis. *Scandinavian Journal of Infectious Diseases*, 35(9):535–544, 2003. ISSN 0036-5548. doi: 10.1080/00365540310015935. URL http://www.tandfonline.com/doi/full/10.1080/00365540310015935.

[72] B L Jones, D S Nagin, and R Kathryn. A SAS Procedure Based on Mixture Models for Estimating Developmental Trajectories. *Sociological Methods & Research*, 29(374):374–393, 2001.

[73] Hao Hong Yiu, Andrea L Graham, and Robert F Stengel. Dynamics of a cytokine storm. *PloS one*, 7(10):e45027, jan 2012. ISSN 1932-6203. doi: 10.1371/journal.pone.0045027. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3462188{&}tool=pmcentrez{&}rendertype=abstract.

[74] P Damas, D Ledoux, M Nys, Y Vrindts, D De Groote, P Franchimont, and M Lamy. Cytokine serum level during severe sepsis in human IL-6 as a marker of severity. *Annals of surgery*, 215(4):356–62, apr 1992. ISSN 0003-4932. doi: 10.1097/00000658-199204000-00009. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage{&}an=

00000658-199204000-00009http://www.ncbi.nlm.nih.gov/pubmed/1558416http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1242452.

[75] Savitri Kibe, Kate Adams, and Gavin Barlow. Diagnostic and prognostic biomarkers of sepsis in critical care. *Journal of Antimicrobial Chemotherapy*, 66(SUPPL. 2):ii33–40, apr 2011. ISSN 03057453. doi: 10.1093/jac/dkq523. URL http://www.ncbi.nlm.nih.gov/pubmed/21398306.

[76] Salih Cesur, Ali Şengül, Yasemin Kurtoğlu, Yasin Kalpakçı, Selcan Arslan Özel, Aybars Bilgetürk, Hakan Erdem, Turan Aslan, Sami Kınıklı, Can Polat Eyigün, and Zeynep Bıyıklı. Prognostic value of cytokines (TNF-α, IL-10, Leptin) and C-reactive protein serum levels in adult patients with nosocomial sepsis. *Journal of Microbiology and Infectious Diseases*, 1(3):101–109, dec 2011. ISSN 21463158. doi: 10.5799/ahinjs.02.2011.03.0024. URL http://jmidonline.org/upload/sayi/5/JMID-00546.pdfhttp://dergipark.gov.tr/doi/10.5799/ahinjs.02.2011.03.0024.

[77] Daniel G Remick, Gerald Bolgos, Shannon Copeland, and Javed Siddiqui. Role of interleukin-6 in mortality from and physiologic response to sepsis. *Infection and Immunity*, 73(5):2751–2757, 2005. ISSN 00199567. doi: 10.1128/IAI.73.5.2751-2757.2005.

[78] Fumitake Takahashi and Shigeo Abe. Decision-Tree Based Multiclass Support Vector Machines. In *International Conference on Neural Information Processing*, number 3, pages 1418–1422. IEEE, 2002.

[79] Carson C Chow, Gilles Clermont, Rukmini Kumar, Claudio Lagoa, Zacharia Tawadrous, David Gallo, Binnie Betten, John Bartels, Gregory Constantine, Mitchell P Fink, Timothy R Billiar, and Yoram Vodovotz. The acute inflammatory response in diverse shock states. *Shock (Augusta, Ga.)*, 24(1):74–84, jul 2005. ISSN 1073-2322. doi: 10.1097/01.shk.0000168526.97716.f3. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage{&}an=00024382-200507000-00012http://www.ncbi.nlm.nih.gov/pubmed/15988324.

[80] Alexander D. Malkin, Robert P. Sheehan, Shibin Mathew, William J. Federspiel, Heinz Redl, and Gilles Clermont. A Neutrophil Phenotype Model for Extracorporeal Treatment of Sepsis. *PLoS computational biology*, 11(10):e1004314, oct 2015. ISSN 1553-7358. doi: 10.1371/journal.pcbi.1004314. URL http://dx.plos.org/10.1371/journal.pcbi.1004314http://www.ncbi.nlm.nih.gov/pubmed/26468651http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4607502.

