Dynamic Mathematical Approaches to Understanding Pain in Sickle Cell Disease

Quindel Jones
Dynamic Mathematical Approaches to Understanding Pain in Sickle Cell Disease

Quindel D. Jones

Dissertation submitted to the Faculty of the Virginia Commonwealth University in fulfillment of the requirements for the degree of

Doctor of Philosophy in Systems Modeling and Analysis

Rebecca Segal, Advisor
Cheng Ly
D’Arcy Mays
Cecelia Valrie

May 07, 2024
Richmond, Virginia

Keywords: sickle cell disease; ODE model; pain; sensitivity analysis; vaso-occlusion; inflammation; sleep;
ACKNOWLEDGEMENTS

I want to thank and acknowledge the people who have supported this journey and kept me excited and engaged with contributing to my community.

To my mother - Sonya Parker, my brother - Dr. Bobby Portis II, my friend - Dr. Chelsea Jones, and the rest of my family:

I would not be who I am, how I am, and where I am without your love and support. Thank you for carrying me through this experience and keeping me encouraged.

To my advisor - Dr. Rebecca Segal, research team (Dr. RB McGee, Dr. Cecelia Valrie, and Dr. Angela Reynolds), and Ph.D. committee (Dr. Cheng Ly, Dr. D’Arcy Mays):

I have learned and grown so much in your presence and I cannot adequately put into words how grateful I have been for your kindness, mentorship, and support. Thank you for all of your contributions within this work and within my life.

And a special thank you the Southern Regional Education Board (SREB) Fellowship and the NIH HEAL grant for financially supporting me through this journey.

I have made it my personal mission to use the skills I have gained to alleviate as much pain and suffering in my community as I can; starting with a better understanding of pain dynamics in Sickle Cell Disease. Growing up in the Deep South of America, specifically Jackson, MS, I got a first hand experience of the health disparities prevalent in this country for Black Americans. In a city of food deserts, high stress policing, and underpaid labor, illness and disease is a common occurrence where I’m from. Many members of my family developed disease conditions due to multiple factors, environment among them, yet had very little understanding or resources to combat their impact. Watching my family and community suffer with little to no recourse sparked my initial interest in pathology, the study of the cause and effects of disease. Once I later realized that my affinity for mathematics and its problem - solving nature could enable me to find solutions for my community, I embarked on this journey to become a mathematical scientist for public health, with a particular interest in under-funded and under-researched diseases prevalent in the Black community.
Abstract

In the last 50 years, pain in Sickle Cell Disease (SCD) has become more widely studied thanks to advancements in technology and awareness. Clinical studies, population surveys, pharmaceutical trials, and computational models have been created and conducted to elucidate the mechanisms, treatments, and prediction of Sickle Cell disease pain episodes. Computational models have become quite useful in unraveling disease pathology [33] with the rise in data collection accessibility and advanced computational power. In particular, dynamic mathematical models have been used to investigate Sickle Cell disease pathology and treatment. In this work we conduct a literature review of mathematical models used in SCD research and present two modeling approaches for understanding pain behavior and treatment in SCD - one behavioral approach and one biological approach.

The third chapter examines and elaborates on the connection between sleep quality and pain in SCD using a differential equations (ODE) model. The ODE model uses patient mobile health (mHealth) data to compute a simulated pain profile that is compared to raw patient data for predictive value and analysis. The results show that the sleep - pain relationship becomes more significant with increasing incidence of pain, which is correlated with patient age, supporting the theory that SCD's progression into a chronic pain state occurs in adolescence.

The fourth chapter switches from a behavioral model of pain in SCD to a biological model to investigate the underlying mechanisms of the painful vaso-occlusive crises - the hallmark occurrence of SCD and cause of the recurrent pain episodes. The recurrence of vaso-occlusive crisis has been revealed to be exacerbated by the chronic inflammatory processes associated with SCD pathophysiology [44] so we have formulated an ODE model of the temporal dynamics of inflammation and adhesion in Sickle Cell disease. The ODE model uses hematological blood data to estimate cell concentrations and their impact on the likelihood of secondary VOC (sVOC). Model formulation allows for the investigation of potential mechanistic pathways to treat or prevent VOC duration and frequency. Our results identify key parameters in the VOC process and provide some mechanistic explanation for the changes in patient state during crisis.

Through these models, we highlight the power of mathematical biology to change the lives of patients in medicine, through little to no cost to the patient or family. As more data is collected and becomes available, models like these can be used to investigate these complex processes more thoroughly with minimal cost and injury to patients.
Contents

1 Introduction
  1.1 Sickle Cell Disease Etymology ............................................. 11
  1.2 AntiBlackness and Health Disparities Worldwide ....................... 19

2 Quantitative Modeling of Sickle Cell Disease Literature Review
  2.1 Introduction ................................................................. 10
     2.1.1 Biological Background ............................................... 8
     2.1.2 Mathematical Background ......................................... 12
  2.2 Quantitative Modeling of SCD ........................................... 13
     2.2.1 Statistical Modeling ................................................ 13
     2.2.2 Machine Learning Models .......................................... 13
     2.2.3 Network Models ..................................................... 15
     2.2.4 Discrete Time Modeling: Difference Equations Models .......... 15
     2.2.5 Differential Equations ............................................. 16
     2.2.6 Partial Differential Equations Modeling ......................... 19
     2.2.7 Hybrid Models .......................................................... 20
  2.3 Choosing a Math Model Type ............................................. 21
  2.4 Models by SCD Research Area Summary .................................. 23
     2.4.1 Modeling Hemodynamics and Vaso-occlusive Crises in SCD .... 23
     2.4.2 Modeling Cellular Dynamics + Rheology of SCD ................ 24
     2.4.3 Modeling Genetic Inheritance and Mutation in SCD ............. 26
     2.4.4 Modeling Pain Behavior and Management in SCD ................. 26
     2.4.5 Modeling Treatment Efficacy in SCD ............................. 29
  2.5 Conclusion: Future of Modeling ......................................... 29
  2.6 Acknowledgements ....................................................... 30
     2.6.1 TABLES ................................................................. 30

3 Leveraging the Sleep-Pain Relationship in an ODE Model to Predict Acute Pain in Pediatric Sickle Cell Disease 43
  3.1 Introduction ................................................................. 43
     3.1.1 Background Application to Pain ................................... 43
  3.2 Materials/Data ............................................................ 45
     3.2.1 Data source: .......................................................... 45
  3.3 Data .......................................................................... 47
  3.4 Methods and Formulation ................................................ 49
     3.4.1 Modeling Framework: ............................................... 48
     3.4.2 ODE Model .......................................................... 49
     3.4.3 Initial Transient Results .......................................... 52
  3.5 Sensitivity Analysis ........................................................ 54
     3.5.1 Quantifying the Cumulative Impact of Sleep and Defining 'Bad Sleep' 54
     3.5.2 Pain onset vs pain level - categorizing pain level through sensitivity testing 56
     3.5.3 Opioid Usage as a Predictability Factor? ........................ 57
  3.6 Threshold Sensitivity Testing Results .................................. 58
     3.6.1 Finding Patient Trends ............................................. 58
4 Temporal Dynamics of Inflammation and Adhesion in Sickle Cell Vaso-Occlusion: An ODE Approach

4.1 Introduction

4.1.1 Overview of Inflammation-Driven Vaso-Occlusive Crises

4.2 History of VOC Modeling and Comparison of Our Approach

4.2.1 Red Blood Cell Models

4.2.2 Models of VOC and Inflammation

4.2.3 Research Aims

4.3 Methods and Formulation

4.3.1 Assumptions on the Subject

4.3.2 Assumptions on Physiological Process

4.3.3 Model Equations and Variables

4.3.4 Model Equations and sVOC Likelihood

4.3.5 Parameters

4.3.6 Model Fitting and Validation: Klouda et al 2020 Data

4.4 Results

4.4.1 Parameter Set Collections: Healthy, Baseline VOC, VOC Crisis

4.4.2 Determining Predictors: Parameter Estimation + Ranking

4.4.3 Understanding the Driving Dynamics

4.4.4 Predictive Analysis of Interventions

4.5 Discussion

4.5.1 Which inflammatory cell population(s) impacted sVOC likelihood the most? (statistically significant?)

4.5.2 Do any inflammatory marker relationships results hint towards treating pain preventatively instead of after?

4.5.3 Which treatment combo can be used for reducing pain?

4.5.4 Model Limitations and Future Work

5 CONCLUSION

5.1 The Benefits of Dynamic mathematical techniques in SCD research

5.2 The Importance of the Sleep-Pain connection in SCD pain prediction

5.3 Potential Mechanistic Pathways for SCD sVOC prevention

5.4 Using Mathematical Techniques for Medicine

6 Appendix

6.1 SCD Sleep-Pain Model Code:

6.2 sVOC ODE Model Code:

6.3 LHS Sampling Code for the sVOC Model from [61]:

[4]
# List of Tables

1. MATH MODELS SUMMARY ........................................... 31
2. Summary of VOC models .......................................... 33
3. Summary of RBC models .......................................... 35
4. Summary of Gene models ........................................ 38
5. Summary of Pain models ......................................... 39
6. Summary of Treatment models ................................. 40
7. Patient Demographic Information .............................. 46
8. Variables in Mechanistic Model ............................... 50
9. Parameters in Mechanistic Model ............................. 51
List of Figures

1. Sickle Cell Disease History[84] ............................. 2
2. US Government Disease Funding 2022[81] ............... 4
3. History of Black Civil Rights in the United States ......... 6
4. Sickle Cell Biology: Genetic Mutation ..................... 9
5. Sickle Cell Biology: Red Blood Cell Sickling or Gelation . 10
7. How to Choose a Math Modeling Technique .............. 21
8. Patient A10 raw sleep and pain data ....................... 47
9. Patient C46 raw sleep and pain data ....................... 47
10. Schematic flowchart showing model framework .......... 49
11. Pediatric patient reported sleep quality (in blue) over 10 days (240 hours) for individual patients ........ 53
12. Pediatric patient pain episode model prediction (in red) compared to the patient's reported pain (black dashes) and reported drug dose times (vertical lines) over 10 days (240 hours) . 53
13. Defining "Bad Sleep" and Sleep Window Accumulation .......... 55
14. Pain Level Threshold Examples ........................... 56
15. Model Performance Predictive Value Table - Opioid Bins . 57
16. Pain Model Sensitivity as Bad Sleep Threshold Lowers ........ 58
17. Pain Model Sensitivity as Pain Threshold Rises ........... 59
18. Pain Model Sensitivity as Sleep Accumulation Window Increases .......... 59
19. Testing for Best Thresholds - Variable Sweep Tree ........ 61
20. 2DAY Sleep Accumulation Window Results: Sleep Quality Bins .... 92
21. 2DAY Sleep Accumulation Window Results: Children .......... 93
22. 2DAY Sleep Accumulation Window Results: Adolescents ........ 93
23. Average Percent of Pain Days ................................ 64
24. 2DAY Sleep Accumulation Window Results: Adolescents No Medication vs Non-Opioids vs Opioids .................................. 65
25. Average Percent of Pain Days by Opioid Usage .............. 66
26. Percentage of Pain Days - Binned by Quarters .............. 67
27. Model Sensitivity Based on Percentage of Pain Days + Age .......... 68
28. Model Sensitivity Based on Percentage of Pain Days + Opioid Usage ........ 68
29. Schematic of Vaso-Occlusive Crisis inside Vessel .......... 73
30. Sickle Cell VOC Model Variables/Cells ................... 74
31. Inflammation Cycle and Pathophysiology of Sickle Cell VOC ....... 75
32. VOC Initialization: HbS Polymerization and Blood Cell Hemolysis .... 76
33. The Vaso-Occlusive Crisis and Sterile Inflammation .......... 78
34. Phases of Pain During Vaso-Occlusive Crisis ............... 79
35. VOC ODE Model Schematic .................................. 82
36. Klouda et al 2020 Inflammatory Marker DataSet Summary[51] .... 92
37. Scatterplot of Baseline Hgb vs RBC count .......... 93
38. Scatterplot of Baseline Platelet vs Baseline WBC ........... 94
39. Scatterplot of Baseline ANC vs Baseline WBC ............. 94
40. WBC Percent Change vs Hgb F % ............................ 95
41. Platelet Percent Change vs Hgb F % .......................... 96
Healthy Patient - no SCD present

Note: This patient has cell well below SCD baseline values as seen in Figure 47. The model transient does not reach a crisis state value either (the red line) since the sickling rate $p_s = 0$, the death rate of Sickle Cells is $\mu_s = 0$, and their source of platelets is minimal $s_P = 6$.

Healthy Patient - all cells (individually) $p_s = 0 \& \mu_s = 0 \& s_P = 6$.

3 Sample Cases of Parameter Set Collections for VOC Crisis as seen in Figure 62.

SCD VOC Baseline to Crisis state (main outcome cells) $p_s = .45, \mu_s = 0.05, s_P = 50$.

Note: This patient has cell baselines set from the data as $Y0$ and the model transient reaches their resulting crisis state value (the red line) through parameter value changes. SCD patient parameters start with sickling rate $p_s = .45$, the death rate of Sickle Cells is $\mu_s = 0.05$, and their source of platelets is elevated from pain history $s_P = 50$.

SCD crisis state (main outcome cells) $p_s = .45 & \mu_s = 0.05 & s_P = 41$.

sVOC Model Transient Using LHS matrix parameter set 6.

sVOC Model Transient Using LHS matrix parameter set 17.

sVOC Model Transient Using LHS matrix parameter set 42.

sVOC Model Transient Using LHS matrix parameter set 229.

PRCC plot: $PA \ vs \ s_P \ | \ p-value = 1.4942e-315$.

PRCC plot: $EA \ vs \ \beta_{ead} \ | \ p-value = 1.1328e-194$.

PRCC plot: $E A \ vs \ s_P \ | \ p-value = 5.4917e-164$.

PRCC plot: $AM \ vs \ \mu_A M \ | \ p-value = 1.1015e-128$.

PRCC plot: $NA \ vs \ s_P \ | \ p-value = 2.2425e-180$.

PRCC table (total size 48x10).

Patient 6 VOC Crisis System Behavior.

Patient 40 VOC Crisis State System Behavior.

Patient 37 VOC Crisis State System Behavior.
1. Introduction

1.1. Sickle Cell Disease Etymology

Sickle cell disease was discovered in 1910 by James B. Herrick, M.D. when he identified the first case of abnormally shaped blood cells of a West Indian dental student, though the disease existed long before that. The disease is primarily found in people of African descent, with the Center for Disease Control (CDC) estimating every 1 in 365 Black Americans as born with the disease [10]. Nearly 40 years later in 1949, Linus Pauling and associates dubbed Sickle Cell Anemia the first "molecular disease" and created the term now known as "molecular medicine" [40]. After another 20 years, the National Sickle Cell Anemia Control Act was established in 1972 and was the first time researchers were provided the authority to establish education, information, screening, testing, counseling, research, and treatment programs. The following year in 1973, 12 public health labs began testing for SCD [84].

With 10 years of research knowledge gained, the NIH consensus conference of 1987 recommended universal new born blood screen testing for SCD, a major stride in the validation and diagnosis of the disease and it's severity. Hydroxyurea, a drug treatment that reduces the sickling rate of red blood cells and keeps them rounder for longer, was then established and approved for adults by 1998. Over the last two decades, there have been more advancements, as seen in Figure [1] thanks to growing recognition from government officials [84].
Some of the most notable new treatments include:

- L-glutamine approved as treatment for children and adults 2017
- Crizanlizumab approved 2019
- Voxelotor approved 2021
- CRISPR gene therapy approved 2023

This timeline shows that it took nearly 50 years between the discovery of SCD to the development of the first treatment of the disease, delaying the progress of SCD research. Given the severity of the disease symptoms, the state of SCD research has been at a huge disadvantage by not having the resources and awareness for discovering finding treatments for this illness.
1.2. AntiBlackness and Health Disparities Worldwide

Despite being the first genetic disease found and impacting populations across continents, SCD has been significantly understudied and under-treated. While many other diseases that also impact the immune system, inflammatory system, and nervous system have been appropriately funded, Sickle Cell Disease has only recently become a major research topic of conversation. The Figure 2 displays the US government funding in 2022 for various disease categories based on grants, contracts, and other funding mechanisms used across the National Institutes of Health (NIH), as well as disease burden data published by the National Center for Health Statistics (NCHS) at the Centers for Disease Control & Prevention (CDC) [81]. Compared to other inflammatory diseases, blood diseases, and diseases with the same high prevalence, Sickle Cell Disease has been comparatively under-funded. It can be seen at the very bottom of each chart, that it tends to be allocated the least amount of funding.
Figure 2: US Government Disease Funding 2022
One proposed reason for this delay in resources can be attributed to worldwide Anti-Blackness that permeates all corners of society, but specifically the fields of public health and medicine \[1\].

The US constitution was written to see Black Americans as \(3/5s\) a person, non-people who were just an extension of their personal property, and "scientific racism" was used to justify the superiority of white Americans over Black people. \[1\] Even after the 13th amendment outlawed slavery as we knew it, Anti-Black racism was still set as the foundation of the cultural and medical landscape \[1\].

AntiBlackness is a term to highlight the ways in which Black people in particular experience racism \[28\]. While many racial and ethnic groups experience racism or prejudice, anti-Black racism permeates across ethnic groups and cultures since it is foundational to the establishment of the US. Any person socialized in the US grows up in this culture built on the enslavement of Black people and the displacement of Native Americans, see Figure 3 \(28\).

Throughout world history, there have been examples of medical professionals and government officials using the subjugation of Black people to make advancements in the field. From the founder of gynecology, James Marion Sims, perfecting his craft on enslaved Black women without anaesthesia or medication \[1\], to the Tuskegee experiment using Black men's blood to study syphilis without aid \[1\], to the unauthorized use of Henrietta Lacks' DNA \[100\], there is a long history of indifference toward Black people's pain and suffering in favor of medical advancements.
Health disparities and resource disparities for Black people exist all around the US, with the treatment of minority communities during the SARS-CoV-2 outbreak as a primary example [28]. Many government, health, and educational programs and entities have foundational systems in place that exacerbate the historical and still present subjugation of Black people [28] [100].

Therefore, one might surmise that since the patient demographic that would benefit from support in SCD is majority Black people, hospitals, companies, and many other institutions are less likely to create treatments due to their lack of interest in the population and the lack of research funding. The treatments that are created often price out the disease demographic entirely. For example, the newest cure for SCD, the CRISPR gene treatment, which gained Food and Drug Administration (FDA) approval in 2023, costs between $450,000 - $2 million per treatment, whereas the highest patient demographic has a median family income of $48,297 a year. [81]

While investigating pain in Sickle Cell Disease in America, we must not negate or ignore the historical implications of disparity and stress on the largest SCD patient demographic, those of African descent. In
the nearly 100 years between SCD discovery and treatment, many Black Americans had to continually endure even more systemic and structural challenges and hardships. Discovering treatments to reduce physiological pain is a good first step, but systemic change will be required for sustainable, long term success of alleviating pain for SCD patients over generations.

In order to do what I can in alleviating the mystery and pain for my people, I’ve studied dynamical approaches to understanding pain in Sickle Cell disease.

In Chapter 2 we discuss how math has been used to investigate Sickle Cell disease through a literature review of math models used in SCD research.

In Chapter 3 we present an ordinary differential equations (ODE) model for pain prediction leveraging the sleep-pain connection in SCD and further investigate the strength of this connection.

In Chapter 4 we develop a mechanistic ODE model of the inflammatory cycle in the vaso-occlusive crisis process in Sickle Cell disease and use it to understand some of the key components to the onset of a sickle cell crisis.
2. Quantitative Modeling of Sickle Cell Disease Literature Review

This work is from a journal submission that is currently under review.

2.1. Introduction

Sickle cell disease (SCD) is a family of genetic blood disorders caused by mutated hemoglobin genes that are responsible for beta-globin production. A single gene mutation ultimately leads to an abnormal hemoglobin, which distorts the erythrocyte, or red blood cell (RBC), from within. The Centers for Disease Control and Prevention (CDC) estimates that there are approximately 100,000 people living with SCD in the US, with 1 of every 365 Black-American births and 1 of every 16,300 Hispanic births affected. Historically, there has been little data collected on SCD patients, with widespread clinical research efforts only being established in the 1970’s. Even today, few treatment options for SCD exist. Population-based studies as well as in vitro studies reveal the dynamic, unpredictable nature of SCD complications. Trying to better predict these outcomes is at the core of SCD prevention and control [48]. The current age of computational power and "big data" analytics has brought new techniques to better predict outcomes of interest, using population-based studies of SCD, along with the traditional basic science studies of RBCs and tissues.

One available tool of both population-based and basic SCD research is the use of mathematical modeling techniques. We discuss modeling frameworks which can expand the scope of biological SCD research in this review.

2.1.1. Biological Background

Hemoglobin is a large, complex protein within RBCs that is responsible for oxygen transport. Adult Hemoglobin (normal hemoglobin A, HbA) is made up of four protein chains: 2 alpha-globin chains and 2 beta-globin chains. The beta chains in SCD are abnormal, caused by a genetic mutation occurring in the $HBB$ gene [48].

The abnormal beta chains result in abnormal hemoglobin molecules (hemoglobin S, HbS) inside the RBCs, causing the RBCs to no longer be flexible and doughnut-shaped, but instead inflexible and
crescent or sickle-shaped. These inflexible RBCs cannot traverse very small blood vessels, and the subsequent backup obstructs blood flow. The term vaso-occlusion has been applied mainly to this red-cell-caused obstruction. But the abnormal RBC obstruction is accompanied by a cascade of white cell, platelet, and blood vessel wall abnormalities, and is more correctly termed vasculopathy. While the primary complication of SCD is severe, acute painful episodes called vaso-occlusive crises (VOCs), the vasculopathy results in destructive anemia, ubiquitous and sometimes chronic pain, inflammation, widespread tissue destruction, organ failure, and ultimately early death [106].

Genetically, SCD is transmitted via autosomal recessive inheritance. In autosomal recessive inheritance, the host is an asymptomatic carrier if only one SCD gene mutation is transmitted from the host’s parents, but a symptomatic patient if a mutated gene is transmitted from both parents and at least one is a SCD gene mutation. Different mutation combinations of the HBB gene produce different alternative symptomatic forms of SCD, with the most severe and common mutation being HbSS (homozygous SCD, also called sickle cell anemia). Other forms of SCD include HbSC, HbSB0, HBSC, HbSB+, HbSD, HbSO, HbSE [48]. See Figure 1 where an HbSS child inherits an abnormal (red) HBB gene from each of their parents.

Pathologically, sickle shaped RBCs are less deformable. Their cellular membranes are weaker. The
cells literally burst (hemolysis) easily and die prematurely, lasting only 10-20 days rather than the usual 120 days. The resulting shortage of total RBCs in the body causes anemia, or low blood counts, usually associated with weakness and easy fatigue. RBCs may not always be sickle-shaped under the microscope. This is because they only assume abnormal shapes in low oxygen conditions. The mutated hemoglobin S protein chains inside the red cell can adhere to each other more easily, making rigid strings or polymers, only when they are not carrying oxygen.\footnote{53} \textbf{See Figure 2.}

Only under low oxygen states can hemoglobin polymerize and sustain long chains, grow, and stretch the red cell membrane, and alter the shape of the red blood cell from doughnut to sickle shaped. This is the essential process or Hemoglobin S gelation.\footnote{80}

The cells' new rigidity and stickiness causes them to clump together and stick to the walls of vessels in the body. These obstructions lead to reduced oxygen supply to various organs. This is the process of vaso-occlusion and is the cause of the painful, hallmark, unpredictable VOCs. Indeed, vaso-occlusion is the underlying genesis of most SCD complications throughout the body. \textbf{See Figure 3} for an example of an obstructed vessel caused by a VOC. VOCs last from hours to weeks and may occur as few as once a year during childhood, but more often for adolescents and adults. The types of SCD complications are vast, including skin ulcers, pulmonary hypertension, retinopathy, cardiomegaly, delayed puberty, and erectile dysfunction.\footnote{78} Acute complications, ones that require immediate medical attention, include rapid spleen swelling due to blood flow obstruction (splenic sequestration), ischemic stroke (occlusion
of a large cerebral artery), and acute chest syndrome (lung sickling; the leading cause of death in adults).

Because SCD complications occur all over the body, systemic treatment is needed to address the full extent of the disease. However, limited options are currently available. Bone marrow transplants and gene therapy can be used to replace diseased cells and are the only functional cure for SCD [31]. Since successful bone transplant matches are a rarity in the disease demographic, and since gene therapy is also invasive and expensive, most SCD treatments aim to relieve pain and symptoms, avoid VOCs, and prevent organ complications. [78] Opioids and other analgesics are often prescribed to relieve pain. Periodic blood transfusions are often given to reduce anemia. Preventative disease modifiers, including hydroxyurea and newer anti-sickling drugs that can be additive to hydroxyurea [60], are given to promote healthy red cells and suppress sickling inside the RBCs. Vaccinations and prophylactic antibiotics are provided to counteract the high rate of infections.

Until more accessible curative options are available, a critical research goal is early detection as well as prediction and prevention of symptoms and complications with drugs and blood [42]. Unfortunately, clinicians often have incomplete risk information, or have difficulty risk-stratifying adults with SCD, in order to quickly intervene on those at greatest risk. Ideal risk stratifiers would be simple, feasible, easily accessible, and available early. They would predict frequent hospitalizations, pulmonary hypertension or renal failure as harbingers of death itself.

Researchers are therefore developing computational models of various biological phenomena in SCD. Mathematical modeling can allow researchers to quantitatively represent multiple components and scales
of a system, and investigate the dynamic behavior of these components and their interactions over time under various conditions. These models enable the analysis of a system on scales that range from inter- and intra-cellular tissue, organ, host, and population. Many mathematical techniques can be used for prediction and simulation of biological systems and events that aid in the decision-making process for potential disease treatment protocols and policies.

Current research efforts include using mathematical modeling to understand the inheritance frequency of the $HBB$ gene mutation amongst different populations. These investigations also aim to provide clarity on the existence of genetic markers for SCD and optimize gene therapy intervention. Other researchers have focused on the role RBC system dynamics play in SCD complications, primarily the sickling process and mechanics of vaso-occlusion. Beyond the mechanics of SCD, computational researchers are interested in understanding the presentation and frequency of pain in individuals with SCD. They are also interested in investigating the feasibility and efficacy of different treatment options.