[81] Daniel G Remick and Peter A Ward. Evaluation of endotoxin models for the study of sepsis. *Shock (Augusta, Ga.)*, 24 Suppl 1(Supplement 1):7–11, dec 2005. ISSN 1073-2322. doi: 10.1097/01.shk.0000191384.34066.85. URL http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage{&}an=00024382-200512001-00002http://www.ncbi.nlm.nih.gov/pubmed/16374366.

[82] Catalin Vasilescu, Klaus Buttenschoen, Mircea Olteanu, and Paul Flondor. Severe acute pancreatitis between systematic inflammatory re-

sponse syndrome and sepsis: insights from a mathematical model of endotoxin tolerance. *The American Journal of Surgery*, 194(4):S33–S38, oct 2007. ISSN 00029610. doi: 10.1016/j.amjsurg.2007.05.002. URL http://linkinghub.elsevier.com/retrieve/pii/S0002961007003509http://www.sciencedirect.com/science/article/pii/S0002961007003509.

[83] Steve E Calvano and Susette M Coyle. Experimental human endotoxemia: a model of the systemic inflammatory response syndrome? *Surgical infections*, 13(5):293–9, oct 2012. ISSN 1557-8674. doi: 10.1089/sur.2012.155. URL http://www.ncbi.nlm.nih.gov/pubmed/23072275.

[84] Anthony F. Suffredini, H. Donald Hochstein, and F. Gilbert McMahon. Dose-Related Inflammatory Effects of Intravenous Endotoxin in Humans: Evaluation of a New Clinical Lot of Escherichia coli O:113 Endotoxin. *The Journal of Infectious Diseases*, 179(5):1278–1282, may 1999. ISSN 0022-1899. doi: 10.1086/314717. URL http://jid.oxfordjournals.org/lookup/doi/10.1086/314717https://academic.oup.com/jid/article-lookup/doi/10.1086/314717.

[85] William E. Hart, Jean-Paul Watson, and David L. Woodruff. Pyomo: modeling and solving mathematical programs in Python. *Mathematical Programming Computation*, 3(3):219–260, sep 2011. ISSN 1867-2949, 1867-2957. doi: 10.1007/s12532-011-0026-8. URL http://link.springer.com/10.1007/s12532-011-0026-8.

[86] William E. Hart, Carl D. Laird, Jean-Paul Watson, David L. Woodruff, Gabriel A. Hackebeil, Bethany L. Nicholson, and John D. Siirola. *Pyomo — Optimization Modeling in Python*, volume 67 of *Springer Optimization and Its Applications*. Springer International Publishing, Cham, 2017. ISBN 978-3-319-58819-3. doi: 10.1007/978-3-319-58821-6. URL http://link.springer.com/10.1007/978-3-319-58821-6.

[87] Robert Robere. Interior Point Methods and Linear Programming. *University of Toronto*, pages 1–15, 2012.

[88] J G Wagner and R a Roth. Neutrophil migration during endotoxemia. *Journal of leukocyte biology*, 66(1):10–24, 1999. ISSN 0741-5400.

[89] Elzbieta Kolaczkowska and Paul Kubes. Neutrophil recruitment and function in health and inflammation. *Nature reviews. Immunology*, 13(3):159–75, 2013. ISSN 1474-1741. doi: 10.1038/nri3399. URL http://www.ncbi.nlm.nih.gov/pubmed/23435331.

[90] Charlotte Summers, Sara M Rankin, Alison M Condliffe, Nanak Singh, A Michael Peters, and Edwin R Chilvers. Neutrophil kinetics in health and disease. *Trends in immunology*, 31(8):318–24, aug 2010. ISSN 1471-4981. doi: 10.1016/j.it.2010.05.006. URL http://www.ncbi.nlm.nih.gov/pubmed/20620114http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2930213.

[91] Thang Ho, Gilles Clermont, and Robert S Parker. A model of neutrophil dynamics in response to inflammatory and cancer chemotherapy challenges. *Computers & Chemical Engineering*, 51(51):187–196, apr 2013. ISSN 00981354. doi: 10.1016/j.compchemeng.2012.07.003. URL http://linkinghub.elsevier.com/retrieve/pii/S0098135412002311.

[92] Tekashi Terashima, Dean English, James C. Hogg, and Stephan F. van Eeden. Release of polymorphonuclear leukocytes from the bone marrow by interleukin-8. *Blood*, 92(3):1062–1069, 1998. URL http://www.bloodjournal.org/content/92/3/1062.

[93] Rebecca C. Furze and Sara M. Rankin. Neutrophil mobilization and clearance in the bone marrow. *Immunology*, 125(3):281–288, 2008. ISSN 00192805. doi: 10.1111/j.1365-2567.2008.02950.x.

[94] Craig L. Semerad, Fulu Liu, Alyssa D. Gregory, Katherine Stumpf, and Daniel C. Link. G-CSF is an essential regulator of neutrophil trafficking from the bone marrow to the blood. *Immunity*, 17(4):413–423, 2002. ISSN 10747613. doi: 10.1016/S1074-7613(02)00424-7.

[95] Lena E Friberg, Anja Henningsson, Hugo Maas, Laurent Nguyen, and Mats O. Karlsson. Model of chemotherapy-induced myelosuppression with parameter consistency across drugs. *Journal of Clinical Oncology*, 20(24):4713–4721, 2002. ISSN 0732183X. doi: 10.1200/JCO.2002.02.140.

[96] Christian D Sadik, Nancy D Kim, and Andrew D Luster. Neutrophils cascading their way to inflammation. *Trends in Immunology*, 32(10):452–460, 2011. ISSN 14714906. doi: 10.1016/j.it.2011.06.008. URL http://linkinghub.elsevier.com/retrieve/pii/S1471490611001141{%}5Cnhttp://dx.doi.org/10.1016/j.it.2011.06.008.

[97] Sarah Fox, Andrew E. Leitch, Rodger Duffin, Christopher Haslett, and Adriano G. Rossi. Neutrophil apoptosis: Relevance to the innate immune response and inflammatory disease. *Journal of Innate Immunity*, 2(3):216–227, 2010. ISSN 1662811X. doi: 10.1159/000284367.

[98] A Aderem and D M Underhill. Mechanisms of phagocytosis in macrophages. *Annual review of immunology*, 17:593–623, 1999. ISSN 0732-0582. doi: 10.1146/annurev.immunol.17.1.593. URL http://www.ncbi.nlm.nih.gov/pubmed/10358769.

[99] Günter Weiss and Ulrich E. Schaible. Macrophage defense mechanisms against intracellular bacteria. *Immunological Reviews*, 264(1):182–203, 2015. ISSN 1600065X. doi: 10.1111/imr.12266.

[100] G Meuret and G Hoffmann. Monocyte kinetic studies in normal and disease states. *British journal of haematology*, 24(3):275–85, 1973. ISSN 0007-1048. URL http://www.ncbi.nlm.nih.gov/pubmed/4713630.

[101] Francesc Miró-Mur, Isabel Pérez-de Puig, Maura Ferrer-Ferrer, Xabier Urra, Carles Justicia, Angel Chamorro, and Anna M. Planas. Immature monocytes recruited to the ischemic mouse brain differentiate into macrophages with features of alternative activation. *Brain, Behavior, and Immunity*, 53:18–33, 2016. ISSN 10902139. doi: 10.1016/j.bbi.2015.08.010.

[102] Jiong Wang, WeiLin L. Huang, Cheng Wang, and RongYu Y. Liu. Dynamic process of phagocytosis and forms of macrophage cell death induced by ingestion of apoptotic neutrophils. *Science China Life Sciences*, 57(10):1018–1023, 2014. ISSN 16747305. doi: 10.1007/s11427-014-4726-y.

[103] Yue Xu, Ouk Kim Sung, Yilei Li, and Jiahuai Han. Autophagy contributes to caspase-independent macrophage cell death. *Journal of Biological Chemistry*, 281(28):19179–19187, 2006. ISSN 00219258. doi: 10.1074/jbc.M513377200.

[104] Clemens Kreutz, Andreas Raue, Daniel Kaschek, and Jens Timmer. Profile likelihood in systems biology. *FEBS Journal*, 280(11):2564–2571, 2013. ISSN 1742464X. doi: 10.1111/febs.12276.

[105] Alan S Cross and Steven M Opal. A new paradigm for the treatment of sepsis: is it time to consider combination therapy? *Annals of internal medicine*, 138(6):502–5, mar 2003. ISSN 1539-3704. URL http://www.ncbi.nlm.nih.gov/pubmed/2639085http://www.ncbi.nlm.nih.gov/pubmed/12639085.