2.1.2. Mathematical Background

When studying models, it is important to understand the broad classifications they can fall into, since this characterizes the essentials of their structure. Mathematical models can be mechanistic or empirical. Models that account for the mechanisms through which changes occur are called mechanistic models. The mechanisms typically describe physical and/or chemical processes taking place in the system. In SCD research, these models typically investigate processes such as RBC sickling and vaso-occlusion. On the other hand, empirical models do not incorporate the mechanisms by which changes occur but instead use mathematical structures to mimic observed behavior patterns. Mathematical models like these are sometimes used to describe genetic mutation frequencies and SCD pain episode severity or give heuristic descriptions of outcomes.

Another major classification category of models indicates whether the equations account for stochastic (random) events. Deterministic models always predict the same outcome from a given starting point, disregarding any random variation. Models that do include variation can predict a distribution of possible outcomes even with a common starting point. These models are known as stochastic models. Of the four classifications above, models can fall into one of each kind, i.e., empirical/mechanistic, and deter-
ministic/stochastic. Below we will briefly introduce how these common modeling techniques are used to study SCD at various levels, ranging from the molecule to the cell to the organ to populations.

2.2. Quantitative Modeling of SCD

2.2.1. Statistical Modeling

Since many researchers collect data to test their hypotheses, statistical methods are commonly used to gain an understanding of relationships between quantities measured in the data. Common examples of these include line of best fit testing and regression. Regression quantifies the impact that variables have on the model output. While researchers of various fields use regression and statistical tools, statistical methods can also be used to supplement other modeling approaches; with regression tools being used to conduct statistical analysis beforehand. A model is a mathematical expression that quantitatively represents the relationships and dynamics of a system based on the data provided. Once an initial model is created, statistical methods can be used to assess model accuracy and provide insight into variable relationships through parameter estimation. For example, Chalacheva et al. [13] utilized multi regression analysis as part of determining underlying physiological mechanisms in the vaso-occlusion process based on patient lab data and blood samples.

Regression is also especially useful for identification of correlation between continuous variables, making it a common tool for investigating collinearity (if multiple predictor variables are correlated - skewing individual variable impact). When studying complex systems such as SCD, regression models often have multiple variables of interest, making multivariate regression a useful tool, for instance, in investigating measurable indicators for VOCs in SCD.

2.2.2. Machine Learning Models

The improved accuracy brought by combining computational modeling alongside statistical analysis has led to an increase in using machine learning (ML) techniques for data-informed research. ML is the computational study of statistical models and algorithms using only pattern recognition and inference to predict outcomes from sample data. ML and statistical techniques are commonly used in pain
and disease progression research. Quantitative models built through ML do not require explicit instruction, but rather sample data and appropriate algorithmic application, allowing researchers to analyze data without needing to know too much about the dynamics of a system. Since pattern recognition is a primary component of classification, SCD researchers may use ML to classify/categorize SCD disease genotype, treatment outcome probabilities, and/or SCD pain severity. Xu et al. [99], for example, created a ML model for sickle RBC classification from patient specific microscopy image data. Since early symptom detection is a major concern in SCD, ML algorithms can be extremely useful in processing patient data and predicting future behavior, especially from image data [99] [77] [22] [79]. Most recently, though, Darrin et al. 2023, used the “fact that the individual movement of an RBC in shear flow is an indicator of its deformability,” along with ML, to classify RBCs through cell motion videos [21]. Other ML tasks include cluster analysis, association analysis, and anomaly detection.

Because of its predictive power, ML is also often used to predict pain patterns and pain levels. Khalaf 2017 tested several ML algorithms to find the best drug dosage classification for a particular SCD pain treatment [50]. Yang et al. 2018 [104] and 2019 [103] produced ML models that showed promising results for predicting pain scores based on physiological patient data. While Yang [104] first used self-reported data and later [103] used wearable data, both models highlighted that the patient cohort plays a major role in the relationship between pain and physiological symptoms. Though ML has improved the ability to predict pain levels, studies of the relationship between pain and physiology are often inconclusive due to the limited size of the patient datasets.

Another example of the utility of ML is the use of physiological measurements for pain prediction. The Johnson 2019 ML model tested the feasibility of using Microsoft bands health data to predict patient reported pain scores [46] while Panaggio 2021 [74] used patient vital sign data. Padhee et al. 2021 [72] also used vital sign data but focused on determining which ML method best predicts pain scores. Patel 2019 used ML in their model but aimed instead to predict hospital readmission rates from raw hospital data, outperforming standard hospital scoring systems [76]. Similarly, Padhee 2022 [71] recently designed a ML classification algorithm to build pain prediction models using electronic health record data.

While machine learning’s ability for prediction with limited knowledge of system dynamics can be a great benefit, the lack of explicit instruction or explanation of network structure often makes them hard
to interpret. Model complexity and data size are additional features of ML models that may decrease their interpret-ability.

2.2.3. Network Models

Another mathematical technique to model systems is network modeling (NM). NM involves creating graphs out of qualitative networks for analysis, where graph nodes (or vertices) are the system elements/-variables and the edges, or lines/links, represent the pairwise relationships and connect the nodes. A system can be thought of as functioning through the relationships and connections between variables. These connections and interactions can be represented through dynamic NM. These models are useful since techniques from graph theory can be used to glean new biological insights and make predictions from the results.

One application of NM in SCD is modeling RBC behavior. Since disease progression depends upon poor RBC flow, a network model with each node representative of a RBC and each link representative of the processes and interactions connecting each RBC can be useful. A biological system's underlying network organization may sometimes be straightforward (linear) but often is incredibly complex.

One example can be seen in Sebastiani et al.'s [86] work developing a predictive model of SCD severity using a Bayesian NM approach - a probability technique that quantifies uncertainty and incorporates prior knowledge into model analysis. Mehraei et al. [63]'s multi-scale network model functions similarly to Lu 2019's treatment model for drug efficacy prediction. In these networks, the quantification of variable relationships are often incorporated through the graph edge definitions used. More recently, Balamanikandan 2022 [3] utilized a quantum graph theory mathematical model to extract elasticity properties and distinguish unsickled RBCs from sickled RBCs with DNA sequence data as input.

2.2.4. Discrete Time Modeling: Difference Equations Models

While statistical and network-based models are great tools for understanding static variable relationships and prediction, discrete dynamical models may incorporate system patterns based on the differences in variable measurements over time. Discrete data (data taken at distinct time points), can be modeled by a difference equation.
\[ x_t = f(x_{t-1}) \]

Difference equations are known as recursively defined equations, since to determine the next value of a variable, it is necessary to know the previous state of the variable.

For example, Chalacheva et al. 2015 used experimental lab data and created a discrete dynamical system (DDS) model to identify key mechanisms that influence observed autonomic responses triggered from stimulation known to aggravate VOC. Different control groups’ data were used to highlight mechanistic differences among individuals with SCD. Equations for heart rate variability and peripheral resistance variability were calculated using blood flow related mechanisms for each group. Statistical regression was then used to see which response variable could be a potential contributing factor to VOC [11].

In Chalacheva 2017 [12], a discrete dynamical systems (DDS) model was employed to elucidate the relationships between physiological responses of interest and the experimentally applied stimulus. While SCD severity is commonly measured by genotype, the model equations based on underlying physiology enabled Chalacheva to decompose the total vaso-occlusive response to pain into 4 measurable mechanisms of interest. The model showed that the biophysical marker neurogenic-vascular interaction BMn-v enhanced pain-induced vasoconstriction in SCD patients. A similar approach to finding biomarkers associated with VOC also used difference equations and statistical regression [13], with head-up tilt being the stimulation applied to induce VOC.

Dynamic discrete models have also been used to elucidate sickle RBC behavior. Li et al. 2017’s [56] multi-scale DDS model equations represented the network of vertices along the cell membrane to understand sickle RBC membrane behavior under gelation. Bazzi et al. 2020 [6] also used a discrete modeling approach to locate systemic causes of changes in SCD blood during high oxygen tensions with a system of finite difference equations. Their combined experimental and computational approach enabled the quantification of sickle blood velocity.

2.2.5. Differential Equations

Discrete models work well for distinct time events. In contrast, differential equations are mathematical expressions that equate the change in an independent (or state) variable \( x \) over continuous time \( t \) to
the dynamics of the rates and relationships between the system variables. Each state variable represents a vital component of the system such as a cell, a chemical concentration, etc.

$$\frac{dx}{dt} = f(x, t)$$

For example, the equation of this form could be something like $\frac{dx}{dt} = ax(t) + b$, where $\frac{dx}{dt}$ is measuring the change in $x$ over the change in time, $t$, where the rate of change of $x$ is proportional to its current size along with a constant source $b$. This is just a simplistic example, though, as these equations can become extensive and complex. Differential equations allow researchers to relate/connect biologically significant components in the system with appropriate relative rates of interactions. Ordinary differential equations (ODEs) and partial differential equations (PDEs, discussed later) allow the predictive outcome to be queried for any time point. Certain model analysis is also facilitated in these continuous models. In complex systems, interdependent differential equations form a dynamical system. In a SCD model, in the example equation above, $dx/dt$ could represent the change in the number of sickled RBCs $x$ over time $t$, where the number of sickle RBCs at any given time $x(t)$ is impacted by a natural cell death rate $a$ and there is a constant source of new made sickle RBCs $b$ from the bone marrow.

While this equation is a simplistic example, ODEs can be used to represent multiple variable interactions over time that involve varying scales.

In Zheng 2021 [107], multi-scale ODEs, (i.e., models with multiple biological scales represented such as micromolecules, organelles, cells, and organism populations), were used to model RBC production and hemoglobin assembly. Zheng's model used ODEs to simulate transplantation of autologous stem cells that produce an anti-sickling hemoglobin to understand how varying treatment parameters impacts the efficacy of different gene therapies. Their model findings reported the minimal dose of LT-HSC (stem cells) that were needed for a stable, long-term source of functional RBCs.

These types of ODE simulations based on real world data are especially useful because of their ability to be analyzed using several computational techniques. For example, sensitivity analysis can allow researchers to identify the most impactful model parameters and assess the associated changes in model outcomes (like above). Bifurcation analysis can elucidate the stability and intrinsic behavior of the system variables in the model. Optimal control methods can be used to hypothesize treatment options.
An early model of the RBC fluid system was created by Dong et al. in 1992 [25] using an ODE model and lubrication theory to describe the pulsatile flow in the gap between a cell and the vessel wall. Modelers can also use ODEs to investigate the intracellular and intercellular dynamics of sickle RBCs. For example, models like Lei et al. 2015 [54] and Lu et al. 2017 [58] use multi-scale ODEs to quantitatively simulate HbS polymerization inside RBCs by investigating cell morphology and fiber orientation, respectively.

Compartmental ODE methods, like the well-known SIR model, were originally developed to model disease transmission in epidemiology as members of a population move from one population compartment to another: susceptible (S) to infected (I) to resistant (R). Using this compartmental based framework, researchers have adapted it to investigate cell interactions in disease systems. In SCD research, models such as Altrock at al. 2016 [2] and Zheng et al. 2021 [107] computationally determine the effect of potential treatment protocols on RBCs at different stages. Altrock et al. [2] created a mathematical model used to investigate RBC populations impacted by a gene therapy tool and then used the chimerism level found to inform future mouse model experimentation. The model was able to identify the number of HSCs required for successful RBC count/level by quantifying the relationship between HSCs and RBCs populations. Similarly, Zheng’s model used multiscale ODEs to predict dynamics of stem cell engraftment on RBCs during gene therapy by modeling RBC production and hemoglobin assembly during treatment [107]. Their model findings reported the minimal dose of stem cells needed for a stable, long-term source of RBCs.

Compartmental models have also been applied to studying genetic inheritance and frequency of individuals with SCD. Liddell et al.’s 2014 box population model [57], for example, tracked the dynamics of genotypes within populations to investigate the impact of malaria selection pressure on sickle cell carrier population size. ODE models have also been used for treatment models, since parameters representing biological processes can be manipulated to reflect the impact of treatment (e.g., death rate/growth rate x population size) within the systems. One ODE treatment model was used in Lu et al. 2019 [59] to monitor the impact of a potential anti-sickling drug on RBCs over time.
2.2.6. Partial Differential Equations Modeling

When spatial dynamics need to be considered, partial differential equations (PDEs) are used in place of ODEs. PDEs are multidimensional and can track different populations throughout a spatial domain over time, making them especially useful for modeling fluid dynamics. For example, Deonikar et al. [24] utilized PDEs to make a flow dependent 2D model of nitric oxide (NO) production and transport in an arteriole and found that even a low presence of cell free Hb can reduce the amount of smooth muscle cell NO produced - potentially playing a role in the development of pulmonary hypertension.

Computational fluid dynamic (CFD) models derive from the Navier-Stokes equations, a set of partial differential equations that describe how "the velocity, pressure, temperature, and density of a moving fluid are related" [67]. These equations allow researchers to simulate SCD blood conditions given the proper data.

To identify triggers and indicators of VOCs, a deeper understanding of the disease hemodynamics (blood flow) is necessary and CFD models are commonly used to model the flow of fluid through space; thus, making PDEs especially useful for modeling RBC movement in the arteries. In 1980, Berger [7] used the Krogh cylinder fluid dynamics model to represent SCD blood flow through the capillaries. The results suggested that feedback interactions between oxygen concentration and flow velocity through the capillary were important in SCD blood. Later, Rivera 2016 [83] used computational fluid dynamics to alter fluid and artery wall properties to simulate scenarios causative of significantly elevated arterial blood velocities.

In 2020, Bazzi et al. [6] made a CFD model to simulate microfluidic movement and investigate VOCs, highlighting the advantages of microfluidic modeling. Szafraniec 2022 [30] used an experimental microfluidic system to separate the cell flow profile into a bulk component and a wall component and then computationally modeled the system to evaluate differential contributions of effective viscosity and wall friction to the overall blood resistance.

To investigate potential reasons for the inflammation of the endothelial cells lining blood vessel walls, Zhang 2020 implemented a suspension flow PDE model and simulated idealized blood flow in SCD. Similarly, a PDE model was developed by Chaturvedi 2021 [14] to model how an RBC (in his case a pellet)
moves through a capillary (or narrow fluid filled cylindrical tube) to investigate the impact of viscosity on blood flow. Sawyer et al. 2022 [85] used CFD modeling to reconstruct the intracranial portion of the internal carotid artery and branches, and extract the geometry to analyze vascular architecture influence on cerebrovascular risk. Most recently, a CFD model of the flow of a single RBC with alterable rheological properties was implemented to test the feasibility of the lab-on-a-chip microfluidic diagnostic tool for SCD [14].

The use of this foundational system CFD along with new computational simulation power has led to a rise in dissipative particle dynamic (DPD) models. DPD is a simulation technique used for molecular or mesoscopic modeling and investigation, also commonly used in chemistry and pharmacology. Lei 2012 and 2013 [53] [54] constructed a 3D model of three typical shapes of sickle cells using experimental SEM observations to observe and quantify adhesion behavior of SS-RBCs inside capillaries. The model was able to validate that vaso-occlusion is a "complex process triggered by the interactions between multiple density groups, where each group contributes differently to the occlusion crisis" [53].

PDEs have also been utilized in studying genetic inheritance (in place of ODEs) when the number of individuals is large enough. Tchuenche at al 2007 [91], for example, created a theoretical PDE population model to track genetic inheritance of SCD.

### 2.2.7. Hybrid Models

Hybrid models are models that contain two or more model types. For example, Clifton et al. [17] utilized a hybrid statistical and mechanistic ODE model to predict pain levels in SCD patients using smartwatch data along with drug mechanics. The statistical component of the model estimated the best estimates of parameter rates from the data (elucidating variable relationships). These parameters are then utilized in an ODE model to capture the relationship between the data and those parameters and how their relationship impacts the system. Clifton’s [17] hybrid model results along with the effectiveness of ML on outcome prediction elucidate the need for more models that can accurately assess the personalized data available today. The increased accessibility of patient monitoring data combined with the low-cost implementation of complex quantitative models will allow for research advancements that save error and injury. Table 1 lists different types of models, what systems they are often used to analyze,
and examples of investigation of that system.

2.3. Choosing a Math Model Type

Mathematical techniques can be used to quantify biological system processes, simulate these systems, elucidate variable relationships, and predict future outcomes. The type of model applied is based on multiple factors that should be considered:

![Diagram: How to Choose a Math Modeling Technique]

(1) the research objective

- classification models: RBC shape, disease progression, pain class
- prediction models: gene inheritance, pain score prediction
- models investigating variable relationships: Hemoglobin S molecule, RBC flow and behavior, VOC triggers, treatment efficacy

(2) the data collected/available:

- aggregate or individual
- system or local
• sample size
• frequency of data collection
• blood image and/or video data
• self-reported/survey data
• physiological data (vital signs + lab data)

(3) what is known and what is unknown about the system

• the relationships between the variables are known (e.g., "oxygen concentration directly impacts gelation")
• there are unknown factors or missing variables in the analysis (e.g., "some aggregate of these vital sign measurements influence pain expression")

The same way many understand mathematics as a language in its own right, mathematical modeling can be seen as an art form. It offers the unique ability to investigate and experiment with biological processes that may otherwise elude us due to their complexity. Because of this complexity, there is not necessarily a right way to model biological phenomena. The type of model chosen is entirely dependent upon the research question, the available data, and the dynamics of interest. Consulting with mathematicians/computational researchers, sharing the data to be analyzed, and collaborating on research questions may lead researchers and clinicians to revelations that save time, errors, and money.

In the following sections, the authors give examples of the way multiple approaches can be used for each problem/research area in SCD. Section 3 illustrates models created to investigate VOCs and hemodynamics in SCD. Section 4 describes models made to represent SCD cellular dynamics and RBC behavior. Section 5 illustrates gene-based models and Section 6 explores pain prediction models. Lastly, Section 7 reviews models created to analyze treatment options in SCD. Each section is associated with an illustrative table. These tables, Tables 2 – 6, each summarize models discussed in that section, that were published for the purpose being discussed in that section (investigate VOCs and hemodynamics, versus predict pain, analyze treatment options, etc.).
2.4. Models by SCD Research Area Summary

2.4.1. Modeling Hemodynamics and Vaso-occlusive Crises in SCD

For models of VOCs, findings indicate potential markers/mechanisms that contribute to the triggering of VOCs, including analysis of blood flow property contributions.

Of the hemodynamic models for VOC investigation, Lei 2013’s discrete dynamics model found that interactions between multiple density groups, where each group contributes differently to the occlusion crisis, is a trigger for VOCs. Rivera 2016’s 3D fluid dynamics PDE model reconstructions revealed "an uneven, internal arterial wall surface in children with homozygous SCA and higher mean velocities in the Middle Cerebral Artery up to 145 cm/s compared to non-SCA reconstructions." Identifying cellular causes of these microstructures of RBCs or how luminal narrowing due to endothelial hyperplasia is induced by disturbed flow would provide new targets to treat children with SCD.

During their investigation for potential triggers of VOC, Chalacheva et al. 2015’s ODE model system found that vascular disturbance when the baroflex is blunting could be a potential factor that contributes to VOCs. In a similar attempt to monitor and identify physiological symptoms that indicate pain induced VOCs, Chalacheva 2017 used multiple regression analysis and found that blood pressure markers showed strong differences between SCD and non SCD individuals. Chalacheva 2019 also found measurable factors that indicate the parasympathetic activity that appears when VOCs occur using their statistical regression model.

A common thread throughout these models is the emphasis on functional mechanisms. Since SCD is a blood disorder, most of the measured factors are related to cardiac activity and blood properties. In 2015, Chalacheva et al. mathematically modeled heart rate variability and peripheral resistance variability to investigate vascular functionality. But Chalacheva’s later models, 2017 and 2019, induced VOCs through stimuli and compared the functional differences in heart rate and blood pressure mechanisms in SCD and non SCD individuals. While the findings from these models provide some insight into what measurements are important to track VOCs, many note that investigating the underlying physiological mechanisms that are responsible for these measurement differences is imperative.

Table 2 lists the different math models of VOC in SCD, the key variables and data used in model...
2.4.2. Modeling Cellular Dynamics + Rheology of SCD

RBCs are at the center of all processes in SCD. HbS polymerization occurs inside RBCs and deforms RBCs from usual doughnut or biconcave disc-shaped cells into sickle-shaped and odd-shaped cells. Therefore, understanding the intercellular dynamics of RBCs and other blood elements is an imperative effort in SCD research. Since SCD is a multi-scale, multi-organ disease, RBC research in SCD has been approached through multiple mathematical methods.

The conclusions from RBC models generally fall under four broad categories: (1) insight into the complex processes in SCD; (2) recognition of the role of cell-cell and cell-endothelial wall interactions in SCD complications; (3) further understanding of sickle RBC behavior, and; (4) classification of RBCs by shape / disease severity.

HbS polymerization is a particular area of interest in the context of SCD as one of the main tasks inside RBCs that contribute to sickling/gelation. Using multi-scale modeling, Lu et al. 2017 \[58\] was able to predict the efficacy of sickling inhibitors and identify additional effects of existing and future drugs. Differential equations and the MAR scheme allowed them to simulate the HbS polymerization process, measure the growth rate and bending stiffness of a single HbS fiber, and extract the interaction forces between HbS fibers that occur during polymerization \[58\]. Multi-scale modeling also allowed for illustration of how VOCs begin and are sustained. Lei at al. 2015's \[54\] stochastic VOC model was used to represent the development of the intracellular aligned sickle hemoglobin polymer domain. The insight gained from this model revealed that perhaps processes other than sickling such as vessel endothelium activation and cell-endothelium adhesion may be legitimate mechanistic targets of treatment and prevention of VOCs and treatment of SCD in general. Indeed, one SCD drug currently on the market utilizes one of these alternative targets. [Kaur et al. 2023 PMID: 37850353.]

Other RBC models illustrated the merit of understanding RBC interactions in the formation of treatment protocols that impact sickle cell development. Altrock et al's \[2\] compartmental model revealed the number of stem cells necessary for a successful non-sickle RBC count/level that was found by quantifying the relationship between stem cell and RBC populations. Similarly, Zheng et al. \[107\] found the
minimal dose of transduced stem cells necessarily engrafted for a stable, long-term source of functional
RBCs after transfusion. Additionally, a model finding that even a low hemoglobin concentration in the
blood can negatively impact smooth muscle cells by Deonikar et al. [24] indicates a pathway for treating
the development of pulmonary hypertension in individuals with SCD. All these models utilized differen-
tial equations (ODEs/PDEs) for simulation. PDEs were also used to create computational fluid dynamics
models. Sawyer et al. 2022 [85]’s computational fluid dynamics approach allowed for the evaluation of
be used to identify patients who are at higher risk for cerebrovascular complications. The results suggest
that sickle blood flow "is altered systemically from the arterial to the venous circulation". Zhang 2020
[106]’s findings from their idealized blood flow model may aid in understanding the pathophysiology of
chronic endothelial inflammation in SCD from a biophysical perspective.

Other RBC models investigated the biodynamics of sickle RBC shape formation. Li 2017’s multi-scale
network model [56] found that biodynamic behavior of sickle RBCs under hypoxia is heavily impacted
by their irregular geometry, decreased cell deformability, and elevated cell volume. Dong 1982 [25] found
a transition from membrane to internal polymer dominance of deformability as oxygen saturation was
lowered. Chaturvedi 2021’s [14] model showed that the higher viscosity of plasma exerted comparatively
higher drag force through their PDE model. Modeling the microfluidic channel, Sawyer 2022 [85] showed
that blood from patients with SCD exhibited elevated frictional and viscous resistances at all physiologic
tensions. From another perspective, RBC shape was studied by classifying sickle RBCs from blood data.
ML models by [99] and [77] allowed classification of RBC types according to ML patterns and defined
how to select the most important features for classification of blood smears.

The remaining models classified RBCs by shape and disease severity. Delgado Font 2020 [22] analyzed
blood sample images to classify RBCs as normal or elongated or having other deformations. Praljak 2021
[79] established a morphology based on classifying the RBC population into subgroups using novel visual
markers that linked to underlying cell biomechanical properties. Xu et al. 2017 [102] reported a ML model
that created an automated algorithm for sickle RBC classification using a convolutional neural network
(CNN). Similarly, Darrin 2023 [21] created a ML algorithm to automatically classify cell motion in videos
with high motion imbalance.
Table 3 lists the different math models of red blood cell behavior in SCD, the key variables and data used in model formulation, and a quick summary of the model.

2.4.3. Modeling Genetic Inheritance and Mutation in SCD

The ODE models utilized to model gene populations investigated selection factors that impacted SCD inheritance. Liddell 2014 [57] compartmental ODE model tracked the dynamics of genotypes within key SCD populations to understand SCD's relationship with malaria. By using a compartmental model, similar to the SIR epidemiology framework, [57] tracked the population size of children vs adult non-carriers, carriers, and afflicted populations. After simulations, stability and sensitivity analyses were done and supported the conclusion that selection pressure from malaria does result in a higher fraction of carriers of the HbS gene. This is likely because malaria protection selectively favors the sickle cell gene. Another supporting finding is that the selective advantage of the HbS gene over the HbA gene does allow SCD inheritance to persist in tropical populations as shown by Balamanikandan's theoretical model. Balamanikandan's 2022 [3]'s quantum network model also aimed to aid in early diagnosis and identification of SCD. The authors were able to input DNA sequence data into a quantum graph theory model and utilize ML mining approaches to distinguish normal RBCs from sickled RBCs.

Table 4 lists the different math models of the genetics in SCD, the key variables and data used in model formulation, and a quick summary of the model.