[106] Mark Bodmer, Michael A Fournel, and Lerner B Hinshaw. Preclinical review of anti-tumor necrosis factor monoclonal antibodies. *Critical care medicine*, 21(10 Suppl): S441–6, oct 1993. ISSN 0090-3493. URL http://www.ncbi.nlm.nih.gov/pubmed/8403982.

[107] Thomas a Moore, Michelle L Perry, Andrew G Getsoian, Christine L Monteleon, Anna L Cogen, and Theodore J Standiford. Increased Mortality and Dysregulated Cytokine Production in Tumor Necrosis Factor Receptor 1-Deficient Mice following Systemic Klebsiella pneumoniae Infection. *Infection and Immunity*, 71(9):4891–4900, 2003. ISSN 00199567. doi: 10.1128/IAI.71.9.4891.

[108] Morgan V DiLeo, John a Kellum, and William J Federspiel. A simple mathematical model of cytokine capture using a hemoadsorption device. *Annals of biomedical engineering*, 37(1):222–9, jan 2009. ISSN 1573-9686. doi: 10.1007/s10439-008-9587-8. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2758484{&}tool=pmcentrez{&}rendertype=abstract.

[109] G Schlag, H Redl, H Gasser, Z Khakpour, and J Davies. Successful Administration of the NO Synthase Inhibitor 546C88 as a Delayed Continuous Infusion in a Baboon Model of Septic Shock. In Günther Schlag and Heinz Redl, editors, *Shock, Sepsis, and Organ Failure*, pages 23–43. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999. ISBN 978-3-642-63672-1. doi: 10.1007/978-3-642-58630-9. URL http://link.springer.com/chapter/10.1007/978-3-642-58630-9{_}2http:

//www.springerlink.com/index/10.1007/978-3-642-58630-9http://link.springer.com/10.1007/978-3-642-58630-9.

[110] Jerome Friedman, Trevor Hastie, and Rob Tibshirani. Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of statistical software*, 33(1):1–22, jan 2010. ISSN 1548-7660. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2929880{&}tool=pmcentrez{&}rendertype=abstract.

[111] Nir Atias, Michal Gershenzon, Katia Labazin, and Roded Sharan. Experimental design schemes for learning Boolean network models. *Bioinformatics*, 30(17):i445–i452, 2014. ISSN 1367-4803, 1460-2059. doi: 10.1093/bioinformatics/btu451. URL http://bioinformatics.oxfordjournals.org/content/30/17/i445.

[112] I E Grossmann and L T Biegler. Part II. Future perspective on optimization. *Computers & Chemical Engineering*, 28(8):1193–1218, 2004. ISSN 00981354. doi: DOI10.1016/j.compchemeng.

[113] Franklin R Toapanta and Ted M Ross. Impaired immune responses in the lungs of aged mice following influenza infection. *Respiratory Research*, 10(1):112, dec 2009. ISSN 1465-993X. doi: 10.1186/1465-9921-10-112. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2785782{&}tool=pmcentrez{&}rendertype=abstracthttp://respiratory-research.biomedcentral.com/articles/10.1186/1465-9921-10-112.

[114] Jorge Nocedal and Stephen J Wright. *Numerical Optimization*. Springer-Verlag, New York, 1999. ISBN 9780387227429.

[115] A. Sokal. Monte Carlo Methods in Statistical Mechanics: Foundations and New Algorithms. In Cecile DeWitt-Morette, Pierre Cartier, and Antoine Folacci, editors, *Functional Integration: Basics and Applications*, pages 131–192. Springer US, 1997. ISBN 978-1-4899-0321-1. doi: 10.1007/978-1-4899-0319-8_6. URL http://www.stat.unc.edu/faculty/cji/Sokal.pdfhttp://link.springer.com/10.1007/978-1-4899-0319-8{_}6.

[116] Persi Diaconis. The Markov Chain Monte Carlo Revolution. *Bulletin of the American Mathematical Society*, 46:179–205, 2009. doi: 10.1090/S0273-0979-08-01238-X.

[117] Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman. emcee : The MCMC Hammer. *Publications of the Astronomical Society of the Pacific*, 125 (925):306–312, mar 2013. ISSN 00046280. doi: 10.1086/670067. URL http://arxiv.org/abs/1202.3665http://iopscience.iop.org/article/10.1086/670067.