2.4.4. Modeling Pain Behavior and Management in SCD

Pain is the primary symptom of SCD. As a symptom, information about pain may only come through observation - i.e., using patient reports. Management of SCD related pain is particularly challenging due to its subjective nature. In clinical practice, medical providers often search for objective indicators, such as vital signs, non-verbal cues, and biomarkers, to guide their assessment and treatment of pain. This exercise often results in misdiagnosis and failure. Many envision instead employing a precision medicine, individualized, integrative treatment approach, based on a whole-person model of pain. [Smith et al. 2023 PMID: 38028431]. As such, the development of an objective automatic pain estimation method would lead to marginal improvements in pain assessment and management but multiple models
might need to be applied to refine prediction. Since the goal of pain related modeling is often to predict pain intensity, pain frequency, pain behavior, changes in pain over time, model conclusions may often fail to predict pain accurately over time. Instead, researchers, now with access to an ever-increasing amount of medical data, should think of more complex hybrid modeling and ML methods in order to design pain therapy. One example of this kind of thinking and ML modeling was Sebastiani's Bayesian network model [86]. It integrated individual disease complications and lab test results and was able to compute personalized severity scores. Severity scores were determined by a network of interactions of 14 individuals. Yang 2018, Yang 2019, and Padhee 2021 experimented with ML methods/algorithms to predict patient pain scores [104] [103] [72]. Yang et al. 2018 were able to predict pain scores based using vital signs, without needing medication information [103]. Padhee also used vital sign data along with self-reported pain data to predict pain scores [72]. They found that the Decision Tree algorithm was the most promising approach for prediction. Consistent with the “big data” hypothesis of increasing accuracy with more complex models, all methods Padhee tested had higher accuracy when they had more training data to feed the model [72]. Yang 2019 conducted a similar study, but also used feature selection algorithms to identify which sensor data features were significant [104]. Sensor data was also used in Clifton's hybrid statistical and mechanistic ODE model, but [17]'s data focused on pain scores and pain medication information.

Clifton et al. noted that their hybrid model would drastically improve in accuracy with the addition of blood pressure, heart rate, and activity level information via Fitbits. This assumption was supported by the conclusions from [62] [63] [46] [56] [102] [104] [72] [86]. All the models mentioned here focused on physiological factors that often centered on heart-related measurements. The models from Padhee 2021 [72] and Yang 2018 [104] used vital sign data (SpO2, systolic BP, diastolic BP, pulse, resp, temp) and pain scores for prediction. Similarly, Yang's feature selection algorithm revealed that galvanic skin temperature, heart rate variability, acceleration, and steps taken were physiological and body movement related features that had high significance in pain prediction [104]. These findings raised the perceived value of employing more ML and mechanistic ODEs for pain prediction in SCD.

Using ML, these authors [30] [48] [47] [46] [74] [71] used electronic health data to predict subjective pain scores from patient data. Electronic health data outperformed hospital prediction systems. Most
recently, Padhee 2022 found that models performed better when provided with medication data along with vital signs data [71].

Table 5 lists the different math models of pain behavior and dynamics in SCD, the key variables and data used in model formulation, and a quick summary of the model.

2.4.5. Modeling Treatment Efficacy in SCD

Mehraei’s 2016 network model was created as an illustrative guide for drug discovery and prediction for β-globin disorders [63]. The authors hypothesized possible treatment of SCD and other β-globin disorders by inducing protective fetal hemoglobin via modulating specific protein targets that induce β-globin. The authors validated the superiority of this treatment approach versus others, using simulation models based on the literature, and comparing the simulations for postulated efficacy at fetal hemoglobin induction. Another model studying drug treatment impact, Lu 2019’s kinetic ODE model, was able to reproduce different experimental results without tuning model parameters, indicating its capability to confirm and analyze laboratory and clinical findings [59]. Since a lot of treatments like hydroxyurea target multiple pathways like sickling, adhesion, and inflammation, Lu 2019’s model may be useful since it can predict the efficacy of sickling inhibitors.

Khalaf 2017’s ML classification algorithms found that the Random Forest Classifier worked best on their dataset to classify the dosage of medication required for the treatment of patients with SCD [50]. Patel 2019 [76] applied ML algorithms to predict 30-day unplanned hospital readmission rates and outperformed the standard hospital readmission risk scoring systems (LACE and HOSPITAL) by a large margin. Zheng 2021’s [107] (QSP) model and dynamical approach to represent RBC production and treatment simulations found the minimal dose of LT-HSC needed for a stable long-term resource of functional RBCs. This will allow researchers to understand how varying specific treatment parameters affect short- and long-term measures of treatment efficacy.

Table 6 lists the different math models of SCD treatment efficacy, the key variables and data used in model formulation, and a quick summary of the model.
2.5. Conclusion: Future of Modeling

This review has demonstrated the utility of using mathematical modeling to gain a deeper understanding of SCD. By quantifying known dynamics and phenomena, fitting models to experimental data, and conducting parallel assays and simulations, researchers have been able to elucidate significant biomarkers, processes, and clinical findings that at least partially contribute to SCD progression.

However, the field of SCD research is relatively new, with official research efforts only starting in 1949 [48] despite its discovery in 1910. There is much to still be understood about the disease’s dynamic behavior. Mathematical and computational efforts have progressed the field greatly, taking advantage of the new abundance of data on individuals with SCD. Like most research, however, these findings mark only the beginning of what can be understood.

The models investigating VOCs in SCD are still aiming to uncover biomarkers and other physiological mechanisms that primarily impact the vascular and other system dysfunctions that together cause VOCs. There are also still many questions about RBCs, as everything from their size and shape to their rigidity and location have major implications on SCD presentation. Li et al. 2017 demonstrated their interest in understanding the complex relationship between RBCs, morphological membrane distortion of RBCs, and deoxy HbS polymer chain formation [56]. By classifying RBCs according to their shape, Xu et al. 2017 have tried to build a gold standard RBC phenotype library within SCD [102]. Petrovic et al. 2020 have tried to optimize RBC classification by adding more features based on PCA and LDA for SCD diagnosis support [77]. Zheng et al. 2021 and Altrock et al. 2016 have suggested extending their models by incorporating diffusion of chemicals and/or drug pharmacokinetics to help understand therapeutic effects on different RBC shape phenotypes and blood flow [107] [2].

The common theme found in studies building cellular, genetic, behavioral, pain, and treatment SCD models is the need to increase model complexity. Parameters and variables have started with simplified models, in order to understand or explain the operation of small galaxies of a universe of SCD dynamics, and only at a basic level. Now, simple models must be either combined, expanded or layered properly, in a branching tree-like effect, to account for the impact of confounding variables and mathematical degrees of freedom at many levels—the molecule, cell, the momentary VOC, the organ, and the biopsy-
chosocial milieu.

For example, Lu 2019’s model parameters must now be varied to incorporate the effect of anti-sickling agents. Similarly, researchers who have successfully predicted short-term SCD pain intensity, must now improve system prediction performance by predicting longer term SCD pain intensity, along with VOC visit behavior [59]. These are the kinds of predictions needed to affect real-life events, modify the disease course, and prevent mortality. In order to accomplish that, researchers must now collect and incorporate into mathematical models more variables, and more layers of variables, simultaneously. Added to models of molecular and cellular behavior, must be simultaneous models of physiological measurements, pain reports, activity level, age, gender, medication usage, and psychosocial and environmental stressors—whole-body models. One investigator stated upon completing a model of red cell behavior: “This is a step towards continuous and non-invasive pain management for SCD patients during hospitalization and after they are discharged from the hospital. Doing so will allow us to create a remote pain management system that can hopefully reduce re-hospitalization and improve the quality of life for patients with SCD” [104]. One hope is that mathematical modeling, among other benefits, will benefit the field by being the ideal platform for testing optimal dosage of medications and hypothesizing treatment effects in order to obviate conducting expensive and time-consuming randomized controlled trials.

Systems biology and mathematical modeling are having a resurgence with the rise in computational power along with an abundance of data. Computational modeling is going to play an important role in science, but models, to be used, must demonstrate reality. In most cases, life and the world are not simple. SCD is not simple.

### 2.6. Acknowledgements

This work was supported by the National Institutes of Health Helping to End Addiction Long-term (HEAL) Initiative, the National Institute of Dental and Craniofacial Research and the National Institute of Neurological Disorders and Stroke (R21DE032583).

### 2.6.1. TABLES
<table>
<thead>
<tr>
<th>Math Model Type</th>
<th>Example</th>
<th>Applications</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistics</td>
<td>$Y_i = f(X_i, \beta) + e_i$</td>
<td>• VOC mechanism investigation</td>
<td>12</td>
</tr>
<tr>
<td><strong>regression</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Machine Learning</td>
<td></td>
<td>• RBC shape/ severity classification</td>
<td>102, 77, 22, 79</td>
</tr>
<tr>
<td><strong>algorithms</strong></td>
<td></td>
<td>• subjective pain score prediction</td>
<td>30, 21, 3, 104</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• medication/treatment dosage/frequency analysis</td>
<td>103, 72, 46, 74</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>71, 50, 76</td>
</tr>
</tbody>
</table>
| ODEs                      | Differential equations | \( \frac{dX}{dt} = x(t) + q \) | - red blood cell modeling (cell shape, cell behavior)
- insight into polymerization process
- pain score prediction
- pharmacology/medication dosage + interactions |
|--------------------------|------------------------|---------------------------------|--------------------------------------------------------------------------------|
| PDEs                     | Computational fluid dynamics | \( \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x_j} [\rho u_j] = 0 \) | - VOC mechanism investigation
- modeling sickle RBC blood flow (re: microfluidic channels)
- cellular interaction impact on complications |
### Difference equations

*discrete time modeling*

\[ x_t = x_{t-1} + b \]

- VOC mechanism investigation
- red blood cell behavior

### Networks

*graph theory*

- biodynamic behavior of sickle RBCs
- cell shape classification/diagnosis
- pain severity prediction
- genetic drug treatment modeling

<p>| Table 2: Summary of VOC models | 11 | 13 | 6 | 30 | 56 | 21 | 3 | 86 | 63 |</p>
<table>
<thead>
<tr>
<th>References</th>
<th>Model Type</th>
<th>Dataset</th>
<th>Key Variables</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berger 1980</td>
<td>PDE model</td>
<td>n/a</td>
<td>oxygen concentration, oxygen partial pressure,</td>
<td>a model of sickle cell blood flow in the capillaries using the Krogh cylinder model (fluid dynamics)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>tissue region</td>
<td></td>
</tr>
<tr>
<td>Lei 2012</td>
<td>DPD model</td>
<td>shapes observed in</td>
<td>RBC free energy, cell adhesion reaction rates</td>
<td>validated multiscale model to quantify the morphology and dynamic properties of SSRBC into 3D model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>experiments by</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>scanning electron</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>microscopy (SEM)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lei 2013</td>
<td>DPD model</td>
<td>ref Lei 2012</td>
<td>RBC free energy, adhesive interaction rates</td>
<td>Observed and quantified adhesion behavior of SS-RBCs in terms of balance of free energies inside capillaries and validated that &quot;vasoocclusion is a complex process triggered by the interactions between multiple density groups&quot;</td>
</tr>
<tr>
<td>Chalacheva 2015</td>
<td>linear ODE</td>
<td>28 patients</td>
<td>heart-rate variability, peripheral resistance</td>
<td>investigates functional mechanisms relating to blood pressure, respiration, and peripheral vascular resistance</td>
</tr>
<tr>
<td></td>
<td>system</td>
<td>age: 10 - 30 years old</td>
<td>variability</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>SS genotype (save 2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rivera 2016</td>
<td>CFD</td>
<td>3 patients</td>
<td>SCA genotype</td>
<td>arterial segments: distal ICA, MCA, ACA</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chalacheva 2017</td>
<td>multiple regression analysis</td>
<td>45 subjects</td>
<td>age: 13 - 55 yrs old HbSS and HbS β0</td>
<td>finger blood volume, temperature, blood pressure, respiration &amp; blood pressure coupling, respiratory coupling, neurogenic thermal pain coupling, neurogenic-vascular interaction</td>
</tr>
<tr>
<td>References</td>
<td>Model Type</td>
<td>Dataset</td>
<td>Key Variables</td>
<td>Summary</td>
</tr>
<tr>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------------</td>
<td>---------</td>
</tr>
<tr>
<td>Chalacheva 2019 [13]</td>
<td>DDS, Regression</td>
<td>66 subjects</td>
<td>heart rate variability, peripheral vascular variability, diabolic &amp; systolic blood pressure</td>
<td>Head Up Tilt (HUT) stimulus was applied and responses analyzed to assess subjects’ cardiac and/or peripheral autonomic function responses</td>
</tr>
<tr>
<td>Chaturvedi and Shah 2023 [15]</td>
<td>DPD model</td>
<td>n/a</td>
<td>pressure, shear viscosity, fluid film thickness</td>
<td>To test the feasibility of the lab-on-a-chip microfluidic diagnostic tool for SCD diagnosis, a CFD model of the flow of a single RBC with alterable rheological properties is created</td>
</tr>
</tbody>
</table>

Table 3: Summary of RBC models

<table>
<thead>
<tr>
<th>References</th>
<th>Model Type</th>
<th>Dataset</th>
<th>Key Variables</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dong 1982</td>
<td>geometric ODEs</td>
<td>Chien et al 1992</td>
<td>cell radius, cell length, plasma velocity, vessel radius, distance between cells, distance between cell and wall</td>
<td>PDE model of flow in gap between cell and the vessel wall</td>
</tr>
<tr>
<td>Deonikar 2012 [24]</td>
<td>PDE</td>
<td>dataset</td>
<td>steady state SCD conditions, NO concentration, NO consumption rates, NO production rate</td>
<td>Model revealed the role of Hb concentration in disturbing smooth muscle cells, thereby also potentially playing a role in pulmonary hypertension</td>
</tr>
<tr>
<td>Lei 2015 [54]</td>
<td>multi-scale DPD model</td>
<td>n/a</td>
<td>red blood cells, leukocytes, cell-endothelium interaction</td>
<td>Found specific processes that occur in polymerization and VOC events that may indicate areas of interest for intervention/ treatment</td>
</tr>
<tr>
<td>References</td>
<td>Model Type</td>
<td>Dataset</td>
<td>Key Variables</td>
<td>Summary</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------------</td>
<td>-------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Altrock 2016 [2]</td>
<td>discrete time compartmental model</td>
<td>Berkeley sickle mice used</td>
<td>hematopoietic stem cells, RBC populations, survival factor, age</td>
<td>A parameterized math model was proposed to predict the number of HSCs required for successful RBC count</td>
</tr>
<tr>
<td>Li 2017 [56]</td>
<td>multi-scale DDS</td>
<td>4 patient blood samples</td>
<td>elastic bond energy, bending energy, corresponding energy</td>
<td>The biodynamic behavior of sickle RBCs under hypoxia is heavily impacted by sickle RBCs' irregular geometry, decreased cell deformability, and elevated cell volume</td>
</tr>
<tr>
<td>Lu 2017 [58]</td>
<td>ODE, MARS</td>
<td>n/a</td>
<td>spring constant, instantaneous and equilibrium distance, equilibrium angle</td>
<td>Model extracts that fiber orientation and interaction range between HbS fibers are key interactions that occur during polymerization</td>
</tr>
<tr>
<td>Xu 2017 [102]</td>
<td>machine learning</td>
<td>SCD patient blood 7 patients</td>
<td>patient specific microscopy image data</td>
<td>The deep CNN helped to classify sickle RBCs into different types according to training, which then allowed them to identify RBC shape factors and parameters</td>
</tr>
<tr>
<td>Bazzi 2020 [6]</td>
<td>finite difference equations</td>
<td>patient red blood samples</td>
<td>viscosity, shear rate</td>
<td>Quantified sickle blood velocity through an experimental + computational approach. Results suggest sickle blood flow is altered systemically</td>
</tr>
<tr>
<td>Delgado-Font 2020 [22]</td>
<td>machine learning algorithms</td>
<td>peripheral blood smear sample images</td>
<td>n/a</td>
<td>Analyzed patient blood sample images and classified erythrocytes as normal, elongated, or deformed</td>
</tr>
<tr>
<td>Petrovic 2020 [77]</td>
<td>machine learning</td>
<td>825 blood smear images 2695 tagged cells</td>
<td>n/a</td>
<td>Defined how to select the most important features for classification in order to reduce their total number to decrease the complexity and training time</td>
</tr>
<tr>
<td>References</td>
<td>Model Type</td>
<td>Dataset</td>
<td>Key Variables</td>
<td>Summary</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Zhang 2020</td>
<td>CFD suspension flow model</td>
<td>n/a</td>
<td>slow velocity, discoid radius, total energy, energy density</td>
<td>Detailed boundary integral simulations are performed to investigate potential reasons for the inflammation of the endothelial cells lining the blood vessel walls</td>
</tr>
<tr>
<td>Chaturvedi 2021</td>
<td>PDE model</td>
<td>n/a</td>
<td>pellet radius, pellet curvature</td>
<td>PDE model developed to investigate the impact of viscosity on how a red blood cell moves through a narrow capillary</td>
</tr>
<tr>
<td>Praljak 2021</td>
<td>machine learning</td>
<td>microfluidic blood smear images</td>
<td>n/a</td>
<td>In an effort to improve blood image analysis, they &quot;establish a morphology based classification scheme to identify two naturally arising sRBC subpopulations, utilizing novel visual markers that link to underlying cell biomechanical properties&quot;</td>
</tr>
<tr>
<td>Sawyer et al 2022</td>
<td>CFD</td>
<td>10 participants age 5-15 yrs old HbSS and HbS β0</td>
<td>3D model of internal carotid artery: ICA, MCA, ACA</td>
<td>Reconstructed the intracranial portion of the internal carotid artery and branches and extracted the geometry using CFD modeling to analyze the vascular architecture influence on cerebrovascular risk in SCD</td>
</tr>
<tr>
<td>Szafraniec 2022</td>
<td>machine learning and difference equations</td>
<td>n/a?</td>
<td>RBC flow rate, velocity, resistance</td>
<td>Math modeling of the microfluidic blood system to evaluate differential contributions of effective viscosity and wall friction to the overall resistance in blood</td>
</tr>
<tr>
<td>References</td>
<td>Model Type</td>
<td>Dataset</td>
<td>Key Variables</td>
<td>Summary</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>----------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Darrin et al 2023</td>
<td>machine learning</td>
<td>CNN</td>
<td>4 patients age 18+ yrs old HbSS, HbS β0</td>
<td>Using “the fact that the individual movement of an RBC in shear flow is an indicator of its deformability”, a machine learning algorithm/ pipeline is proposed to automatically classify cell motion in videos with high class imbalance.</td>
</tr>
<tr>
<td>Tchuenche 2007 [91]</td>
<td>PDE</td>
<td>n/a</td>
<td>males, females, age, growth rate, birth rate, death rate, mating rate,</td>
<td>the persistence of SCA is probably due to the role played by the selective advantage of the abnormal S gene over the normal haemoglobin A in tropical regions</td>
</tr>
<tr>
<td>Liddell 2014 [57]</td>
<td>ODE</td>
<td>n/a</td>
<td>carrier adults, non-carrier adults, carrier children, non-carrier children, afflicted adults, afflicted children</td>
<td>the selection pressure from malaria results in a higher fraction of sickle cell carriers and that malaria protection selectively favors the sickle cell gene</td>
</tr>
<tr>
<td>Balamanikandan 2022 [3]</td>
<td>quantum graph</td>
<td>1387 patients</td>
<td>input DNA sequence</td>
<td>To aid in early diagnosis and identification of SCD, modelers input DNA sequence information into a Quantum graph theory model and use machine learning mining approaches to distinguish normal cells from sRBCs.</td>
</tr>
</tbody>
</table>

Table 4: Summary of Gene models
<table>
<thead>
<tr>
<th>References</th>
<th>Model Type</th>
<th>Dataset</th>
<th>Key Variables</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clifton 2017</td>
<td>hybrid ODE + statistics</td>
<td>47 patients age: 18+ SCD type: HbSC, HbSS, HbSB+ thalassemia, HbB0</td>
<td>patient pain level, amount of drugs (LA,SA,NO) within patient, pain level probability distribution, unmitigated pain level, drug relaxation rate, drug rate of decay, drug dosage times, number of drug doses taken</td>
<td>model shows predictive value since it allows for forecasting the probability distribution of pain for a patient at a point in the near future</td>
</tr>
<tr>
<td>Yang 2018</td>
<td>machine learning</td>
<td>5363 records from 40 inpatient participants</td>
<td>pain score, vital signs (SpO2, systolic BP, diastolic BP, heart rate pulse, respiratory rate, temp)</td>
<td>the model used multiple imputation to take advantage of missing data and estimate pain score predictions based solely on physiological measurements without including patient medication information</td>
</tr>
<tr>
<td>Yang 2019</td>
<td>machine learning</td>
<td>29 patients</td>
<td>heart rate, RR interval, galvanic skin response, skin temp, steps</td>
<td>model showed that subjective pain scores can be estimated using objective wearable sensor data with high precision</td>
</tr>
<tr>
<td>Padhee 2021</td>
<td>machine learning</td>
<td>67927 records from 50 participants over 5 years</td>
<td>pain score, vital signs (SpO2, systolic BP, diastolic BP, pulse, resp, temp)</td>
<td>kNN, SVM, MLR, DT, and RF algorithms were employed to predict pain scores, which all showed higher accuracy with increased data. The Decision Tree algorithm seems to be the most promising model</td>
</tr>
<tr>
<td>References</td>
<td>Model Type</td>
<td>Dataset</td>
<td>Key Variables</td>
<td>Summary</td>
</tr>
<tr>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------------</td>
<td>---------</td>
</tr>
<tr>
<td>Panaggio 2021</td>
<td>machine learning</td>
<td>46 patients over 3 years</td>
<td>vital sign data including respiratory rate, heart rate, and blood pressure</td>
<td>Gaussian Bayes (GBN) classifiers and hidden Markov (HMMs) machine learning models outperform baseline models in estimating subjective pain, distinguishing between typical and atypical pain levels, and detecting changes in pain</td>
</tr>
<tr>
<td>Padhee 2022</td>
<td>machine learning</td>
<td>5363 records from 40 participants</td>
<td>6 vital signs, patient self-reported pain, medication type, medication status, total medication dosage</td>
<td>Using electronic health data from Duke University Medical Center, they designed a classification model using raw data and deep representational learning to predict subjective pain scores. It was observed that at varying Likert scales, the models performed better when provided with medication data along with vital signs data</td>
</tr>
</tbody>
</table>

**Table 6: Summary of Treatment models**

<table>
<thead>
<tr>
<th>References</th>
<th>Model Type</th>
<th>Dataset</th>
<th>Key Variables</th>
<th>Summary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mehraei 2016</td>
<td>hybrid functional petri network (HFPN) model</td>
<td>qPCR data</td>
<td>therapy targets</td>
<td>Created a quantitative model of human fetal-to-adult hemoglobin switch network that results in maximum ( \beta )-globin gene induction to test and compare six protein and multiprotein target-based strategies for hemoglobin drug treatment</td>
</tr>
<tr>
<td>References</td>
<td>Model Type</td>
<td>Dataset</td>
<td>Key Variables</td>
<td>Summary</td>
</tr>
<tr>
<td>------------</td>
<td>------------</td>
<td>---------</td>
<td>---------------</td>
<td>---------</td>
</tr>
<tr>
<td>Khalaf 2017 [50]</td>
<td>machine learning</td>
<td>commissioned specifically for the purposes of this study and was collected within a five-year period from the Alder Hey Children's hospital</td>
<td>weight, Hb, MCV, platelets, neutrophils, reticulocyte count, alanine aminotransferase, body bio blood, HbE, Bilirubin, LDH, AST</td>
<td>Tested different machine learning architectures to test which best classifies the dosage of medication required for SCD patient treatment.</td>
</tr>
<tr>
<td>Lu 2019 [59]</td>
<td>hybrid stochastic and mechanistic ODE</td>
<td>microfluidic channel observations</td>
<td>temperature, de-oxygenation time, oxygen pressure, RBC volume, Hb concentration</td>
<td>Created a model of HbS polymerization and growth against the RBC membrane. The outputs are numbers of nuclei generated through homogeneous and heterogeneous pathways, respectively, the lengths of HbS fibers, the configuration of fiber domains, and the sickling of RBCs.</td>
</tr>
<tr>
<td>Patel 2019 [76]</td>
<td>machine learning</td>
<td>3299 admissions data comprising of 446 adult SCD patients</td>
<td>n/a</td>
<td>Machine learning algorithms were applied to predict for 30-day unplanned hospital readmissions in SCD. Machine learning algorithms outperformed the standard hospital readmission risk scoring systems, LACE and HOSPITAL, by a large margin in a real world data set of SCD patients at a single institution and identified most important variable predictors.</td>
</tr>
<tr>
<td>References</td>
<td>Model Type</td>
<td>Dataset</td>
<td>Key Variables</td>
<td>Summary</td>
</tr>
<tr>
<td>------------</td>
<td>------------------</td>
<td>---------</td>
<td>--------------------------------------</td>
<td>------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Zheng 2021</td>
<td>multi-scale ODE</td>
<td>n/a</td>
<td>endogenous LT-HSC, transduced LT-HSC</td>
<td>Developed a quantitative systems pharmacology (QSP) model to understand how varying specific treatment parameters affects short and long term measures of treatment efficacy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>LT-HSC into ST-HSC, fold-amplification, mean residence time</td>
<td></td>
</tr>
</tbody>
</table>
3. Leveraging the Sleep-Pain Relationship in an ODE Model to Predict Acute Pain in Pediatric Sickle Cell Disease

3.1. Introduction

Sickle cell disease (SCD) is a genetic blood disorder that affects over 20 million people worldwide, with approximately 100,000 people in the US living with the disease. According to the CDC, every 1 of 365 Black Americans and 1 of every 16,300 Hispanic Americans are diagnosed with the disease at birth. The most prevalent complication associated with SCD is pain, often occurring in acute episodes ranging from hours to days and weeks due to vaso-occlusion (i.e. obstructed blood vessels), that result in tissue death and inflammation. These vaso-occlusive crises (VOC) are the landmark occurrence of SCD as they lead to several complications throughout the body from ischemic stroke to pulmonary hypertension. The pain episodes and complications vary greatly throughout a patient's lifespan starting with unanticipated pain periods in childhood and adolescence that later progress into a chronic pain condition.

3.1.1. Background Application to Pain

Sickle cell disease is an inherited blood disorder where some red blood cells' shape is no longer round but crescent shaped - or sickle. Sickle cells inherent shape distortion causes RBCs to stick together against vessel walls and obstruct oxygen supply, known as the vaso-occlusive crises (VOC). The most reported complication caused by VOC in sickle cell disease is periodic episodes of pain. Pain is the primary factor linked to poor health outcomes, increased medical costs, increased health care use, and poorer quality of life. Since pain is a primary complication from SCD, information about pain primarily comes from patient reports.

Pain is a subjective experience, making it difficult to quantify and monitor. Mobile technology has been a major aid in this quest since it allows continuous reporting and monitoring of several patient factors. The accessibility of mobile technology has lead to an increase in patient data and may aid in the quest for more prophylactic pain management. In order to analyze this data, mathematical models have
been employed. Common mathematical methods for predicting SCD pain include statistical analysis and machine learning (ML) \cite{104,103,17}. Most of these models used to predict pain input physiological data collected from patients and once processed, employ regression and machine learning algorithms to outright predict pain scores/levels.

Yang et al 2018 \cite{104} and Yang et al 2019 \cite{103} experimented with machine learning methods and were able to predict pain scores based off patient vital signs, without medication information and identify which tracker/sensor data features were significant. Both papers noted that \textit{the patient cohort plays a major role} in the relationship between pain and physiological symptoms.

Padhee et al 2021 \cite{72} also used machine learning methods for SCD pain modeling. ML classification algorithms were used to investigate which method best predicts patients’ pain scores based on patient vital signs. While these measures are successful and revealing, the methodology limits the ability to predict time dependent dynamics with interventions. To address this concern, Clifton et al \cite{17} created a hybrid statistical and mechanistic ODE model for pain prediction that incorporates the impact of time dependent dynamics on the model outcome.

These studies use mobile technology and clinic data to identify connections between physiological factors and pain. Previous research established specific risk factors for SCD pain such as age, sex, SCD genotype, and medication use. Another related factor has been identified, as individuals with SCD also commonly report poor sleep, and poor sleep is often co-morbid with pain for SCD patients and other chronically ill populations.

In a previous study by Valrie et al 2019 \cite{95}, statistical models were employed to examine the temporal relationship between multiple sleep indicators and daily self-reported pain levels, and to investigate the interactive influences of sleep and daily pain on functional outcomes. A multilevel model was calculated to predict daily SCD pain severity using daily self-reported sleep quality; with actigraphy indicators of sleep duration, sleep efficiency, and sleep latency averaged across the diary period. Control factors included the following variables: sex, SCD genotype, and whether the youth was currently prescribed hydroxyurea. Findings indicated that self-reported poor sleep quality predicted high pain severity the following day while mean actigraphy measures were unrelated to daily SCD pain severity \cite{95}.

We aim to investigate this relationship between sleep and pain further with a dynamical systems
approach that incorporates the impact of time dependent dynamics onto the model outcome. Given the similarity of patient data collected, the closest modeling approach is the Clifton et al model \cite{17} that utilizes e-diary data from 39 adults with SCD and a dynamic component incorporated recurrently collected daily pain medication use. Their model was found to be accurate in predicting individual pain severity during a pain crisis.

In order to address the pediatric population surveyed in Valrie et al 2019 we adapted the pain dynamic setup from the Clifton model \cite{17} in our initial formulation but we do not assume a chronic background level of pain since it is physiologically incompatible with the variability of individual SCD pain patterns seen in pediatric patients. Our model also functions differently since pain dynamics are modulated based on sleep quality and pain medication.

The ultimate goal of this work is a model predictive enough for future health app integration that allows accessible prophylactic information for patients.

3.2. Materials/Data

3.2.1. Data source:

Valrie et. al \cite{95} conducted a prospective study of 88 pediatric SCD patients aged 8 to 17 years using twice daily electronic surveys (e-diaries) for 4 to 8 weeks with concurrent sleep actigraphy for 2 weeks. This resulted in 4473 total e-diary assessments completed across the sample. Dr. Valrie investigated the temporal relationship between SCD pain and sleep. Morning e-diary questions assessed sleep quality, night awakenings, and nighttime pain. Evening e-diary questions assessed daytime pain, positive and negative affect, stress, and daytime tiredness. Actigraphy assessed sleep duration, sleep efficiency, and daytime and nighttime activity levels.

Youth with SCD and their guardians were recruited from three regional pediatric SCD clinics in the southeastern US during the youth's scheduled appointments.

The final sample consisted of 88 youth aged 8 to 17 years (M = 11.66, SD = 2.99) out of the total 123 youth who participated in the larger study (72% of the total sample).

The final sample of youth did not significantly differ in relation to age, sex, SCD genotype severity, or
whether they were currently prescribed hydroxyurea when compared to the youth in the larger study.

<table>
<thead>
<tr>
<th>Sex</th>
<th>N(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>52 (59.0)</td>
</tr>
<tr>
<td>Male</td>
<td>36 (41.0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age</th>
<th>N(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 - 12 years old</td>
<td>52 (59.0)</td>
</tr>
<tr>
<td>13 - 17 years old</td>
<td>36 (41.0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SCD type</th>
<th>N(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sickle cell anemia</td>
<td>44 (50.0)</td>
</tr>
<tr>
<td>sickle thalassemia syndrome 0</td>
<td>3 (3.4)</td>
</tr>
<tr>
<td>sickle thalassemia syndrome plus</td>
<td>12 (13.6)</td>
</tr>
<tr>
<td>Hemoglobin SCD</td>
<td>27 (30.7)</td>
</tr>
<tr>
<td>Other</td>
<td>2 (2.3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pain Medication</th>
<th>N(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Long-acting opioids</td>
<td>9 (10.2)</td>
</tr>
<tr>
<td>Short-acting opioids</td>
<td>68 (77.2)</td>
</tr>
<tr>
<td>Non-opioids</td>
<td>66 (75.0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Hydroxyurea Users</th>
<th>N(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>40 (45.5)</td>
</tr>
<tr>
<td>No</td>
<td>47 (53.4)</td>
</tr>
</tbody>
</table>

Table 7: Patient Demographic Information

A multilevel model was then calculated to predict daily SCD pain severity using daily self-reported sleep quality; and actigraphy indicators of sleep duration, sleep efficiency, and sleep latency averaged across the diary period. Control factors included the following variables: sex, SCD genotype, and whether the youth was currently prescribed hydroxyurea. Findings indicated that self-reported poor sleep quality predicted high pain severity the following day. These findings led to our hypothesis that daily variations in mHealth data (i.e., e-diary reported sleep quality) precede changes in pediatric pain severity. Also, these analyses examined same day or previous night e-diary data, and averaged actigraphy sleep variables as they relate to pediatric SCD pain severity. This study will expand beyond the previous work by modeling the dynamic relationship of self-reported sleep quality and pain medication as indicators of risk for pediatric SCD pain over time.
3.3. Data

Figure 8: Patient A10 raw sleep and pain data

Figure 9: Patient C46 raw sleep and pain data
Data sets from 35 of the 123 patients were excluded because of excessive sparsity in each of those patients’ total number of reports. See Table 7 for demographic details of included patients. We denote the sample size $n = 88$. We use the patient self-reported pain data, self-reported sleep quality data, and medication usage during the survey period to generate individual parameters for each patient.

We began by visually analyzing the data for evidence of the sleep-pain relationship. We can see in the two different patient plots above, Figure 8 and 9, that there is often a rise in pain following a drop in sleep quality. This holds true for both patients despite their different pain presentations. Patient A10, Figure 8, has more frequent, high pain episodes compared to Patient C46, Figure 9, with less frequent and less intense episodes yet the dip in sleep quality, the blue line, for both charts shows a high pain episode the next day, the red line. With this visual confirmation of the sleep-pain relationship in the data, we then determined our mathematical framework.

3.4. Methods and Formulation

3.4.1. Modeling Framework:

The mathematical modeling framework consists of a dynamic ordinary differential equations mathematical model and time-varying mHealth data to capture individual pain patterns and predict pain onset. We began by processing the data down to reported pain severity, normalized sleep quality with specific bad sleep quality times identified, and patient reported drug usage for the mechanistic ode model. The model uses these inputs to compute the patient pain level over time with the drug level and sleep quality as pain factors. See Figure 10 for the flow of this framework.

In this framework, pain, sleep quality, and pain medications are tracked over time via differential equations. The equation tracking pain level modulates in response to change in the mHealth variable sleep and pain medication, seen by equation 1. The structure of the equations remains the same for each individual, however, the individual subjects’ data were used to fit the model parameters (e.g. the rates controlling the pain response to opioids and mitigation of pain through sleep).

Each patient had a unique sleep profile, differing drug dosing, and personal pain reports. Therefore, we fit the model to each individual patient and then measured how well the model was able to be fit in
different sub-populations. To create the sub-populations, we post-processed Dr. Valrie’s mHealth data to bin patients into subgroups based on medication usage, sleep quality, age, and percent of pain days. Since opioids, non-opioid NSAIDs, and prescription medications can impact pain directly, we did subsets based on medication type(s) used. We create subsets for each level of sleep quality reported by sets of 10 to investigate the relationship to pain as sleep quality reduces for everyone. We also subset by age to account for the potential difference in pain presentation since SCD is known to be a progressive disease that worsens over time. We then subset by percentage of pain days to assess the impact of overall pain on model performance of these subgroups. This will be explained further in the following sections.

To make this approach more concrete, we have presented a flow chart of our modeling approach in Figure 10.

![Figure 10: Schematic flowchart showing model framework](image)

3.4.2. ODE Model

This mechanistic model framework is comprised of a deterministic set of ODEs. For a single patient, we propose the following deterministic ODE model:
\[
\frac{dP}{dt} = -(k_0 + k_1 D_1 + k_2 D_2 + k_3 D_3)P + \beta \cdot (100 - S_q) \cdot \frac{|S_b(t) - t|}{\epsilon_{S_q}}
\]  

This model predicts pain level by sleep quality and drug kinetics. The first equation shows that patient pain level \( P \) (on a scale of 1 to 100) decreases relative to the total drug profile in their system along with \((k_i)\) when drugs \((D_i)\) are present, with a natural pain relaxation rate of \( k_0 \). Pain also increases at a rate \( \beta \) if a patient experiences low quality sleep \( S_q \) within a period of \( \epsilon_{S_q} \) of the bad sleep event.

### Table 8: Variables in Mechanistic Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P )</td>
<td>Instantaneous pain level on 0–10 scale</td>
</tr>
<tr>
<td>( S_b )</td>
<td>Bad sleep times, i.e., when sleep quality below fixed threshold</td>
</tr>
<tr>
<td>( S_q )</td>
<td>Patient input sleep quality</td>
</tr>
<tr>
<td>( D_1 )</td>
<td>Concentration of drug 1 (long-acting opioid) in the body</td>
</tr>
<tr>
<td>( D_2 )</td>
<td>Concentration of drug 2 (short-acting opioid) in the body</td>
</tr>
<tr>
<td>( D_3 )</td>
<td>Concentration of drug 3 (non-opioid) in the body</td>
</tr>
</tbody>
</table>
The previous three equations\textsuperscript{2}\textsuperscript{3}\textsuperscript{4}D.\textsubscript{i} capture the amount of standard drug \textit{i} doses are within the patient, assuming that the drugs eliminate at constant metabolized rates \textit{k}\textsubscript{i} at drug dosage times \{\textit{τ}_{i,j}\}\textsubscript{j=1}^N. \textit{δ} represents the Dirac delta function that shows the onset of medication concentration as an approximation of the fast rises caused by the typical medications under consideration: long acting (LA) opioids, short acting (SA) opioids, and non-opioid (NO) medications \textsuperscript{17}. Opioids include tramadol (LA), morphine (LA), oxycontin (LA), lortab (SA), and tylenol with codeine (SA). Non opioid (NO) medications such as tylenol, acetaminophen, and advil are used for pain relief and medicinal therapy.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>\textit{k}_0</td>
<td>Pain relaxation rate without drugs</td>
</tr>
<tr>
<td>\textit{k}_i</td>
<td>Effect of drug \textit{i} on pain relaxation rate</td>
</tr>
<tr>
<td>\textit{β}_0</td>
<td>Strength of pain/sleep interaction</td>
</tr>
<tr>
<td>\textit{ε}_S</td>
<td>Tolerance Window of sleep's impact on pain</td>
</tr>
<tr>
<td>\textit{k}_{D1}</td>
<td>Rate of decay of drug 1 (long-acting opioid) in body due to metabolism</td>
</tr>
<tr>
<td>\textit{k}_{D2}</td>
<td>Rate of decay of drug 2 (short-acting opioid) in body due to metabolism</td>
</tr>
<tr>
<td>\textit{k}_{D3}</td>
<td>Rate of decay of drug 3 (non-opioid) in body due to metabolism</td>
</tr>
<tr>
<td>\textit{N}_i</td>
<td>Number of standard drug \textit{i} doses taken</td>
</tr>
<tr>
<td>{\textit{τ}_{i,j}}</td>
<td>Drug \textit{i} dose times (indexed by \textit{j})</td>
</tr>
</tbody>
</table>

The natural pain relaxation rate without drugs \textit{k}_0 was set with a rebound half life of half an hour or \(\frac{\log(2)}{8}\). The drug relaxation rates \textit{k}_i were based on pharmacokinetic models of the medications used. Therefore, we assumed long acting opioids have a 10 hr half life while short acting and non-opioids have 4 hr half lives. Thus, \(k_1 = \frac{\log(0.5)}{-10.0}, k_2 = \frac{\log(0.5)}{-4.0}, k_3 = \frac{\log(0.5)}{-4.0}\).

The pain sleep strength \textit{β} initial value is fixed at 0.23 based on the findings in \textsuperscript{95}. These initial parameters did not yield accurate fit so we implemented an algorithmic tool that utilizes constrained nonlinear optimization to find the constrained minimum of the scalar function (parameters) of multiple variables starting at an initial estimate. The value for \textit{ε}_S is the tolerance window on how long sleep should impact pain and is set to 18 hours.

This parameter optimization routine is personalized for each patient and is updated through each simulation iteration, resulting in each patient having their own parameter values for \textit{β}, \textit{k}_0, & \textit{k}_i that fit best based on their specific initial sleep quality and pain states. We plot the simulated pain level against...
the patient reported pain level and measure the model's predictive value through sensitivity, specificity, and accuracy changes.

3.4.3. Initial Transient Results

Some patient behavior is captured by the model transient as shown in the Figure 12, but it was not as accurate as we would like. The pain response is similar in shape but the timing and peaks are not aligned. It is likely that the medications are having more of an impact on pain than sleep quality based on the current data, (i.e. bad quality sleep not triggering enough of a pain response to match the amplitude and timing of patient data). It is also likely that other health parameters will be needed to better predict pain episodes.

The initial parameter values yielded the model fit seen in Figures 11 and 12 for 10 days (240 hours).
Figure 11: Pediatric patient reported sleep quality (in blue) over 10 days (240 hours) for individual patients

Figure 12: Pediatric patient pain episode model prediction (in red) compared to the patient’s reported pain (black dashes) and reported drug dose times (vertical lines) over 10 days (240 hours)
These initial fittings made us ask "is sleep enough of a predictor to anticipate pain episodes thru a dynamical systems approach? How can we use the data set we have to optimize/understand the sleep pain relationship?" Yang et al 2018 and 2019’s ML predictive pain models [104] [103] highlighted how the patient cohort plays a major role in the relationship between pain and physiological symptoms so we decided to address our three initial assumptions about the sleep - pain relationship for this data and conduct a parameter sweep to identify clinically relevant thresholds.

3.5. Sensitivity Analysis

We noticed that some patients had relatively well matched pain prediction curves and others did not. This led us to ask further questions. We began with three initial questions/hypotheses about the sleep-pain relationship.

(1) The drop in sleep quality that we saw in the raw data (Figures 8 & 9) can be thought of as "a night of bad sleep" but how do we quantify what is bad sleep for this population?

(2) It is known that sleep deprivation or sleep debt can have a delayed and cumulative impact on daily symptoms [36] so we assumed just one night of sleep is not significant and that there’s a cumulative effect. But how much cumulative sleep is most predictive? 2 nights worth? 3 nights? 4 nights? This cumulative effect will be represented by moving mean averages of the sleep quality data.

(3) The subjectivity of pain makes matching daily changes difficult without more biological/ physiological data, as we saw in the initial transients (Figure 12), so we decided to measure the onset of pain instead of pain level. This means we are concerned with an on/ off presentation of pain and moved to a category based fitting. Can we match the switch from low to high pain at a certain threshold? And What threshold is most predictive for this population?

We decided to test these assumptions by conducting a parameter sweep of best potential values for our variable thresholds. By setting these thresholds, we aim for a better understanding of the sleep - pain relationship in pediatric SCD patients as well as identifying the ideal data structure/ requirements for this dynamic model formulation.
3.5.1. Quantifying the Cumulative Impact of Sleep and Defining 'Bad Sleep'

Valrie et al 2019 [95] found that a drop in sleep quality or low sleep quality correlates with high pain the next day. Therefore, there must be a cutoff between good quality sleep and low quality sleep. What is this threshold that signifies bad sleep? In order to identify the relevant time points reported from sleep quality we must vary the threshold for "bad sleep times". A visual example of this threshold is given in the figure above [13] where the value of the dotted line determines which sleep time points will be included in the modulation of pain by sleep.

The time points included are then used to calculate the accumulation windows of sleep. Given the known delayed consequences from sleep deprivation or sleep debt, it is possible that just one night worth of sleep is not significant. [36] We adjust the model to use multiple days of sleep to impact pain for a cumulative effect from sleep quality. In order to quantify this cumulative impact of sleep, we reformatted the sleep quality data into moving mean averages. An example of this formatting can be seen above in Figure 13. Moving mean averages are based on n-days worth of previous data and are often used to define trend direction. Because these averages are based on past data, the larger the n, the larger the lag in data

<table>
<thead>
<tr>
<th>Date</th>
<th>SQ</th>
<th>M.M.A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>sun</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>mon</td>
<td>85</td>
<td></td>
</tr>
<tr>
<td>tues</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>wed</td>
<td>52</td>
<td>56,66,72</td>
</tr>
<tr>
<td>thr</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>fri</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>sat</td>
<td>70</td>
<td></td>
</tr>
</tbody>
</table>
so we only tested model sensitivity up to a 4-day moving average.

3.5.2. Pain onset vs pain level - categorizing pain level through sensitivity testing

![Figure 14: Pain Level Threshold Examples](image)

Given the subjectivity of pain, many models have found issues in finding the necessary factors for fitting to patient pain level exactly [104] [72]. There has been some success in predicting patient pain level with machine learning based on physiological data, but that often requires hospital recorded data. Utilizing mHealth data instead of (or in addition to) hospital data will allow for more accessible prophylactic pain management. Given our reliance on self-reported data, we decided to focus model accuracy on pain onset rather than pain level.

Pain onset conceptualizes pain as an on/off disruption of life activities based on a specified threshold instead of as a subjective scale from 0-100. In Figure 14 we visualize the difference in counting more levels as "high" (red pain) versus the "low" (orange pain). Given the sparsity of pain episodes experienced by children, predicting pain onset requires less nuance and will maximize the use of the low amount of data. In order to define this threshold between low and high pain for model performance, we tested model sensitivity as the threshold varies to subset the data.
3.5.3. Opioid Usage as a Predictability Factor?

<table>
<thead>
<tr>
<th>MODEL PERFORMANCE</th>
<th>Overall</th>
<th>NOMED</th>
<th>NONOP</th>
<th>OPIOIDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>60.16%</td>
<td>41.66%</td>
<td>59.45%</td>
<td>61.78%</td>
</tr>
<tr>
<td>Specificity</td>
<td>76.25%</td>
<td>84.66%</td>
<td>79.48%</td>
<td>67.28%</td>
</tr>
<tr>
<td>Accuracy</td>
<td>73.23%</td>
<td>83.13%</td>
<td>75.93%</td>
<td>65.84%</td>
</tr>
<tr>
<td>PPV</td>
<td>36.90%</td>
<td>9.09%</td>
<td>38.46%</td>
<td>40.11%</td>
</tr>
<tr>
<td>NPV</td>
<td>89.23%</td>
<td>97.52%</td>
<td>90.09%</td>
<td>83.23%</td>
</tr>
</tbody>
</table>

Figure 15: Model Performance Predictive Value Table - Opioid Bins

Patients also report pain medication used during the survey, along with sleep quality and pain. Given the severity of most sickle cell pain episodes, many children and adolescents are prescribed opioid medication amongst others. Long and short acting opioids impact pain for much longer than NSAIDs and other non-opioids and can therefore skew the patient's self-reported pain score, disguising otherwise measurable pain that has unknowingly returned. With this in mind, we also tracked the model's sensitivity to the drug class taken by the patient. Drug parameters were defined using known drug kinetics and patient response was fit in early model optimization. As we determine the optimal sleep and pain thresholds, we also monitor model performance based on drug class subgroups.

We assessed model performance based on the quantified measurements of model accuracy as seen in the Figure 15 above:

- sensitivity := the true positive fraction (correct about diagnosis)
- specificity := the true negative fraction (correct about no diagnosis)
- accuracy := diagnostic evaluation
- negative predictive value (NPV) := probability of true diagnosis as correct
- positive predictive value (PPV) := probability of no diagnosis as correct
Model analysis going forward will highlight sensitivity, the true positive fraction since our aim is to correctly identify when patients are in pain.

3.6. Threshold Sensitivity Testing Results

3.6.1. Finding Patient Trends

Figure 16: Pain Model Sensitivity as Bad Sleep Threshold Lowers
We varied state variable thresholds for a small range of values and recorded model sensitivity in response to these changes. Sensitivity is the true positive fraction used to quantify how often diagnostic and
classification models are correct about the diagnosis. Since we are interested in when pain is switched on/switches to high, we use sensitivity to measure how often our model accurately assesses the patient is in pain. This initial range of values tested (or sweep) was used to identify variable trends in patient populations. We began with our initial assumptions and tested the values right above and below. Initial assumption for pain level uses the lowest $P = 10$ as best and tested sensitivity as that threshold rises. Thresholds for bad sleep are tested next, starting at the highest threshold $< 80$ and testing model sensitivity as that threshold lowers. Finally the ideal sleep window is tested, starting at with the 3 day window accumulation.

Figure 16 shows the results of the Bad Sleep Times threshold testing. In order to define “low quality” sleep, the threshold for where bad sleep begins starts at $x = 80$ - meaning any value of sleep quality $< 80$ is classified as bad sleep. At this threshold, model sensitivity is 59.9% for the overall population blue, 72.5% for the adolescent population purple, and 44% for the child population gold. This threshold value then changes to $x = 70, 60$ and the sensitivity drops to 55% and 42% for the overall population, respectively. This tells us that as the threshold increases, sensitivity decreases; making $BST < 80$ the best threshold value thus far.

Figure 17 shows the results of the pain threshold testing. In order to define “low pain” vs “high pain”, the threshold for where high pain begins starts at $x = 10$ - meaning any value of pain severity $> 10$ is classified as “high pain”. At this threshold, model sensitivity is the same as above with 59.9% for the overall population blue. This threshold value then changes to $x = 10, 20, 30, 40, 50, 60$ and the sensitivity drops to 55.6%, 55.1%, 55.4%, 54.1%, and 52.2% for the overall population, respectively. This tells us that as the threshold increases, sensitivity decreases; making pain severity $> 10$ the best threshold value thus far.

Figure 18 shows the results of the Cumulative Sleep Windows threshold testing. In order to determine which accumulation window of sleep is the most impactful on pain prediction, the n-day moving mean average starts at $n=3$ meaning 3 days of of sleep quality data is averaged to count as one time point of sleep. The sleep quality is later classified as good or bad when inputted into the model. The raw sleep data $n = 1$ model sensitivity is 56.6% for the overall population blue, 70.1% for the adolescent population purple, and 40% for the child population gold. This threshold value then changes to $n= 2, 3, 4$ and the sensitivity changes to 59.9%, 60%, and 55% for the overall population, respectively. This tells us that the
accumulation windows are all more predictive than the raw sleep quality data, though the 4 day moving mean loses too much data to be as effective. We see that the 2 and 3 day moving mean averages are quite comparable and change depending on sub-population.

For all variable thresholds, we see the model performs significantly differently for the adolescent and child sub-populations compared to the overall cohort. This observation, along with our interest in model performance for different drug class sub-populations, led us to extending the sensitivity analysis to broader variable values in order to determine optimal thresholds and model performance for each of the 5 subgroups (adolescents, children and opioid, non-opioids, and no medication).

3.6.2. Subgroup Model Performance

The variable sweep for patient trends revealed pain severity >10 as the best pain threshold, sleep quality <80 as the best “bad sleep” threshold, and n=2,3 as the best accumulation windows for sleep quality for the overall population. Given the difference in model performance for the sub-populations though, defining these thresholds for the sub-populations became imperative. We then extended the variable sweep above and below the established thresholds to test a wider combination of values. All combinations for bad sleep quality <90, 80, 70 | accumulation windows n = 2, 3 | pain severity > 10,20,30 were tested and the corresponding model sensitivity for each subgroup was recorded. Overall, the 2 day moving mean outperformed the 3 day window for all sub-populations so only the 2 day window results are shown below.
The extended sweep shows that (left to right) as the pain severity threshold increases from 10 to 20 to 30 and includes less pain episodes, model sensitivity decreases. We can also see that as the bad sleep quality threshold rises to include more sleep events, model sensitivity rises.

**Children vs Adolescents**
The trends from the overall population are the same for the subgroups (pain severity threshold negatively correlated with model sensitivity while bad sleep quality threshold is positively correlated) but the model still performs significantly better for the adolescent population with 72.8% sensitivity (Figure 22) compared to 47.2% (Figure 21).

Why does the model perform better for adolescents versus children?
We assume this difference in model sensitivity by age is due to the increased number of pain events documented by the adolescents in the study. On average, adolescents experienced pain episodes on 26% of the days reported, compared to children who only had pain episodes on average 21% of the days reported. Children also experienced less pain days compared to the entire patient demographic. Overall, adolescents had more pain days, giving the model more pain history to use for prediction.

This finding may support the theory that the acute pain episodes in childhood slowly increase in adolescence to become a chronic pain condition with a new measurable baseline in adulthood.

**Opioids vs Non-Opioids vs No Medication**
The trends from the overall population are the same for the subgroups (pain severity threshold negatively correlated with model sensitivity while bad sleep quality threshold is positively correlated) but the model does perform significantly better for the opioid population with 61.8% sensitivity and 63.8% for non-opioids compared to 45.8% for no medication, respectively. The opioid and
non-opioid classes are quite similar in model performance but the model performs much worse for those with no medication taken.

**Why does the model perform best with the opioid population?**

We assume this difference in model sensitivity by drug class may be due to the fact that those who took no medication, often had little to no pain reported and therefore little to no pain to predict.