[118] Arnaud Doucet, Simon Godsill, and Christophe Andrieu. On sequential Monte Carlo sampling methods for Bayesian filtering. *Statistics and Computing*, 10 (3):197–208, 2000. ISSN 09603174. doi: 10.1023/A:1008935410038. URL

http://link.springer.com/article/10.1023/A:1008935410038{%}5Cnhttp://link.springer.com/10.1023/A:1008935410038.

[119] Christophe Andrieu, Nando de Freitas, Arnaud Doucet, and Michael I. Jordan. An introduction to MCMC for machine learning. *Machine Learning*, 50(1):5–43, 2003. ISSN 1573-0565. doi: 10.1023/A:1020281327116. URL https://doi.org/10.1023/A:1020281327116.

[120] David J. Earl and Michael W. Deem. Parallel Tempering: Theory, Applications, and New Perspectives. *Physical Chemistry Chemical Physics*, 7 (23):3910, aug 2005. ISSN 1463-9076. doi: 10.1039/b509983h. URL http://arxiv.org/abs/physics/0508111http://dx.doi.org/10.1039/b509983hhttp://xlink.rsc.org/?DOI=b509983h.

[121] Gautam Altekar, Sandhya Dwarkadas, John P Huelsenbeck, and Fredrik Ronquist. Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics*, 20(3):407–415, feb 2004. ISSN 1367-4803. doi: 10.1093/bioinformatics/btg427. URL https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btg427.

[122] Jonathan Goodman and Jonathan Weare. Ensemble samplers with affine invariance. *Communications in Applied Mathematics and Computational Science*, 5(1):65–80, 2010. ISSN 1559-3940. doi: 10.2140/camcos.2010.5.65.

[123] Andrew Gelman and Kenneth Shirley. Inference from simulations and monitoring convergence. In S. P. Brooks, A. Gelman, G. Jones, and X-L Meng, editors, *Handbook of Markov Chain Monte Carlo*, chapter 6, pages 163–174. Chapman Hall, Boca Raton, 2011. ISBN 9781420079418. doi: 10.1201/b10905-7.

[124] Alan C Hindmarsh, Peter N Brown, Keith E Grant, Steven L Lee, Radu Serban, Dan E Shumaker, and Carol S Woodward. SUNDIALS: Suite of Nonlinear and Differential/Algebraic Equation Solvers. *ACM Trans. Math. Softw.*, 31(3):363–396, sep 2005. ISSN 0098-3500. doi: 10.1145/1089014.1089020. URL http://doi.acm.org/10.1145/1089014.1089020.

[125] Heiko Bauke and Stephan Mertens. Random numbers for large-scale distributed Monte Carlo simulations. *Physical Review E*, 75(6):1–14, 2007. doi: 10.1103/PhysRevE.75.066701.

[126] Momin Jamil and Xin-She Yang. A Literature Survey of Benchmark Functions For Global Optimization Problems. *International Journal of Mathematical Modelling and Numerical Optimisation*, 4(2):150–194, aug 2013. ISSN 2040-3607. doi: 10.1504/IJMMNO.2013.055204. URL http://arxiv.org/abs/1308.4008http://dx.doi.org/10.1504/IJMMNO.2013.055204.

[127] H. Scott Fogler. Bioreactors. In *Elements of Chemical Reaction Engineering*, chapter 7, pages 432–434. Pearson Prentice Hall, Upper Saddle River, NJ 07458, fourth edition, 2013. ISBN 9780131278394.

[128] Jason P Fine, Robert J Gray, Jason P Fine, and Robert J Gray. A Proportional Hazards Model for the Subdistribution of a Competing Risk Stable URL : http://www.jstor.org/stable/2670170 All use subject to http://about.jstor.org/terms A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association*, 94(446):496–509, 1999.

[129] Thang Ho. *A Model-Based Clinically-Relevant Chemotherapy Scheduling Algorithm for Anticancer Agents*. PhD thesis, University of Pittsburgh, 2007.

[130] Sven Zenker, Jonathan Rubin, and Gilles Clermont. From inverse problems in mathematical physiology to quantitative differential diagnoses. *PLoS computational biology*, 3(11):e204, nov 2007. ISSN 1553-7358. doi: 10.1371/journal.pcbi.0030204. URL http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2065888{&}tool=pmcentrez{&}rendertype=abstract.