![Figure 25: Average Percent of Pain Days by Opioid Usage](image)

On average, long acting opioid users experienced pain episodes on 39% of the days reported, compared to short acting opioid users who only had pain episodes on average 35% of the days reported and non-opioid users who 30%. This makes sense logically, as those in more pain would need stronger pain medication. Overall, those who took long acting opioid medication had more pain days on average, giving the model more pain history to use for prediction.

3.6.3. **Subset SCD Population by Age or Drug Class?**

The hybrid model from Clifton [17] was quite predictive with pain and medication data alone due to the unmitigated pain level experienced by adults as a driving factor for pain. The unmitigated pain level was calculated using the patient's age, disease type, vitamin use, and opioid-vs-non opioid usage where unmitigated pain levels increase with patient age and certain SCD types and unmitigated pain decreases
When non-opioids are used.

Since there is no unmitigated pain level for children, pain in our model is driven by sleep alone. Given their proximity to the shift to chronic pain, we surmise that our model performs best for adolescents because they are actively “building” a pain level baseline through their more frequent episodes (compared to children).

When looking at model sensitivity based on percentage of pain days first, we can see that the more percent days of pain, the better the model performs for all groups (age-wise). Therefore, our model seems to run best based on the amount of pain data present.
Similarly, looking at sensitivity based on percentage of pain days first (medication-wise), we see that the model performs best for non opioid users in very little pain of for long acting opioid users in a lot of pain. This again confirms that the amount of pain days plays a major role in model sensitivity.

This observation supports the theory that pain days increase as SCD patients age, improving predic-
tive value (because of pain history) and requiring more pain medication over time in response.

This finding supports subsetting the pain population into cohorts based on percentage of pain days to effectively capture the progression of pain dynamics in sickle cell disease, regardless of age or medication usage.

3.7. Discussion

3.7.1. Is self reported sleep enough of a predictor?

In this work, we present a mechanistic ODE model for predicting pediatric pain episodes. We incorporated self-reported sleep quality as a primary factor in pain levels the next day along with the usage of common pain medication. The model output displays some evidence of a pain response influenced by sleep, but with model sensitivities below 80%, not enough to be predictive on its own.

When plotting sleep quality bins against the pain response, our team found that there is a higher sleep quality correlation with pain in adolescents, compared to children. Perhaps this highlights how both increased pain days and increased medication usage impacts the importance of patient sleep quality as patients age - becoming more significant as the patient gets older and the pain baseline rises, requiring better quality sleep to offset the growing effect of chronic pain.

It should be noted that there are many factors known to impact pain such as temperature, heart rate, and activity level. Incorporating these and other additional factors may improve the pain prediction by increasing the complexity of the model.

3.8. Conclusion

3.8.1. Model limitations and Future Work

In this work, we present a mechanistic ODE model for predicting pediatric pain episodes. We find that self-reported sleep quality alone is not sufficient to predict pain onset. The usage of self-reported sleep quality, as opposed to sleep actigraphy, sleep duration, or other sleep measurement variations, plays a significant role in this finding. Perhaps the incorporation of additional or alternative sleep measurements for the sleep variable would perform differently in this model framework. However, the analysis
of the model results did reveal some interesting patterns.

The sensitivity analysis revealed that 2 days accumulation of sleep quality to be more predictive than 1, 3, or 4 days accumulation and could inform data collection protocol for future clinical SCD research. Similarly, data analysis revealed that sub-setting SCD patient cohorts by percent or amount of pain experienced rather than behavioral metrics to be the most effective for future pain prediction.

Consider our model limitations to avoid pitfalls in new data collection. One limitation of our model is that the amount of time the data was collected was discrete and over a short period of time, decreasing the amount of pain history available for prediction. Since pediatric SCD patients experience fewer episodes throughout the year, longer data collection windows may be useful for more accurate prediction.

Another limitation was the size of the data cohort. Having only 88 patients didn’t allow for much room for training and fitting the model based on the pain history. The larger the cohort size, the more valuable the findings. The sub population findings here are severely impacted by how few people are in each sub group because of the overall cohort size.

A primary limitation was the lack of time series data for our mHealth variable sleep quality. Since sleep quality was subjective, we may be missing physiological impacts happening before, during, and after sleep that impact its quality. Sleep quality could also be replaced with another mHealth variable (positive affects, heart rate avg, etc.) to predict pain or multiple mHealth variables could be cumulatively more effective. Therefore many considerations for data collection should be taken when setting up future clinical SCD research studies for future analysis.

Increasing the complexity of the model is imperative to increase model accuracy. One way of doing so will be to add more sleep history of different time windows along with sleep quality to impact pain. Also, incorporating additional mHealth variables may also improve the accuracy of the model. Machine learning algorithms are particularly useful in this, as they allow for best fit feature selection amongst the variables. We expect the results of these extensions to aid in informing future data collection for Sickle Cell studies.

This project has the potential to change clinical care by providing models that can assist pediatric SCD patients, their caregivers, and their health care providers. Thee model could also be extended to
youth with other illnesses that cause pain. The long-term benefits of implementing mechanistic models into patient care include reducing the suffering of youth with SCD and their families as well as the societal burden of the substantial health care costs of managing SCD pain.
4. Temporal Dynamics of Inflammation and Adhesion in Sickle Cell Vaso-Occlusion: An ODE Approach

4.1. Introduction

4.1.1. Overview of Inflammation-Driven Vaso-Occlusive Crises

Sickle Cell Disease (SCD) can be characterized as a chronic inflammatory disease where multiple cell types interact, as the polymerization of deoxygenated HbS creates numerous pathological responses in the body. Deoxygenated HbS polymerizes inside red blood cells, making the membrane deformable and sickle shaped with more adhesive molecules now available. These sickle shaped cells (sRBCs) also have shorter lifespans than the average Red Blood Cells (RBCs), causing a constant shortage of red blood cells, known as anemia, and a constant influx of cell-free heme from sRBC hemolysis. The increased oxidative stress from heme activates healthy endothelial cells (ECs), bringing stored P-selectin adhesion molecules to the surface and enabling the highly adhesive, sickle shaped cells to stick the vascular endothelial wall, as seen in part A of the Figure.
Sickle blood cell adhesion to the endothelium causes more inflammation and activated ECs, signaling the start of the biological immune response known as the "coagulation cascade" where activated endothelial cells recruit platelets, neutrophils, and macrophages to the site of the obstruction for assistance [65], as seen in part B of the Figure. In SCD, this "solution" from the body becomes the cause of larger rolling aggregates on the vascular wall due to the adhesion molecules present from endothelial cells, platelets, sickle cells, and neutrophils.
Figure 30: Sickle Cell VOC Model Variables/Cells

These cells form aggregates that cause blood flow obstructions known as vaso-occlusive crises and are the hallmark symptom of SCD \(^\text{18}\), as seen in part C of the Figure 30. This pro-inflammatory state is continually promoted, as the Sickle cells die every 10-20 days and the hemolysis from both RBCs and sRBCs releases heme into the blood and activates endothelial cells \(^\text{47}\). The constant increased heme causes endothelial dysfunction/sterile inflammation that increases oxidative stress on the blood, increasing internal hypoxia and ischemic organ damage \(^\text{89}\). This oxidative stress causes more HbS molecules to polymerize and form sickle RBCs within the blood, leading to the beginning of the cycle again.
This process of HbS polymerization and hemolysis causing SCD adhesion to the endothelium causes vascular inflammation and the resulting immune responses lead to more sRBC adhesion and sterile inflammation. While much is known from assays and the literature about these pathways, there is little knowledge about which mechanistic pathways are significant and treatable.\cite{44} Given the cyclic pathway of these cell populations\cite{31}, interruptions or inhibitions of one group may shift the resulting symptoms of the system, making VOC and Red Blood Cell behaviour a major focus in SCD research.

4.2. History of VOC Modeling and Comparison of Our Approach

Dynamic math modeling allows us to model and simulate this process and gain insight about ways to prevent and/or treat VOC pain in SCD with minimal cost and injury to patients. Given the multiple si-
multaneous processes involved in the vaso-occlusive crises, many modeling approaches have been used to understand cell behavior and VOC triggers. Each phase of the vaso-occlusive process offers multiple techniques for modeling and analysis.

### 4.2.1. Red Blood Cell Models

The primary triggers of VOCs are RBC hemolysis, as seen in Figure 32, increasing oxidative stress and HbS polymerization causing gelation of cells that have more adhesion molecules on the surface. Therefore, many researchers have created models of this initial process to glean potential ways of preventing sickling or hemolysis and reducing VOC frequency [44] [62] [107].

HbS polymerization, also known as sickling or gelation, initializes the VOC process through sRBC membrane changes, making it one of the first priorities for computational research. Dong 1992 [25] studied the rheological property changes that polymerization has on membrane deformability using an ODE Model. Other computational models of polymerization investigate causes of polymerization like hypoxia (Li et al 2017 [56]), and blood cell mechanics (MedKour et al 2008 [62], Lu et al 2017 [58], Daniel-
Understanding the blood flow and fluid dynamics in SCD has also been an imperative part of VOC modeling. Many of these models use partial differential equations to account for spatial impact on cell movement and diffusion. Models such as Bazzi et al 2020 [6], Zhang et al 2020 [106], Sawyer et al 2022 [85], and Szafraneic 2022 [30] track red blood cell behavior, movement in vessel flow and validate through in vivo experiments and assays. Fluid based models have also been used to evaluate cell adhesive behavior in flow (Lei 2013, Blakely 2020) and the influence of cell shape on obstructions (Chaturvedi et al 2021 [14]).

PKPD and QSP models have also been used in modeling VOC to understand treatment efficacy and parameters (Altrock 2016 [2], Zheng et al 2021 [107], Sy et al 2022 [90], Pandey et al 2023 Pandey 2023 [75]). The model in this study is most similar to these in its aim and compartmental-type structure, though our model tracks temporal dynamics and local concentrations instead of time specific compartment changes.

4.2.2. Models of VOC and Inflammation

Other researchers have modeled active vaso-occlusive crises in order to understand cell behavior and flow under different conditions. The vaso-occlusive crisis is identified by the adhesion and aggregation of sRBCs and leukocytes on the endothelial vessel wall, as seen in the Figure 33 below, so researchers have explored the mechanisms of this process through math modeling.

These models often measure the impact of VOC through measures of endothelial dysfunction and/or adhesive molecule dynamics. Deonikar 2012 [24]'s model, for example, investigated the impact of heme on endothelial dysfunction as a trigger for VOC. Chalacheva et al 2017 [12] measured autonomic responses to pain as a potential trigger for endothelial dysfunction and VOC.

The flow models mentioned previously also measure endothelial and adhesive dynamics during VOC (Lei et al 2013 [53], Blakely & Horton et al 2020 [8], Sy et al 2022 [90], Chaturvedi et al 2023 [15]).
More recently, with assays revealing the importance of leukocytes, platelets, and other immune cells in the vaso-occlusive process, researchers have been considering SCD an chronic inflammatory disease. These studies have established increased inflammation present during SCD crises. 

While there have been many microfluidic experiments to investigate this further
there are little to no mathematical models of the temporal relationships between sRBCs, endothelial cells, and leukocytes in Sickle Cell disease specifically. There have been studies of the relationships between blood cells and immune cells in other diseases, laying the foundation for further exploration of SCD through an inflammatory lens. We have identified key inflammatory and immune variables from the literature and created an ODE model to understand the molecular pathways in SCD and elucidate how they influence disease pathophysiology.

In this study, we present a model of the temporal dynamics of VOC as visualized in Figure 31 as described by the literature, particularly informed by Jang et al 2021.

Jang et al 2021 discusses the phases of pain during the vaso-occlusive crisis in SCD in relation to the inflammatory state of the patients during crisis. It explains that hospitals generally recognize four phases of VOC as seen in the figure below.

Phase 1 is associated with low aching intense pain. Phase 2 is associated with max aching pain from the ischemia re-perfusion damage of VOC. Phase 3 is associated with constant severe pain from the immune response to inflammation. This phase can be considered a tipping point towards what Jang et al refers to as secondary VOC (sVOC), the cyclic downstream effect pictured in figure that causes more VOC instead of the assumed resolution of pain. The model in this work represents this cyclic VOC process as influenced by inflammation through an ODE system. We aim to predict sVOC likelihood over time and
elucidate potential inflammatory pathways for pain reduction in SCD.

4.2.3. Research Aims

There are a variety of treatments available to inhibit different steps of the VOC process without the full picture of the system's response. A broader understanding of the mechanisms involved may highlight key research areas that may be new or otherwise previously dismissed. The aim of this model is to elucidate significant pathways for SCD treatment to decrease overall VOC pain duration by modeling changes in significant cell populations in response to known VOC triggers and/or treatments.

The primary research questions are:

- Does this model accurately represent the known sterile inflammation, ischaemia–reperfusion injury or chemically induced injury typically occurs in the absence of any microorganisms [16], present in pediatric sickle cell patients? 
  
  Parameter Estimation: Latin Hypercube Sampling

- Which parameters are most significant? What mechanistic pathways impact sVOC likelihood transition from steady state to crisis?
  
  Sensitivity Analysis + Identifiability

- Which treatment options and combinations reduce sVOC likelihood over time and/or return patients to a healthy equilibrium?
  
  Predictive Analysis of Interventions: testing known SCD Therapies

4.3. Methods and Formulation

4.3.1. Assumptions on the Subject

- Healthy vs Steady State vs Crisis. The aim of the proposed model is to describe the cell dynamics of SCD patients at steady state versus in crisis, in comparison to healthy controls. In vivo and clinical studies show inflammatory cell concentration differences for each patient state. [96]
• **Existence of Sterile Inflammation at Baseline.** It is assumed that SCD patients have *at least one* history of VOC in their lifetime. In this case, the initial conditions and source rates should be $>0$ for all non-healthy patients and will use patient data to benchmark these values.

### 4.3.2. Assumptions on Physiological Process

• **Average Cell.** We are using a non-linear cell-population ODE modeling approach to understand system dynamics.

• **Only Most Necessary Processes.** In order to reduce model complexity, only the primary mediators are considered at this time. It should be noted that there are other immune and nerve cells involved in the triggering and pain response to VOC not counted in this model such as T cells, mast cells, reactive oxidation species (ROS), and astrocytes. These cells can be added in future work to explore a more nerve based model along with these temporal dynamics.

• **Concentration in One Area.** We are not accounting for the movement of cells or blood flow and are instead focusing on the average concentration of that cell population over time in the region. Our concern is the timing of treatment and ability to turn off rampant inflammation. We assume the environment is spatially averages so we use an ODE model instead of a PDE model, as are found in other VOC math models.

• **Individual Features of VOC.** Individuals have different SCD types that may impact their HbS count and sickling rate naturally or in response to stress. This is covered by two gelation rate parameters, $p_s$ and $p_o$, in the model.
4.3.3. Model Equations and Variables

The model, defined by the schematic in Figure 35, developed in this work tracks the resulting immune response within the vessel during VOC for SCD patients. We do not explicitly model the blood component and all variables represent local levels. To create this model, a previous model of the inflammatory process using immune cells, neutrophils, and macrophages [93] has been adapted to include endothelial cell activation and damage, the resulting activation of platelets, and the build-up of adhesion molecules from the immune response that is indicative of VOC in SCD patients.

We assume Red Blood Cells (R) (and thereby sickle RBCs (S)) to be at a steady state. We use the following two auxiliary equations as substitutions inside the system equations.

\[ R' = BM - ps \cdot R - \mu_r \cdot R \]

\[ S' = \frac{ps \cdot R}{(\mu_s \cdot S - ps \cdot R)} \]
A healthy subject is assumed to have nearly no inflammation prior to stimulus, with very low levels of gene, and so has no immune cell influx. Therefore all of our immune cell variables have an initial condition of zero. The presence of sickle cells and resulting heme is assumed to stimulate a very rapid increase in inflammation through activated Endothelial cells that grows instead of subsides, as heme and neutrophil removal by macrophages is inhibited in SCD patients. The system is assumed to progress to a chronic state with no intervention or inhibition.

To capture the cell-to-cell interactions during VOC, we must track the cell populations listed below as variables. As in Torres et al [93], we assume immune cells are activated and influx into the local environment rapidly compared to other dynamics, so a quasi-steady state assumption is used to create the Michaelis-Menten type influx terms for equations 5, 9, 11, 15, and 16.

We do not track cytokines and chemokines explicitly but let the production of immune cells act as an indicator of the associated cytokine level. We then track the total cells for each population with units of number cells. Model parameters for activation, decay, and transition are listed in the Table below. Many of the model parameters are representative of immune functions such as cell signaling and mediation so their units are given in context-specific values. The following cell relationships are modeled in this system:

1. Activated Endothelial Cells - $E_A$

   The first step of the vaso-occlusive process begins with the activation of the endothelium. Healthy endothelial cells may become activated through several processes, including the heme-activated generation of ROS (reactive oxygen species) by sRBCs or endothelial cells, sRBC adhesion and signaling to endothelial cells (ECs), or infections/triggers that lead to the secretion of inflammatory cytokines. These activating factors often cause endothelial cells to translocate P-selectin from its granule to the surface due to the increased ROS and adhere to the PSGL-1 like receptor on sickle RBCs [34].

   This subsequently recruits immune cells (platelets, neutrophils, monocytes, T cells, and mast cells) to the site of inflammation through the TLR4 pathway. The platelets and neutrophils bring more adhesive molecules (P-selectin and MAC-1 respectively) to the already sticky site. Endothelial ad-
hesive molecules along with sickle cells, platelets, and neutrophils aggregate the cells together to form even more occlusions in the vessel [105] [41].

(2) Damaged Endothelial Cells - $E_D^{6}$
Activated endothelial cells become damaged after prolonged time under inflammation with no resolution. Damaged endothelial cells die off to be later replaced by floating progenitor ECs upon their maturation. [88] [49]

(3) Adhesion Molecules - $AM^{19}$
Adhesion molecules are stored in granule of endothelial cells, platelets, and neutrophils. These molecules are also expressed on the surface of sRBCs due to the membrane changes during HbS polymerization. Endothelial cells with expressed P-selectin adhere to the PSGL-1 like receptor on sickle RBCs, causing inflammation and initiating the immune coagulation cascade. Activated platelets and neutrophils are recruited to the site of damage where platelets and neutrophils adhere to the endothelium (or together) to repair the damage. The amount of adhesion molecules present in total, from the cascade and local cells, increase VOC likelihood and is correlated with the clinical severity score [37] [94].

(4) Heme from Dead sRBCS - $AS^{8}$
Cell-free heme is released more quickly by sRBCs, as HbS polymerization promotes hemolysis. As sickle cells die every 10-20 days, cell-free heme is released into the blood stream, promoting endothelial dysfunction by depleting the endothelial nitric oxide (NO) reserves. Heme-mediated TLR4 signaling produces reactive oxygen species (ROS) inside ECs and promotes endothelial dysfunction that consistently activates inflammasomes, ECs, and platelets. Hemin is also released during hemolysis, activating platelets. The byproducts of heme also alter the function of recruited macrophages, creating more M1 than M2 macrophages and disrupting the inflammatory resolution process [87] [97].

(5) Activated Platelets - $PA^{5}$
Platelets that are resting in circulation become activated when ECs express surface adhesion molecules and are recruited to the site to manage endothelial inflammation. Platelets gather at the dam-

84
age site and adhere to white blood cells to create the "platelet plug" and restore endothelial function. These activated platelets also release cytokines and chemokines that are chemotactic to neutrophils and monocytes [96] [101] [43].

(6) Neutrophils - **NA** [11]

Neutrophils resting in circulation are the first responders to acute inflammation and are continuously activated and recruited to the site of damage. They release serine proteases chemicals, attach to the adhesive environment, and activate macrophages for repair. Platelet-neutrophil interaction and adhesion consequently creates more obstructions with the sRBCs in circulation [39] [105].

(7) Neutrophil NETS - **NT** [12]

Neutrophils can also form NETs (extracellular traps) that promote inflammation with multiple ends to attach to the adhesive environment. These pro-inflammatory neutrophils create more cell-cell to aggregates that create more obstructions with the sRBCs in circulation. They also release pro-inflammatory cytokines that activate ECs [105].

(8) Apoptotic Neutrophils - **NP** [13]

Neutrophils apoptosis is a programmed cell death process that allows for the removal of aged or activated neutrophils. Since neutrophils release cytokines that promote inflammation, apoptosis prevents the release of the harmful contents of dying cells. Neutrophils are cleared by macrophages upon signal and contribute to the resolution of inflammation [39].

(9) M1 Macrophages - **M1** [16]

M2 macrophages are dysregulated in SCD due to the increased heme and ROS in the blood, causing M2 macrophages to behave like M1 macrophages. M1 macrophages have pro-inflammatory functions that promote an environment of sterile inflammation [27] [26].

(10) M2 Macrophages - **M2** [15]

Monocytes have many cell types but are commonly known to become macrophages during inflammation. Macrophage recruitment is mediated by granule proteins released from neutrophils and supported through platelet chemokine signaling. Macrophages tend to extend the lifespan of
neutrophils through the secretion of GM-CSF and also eat away much of the dead cells in circulation. M2 macrophages have anti-inflammatory functions that start the resolution of inflammation through its chemokines and clearing of the toxins in the environment [27] [26] [97].

The model is summarized in the previous figure and is described by Eqs [5-19].

**Endothelial Cells:**

\[
\frac{dE_A}{dt} = \frac{s_{E_A} R_E(AS, NT, NA, M1)}{\mu_{ec} + R_E(AS, NT, NA, M1)} - b_{ead} * E_A
\]  

(5)

\[
\frac{dE_D}{dt} = \frac{b_{ead} * E_A - b_{edh} * E_D}{\mu_{edh} * E_D}
\]  

(6)

where the activation/influx rates for EA are given by:

\[
R_E = \frac{s_{E_A} R_E(AS, NT, NA, M1)}{\mu_{ec} + R_E(AS, NT, NA, M1)}
\]

(7)

**Heme:**

\[
\frac{dAS}{dt} = \frac{\mu_s * (p_s * BM_p - \mu_r)}{p_s - \mu_r} - \mu_{mac} * AS * NA - b_{mas} * AS * (M1 * M2)
\]

(8)

**Platelets:**

\[
\frac{dPA}{dt} = \frac{s_P * R_P(AS, EA)}{\mu_p + R_P(AS, EA)} - \frac{\mu_p * PA}{\mu_p * PA}
\]

(9)

where the activation of platelets is:

\[
R_P = k_{asP} * AS + k_{ep} * EA
\]

(10)
Neutrophils:

\[
\frac{d N A}{d t} = \frac{R_N(A S, P A)}{\mu_N + R_N(A S, P A) + b_{nat} * N A} - \frac{b_{nat} * N A}{\mu_N * N A} - \frac{b_{nat} * N A}{\mu_N * N A} \tag{11}
\]

\[
\frac{d N T}{d t} = \frac{b_{nat} * N A - \mu_N * N T}{\mu_N * N T} \tag{12}
\]

\[
\frac{d N P}{d t} = \mu_N * N A - k_{m1np} * NP f_i(M_1, N A) - k_{m2np} * NP f_i(M_2, N A) - k_{np} * N A - \mu_N * NP \tag{13}
\]

where the activation/influx rates for neutrophils is:

\[
R_N = k_{asn} * AS + k_{pn} * E A \tag{14}
\]

Macrophages:

\[
\frac{d M_2}{d t} = \frac{s_M * R_{M2}(N A, N P, M_2, E A)}{\mu_m + R_{M2}(N A, N P, M_2, E A) + R_{M1}(M_1)} - \frac{b_{m2m1} * k_{npm2} * M_2 * A S}{\mu_{M2} * M_2} + \frac{b_{m2m1} * M_2}{\mu_{M2} * M_2} \tag{15}
\]

\[
\frac{d M_1}{d t} = \frac{s_M * R_{M1}(M_2, M_1)}{\mu_m + R_{M2}(N A, N P, M_2, E A) + R_{M1}(M_1)} + \frac{b_{m2m1} * k_{npm1} * M_2 * A S}{\mu_{M1} * M_1} - \frac{b_{m2m1} * M_2}{\mu_{M1} * M_1} \tag{16}
\]

where the activation/influx rates for M1 and M2 are given by:
activation by NA  activation by necrotic NP  activation by M2s and their cytokines  activation by byproducts of EA
\[ R_{M2} = k_{mn2} \cdot NA \cdot M2 + k_{npn2} \cdot NP \cdot M2 + k_{m2m2} \cdot M2 + k_{em2} \cdot EA \cdot M2 \] (17)

activation by M1s and their cytokines  switch from M2 to M1  background pro-inflammamy cytokines
\[ R_{M1} = k_{m1m1} \cdot M1 + b_{m2m1} \cdot M2 + k_{c} \] (18)

Adhesion Molecules:
\[
\frac{d AM}{dt} = \frac{s_{AM} \cdot R_{AM}(S, PA, NA, EA)}{\mu_{AM} + R_{AM}(S, PA, NA, EA)} - \mu_{AS} \cdot AM
\] (19)

where the activation/influx rate for adhesion molecules is given by:
\[
R_{AM} = k_{sam} \cdot \left( p_s \cdot \frac{BM}{p_s - \mu_r} \right) + k_{pam} \cdot PA + k_{nam} \cdot (NA + NT) + k_{eam} \cdot EA
\] (20)

These equations are an extension of the inflammatory response formulation found in Torres et al 2019 [93].

4.3.4. Model Equations and sVOC Likelihood

By tracking these key blood and immune cell populations over time, we are able to quantify the inflammation present at time \( t \) in the body. The vaso-occlusive crisis is directly influenced by the cell aggregates present in a given area at time \( t \). Since the "susceptibility to developing VOC among different individuals may depend on the size of the adaptive reserves in the compensated state" [44], these cell populations can be used to determine sVOC likelihood 4.3.4.
As the amount of the heme (Eq. 8), platelets (Eq. 9), neutrophils (Eq. 11), adhesion molecules (Eq. 19), and activated endothelial cells (Eq. 5) increases, the likelihood of larger aggregates and occlusions occurring increases. The distinctions for sVOC as listed above in Figure 4.3.4 are arbitrary and only meant to be representative of an increasing scale of damage. Steady state patient sVOC likelihood lies between phases 1 and 2 at $sVOC = 0.5$ and crisis patient state begins at the tipping point, phase 3, at $sVOC = 0.8$. Each patient state has its own parameter sets based on baseline patient history.

These equations form a system of ODEs that captures the most important aspects of the immune response to VOC and the occurrence of sVOC. In the following sections, we use computational approaches to explore parameter space, determine the parameters the model is most sensitive to, and establish influential predictors of model outcomes.

We end with case studies in which we modulated particular parameters to evaluate the impact of potential treatment combinations.

### 4.3.5. Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>I.C.</th>
<th>Bounds</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_r$</td>
<td>decay rate of Red Blood Cells</td>
<td>.0083</td>
<td>fixed</td>
<td>Sundd et al 2019</td>
</tr>
<tr>
<td>BM</td>
<td>bone marrow RBC influx rate</td>
<td>10e6</td>
<td>[10e4, 10e8]</td>
<td>Sundd et al 2019</td>
</tr>
<tr>
<td>$p_s$</td>
<td>sickling/gelation transformation rate</td>
<td>.45</td>
<td>[0.25, 0.9]</td>
<td>estimated</td>
</tr>
<tr>
<td>$\mu_s$</td>
<td>decay rate of Sickle Red Blood Cells</td>
<td>.05</td>
<td>fixed</td>
<td>Sundd et al 2019</td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Value</td>
<td>Source</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>-------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>$\mu_{mas}$</td>
<td>Decay rate of sRBC heme eaten by macrophages</td>
<td>6.11</td>
<td>Torres et al 2020 [93]</td>
<td></td>
</tr>
<tr>
<td>$\mu_{m}$</td>
<td>Decay rate of resting macrophages</td>
<td>6.956</td>
<td>Torres et al 2020 [93]</td>
<td></td>
</tr>
<tr>
<td>$b_{eha}$</td>
<td>Transition rate from healthy EC to activated EC</td>
<td>0.3</td>
<td>Ezaki et al 2001 [29]</td>
<td></td>
</tr>
<tr>
<td>$b_{ead}$</td>
<td>Transition rate from activated EC to damaged EC based on time passed</td>
<td>0.215</td>
<td>Montezano et al 2017 [64]</td>
<td></td>
</tr>
<tr>
<td>$k_{ne}$</td>
<td>Activation rate of ECs from neutrophils</td>
<td>0.01</td>
<td>Estimated</td>
<td></td>
</tr>
<tr>
<td>$b_{edh}$</td>
<td>Healing rate of endothelial cells back to healthy cells</td>
<td>0.0001</td>
<td>Estimated</td>
<td></td>
</tr>
<tr>
<td>$k_{m1ea}$</td>
<td>Activation rate of ECs from M1 macrophages</td>
<td>0.01</td>
<td>Estimated</td>
<td></td>
</tr>
<tr>
<td>$k_{ase}$</td>
<td>Activation rate of ECs by sRBC heme</td>
<td>0.333</td>
<td>Baselet et al 2019 [4]</td>
<td></td>
</tr>
<tr>
<td>$\mu_{ec}$</td>
<td>Decay rate of endothelial cells</td>
<td>0.0027</td>
<td>Estimated</td>
<td></td>
</tr>
<tr>
<td>$b_{pha}$</td>
<td>Transition rate of platelets from inactive circulating to active platelets</td>
<td>0.02</td>
<td>Estimated</td>
<td></td>
</tr>
<tr>
<td>$k_{ep}$</td>
<td>Activation rate of platelets mediated by endothelial cells</td>
<td>3.025</td>
<td>Estimated</td>
<td></td>
</tr>
<tr>
<td>$k_{asp}$</td>
<td>Activation rate of platelets mediated by sRBC heme</td>
<td>3</td>
<td>Estimated</td>
<td></td>
</tr>
<tr>
<td>$\mu_{p}$</td>
<td>Decay rate of Platelets</td>
<td>0.1</td>
<td>LeBrasseur et al 2007 [52]</td>
<td></td>
</tr>
<tr>
<td>$b_{nha}$</td>
<td>Transition rate of neutrophils from inactive circulating to activated</td>
<td>0.607</td>
<td>Torres et al 2020 [93]</td>
<td></td>
</tr>
<tr>
<td>$k_{pn}$</td>
<td>Activation rate of neutrophils by platelets</td>
<td>0.1</td>
<td>Estimated</td>
<td></td>
</tr>
<tr>
<td>$k_{asn}$</td>
<td>Activation rate of neutrophils mediated by sRBC heme</td>
<td>3.703</td>
<td>Estimated</td>
<td></td>
</tr>
<tr>
<td>$\mu_{n}$</td>
<td>Decay rate of neutrophils</td>
<td>4.773</td>
<td>No source</td>
<td></td>
</tr>
<tr>
<td>$\mu_{nas}$</td>
<td>Decay rate from eating heme</td>
<td>3.025</td>
<td>Torres et al 2020 [93]</td>
<td></td>
</tr>
<tr>
<td>Parameter</td>
<td>Description</td>
<td>Value</td>
<td>Confidence Interval</td>
<td>Reference</td>
</tr>
<tr>
<td>-------------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------</td>
<td>---------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>$b_{nat}$</td>
<td>transition rate of neutrophils from active to NETs</td>
<td>.1667</td>
<td>$[1e-5, 0.9]$</td>
<td>Zhong et al 2023</td>
</tr>
<tr>
<td>$k_{m1np}$</td>
<td>decay rate of apoptotic neutrophils eaten by M1</td>
<td>2.898</td>
<td>$[.01, 5]$</td>
<td>Torres et al 2020</td>
</tr>
<tr>
<td>$k_{m2np}$</td>
<td>decay rate of apoptotic neutrophils eaten by M2</td>
<td>87.08</td>
<td>$[10, 90]$</td>
<td>Torres et al 2020</td>
</tr>
<tr>
<td>$\mu_{np}$</td>
<td>secondary apoptosis of neutrophils</td>
<td>1.309</td>
<td>$[.001, 3]$</td>
<td>Torres et al 2020</td>
</tr>
<tr>
<td>$k_{sam}$</td>
<td>activation of adhesion molecules from sRBCs</td>
<td>.2</td>
<td>$[.001, 5]$</td>
<td>estimated</td>
</tr>
<tr>
<td>$k_{nam}$</td>
<td>activation of adhesion molecules from neutrophil activation</td>
<td>.1</td>
<td>$[.001, 5]$</td>
<td>estimated</td>
</tr>
<tr>
<td>$k_{pam}$</td>
<td>activation of adhesion molecules from platelet activation</td>
<td>.1</td>
<td>$[.001, 5]$</td>
<td>estimated</td>
</tr>
<tr>
<td>$k_{eam}$</td>
<td>activation of adhesion molecules from endothelial cells</td>
<td>.16</td>
<td>$[.001, 5]$</td>
<td>estimated</td>
</tr>
<tr>
<td>$k_{nm2}$</td>
<td>activation of M2 macrophages by active neutrophils</td>
<td>.025</td>
<td>$[.0001, 0.9]$</td>
<td>Torres et al 2020</td>
</tr>
<tr>
<td>$k_{npm2}$</td>
<td>activation of M2 macrophages by apoptotic neutrophils</td>
<td>.045</td>
<td>$[.0001, 0.9]$</td>
<td>Torres et al 2020</td>
</tr>
<tr>
<td>$k_{m2m2}$</td>
<td>activation of M2 macrophages by itself and its chemokines</td>
<td>1.624</td>
<td>$[.0001, 2.0]$</td>
<td>Torres et al 2020</td>
</tr>
<tr>
<td>$k_{em2}$</td>
<td>activation of M2 macrophages by active endothelial cells</td>
<td>1</td>
<td>$[.0001, 2.0]$</td>
<td>Torres et al 2020</td>
</tr>
<tr>
<td>$b_{m2m1}$</td>
<td>transition rate from anti M2 to pro M1; function of heme</td>
<td>8.281</td>
<td>$[1, 10]$</td>
<td>Torres et al 2020</td>
</tr>
<tr>
<td>$k_{m1m1}$</td>
<td>activation of M1 macrophages by itself and its cytokines</td>
<td>.001</td>
<td>$[1e-4, .9]$</td>
<td>Torres et al 2020</td>
</tr>
<tr>
<td>$k_c$</td>
<td>background anti-inflammatory cytokines</td>
<td>.0125</td>
<td>$[.0001, 0.5]$</td>
<td>Torres et al 2020</td>
</tr>
<tr>
<td>$\mu_{m1}$</td>
<td>decay rate M1 macrophages</td>
<td>6.956</td>
<td>$[1, 10]$</td>
<td>Torres et al 2020</td>
</tr>
</tbody>
</table>

91
4.3.6. Model Fitting and Validation: Klouda et al 2020 Data

While the initial parameter estimation was found using a heuristic approach from literature analysis, in order to improve model accuracy, we fit the parameters based on population data, with the summary data from the Klouda et al 2020 data set [51].

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
<th>Confidence Interval</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_m$</td>
<td>decay rate M2 macrophages</td>
<td>8.271</td>
<td>[1, 10]</td>
<td>Torres et al 2020 [93]</td>
</tr>
<tr>
<td>$ad$</td>
<td>overall adhesion rate</td>
<td>.3</td>
<td>[.001, 3]</td>
<td>estimated</td>
</tr>
<tr>
<td>$s_E$</td>
<td>source of activated endothelial cells</td>
<td>36</td>
<td>[10, 125]</td>
<td>estimated</td>
</tr>
<tr>
<td>$s_P$</td>
<td>source of resting platelets</td>
<td>50</td>
<td>[10, 100]</td>
<td>estimated</td>
</tr>
<tr>
<td>$s_N$</td>
<td>source of resting neutrophils</td>
<td>25</td>
<td>[10, 75]</td>
<td>estimated</td>
</tr>
<tr>
<td>$s_{AM}$</td>
<td>source of adhesion molecules</td>
<td>30</td>
<td>[10, 50]</td>
<td>estimated</td>
</tr>
<tr>
<td>$s_M$</td>
<td>source of resting macrophages</td>
<td>30</td>
<td>[10, 50]</td>
<td>estimated</td>
</tr>
</tbody>
</table>

Table 1: Numerical change and analysis from steady state between VOC admission and VOC admission complicated by ACS.

<table>
<thead>
<tr>
<th>Lab</th>
<th>Sample size</th>
<th>Mean difference in ACS admission and steady state</th>
<th>Mean difference in VOC admission and steady state</th>
<th>Difference between ACS and VOC</th>
<th>95% confidence interval</th>
<th>p value</th>
<th>B-H p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WBC count (K cells/µL)</td>
<td>43</td>
<td>6.8</td>
<td>4.1</td>
<td>2.71 (±7.62)</td>
<td>[0.36, 5.05]</td>
<td>0.025</td>
<td>0.058*</td>
</tr>
<tr>
<td>ANC (K cells/µL)</td>
<td>39</td>
<td>6.0</td>
<td>3.7</td>
<td>2.27 (±6.77)</td>
<td>[0.77, 4.77]</td>
<td>0.043</td>
<td>0.058*</td>
</tr>
<tr>
<td>Platelet count (K cells/µL)</td>
<td>40</td>
<td>−108.0</td>
<td>−44.0</td>
<td>−63.90 (±164.4)</td>
<td>[−116.44, −11.31]</td>
<td>0.019</td>
<td>0.058*</td>
</tr>
<tr>
<td>AEC (K cells/µL)</td>
<td>37</td>
<td>−0.07</td>
<td>−0.16</td>
<td>0.89 (±0.42)</td>
<td>[−0.05, 0.23]</td>
<td>0.212</td>
<td>0.294</td>
</tr>
<tr>
<td>Hgb</td>
<td>41</td>
<td>−0.67</td>
<td>−0.35</td>
<td>−0.32 (±1.75)</td>
<td>[−0.87, 0.23]</td>
<td>0.252</td>
<td>0.294</td>
</tr>
<tr>
<td>ARC (K cells/µL)</td>
<td>34</td>
<td>0.06</td>
<td>0.05</td>
<td>0.006 (±0.15)</td>
<td>[−0.05, 0.06]</td>
<td>0.827</td>
<td>0.827</td>
</tr>
<tr>
<td>Oxygen saturation (%)</td>
<td>27</td>
<td>−2.20</td>
<td>−0.26</td>
<td>−1.96 (±3.73)</td>
<td>[−3.44, −0.49]</td>
<td>0.011</td>
<td>0.058*</td>
</tr>
</tbody>
</table>

*Significant by Benjamini-Hochberg false discovery rate.

Figure 36: Klouda et al 2020 Inflammatory Marker Data Set Summary [51]

The dataset includes $n = 45$ patients (age 2 - 26) with median age 12.4 years old and was compiled from 2 hospital record visits: routine hematology appointment and when experiencing a VOC due to acute chest syndrome (ACS). While there are only 2 time points for this data set, it includes blood cell
counts for individuals at each time for hemoglobin (Hgb), white blood cell count (WBC), platelet count, absolute neutrophil count (ANC), and red blood cell (RBC) count. These blood cell counts at baseline and after crisis are used to benchmark the model outcome variables for model fitting. The model will be considered ‘fit to the data’ when the simulated sVOC model transients meet or pass through the benchmarks for the patient data. The ability to compare the difference in values at steady state versus crisis may give insight into what pathways are triggered during inflammatory VOC.

Initial data analysis was conducted to understand trends or groups present in the dataset that could inform model formulation or parameter estimation. Multiple data visualization approaches were made to investigate data variable relationships. In the Figures 37, 38, 39 scatter-plots of variables against each other reveal relationships between different cell types. There is evidence of a positive relationship between RBC count and Hgb levels, WBC count and Platelets, and WBC count and Neutrophils. This aligns with what the literature about immune cells and their respective roles. The WBC and Platelet relationship, specifically, is an important relationship that can be exploited in model fitting and validation later.

Figure 37: Scatterplot of Baseline Hgb vs RBC count
Bar plots were then created to identify any trends in blood count percent change from baseline to crisis, based on hemoglobin F (Hgb F) % levels, fetal hemoglobin. In SCD, a higher Hgb F % is expected to be associated with decrease in pain frequency due to less polymerization. By plotting the percent
change based on this category to see if there is a significant impact through this measurement; since there are medications available to improve this number. Percent change is calculated by \( \left( \frac{\text{new value} - \text{old value}}{\text{old value}} \right) \times 100 \) and measures the relative magnitude of change between two values. By comparing across the Hgb categories, we can assess if any clusters or groups experienced significant changes. For the four blood count percent changes shown in Figures 40, 41, 42, and 59, there were no significant patterns that indicate the need for grouping or categorization since there were similar changes throughout the population. It should be noted that these plots specifically only show \( n = 23 \) out of 45 patients since the remainder did not have Hgb F % levels logged (=NA).

![WBC Bar Plot Percent Change](image)

Figure 40: WBC Percent Change vs Hgb F %
Figure 41: Platelet Percent Change vs Hgb F %

Figure 42: ANC Percent Change vs Hgb F %
It can be seen that across the populations, the percent change is consistently increasing or decreasing with minimal outliers. From this data analysis, we can see a clear trend in Platelet Count decreasing after crisis 41, WBC count increasing after crisis 40, and neutrophils increasing after crisis 42 for most of this population. In the following sections, we will use these insights to make informed parameter estimations and benchmark model validity on how well it matches the data.

4.4. Results

4.4.1. Parameter Set Collections: Healthy, Baseline VOC, VOC Crisis

In this and following sections, we establish and investigate the appropriate values for the initial conditions and parameter values that correspond with each patient state. There are different parameter sets for each of following patient states:

- Healthy := no Sickle Cell Disease

- Baseline VOC := sterile inflammatory state

- VOC crisis := chronic / exacerbated inflammatory state

A healthy subject is assumed to have minimal inflammation prior to stimulus, with very low levels of heme, and so has no immune cell influx. Therefore all of our immune cell source rate are reduced proportionate to the SCD steady state (based on the literature).
Figure 44: Healthy Patient - no SCD present (main outcome cells) Note: This patient has cell well below SCD baseline values as seen in Figure 47. The model transient does not reach a crisis state value either (the red line) since the sickling rate $p_s = 0$, the death rate of Sickle Cells is $\mu_s = 0$, and their source of platelets is minimal $s_P = 6$.
The healthy patient source rates are: activated endothelial cells $s_E = 36$, circulating platelets $s_P = 6$, resting neutrophils $s_N = 25$, local adhesion molecules $s_AM = 30$, and resting macrophages $s_M = 30$.

The initial conditions for the healthy patient is: $EA(0) = 10; ED(0) = 0; PA(0) = 2; NA(0) = 4.22; NT(0) = 2, AS(0) = 0; AM(0) = 1; NP(0) = 1; M1(0) = 1.5; M2(0) = 1$;

Since Sickle Cell patients have elevated inflammatory markers at steady state, we set their source variable rates as 2-3x quicker for immune cells and 10x more historical endothelial activation/damage: activated endothelial cells $s_E = 36$, circulating platelets $s_P = 50$, resting neutrophils $s_N = 25$, local adhesion molecules $s_AM = 30$, and resting macrophages $s_M = 30$.

The initial conditions for the VOC patient is determined by their baseline given in the blood count values from [51] $EA(0) = \text{(baseline platelets/3)}; ED(0) = .2 \times EA0; PA(0) = \text{baseline platelets}; NA(0) = \text{baseline ANC}; NT(0) = \text{...}$. 

Figure 45: Healthy Patient - all cells (individually) $p_x = 0 & \mu_x = 0 & s_P = 6$
$NA0/5; AS(0) = RBC \times 100 \times \mu_s; AM(0) = AS0; NP(0) = NA0/3; M1(0) = 0.04 \times$ baseline WBC; $M2(0) = 0.04 \times$ baseline WBC

To be biologically sound, the model must show a steady state of sterile inflammation associated with SCD. Sterile inflammation occurs when physical, chemical, or metabolic stimuli trigger inflammatory cells in response to stress where the stimuli persists / cannot be eliminated, resulting in a chronic inflammatory state. Given the cyclic response between sRBC adhesion, inflammation, and Ischemia-Reperfusion injury, we assume there is always an elevated presence of immune cells in SCD.

<table>
<thead>
<tr>
<th>Parameter Set Collections</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patient 37</strong></td>
</tr>
<tr>
<td><strong>Initial Conditions</strong></td>
</tr>
<tr>
<td>EA0</td>
</tr>
<tr>
<td>ED0</td>
</tr>
<tr>
<td>PA0</td>
</tr>
<tr>
<td>NA0</td>
</tr>
<tr>
<td>NT0</td>
</tr>
<tr>
<td>AS0</td>
</tr>
<tr>
<td>AM0</td>
</tr>
<tr>
<td>NP0</td>
</tr>
<tr>
<td>M10</td>
</tr>
<tr>
<td>M20</td>
</tr>
</tbody>
</table>

Primary cells indicative of this sterile inflammatory state are platelets, neutrophils, adhesion molecules, and activated endothelial cells. In the non-stressed state above, we can note increased endothelial damage and increased adhesion molecules present, indicative of an inflamed environment even without ac-

Figure 46: 3 Sample Cases of Parameter Set Collections for VOC Crisis as seen in Figure 62
tive crisis. Hebbel 1980 [38] found sRBC adherence to the endothelium to be positively correlated with the clinical severity score [55], so we associate VOC severity with the number of activated endothelial cells.

We anticipate this severity reaching an elevated baseline over long periods of time since we know SCD tends to become chronic pain in adulthood for about 50% of SCD patients. We reflected this change in the inflammatory cell concentrations during crisis by adjusting the parameters from 48 to meet the benchmark of the crisis data.
Figure 47: **SCD VOC Baseline to Crisis state** (main outcome cells) $p_s = .45, \mu_s = 0.05, sp = 50$  

**Note:** This patient has cell baselines set from the data as $Y0$ and the model transient reaches their resulting crisis state value (the red line) through parameter value changes. SCD patient parameters start with sickling rate $p_s = .45$, the death rate of Sickle Cells is $\mu_s = 0.05$, and their source of platelets is elevated from pain history $sp = 50$
We can see from the graph above that there are noticeable differences in the baseline average and the crisis average benchmarks. Section 4.7 will discuss the parameter adjustments made for this fitting.

4.4.2. Determining Predictors: Parameter Estimation + Ranking

The aim of this work is to understand the impact of inflammatory and immune pathways of SCD, represented by initial conditions and unique parameter sets, on disease pathology and VOC severity. Therefore we must start our simulation at initial conditions associated with a steady state individual so that we may see changing dynamics primarily from the presence and byproducts of Sickle cell crisis.

Given the large number of parameters $k = 48$ and variables $y = 10$, statistical techniques need to be used to analyze the system and generate a variety of outcomes and dynamics of the blood and immune cell populations in this model.

Given the lack of robust time data, we will be conducting global sensitivity analysis, as it is indepen-
dent of parameter estimates. We will be utilizing **Latin Hypercube Sampling**, a stratified Monte Carlo sampling approach, for estimating these parameter sensitivities using MATLAB functions adapted from Kirschner [61]. The Latin Hypercube Sampling (LHS) approach uses the analytical inversion method to generate a uniformly distributed random sample of $n$ values. This sampling divides the parameter space into equitable regions, without replacement, and produces an unbiased selection of parameter values. We chose this method over Monte Carlo sampling since it provides a more representative sample with less variability and less computational time needed.

Ranges for the parameters tested in this sampling can be seen above in the Parameter Table that resulted in a multiple system dynamics and outcomes. We generated $K = 250$ parameter sets from this sampling and plotted the resulting transients to accept or reject those parameters. A parameter set was rejected if the outcome values went negative or became too large, as that’s not biologically sound.

![Figure 49: sVOC Model Transient Using LHS matrix parameter set 6](image)
Figure 50: sVOC Model Transient Using LHS matrix parameter set 17

Figure 51: sVOC Model Transient Using LHS matrix parameter set 42
Based on the rejection criteria, Figures 49 and 51 showed parameter sets that were accepted, while Figures 50 and 52 are rejected. Truncating the parameter ranges based on these criteria allows for plotting over a better parameter range region and informed analysis. The resulting biologically feasible model outputs and ranges are then used to create PRCC plots of the relationships between each model outcome and all parameters to visualize the parameter’s influence on the system.

4.4.3. Understanding the Driving Dynamics

We conducted a correlation matrix or PRCC (partial rank correlation coefficient) analysis to measure the strength of the statistical relationship between each parameter and state variable, without the impact of the other parameters on the output. PRCC values range from -1 to 1, with the sign indicating negative or positive correlation, the magnitude indicating the sensitivity, and its p-value indicative of its significance. A variable is considered sensitive to a parameter if the $\text{PRCC} > 5$ and the p-value is less than .05. In this work, we will make decisions about parameters based on p-value significance.

We vary model parameters to estimate their impact on VOC severity since a small change in parameter value could determine convergence to either a “healing” or “unhealed/chronic” state. A parameter
whose change has a minimal impact on VOC severity is said to be "insensitive" and the ones whose change has a large impact on the model are said to be "sensitive". Identifying these sensitive parameters will reveal their measure of influence on the system and highlights significant factors that impact sVOC likelihood for SCD patients.

Figure 53: PRCC plot: $PA$ vs $sp$ | p-value = $8.084e-139$
Figure 54: PRCC plot: $PA$ vs $\mu_P$ | p-value $= 1.4942e - 315$

Figure 55: PRCC plot: $EA$ vs $b_{ead}$ | p-value $= 1.1328e - 194$
Figure 56: PRCC plot: $EA$ vs $s_p$ | $p$-value = $5.4917e-16$

Figure 57: PRCC plot: $AM$ vs $\mu AM$ | $p$-value = $1.1015e-128$
Though only 6 are shown here, 28 out of the 480 PRCC plots created revealed significant relationships. Interpreting these relationships gives some insight on what may drive the immune cell changes during VOC. Since all 28 correlations are significant and many highlight obvious relationships (e.g. platelet death rate negatively correlating to Platelet concentration), the plots are been omitted from this document.
These parameters can now be specifically targeted and modulated to change system output dynamics. The sVOC likelihood is determined by the aggregation of large cell concentrations in one area over time for platelets (PA), neutrophils (NA + NT), adhesion molecules (AM), and endothelial cells (EA + ED) compared to the crisis state cell averages. Determining which parameters are influential on these variable specifically allows us to make systematic changes to the system that will directly impact sVOC likelihood. We determine changes we need to see in the outcome during crisis and then modulate the correlated parameters based on their relationship (e.g. "since increasing $s_P$ also increases $PA$, I know to decrease that parameter value if my aim is to fit a decreasing PA benchmark")

4.4.4. Predictive Analysis of Interventions

The literature and model have shown that Sickle Cell Disease has a chronic pro-inflammatory state that is so self sustaining that it causes chronic pain over a lifetime. Therefore any inhibition or interruption of the molecular pathways for VOC onset can lead to a decrease in VOC severity and corresponding pain symptoms. To this end, we conduct predictive analysis for selected SCD therapies based on variable impact.
In the initial data analysis we investigated potential bins (or categories) of patients but they didn't give any useful patterns. The additional refinement of the mathematical model may suggest other data collection points that would be more useful.
Figure 60: **Patient 6** VOC Crisis System Behavior

Figure 61: **Patient 40** VOC Crisis State System Behavior
Each treatment targets specific processes and/or impacts state cell behavior. In order to match the changes seen between each state in part A of Figure ?? with the rise in platelets from 283 to 413, the activated endothelial cells rise from 94 to 137, and the increase in neutrophils from 3.15 to 11.2, specific parameters had to be adjusted. The following adjustments were made and inform the differences in parameter sets for steady state VOC and VOC crisis:

1. $\mu_p$ : decay rate of Platelets
2. $sp$ : source of resting platelets
3. $bead$ : transition rate from activated EC to damaged EC over time
4. $se$ : source of activated Endothelial cells
(5) $\mu_{am}$: natural cell death rate of Adhesion molecules

(6) $\mu_N$: decay rate of Neutrophils

(7) $s_N$: source of resting Neutrophils

Some patient variability exists in adjusting these parameters to account for different initial conditions of their baseline. To account for patient differences, all values from above could remain the same during crisis state except for:

- $\mu_p$: decay rate of Platelets - patient specific
  
  **Patient 6 value** = 0.09 | **Patient 40 value** = 0.18

- $s_P$: source of resting platelets - based on patient history of Platelet activation
  
  **Patient 6 value** = 50 | **Patient 40 value** = 30

- $b_{ead}$: transition rate from activated EC to damaged EC - patient specific
  
  **Patient 6 value** = 0.25 | **Patient 40 value** = 0.39

- $s_E$: source of activated endothelial cells - based on patient history of EC activation
  
  **Patient 6 value** = 44 | **Patient 40 value** = 25

Based on the influential parameters listed above, the following treatments for these parameter targets already exist:

(1) Hydroxyurea:= produces HbF to reduce sickling + adhesion)

  *will impact sickling rate $p_s$*

(2) Voxelotor:= decreases RBC polymerization

  *will impact sickling rate $p_s$*

(3) Crizanilizamub:= adhesion site inhibitor

  *will impact viability of adhesion molecules in the environment $\mu_{am}$*

(4) Aspirin:= anti-platelet medication

  *will impact platelet decay and resting source $\mu_p$ & $s_P$*
(5) Tocilizumab:=$\text{anti-cytokine IL-6}$

\textit{impacts endothelial dysfunction and activation source level $b_{ead}$ & $s_E$}

(6) NAC:=$\text{antioxidant}$

\textit{will impact endothelial dysfunction rate $b_{ead}$ \cite{9}}

A treatment plan that includes some combination of these medications is likely to reduce sVOC likelihood. Testing these treatment interventions and comparing to patient data at multiple time points would be a great project for future work.

\section*{4.5. Discussion}

\subsection*{4.5.1. Which inflammatory cell population(s) impacted sVOC likelihood the most? (statistically significant?)}

In this work we created a mechanistic ODE model to capture the most important aspects of the immune response to VOC and the occurrence of sVOC. The blood cell count data from Klouda et al 2020 \cite{51} allowed us to benchmark the start (baseline data subset) and ending (crisis data subset) values of these blood and immune cells and vary system parameters to meet these thresholds. In varying these parameters through sampling the parameter space with \textit{Latin Hypercube Sampling}, the influence (or correlation) of each parameter to these changes can be quantified.

The PRCC plots visualized these relationships and revealed the statistically significant parameters. The influential parameters highlight the impact of \textit{platelets, endothelial cells, and neutrophils} on the likelihood of prolonged inflammation due to VOC, known as \textit{sVOC}. The source rates and decay rates of these cells have a major impact on model outcome and patient state change.

The impact of the source rates of platelets, ECs, and neutrophils on the temporal dynamics may spotlight the importance of patient history with sVOC likelihood. The source rates represent the resting amount of these cells in the body and are specific to patient history. One could assume a patient with frequent or prolonged pain crises has built up a baseline level of inflammation in the body, resulting in elevated immune cells sitting in waiting.
The impact of the decay rates of the platelets, ECs, and neutrophils on the temporal dynamics could be potential target areas for SCD VOC treatment. Increasing the decay rate of these cells would require therapies that directly remove them or make them die at a faster rate.

4.5.2. Do any inflammatory marker relationships results hint towards treating pain preventatively instead of after?

Given the 4 parameters that must be adjusted based on patient history, lowering these values through treatment could potentially reduce sVOC likelihood. The decay rate $\mu_p$ and source rate $s_P$ of platelets can be impacted by anti-platelet medications. The rate of endothelial dysfunction $b_{ead}$ can be improved through antioxidant medications and other anti-inflammatory cytokine therapies. Reducing the source and activation of endothelial cells and it's transition rate to damage could make major differences on inflammatory pain.

4.5.3. Which treatment combo can be used for reducing pain?

The previous section listed medications that already exist that could impact the influential parameters and variables in this system. They include Hydroxyurea, Voxelotor, Crizanilizamub, and Asprin and could be used together or separately for maximum prophylactic VOC reduction. Future work for this model includes testing the impact of multiple combinations of treatments. The main cells that cause aggregates and initiate VOC are NA, PA, EA, and AM so impacting these directly could reduce inflammation in patients over time after treatment.

4.5.4. Model Limitations and Future Work

We have presented a system of ODEs that tracks the resulting immune response within the vessel during VOC for SCD patients. One limitation of this formulation is the lack of blood component options. If available, this model could be adjusted to include direct measurements of pro- and anti-inflammatory cytokines present in the patient over time. Another limitation of this model is the lack of predictive power due to data availability. Given that the data only includes 2 time points, it is hard to optimize for a longer change over time in order to predict specific changes in VOC likelihood over time.
The benefit, however, in using blood cell count data is its relative accessibility / frequency of collection in comparison with cytokine level information. Many patients have blood drawn at hematology visits and that hospital record data could be mined to create multiple blood cell values over time.

This model also allows for treatment intervention analysis in the future work. Investigating the influential pathways for treatment and the impact of on the market medications comes at minimal cost and injury to patients and manufacturers.
5. CONCLUSION

5.1. The Benefits of Dynamic mathematical techniques in SCD research

We have demonstrated the utility mathematical modeling to gain a deeper understanding of SCD. By quantifying known dynamics and phenomena, fitting models to experimental data, and conducting parallel assays and simulations, researchers have been able to elucidate significant biomarkers, processes, and clinical findings that at least partially contribute to SCD progression.

However, the field of SCD research is relatively limited. There is much to still be understood about the disease's dynamic behavior. Mathematical and computational efforts have progressed the field greatly, taking advantage of the new abundance of data on individuals with SCD. Like most research, however, these findings mark only the beginning of what can be understood.

The models investigating VOCs in SCD are still aiming to uncover biomarkers and other physiological mechanisms that primarily impact the vascular and other system dysfunctions that together cause VOCs. There are also still many questions about RBCs, as everything from their size and shape to their rigidity and location have major implications on SCD presentation. Li et al. 2017 demonstrated their interest in understanding the complex relationship between RBCs, morphological membrane distortion of RBCs, and deoxy HbS polymer chain formation [56]. By classifying RBCs according to their shape, Xu et al. 2017 have tried to build a gold standard RBC phenotype library within SCD [102]. Petrovic et al. 2020 have tried to optimize RBC classification by adding more features based on PCA and LDA for SCD diagnosis support [77]. Zheng et al. 2021 and Altrock et al. 2016 have suggested extending their models by incorporating diffusion of chemicals and/or drug pharmacokinetics to help understand therapeutic effects on different RBC shape phenotypes and blood flow [107] [2].

The common theme found in studies building cellular, genetic, behavioral, pain, and treatment SCD models is the need to increase model complexity. Parameters and variables have started with simplified models, in order to understand or explain the operation of small galaxies of a universe of SCD dynamics, and only at a basic level. Now, simple models must be either combined, expanded or layered properly, in a branching tree-like effect, to account for the impact of confounding variables and mathematical
degrees of freedom at many levels—the molecule, cell, the momentary VOC, the organ, and the biopsychosocial milieu. One hope is that mathematical modeling, among other benefits, will benefit the field by being the ideal platform for testing optimal dosage of medications and hypothesizing treatment effects in order to obviate conducting expensive and time-consuming randomized controlled trials.

Systems biology and mathematical modeling are having a resurgence with the rise in computational power along with an abundance of data. Computational modeling is going to play an important role in science to quantitatively demonstrate reality. In most cases, life and the world are not simple. SCD is not simple.

5.2. The Importance of the Sleep-Pain connection in SCD pain prediction

In this work, we present a mechanistic ODE model for predicting pediatric pain episodes. We find that self-reported sleep quality alone is not sufficient to predict pain onset. However, the analysis of the model results did reveal some interesting patterns. The sensitivity analysis revealed that 2 days accumulation of sleep quality to be more predictive than 1, 3, or 4 days accumulation and could inform data collection protocol for future clinical SCD research.

Consider our model limitations to avoid pitfalls in new data collection. One limitation of our model is that the amount of time the data was collected was discrete and over a short period of time (30 days), decreasing the amount of pain history available for prediction. Since pediatric SCD patients experience fewer episodes throughout the year, longer data collection windows may be useful for more accurate prediction.

Similarly, having only 88 patients didn't allow for much room for training and fitting the model based on the pain history. The larger the cohort size, the more valuable the findings. The sub population findings here are severely impacted by how few people are in each sub group because of the overall cohort size. We expect the results of these extensions to aid in informing future data collection for Sickle Cell studies.

This project has the potential to change clinical care by providing models that can assist pediatric SCD patients, their caregivers, and their health care providers. The model could also be extended to youth with other illnesses that cause pain. The long-term benefits of implementing mechanistic models
into patient care include reducing the suffering of youth with SCD and their families as well as the societal burden of the substantial health care costs of managing SCD pain.

5.3. Potential Mechanistic Pathways for SCD sVOC prevention

The sVOC model was calibrated to patient data and shows model validity through benchmarking patient averages. Analyzing this system has given insight to the processes that impact the transition from VOC steady state to VOC crisis. Parameter Estimation revealed the impact of the decay rates of the platelets, ECs, and neutrophils on the temporal dynamics could be potential target areas for SCD VOC treatment. Increasing the decay rate of these cells would require therapies that directly remove them or make them die at a faster rate.

The impact of the source rates of platelets, ECs, and neutrophils on the temporal dynamics may also spotlight the importance of patient history with sVOC likelihood. The source rates represent the resting amount of these cells in the body and are specific to patient history. One could assume a patient with frequent or prolonged pain crises has built up a baseline level of inflammation in the body, resulting in elevated immune cells sitting in waiting.

Given the 4 parameters that must be adjusted based on patient history, lowering these values through treatment could potentially reduce sVOC likelihood. The decay rate \( \mu_p \) and source rate \( s_p \) of platelets can be impacted by anti-platelet medications. The rate of endothelial dysfunction bead can be improved through antioxidant medications and other anti-inflammatory cytokine therapies. Reducing the source and activation of endothelial cells and it’s transition rate to damage could make major differences on inflammatory pain.

The benefit in using blood cell count data is it’s relative accessibility / frequency of collection in comparison with cytokine level information. Many patients have blood drawn at hematology visits and that hospital record data could be mined to create multiple blood cell values over time. This model also allows for treatment intervention analysis in the future work. Investigating the influential pathways for treatment and the impact of on the market medications comes at minimal cost and injury to patients.
5.4. Using Mathematical Techniques for Medicine

We live in the Information Age and should use the technological advances of the 21st century to finally capture more data on SCD patients. With wearable devices, mobile doctor apps, and the accessibility for healthcare at home, health data is more accessible than ever. With this rise in big data collection and computational power, mathematical tools can be used to examine the relationships between numerous variables within this disease system. Since pain is the most common symptom, I’ve concentrated my efforts on better understanding it’s origins. Until there is a more accessible cure for SCD than bone marrow transplant and gene therapy, any effort to mitigate or anticipate the pain for SCD patients could truly change lives. I hope to continue to use mathematical techniques to understand the pathology of (many) diseases and do my part to decrease the healthy disparity gap for all Black Americans and other racial/ethnic minorities.

6. Appendix

6.1. SCD Sleep-Pain Model Code:

```matlab
function run_scd_model_full_with_categories

% Compare fake patient parameters with fitted parameters
% If methods are working, they should be close

%% This script finds parameters to fit model to patient data
clear;
% seed random number generator
%rng(1); % comment out when running each sim for different init cond.
%having this on will give the same results each time

%set the directory where your Excel files are located
```
fileDirectory = 'C:\Users\macuser\Desktop\scd-pain-modeling\data\PatientFiles';

% Add directories to search path
addpath ../
addpath ../util
addpath ../plotting
addpath ../results
addpath ../../data/

% load patient data from mobile health app database
load('A-4_data.mat','drugs_taken','our_data','patient_ID')

% renaming data to use:
mydata = our_data;

% load modeling frameworks
load('model_names.mat')

% number of proposed models
num_models = length(model_names);

% select model to fit this round
model_index = 1;
% get the model name (e.g. 'threshold model')
modelname = model_names(model_index);
% number of patients

% num_patients = 1; length(patient_names_fake);
% number of hours we want to inspect (336 is two weeks)
total_time = mydata(end,1);
% time resolution for integration and plotting (0.5 is half hour)
time_res = 0.5;
% number of initial grid search locations for each patient
num_runs = 10;
% fixed k0 for patients (how fast patients rebound from reduced or increased pain, absent drugs)
% we take rebound half life to be half an hour
fixed_k0 = log(2)/8;
% indicator to merge all drug classes together (1 for merge, 0 for not)
merge_all_drugs = 0; common_drug_info = [log(0.5)/(-4.0),1];
% indicator of drug classes to include in fitting (LA,SA,NO)
% drugs_to_include = [0 0 0];
drugs_to_include = drugs_taken;

% threshold for drug dose times to be included in model (default is 1)
drug_threshold = 1;
% indicator that pain reporting probability is skewed towards higher pain
report_skew = 0;
% tolerance on how long sleep should impact pain
sleepqual_tol = 18;
% $[P_0, u, \epsilon, k_0, k_1, k_2, k_3, \beta]$%
% $\text{params_to_fit} = [2, 4, 8];$ % set this up to auto work with
% $\text{drugs_to_include}$

% stop optimizing $k_0$: only optimize drug rates and $\beta$
if drugs_to_include(1,1) == 1 & drugs_to_include(1,2) == 0 &
   drugs_to_include(1,3) == 0 % 1 0 0
   params_to_fit = [2, 4, 5, 8];
elseif drugs_to_include(1,1) == 1 & drugs_to_include(1,2) == 1 &
   drugs_to_include(1,3) == 0 % 1 1 0
   params_to_fit = [2, 4, 5, 6, 8];
elseif drugs_to_include(1,1) == 1 & drugs_to_include(1,2) == 0 &
   drugs_to_include(1,3) == 1 % 1 0 1
   params_to_fit = [2, 4, 5, 7, 8];
elseif drugs_to_include(1,1) == 0 & drugs_to_include(1,2) == 0 &
   drugs_to_include(1,3) == 1 % 0 0 1
   params_to_fit = [2, 4, 7, 8];
elseif drugs_to_include(1,1) == 1 & drugs_to_include(1,2) == 1 &
   drugs_to_include(1,3) == 1 % 1 1 1
   params_to_fit = [2, 4, 5, 6, 7, 8];
else % 0 0 0
   params_to_fit = [2, 4, 8]; % idea: maybe optimize $k_0$ when there's
   % no drugs?? not doing rn tho
end
% TODO: Test different drugs to include

% run through each patient's data

tic

for jj=1:num_patients

  % print patient number so we know where we're at
  % jj
  % Long acting (t_halflife = 10 hr); Short acting (t_halflife = 4 hr); Non-opioid (t_halflife = 4 hr)
  drug_info = [log(0.5)/(-10.0),1; log(0.5)/(-4.0),1; log(0.5)/(-4.0),1];
  % load patient data into convenient form (just useful info on specified time)
  filename = patient_ID % patient_names_fake{jj};
  % load patient data - use [1:2,4:end] for noisy data and [1,3:end] for deterministic data
  % patient_data = reshape(patient_data_fake(jj,:),[1,3:end]), length(patient_data_fake(1,:),1), 5);
  patient_data = [mydata(:,1) mydata(:,2) mydata(:,3:end)];
  % patient_data = [mydata(:,1) 0.1*mydata(:,2) mydata(:,3:end)]

  % indicator time vectors of drugs taken (1 for drug taken, 0 for not)
la = patient_data(:,4); sa = patient_data(:,5); no = patient_data(:,6);

% drug time vectors for each class of drug (empty vector if no drugs taken)
la_drugtimes = patient_data(la>0,1); sa_drugtimes = patient_data(sa>0,1); no_drugtimes = patient_data(no>0,1);

% indicator time vectors for nights of bad sleep
%sq = patient_data(:,3);
%bad_sleep_times = patient_data(sq<70,1);

unmitigated_pain = 1e-10;

% indicator time vectors for nights of bad sleep
tempsq = patient_data(:,3); % NOTE: doesn't start at NA for each patient
nonnan_idx = find(~isnan(tempsq)); % couldn't figure out how to skip the NAs in the vector

%ONE DAY WORTH OF SLEEP
%using regular sleep quality data
%bad_sleep_times = patient_data(tempsq<70,1);

%Compute average of of current and PREVIOUS entries ([0 1]),
%don't force it when you don't have enough data and put NAs ('Endpoints','fill'),
% TWO DAY MOVING AVERAGE

doody_avg_sq = NaN*ones(size(tempsq));
doody_avg_sq(nonnan_idx) = movmean(tempsq(nonnan_idx),[1 0],'
  Endpoints','fill'); % padding the ends of the input with NaN.
%bad_sleep_times = patient_data(doody_avg_sq<80,1);

%THREE DAY MOVING AVERAGE

% Compute average of current and last two entries ([0 2]),
 tresday_avg_sq = NaN*ones(size(tempsq));
tresday_avg_sq(nonnan_idx) = movmean(tempsq(nonnan_idx),[2 0],'
  Endpoints','fill'); % padding the ends of the input with NaN.
bad_sleep_times = patient_data(tresday_avg_sq<70,1);

%FOUR DAY MOVING AVERAGE

% Compute average of current and last two entries ([0 3]),
 fourday_avg_sq = NaN*ones(size(tempsq));
fourday_avg_sq(nonnan_idx) = movmean(tempsq(nonnan_idx),[3 0],'
  Endpoints','fill'); % padding the ends of the input with NaN.
%bad_sleep_times = patient_data(fourday_avg_sq<80,1);

% collect and reorder drugs taken to eliminate those not taken;
  drug_info must also change to reflect reordering of drugs
[dose_times1, dose_times2, dose_times3, drug_info] = get_dose_times
(drug_info, la_drugtimes, sa_drugtimes, no_drugtimes, 
merge_all_drugs, drugs_to_include, common_drug_info, 
drug_threshold);

% time vector (add an extra hour to end for interp1 - it freaks 
out about rounding errors)
tvec = (patient_data(1,1):time_res:(patient_data(end,1)+1))';
% pain report contains time and pain info from patient (remove 
times when patient does not report actual pain)
pain_report_real = patient_data(:,1:2)'; pain_report_real(:, 
isnan(pain_report_real(2,:))) = [];
% sleep report contains time and pain info from patient (remove 
times when patient does not report actual pain)
sleep_report_real = patient_data(:,[1,3])'; sleep_report_real(:, 
isnan(sleep_report_real(2,:))) = [];

% total number of drug classes taken (between 0 and 3; 0 or 1 if 
merged)
num_drugs_taken = (~isempty(dose_times1)) + (~isempty(dose_times2) 
) + (~isempty(dose_times3));

% use pseudo grid search to find best fit parameters for patient
% [params, err] = find_patient_params_sleep_fixed_k0_via_fmincon(
% tvec, dose_times1, dose_times2, dose_times3, pain_report_real, 
% patient_data, drug_info, fixed_k0, num_drugs_taken, num_runs, 
% report_skew, unmitigated_pain, bad_sleep_times, 
% sleep_report_real, sleepqual_tol);
[params, err] = find_patient_params_via_fmincon(tvec, dose_times1, 
    dose_times2, dose_times3, pain_report_real, patient_data, 
    drug_info, fixed_k0, drugs_to_include, unmitigated_pain, 
    bad_sleep_times, sleep_report_real, sleepqual_tol, params_to_fit);

% best fit patient info
patient_info = [patient_data(1, 2), params(1), params(2), fixed_k0, 
    params(3: length(params) - 1)']'; % cuts off that unnecessary, repeated residual
patient_info = [patient_data(1, 2), params(1:length(params) - 1)'];
    % cuts off that unnecessary, repeated residual
% generate patient report (both real and simulated pain
report = run_painsim_drugs3_ode_sleep(tvec, patient_info, 
    drug_info, dose_times1, dose_times2, dose_times3, 
    pain_report_real, bad_sleep_times, sleep_report_real, 
    sleepqual_tol);

% plot results of model fit -> comment out if no plot desired
P = plot_model_fit_sleep(tvec, params, err, patient_data, drug_info, 
    dose_times1, dose_times2, dose_times3, pain_report_real, fixed_k0, 
    bad_sleep_times, sleep_report_real, params(end), sleepqual_tol, 
    patient_ID);

% categorize pain data and pain simulation
pain_categ_function(tvec, patient_data, mydata, P, pain_report_real, 
    sleep_report_real, sleepqual_tol, patient_ID)
% cat_result = pain_categ_function(tvec, patient_data, mydata, P, pain_report_real, sleep_report_real, sleepqual_tol, patient_ID)

end

toc
patient_ID
params
err

assignin('base','P',P);
end

function [params, err] = find_patient_params_via_fmincon(tvec, dose_times1, dose_times2, dose_times3, pain_report_real, patient_data, drug_info, fixed_k0, drugs_to_include, uhat, bad_sleep_times, sleep_report_real, sleepqual_tol, params_to_fit)

% This function uses fmincon to find best fit of model to patient data.
% Parameters to vary are specified via "params_to_fit"

% matrix to hold predicted patient parameters for each search
% patient_info is always the length 8 for now. Will be reduced down in err_est2 before going to run_sim...
\[ \text{patient_info} = [\text{patient_data}(1,2), \text{uhat}, 0, \text{fixed_k0}, \text{abs}((0.75*\text{randn}(1, \text{length} (\text{drugs_to_include} )) + 0.25)) .* \text{drugs_to_include}, 0.23]; \]
\[ \text{LEANING: init vector from function "find_patient_params_sleep"} \]
\[ \% \text{patient_info} = [\text{patient_data}(1,2), \text{uhat}, 0, \text{fixed_k0}, (0.25*\text{randn}(1, \text{length} (\text{drugs_to_include} )) + 0.75)) .* \text{drugs_to_include}, 0.23]; \% \text{LEANING} : \text{init vector from function "find_patient_params_sleep"} \]
\[ \% \% \text{LEANING: first lines from function "find_patient_parameters_fixk0_sleep", to rename inputs / REDUNDANT} \]
\[ \% \text{patient information} \%
\]
\[ P0 = \text{patient_info}(1); \% \text{initial pain level (first pain report)} \]
\[ u = \text{patient_info}(2); \% \text{unmitigated pain level (initial guess)} \]
\[ \text{eps} = \text{patient_info}(3); \% \text{noise in unmitigated pain level (initial guess - not used)} \]
\[ k0 = \text{patient_info}(4); \% \text{relaxation rate (fixed)} \]
\[ k = \text{patient_info}(5:end-1); \% \text{marginal affect of drug1 (initial guess)} \]
\[ \text{beta} = \text{patient_info}(\text{end}); \% \text{strength of pain/sleep interaction} \]
\[ \% \text{LEANING: use fmincon instead of fminsearch, lines from function "find_patient_parameters_fixk0_sleep"} \]
\[ \% \text{func_handle} = @(xhat)(\text{est_err}(xhat,tvec,P0,k0,\text{drug_info}, \text{dose_times1}, \text{dose_times2}, \text{dose_times3}, \text{pain_report_real}, \text{bad_sleep_times}, \text{sleep_report_real}, \text{sleepqual_tol})); \]
\[ \text{func_handle} = @(xhat)(\text{est_err2}(xhat,\text{patient_info},tvec,\text{drug_info}, \text{dose_times1}, \text{dose_times2}, \text{dose_times3}, \text{drugs_to_include}, \text{pain_report_real}, \text{bad_sleep_times}, \text{sleep_report_real}, \text{sleepqual_tol}), \]
params_to_fit));

xhat_init = patient_info(params_to_fit);  % variable number of drugs
could be an issue / maybe put a 0 in for patient info if that
drug isn't present

[xhat,"] = fminsearch(func_handle, [u,k,beta], optimset('TolX'
','0.05,'TolFun','0.05));

% opts = optimoptions('fmincon','Display','iter-detailed','TolX'
',0.005,'TolFun',0.005);

opts = optimoptions('fmincon','Display','iter-detailed','TolX',1e
-10,'TolFun',1e-10,'MaxIterations',250);

% opts = optimoptions(@lsqnonlin,'Display','iter-detailed','Algorithm'
','levenberg-marquardt','TolX',1e-10,'TolFun',1e-10,''
MaxIterations',50);

fprintf('FMINCON activate!!!

')

[xhat, obj_val] = fmincon(func_handle, xhat_init
 ,[],[],[],[],[],[],[],opts);

disp(num2str(obj_val))

xhat

% create params from the initial values and the fit values
params = patient_info;

params(params_to_fit) = xhat;  % update with fit parameters
if any(drugs_to_include==0)  % are any drugs not included
    remove_k_idx = find(~drugs_to_include)+4;  % drugs are in spaces
      5,6, and 7
params(remove_k_idx) = []; % gets rid of placeholder for missing
drugs. continues as previously done now.

end

uhat = xhat(1); khat = xhat(2:end-1); betahat = xhat(end);
uhat = params(2);
k0hat = params(4);

% LEANING: last lines from function 
find_patient_parameters_fixk0_sleep
% find "best" noise strength epshat using Fokker-Planck PDE
% use report stdev to compute eps assuming equilibrium stdev; will
% always overestimate eps (but method is fast)
res_std = est_err2(xhat,patient_info,tvec,drug_info,dose_times1,
dose_times2,dose_times3,drugs_to_include,pain_report_real,
bad_sleep_times,sleep_report_real,sleepqual_tol,params_to_fit);
epshat = res_std*sqrt(2/(uhat^2*k0hat));

params(1) = []; % get rid of P0 from patient info before passing out
  , easiest way to reuse params_to_fit
params = [params, res_std]';
%params = [uhat, epshat, khat, betahat, res_std]'; %LEANING: output of
  function "find_patient_parameters_fixk0_sleep"
err = res_std;

% Create data for the 3D surface plot
X = k0hat;
Y = xhat(end);
Z = err;
% Create a 3D plot
figure;
scatter3(X, Y, Z);
xlabel('k0');
ylabel('beta');
zlabel('error');
%title('3D Surface Plot');
end

function err = est_err2(xhat, all_parameters, tvec, drug_info, dose_times1, dose_times2, dose_times3, drugs_to_include, pain_report_real, bad_sleep_times, sleep_report_real, sleepqual_tol, params_to_fit)
% if any parameters are less than zero, impose large error
if sum(xhat<0)>0
  err = 1000.0;
else
  % subset the all_parameters vector and create patient info
  patient_info = all_parameters; % always length 8 to start
  patient_info(params_to_fit) = xhat; % puts the parameters being varied in their corresponding spot
  if any(drugs_to_include==0) % are any drugs not included
    remove_k_idx = find(~drugs_to_include)+4; % drugs are in spaces 5,6, and 7
    patient_info(remove_k_idx) = []; % gets rid of placeholder for missing drugs. continues as previously done now.
end
end

% could adjust the patient info here --- if there's a 0 in drugs to
% include place a [] in it before passing it in

% set patient info according to inputs
patient_info = [P0,xhat(1),0,k0,xhat(2:end)];

% run simulation with provided parameters
report = run_painsim_drugs3_ode_sleep(tvec,patient_info,
drug_info,dose_times1,dose_times2,dose_times3,
pain_report_real,bad_sleep_times,sleep_report_real,
sleepqual_tol);
report_times = report(1,:);
real_pain = report(2,:);
sim_pain = report(3,:);

%insert pain thresholds here for real and simpain

%could add in scaling here
r_scl = max(real_pain);
s_scl = max(sim_pain);
scale = r_scl / s_scl;
sim_pain_new = scale*sim_pain;

%update err with new scaled sim pain
% calculate pain prediction error (42.58,41.42,38.68))
err = sqrt(1/(length(report_times)-1)*sum((real_pain-sim_pain).^2));
296  ERROR COMES OUT THE SAME (41.70, 39.6, 44.9) BUT P IS MUCH WORSE
297  err = sqrt(1/(length(report_times)-1)*sum((real_pain-
298   sim_pain_new).^2));
299  end
300  end
301
302
303  function report = run_painsim_drugs3_ode_sleep(tvec, patient_info,
304     drug_info, dose_times1, dose_times2, dose_times3, pain_report_real,
305     bad_sleep_times, sleep_report_real, sleepqual_tol)
306     % read in real patient data
307     report_times_real = pain_report_real(1,:);
308     report_pain_real = pain_report_real(2,:);
309
310     % run pain simulation with three drug interventions
311     % [tvec,P,D] = pain_sim_interventions3(tvec, patient_info, drug_info,
312         dose_times1, dose_times2, dose_times3);
313     [tvec,P,D] = pain_sim_interventions3_sleep(tvec, patient_info,
314         drug_info, dose_times1, dose_times2, dose_times3, bad_sleep_times,
315         sleep_report_real, sleepqual_tol);
316
317     % report pain levels
318     report_pain_sim = interp1(tvec, P, report_times_real);
319     report = [report_times_real; report_pain_real; report_pain_sim];
320  end
function [tvec,P,D] = pain_sim_interventions3_sleep(tvec, patient_info, drug_info, dose_times1, dose_times2, dose_times3, bad_sleep_times, sleep_report_real, sleepqual_tol)

% This function runs the continuous-time part of the simulation of pain versus time.

% Inputs:
% tvec: time points at which to evaluate variables
% patient_info: vector containing all patient information
% drug_info: matrix containing all drug information
% dose_times: vectors of all three drug dose times

% Outputs:
% tvec: time vector at which variables have been evaluated
% P: Estimated current level of pain
% D: Estimated drug level in body

% patient information
P0 = patient_info(1); % initial pain level
u = patient_info(2); % unmitigated pain level
eps = patient_info(3); % noise in unmitigated pain level
k0 = patient_info(4); % relaxation rate
k = patient_info(5:end-1);
beta = patient_info(end);

% k1 = patient_info(5); % marginal affect of drug1
% k2 = patient_info(6); % marginal affect of drug2
% k3 = patient_info(7); % marginal affect of drug3

kD = drug_info(:,1);
a = drug_info(:,2); % drugs_to_include

% % extreme (long-acting) drug information
% kD1 = drug_info(1,1);
% a1 = drug_info(1,2);

% % severe (fast-acting) drug information
% kD2 = drug_info(2,1);
% a2 = drug_info(2,2);

% % mild (non-opioid) drug information
% kD3 = drug_info(3,1);
% a3 = drug_info(3,2);

% Define the function to numerically integrate, incorporating parameters
% (assuming eps = 0)
% TODO: use stochastic model when eps > 0
func_handle = @(t,Y)( ODEmodel(t,Y,k0,u,k1,k2,k3,kD1,kD2,kD3,a1,a2,
a3,dose_times1,dose_times2,dose_times3));

func_handle = @(t,Y)( ODEmodel(t,Y,k0,u,k,kD,a,dose_times1,
dose_times2,dose_times3,bad_sleep_times,sleep_report_real,beta,
sleepqual_tol));

% Set IC: Set initial pain and drug levels
% Y0 = [P0;0;0;0];
Y0 = [P0;zeros(length(kD),1)];

% Set accuracy of numerical integration as well as max step size (note
% that max step is important to capture delta function events)
% options = odeset('RelTol',1e-4, 'AbsTol',1e-4, 'MaxStep', 0.5);
options = odeset('RelTol',1e-8, 'AbsTol',1e-8, 'MaxStep', 0.1);
% Do numerical integration
[tvec, Y] = ode45(func_handle, tvec, Y0, options);
[tvec, Y] = ode45(func_handle, tvec, Y0, options);

% Pain and drug levels over time
P = Y(:,1); D1 = Y(:,2); D2 = Y(:,3); D3 = Y(:,4);
P = Y(:,1); D = Y(:,2:end);
size(P)
end

% ****************************************

% Right hand side of function to integrate
% ****************************************

function dY = ODEmodel(t,Y,k0,u,k,kD,a,dose_times1,dose_times2,
dose_times3,bad_sleep_times,sleep_report_real,beta,sleepqual_tol)

temp_sleep_times = bad_sleep_times;

% 'ODEmodel' This function implements the differential equations
% k0, rate of decrease of pain without drugs or disease
% unmitigated pain, equilibrium for patient in absence of drugs
% dY = zeros(4,1); % delete after test
dY = 0.0.*Y;
% Note that rate constants could vary from patient to patient
k_tot = k0 + k1*Y(2) + k2*Y(3) + k3*Y(4); % delete after test
\[ k_{\text{tot}} = k_0 + \text{dot}(k, Y(2::\text{end})) \]

% size of \( Y \) is determined by \( kD \)

\[ k_{\text{tot}} = k_0 + \text{dot}(k, Y(2::\text{end})) \]

% find index associated with \( t \)

\[ \text{index}_t = \text{find}(\text{temp\_sleep\_times}-t >=0,1) ; \]

% passing in now / 6-28-21

% \text{sleepqual\_tol}=18; % testing out a wider tolerance so that the pulse doesn't resolve as quick

\[ \text{if} \ \sim \text{isempty}(\text{index}_t) \ \&\& \ (\text{abs}(\text{temp\_sleep\_times}(\text{index}_t)-t)<\text{sleepqual\_tol}) \]
\[ \text{if} \ t>\text{temp\_sleep\_times}(\text{index}_t)+\text{sleepqual\_tol} \]
\[ \text{temp\_sleep\_times}(\text{index}_t)= \text{inf} ; \]
\[ \text{RHS\_sleep}=\beta(100-\text{sleep\_report\_real}(2,\text{index}_t))*(\text{abs}(\text{temp\_sleep\_times}(\text{index}_t)-t)/\text{sleepqual\_tol}) ; \]
\[ \text{else} \]
\[ \text{RHS\_sleep}=0 ; \]
\[ \text{end} \]

% \text{RHS\_sleep} = \beta(100-\text{sleep\_report\_real}(2,\text{index}_t))*\text{dosefunc}(t, \text{bad\_sleep\_times}, [], [], 1) ;

\[ dY(1) = -k_{\text{tot}}Y(1) + \text{RHS\_sleep} ; \]
%dY(1) = -k_tot*Y(1) + beta*(100 - sleep_report_real(2, index_t))*
sleepfunc(t, temp_sleep_times); % need to match up the sleep
report index with current bad sleep time
% dY(2) = -kD1*Y(2) + a1*dosefunc(t, dose_times1); % delete after
test
% dY(3) = -kD2*Y(3) + a2*dosefunc(t, dose_times2); % delete after
test
% dY(4) = -kD3*Y(4) + a3*dosefunc(t, dose_times3); % delete after
test
if length(kD)>0
    dY(2:end) = -kD.*Y(2:end) + a.* dosefunc(t, dose_times1, 
                   dose_times2, dose_times3, length(Y)-1);
end
end % function ODEmodel

function P = plot_model_fit_sleep(tvec, params, err, patient_data, 
    drug_info, dose_times1, dose_times2, dose_times3, pain_report_real, 
    fixed_k0, bad_sleep_times, sleep_report_real, beta, sleepqual_tol, 
    patient_ID)
% This function plots the raw data with best fit for the model
% run model using provided parameters (patient_info is P0, u, eps, k0, 
    k_drugs)
[tvec,P,D] = pain_sim_interventions3_sleep(tvec,[patient_data(1,2),
   ...
   params(1:length(params)-1)',...
   drug_info,dose_times1,dose_times2,dose_times3,bad_sleep_times,
   sleep_report_real,sleepqual_tol);

% [tvec,P,D] = pain_sim_interventions3_sleep(tvec,[patient_data(1,2)
   %,...
%   params(1),params(2),fixed_k0, params(3:length(params)-1)'],...
%   drug_info,dose_times1,dose_times2,dose_times3,bad_sleep_times,
%   sleep_report_real,sleepqual_tol);
report_pain_real = pain_report_real(2,:);
report_times_real = pain_report_real(1,:);

% place figure in nice spot on screen, and fix ratio
figure('rend','painters','pos',[10 10 800 800])
subplot(3,1,[1 2]); %original
subplot(3,1,1)

% plot model prediction with shaded error regions
%shadedErrorBar(tvec,P, err*ones(length(P),1), 'r');
plot(tvec,P,'r');
ylim([0,100]);

% reorder layers so that error shading doesn't cover up axes
set(gca,'Layer','top')
hold on

% plot real pain report data
plot(report_times_real, report_pain_real, 'ok')
set(gca,'Xticklabel',[])
% to just get rid of the numbers but leave the ticks
ylabel('Pain level (0 to 100)');
title( patient_ID + " Pain Simulation Graph")

% number of drug classes taken
num_drugs = length(D(1,:));

% create list of colors to cycle through (for drugs)
color_ind = {'r','b--','g-.'};

% subplot(3,1,3); original
subplot(3,1,2)
% for each drug class taken, plot drug in bloodstream
for kk = 1: num_drugs
    plot(tvec, D(:,kk), color_ind{kk});
    hold on
end
ylabel('Std. drug doses in bloodstream');
subplot(3,1,3)
plot(sleep_report_real(1,:), sleep_report_real(2,:), 'c-*');
% Add text to the plot
% text(1, 1, ['Patient Params: ' params], 'Color', 'red');
ylabel('Reported sleep quality');
xlabel('Time (hours)');

T = table(params);

% Convert Table to cell to char array
tableCell = table2cell(T);
tableCell(cellfun(@isnumeric,tableCell)) = cellfun(@num2str, tableCell(cellfun(@isnumeric,tableCell)),'UniformOutput',false);
tableChar = splitapply(@strjoin, pad(tableCell),[1]);

% Add axes (not visible) & text (use a fixed width font)
% axes('position',[.1,.1,.8,.2], 'Visible','off')
% text(.2,.95,tableChar,'VerticalAlignment','Bottom','
% HorizontalAlignment','Left','FontName','Consolas');

axes('position',[0.65 0.65 0.28 0.28], 'Visible','off')

text(-.75,.1,tableChar,'VerticalAlignment','Bottom','
HorizontalAlignment','Left','FontName','Consolas');

end

function [params,err] = 
find_patient_params_sleep_fixed_k0_via_fmincon(tvec,dose_times1,
dose_times2,dose_times3,pain_report_real,patient_data,drug_info,
fixed_k0,num_drugs_taken,num_runs,report_skew,uhat,
bad_sleep_times,sleep_report_real,sleepqual_tol)

% This function uses pseudo grid search to find best fit of model to
% patient data

%probably have to pass what's fixed and what's not to the function
handle

% pass in a general parameter vector / patient info to function
handle

% along with index vector?
for initial conditions for fmincon, do patient info of index

matrix to hold predicted patient parameters for each search

PARAMS contains [u,eps,ki,err]
predicted_params = zeros(3+num_drugs_taken+1,num_runs);

patient_info = [patient_data(1,2),uhat,0,fixed_k0,0.25*randn(1,
    num_drugs_taken)+0.75,0.23]; %LEANING: init vector from function
    "find_patient_params_sleep"

% % LEANING: first lines from function 
    "find_patient_parameters_fixk0_sleep", to rename inputs

% patient information % REDUNDANT FOR NOW JUST TO GET IT WORKING QUICK
P0 = patient_info(1); % initial pain level (first pain report)
u = patient_info(2); % unmitigated pain level (initial guess)
eps = patient_info(3); % noise in unmitigated pain level (initial guess - not used)
k0 = patient_info(4); % relaxation rate (fixed)
k = patient_info(5:end-1); % marginal affect of drug1 (initial guess)
beta = patient_info(end); % strength of pain/sleep interaction

% LEANING: use fmincon instead of fminsearch, lines from function "
    find_patient_parameters_fixk0_sleep"
func_handle = @(xhat)(est_err(xhat,tvec,P0,k0,drug_info,dose_times1,
    dose_times2,dose_times3,pain_report_real,bad_sleep_times,
    sleep_report_real,sleeppqual_tol));
\[
[x_{\text{hat}}, \ldots] = \text{fminsearch}(\text{func\_handle}, [u, k, \beta], \text{optimset}('\text{TolX}', 0.05, '\text{TolFun}', 0.05));
\]

\[
\text{opts} = \text{optimoptions}('\text{fmincon}', '\text{Display}', '\text{iter\_detailed}', '\text{TolX}', 0.05, '\text{TolFun}', 0.05);
\]

\[
\text{fprintf}(''\text{FMINCON activate!!!} \n \n'');
\]

\[
[x_{\text{hat}}, \ldots] = \text{fmincon}(\text{func\_handle}, [u, k, \beta], [], [], [], [], [], [], \text{opts});
\]

\[
u_{\text{hat}} = x_{\text{hat}}(1); k_{\text{hat}} = x_{\text{hat}}(2:\text{end-1}); \beta_{\text{hat}} = x_{\text{hat}}(\text{end});
\]

\[
\% \text{the output of x_{\text{hat}} will go into particular spots based on indices}
\]

\[
\% \text{LEANING: last lines from function }''\text{find\_patient\_parameters\_fixk0\_sleep}''
\]

\[
\% \text{find } ''\text{best}'' \text{ noise strength epshat using Fokker-Planck PDE}
\]

\[
\% \text{use report stdev to compute eps assuming equilibrium stdev; will}
\]

\[
\% \text{always overestimate eps (but method is fast)}
\]

\[
\text{res\_std} = \text{est\_err}([u_{\text{hat}}, k_{\text{hat}}, \beta_{\text{hat}}], \text{tvec}, P_0, k_0, \text{drug\_info},
\]

\[
\text{dose\_times1, dose\_times2, dose\_times3, pain\_report\_real,}
\]

\[
\text{bad\_sleep\_times, sleep\_report\_real, sleepqual\_tol});
\]

\[
\text{epshat} = \text{res\_std} \cdot \sqrt{2/(u_{\text{hat}}^2 \cdot k_0))};
\]

\[
\text{params} = [u_{\text{hat}}, \text{epshat}, k_{\text{hat}}, \beta_{\text{hat}}, \text{res\_std}]; \% \text{LEANING: output of function } ''\text{find\_patient\_parameters\_fixk0\_sleep}''
\]

\[
\text{err} = \text{res\_std};
\]

\[
\text{end}
\]

\[
\% \text{pain category function}
\]
function pain_categ = pain_categ_function(tvec, patient_data, mydata, P, pain_report_real, sleep_report_real, sleepqual_tol, patient_ID)

% CATEGORIZES PAIN DATA AND PAIN SIMULATION INTO HI/LO CATEGORIES

% renaming data to use:
mydata = patient_data;

%CATEGORIZING REAL DATA : from patient file
painsev = mydata(:,2);
% time vector
tvec = (mydata(:,1));

% sleep report contains time and pain info from patient (remove times when patient does not report actual pain)
sleep_report_real = mydata(:,[1,3])'; sleep_report_real(:,isnan(sleep_report_real(2,:))) = [];

% patient pain threshold
painthresh = 20;

% finding lo hi pain
lowpain_idx = (painsev < painthresh);
hipain_idx = (painsev >= painthresh);
% categorizes the positions as low pain
painsev(lowpain_idx) = 0;
painsev(hipain_idx) = 1;

% CATEGORIZING SIMULATED DATA: our P from the model

% renaming data to use:
painsevsim = 10.*P;

% it performs horribly without the times 10!!!! what does that do exactly
% tho?

% time vector
tlen = length(P);
tvec = (mydata(:,1));

% patient pain threshold
painthresh2 = 20;

% finding lo hi pain
lowpain_idx = (painsevsim < painthresh2);
hipain_idx = (painsevsim >= painthresh2);
% categorizes the positions as low pain
painsevsim(lowpain_idx) = 0;
painsevsim(hipain_idx) = 1;

% new time vector
tpoints = [0:tvec(end)/(length(P)-1):tvec(end)];

% OUTPUTS NEW CATGORY PLOT, CONFUSION MATRIX, AND TRUE VALUES

% plotting
% place figure in nice spot on screen, and fix ratio
figure('rend','painters','pos',[15 15 900 900])
% plot categorized pain report data
subplot(2,1,1)
plot(tpoints, painsevsim, 'ok-')
hold on
plot(tvec, painsev, 'diamondr', 'MarkerFaceColor','r');
set(gca,'Xticklabel') % to just get rid of the numbers but leave the ticks
set(gca,'Yticklabel')
title(patient_ID + " Outside Category Optimizing Pain Plot" );
ylim([0,1])
xlabel('TimeSteps (by hour)');
ylabel('Pain level category (LO/HI)');
subplot(2,1,2)
plot(sleep_report_real(1,:), sleep_report_real(2,:),'c-');
ylabel('Reported sleep quality');
xlabel('Time (hours)');

Psim2 = painsevsim(1:(length(painsevsim)/length(painsev)):end,1:(
    length(painsevsim)/length(painsev)):end);

% place figure in nice spot on screen, and fix ratio
figure('rend','painters','pos',[15 15 900 900])

    % plot confusion chart
    cm = confusionchart(painsev,Psim2);
    cm.RowSummary = 'row-normalized';
    cm.ColumnSummary = 'column-normalized';
    cm.Title = (patient_ID + " HI/LO Category Confusion Chart");

% sensitivity analysis
    TP=0; FP=0; TN=0; FN=0;
    raw_pain_data = transpose(painsev);
    for i=1:length(Psim2)
if(raw_pain_data(i)==1 & Psim2(i)==1)
    TP=TP+1;
elseif (raw_pain_data(i)==0 & Psim2(i)==1)
    FP=FP+1;
elseif (raw_pain_data(i)==0 & Psim2(i)==0)
    TN=TN+1;
elseif (raw_pain_data(i)==1 & Psim2(i)==0) % pain is 1 and sim is 0
    FN=FN+1;
end

outputData = [ TP, FP, TN, FN]; % Replace this with your own data

% Specify the file name and sheet name
true_fileName = '%d_confusionmat_vals.xlsx'; % Change the file name as needed
sheetName = 'True vs False Values'; % Change the sheet name as needed

% Write the data to the Excel file
xlswrite(true_fileName, outputData, sheetName);

end

function plot_model_fit_sleep_fixed_k0(tvec, params, err, patient_data,
    drug_info, dose_times1, dose_times2, dose_times3, pain_report_real,
% This function plots the raw data with best fit for the model

% run model using provided parameters (patient_info is P0,u,eps,k0, k_drugs)
[tvec,P,D] = pain_sim_interventions3_sleep(tvec,[patient_data(1,2),
... params(1),params(2),fixed_k0,params(3:length(params)-1)'],
... drug_info,dose_times1,dose_times2,dose_times3,bad_sleep_times,

sleep_report_real,sleepqual_tol);

report_pain_real = pain_report_real(2,:);
report_times_real = pain_report_real(1,:);

% place figure in nice spot on screen, and fix ratio
figure('rend','painters','pos',[10 10 800 800])
% subplot(3,1,[1 2]); %original
subplot(3,1,1)

% plot model prediction with shaded error regions
shadedErrorBar(tvec,P,err*ones(length(P),1),'r');
ylim([0 ,10]);

% reorder layers so that error shading doesn't cover up axes
set(gca,'Layer','top')
hold on

% plot real pain report data
plot(report_times_real, report_pain_real, 'ok')
set(gca,'Xticklabel',[]) %to just get rid of the numbers but leave
 the ticks
ylabel('Pain level (0 to 10)');

% number of drug classes taken

153
num_drugs = length(D(1,:));
%
create list of colors to cycle through (for drugs)
color_ind = {'r','b--','g-.'};
%
subplot(3,1,3); original
subplot(3,1,2)
%
for each drug class taken, plot drug in bloodstream
for kk = 1: num_drugs
    plot(tvec, D(:,kk), color_ind{kk});
    hold on
end
ylabel('Std. drug doses in bloodstream');
subplot(3,1,3)
plot(sleep_report_real(1,:), sleep_report_real(2,:), 'c-*');
ylabel('Reported sleep quality');
xlabel('Time (hours)');
end
%
Approximate a Dirac delta function as response to taking drug
%(Note that this is instantaneous dose at time t)
%******************************************************************************
function inc = dosefunc(t, dosetimes1, dosetimes2, dosetimes3,
    num_drugs)
%
We want this to look like a narrow tall rectangular function centered
%
% at each does time in the list.
'inc' is the incremental increase in the drug concentration in the body.

Here's a tunable small parameter. It's going to have units of time.

\[
\epsilon = 0.5;
\]

Properties of each delta-function approximant:

\[
\text{width} = \epsilon; \quad \text{in hours}
\]

\[
\text{height} = \frac{1}{\text{width}}; \quad \text{so width} \times \text{height} = 1
\]

Compute delta function approximation

Handle cases of 1 drug taken, 2 drugs taken, and 3 drugs taken

switch num_drugs
  case 0
    inc = 0;
  case 1
    inc = height * \sum (abs(t - dosetimes_1 - 2.0 * \text{width}) < (\text{width}/2));
  case 2
    inc(1,1) = height * \sum (abs(t - dosetimes_1 - 2.0 * \text{width}) < (\text{width}/2));
    inc(2,1) = height * \sum (abs(t - dosetimes_2 - 2.0 * \text{width}) < (\text{width}/2));
  case 3
    inc(1,1) = height * \sum (abs(t - dosetimes_1 - 2.0 * \text{width}) < (\text{width}/2));
    inc(2,1) = height * \sum (abs(t - dosetimes_2 - 2.0 * \text{width}) < (\text{width}/2));
    inc(3,1) = height * \sum (abs(t - dosetimes_3 - 2.0 * \text{width}) < (\text{width}/2));
  otherwise

155
print 'error with drug count in dosefunc (pain_sim_interventions3)';

end

% inc = height*sum(abs(t-dosetimes-2.0*width)<(width/2)); % top hat
% inc = sum(1.0/(width*sqrt(2*pi))*exp(-(t-dosetimes).^2/(2.0*width^2))); % gaussian
end % function dosefunc

%****************************************
% Approximate a Dirac delta function as response to taking drug
% (Note that this is instantaneous dose at time t)
%****************************************
function inc = sleepfunc(t, dosetimes1)

% We want this to look like a narrow tall rectangular function centered
% at each does time in the list.
% 'inc' is the incremental increase in the drug concentration in the body
% Here's a tunable small parameter. It's going to have units of time.
epsilon = 0.5;

% Properties of each delta-function approximant:
width = epsilon; % in hours
height = 1/width; % so width*height = 1
num_drugs = 1;

% compute delta function approximation
% % Handle cases of 1 drug taken, 2 drugs taken, and 3 drugs taken
case 0
    inc = 0;

case 1
    inc = height*sum(abs(t-dosetimes1-2.0*width)<(width/2));

case 2
    inc(1,1) = height*sum(abs(t-dosetimes1-2.0*width)<(width/2))
    inc(2,1) = height*sum(abs(t-dosetimes2-2.0*width)<(width/2))

case 3
    inc(1,1) = height*sum(abs(t-dosetimes1-2.0*width)<(width/2))
    inc(2,1) = height*sum(abs(t-dosetimes2-2.0*width)<(width/2))
    inc(3,1) = height*sum(abs(t-dosetimes3-2.0*width)<(width/2))

otherwise
    print 'error with drug count in dosefunc (pain_sim_interventions3)';
end

% inc = height*sum(abs(t-dosetimes-2.0*width)<(width/2)); % top hat
% inc = sum(1.0/(width*sqrt(2*pi))*exp(-(t-dosetimes).^2/(2.0*width^2))); % gaussian
end % function dosefunc

%******************************************************************************
6.2. sVOC ODE Model Code:

```matlab
%QJ SCD INFLAMMATION MECHANISTIC MODEL
%THE LIKELIHOOD OF VOC HAPPENING BASED ON UNDERLYING PHYSIOLOGICAL
%MECHANISMS

% system variables
% sym R S AR AS N T P M A o2 ns nm no m1 m2 kr ks wsn wasn dmn dnm
  wnt wtp wasm wma wasa wara mu_pn mu_pm mu_ms

% global R S AR AS N T P M A o2 ns nm no m1 m2 kr ks wsn wasn dmn dnm
  wnt wtp wasm wma wasa mu_pn mu_pm mu_ms

%%%%%%% PARAMETERS %%%%%%%% 
% [assumptions: ks > kr; no + nm < 1] %%%%%%

% make parameters global

global BM o2 po pm ps scd_type mu_am mu_t mu_mc mu_r mu_s mu_p mu_n
  mu_m mu_nar mu_mar mu_m mu_a mu_nar mu_mar mu_nas mu_mas keam
  knm2 knpm2 km2m2 kem2 bm2m1 mu_m1 mu_m1 km1m1 kc

global kps kse kpe kne karp kasp kep kns kpn ken karn kasn mu_e
  mu_am ad ksam keam kpam knam bedh beha bead bpha bnha bnat
  km1ea mu_mar mu_mas kminp km2np knpna mu_np mu_m mu_ec kare kase
```
% INITIAL CONDITIONS

% TIME VECTOR

tspan = 0:.01:15; % days!

lengthof = tspan(end);

% Initial states - auxiliary equilibirum

% R = BM / (ps - mu_r);
% S = ps*R;

% R0 = 2924403; % source from bone marrow
% S0 = 123551;

% load file

voc_data = xlsread('VOC raw data.xlsx',1,'A2:AU46');

save('voc_data.mat','voc_data');
% load(voc_data.mat)

% renaming data to use:
mydata = voc_data;
% choose rows (patient)
%gained platelets (negative number difference) :: 6, 14, 27, 35, 37, 44

%lost platelets (positive number difference) :: 4, 12, 17, 21, 29, 30, 40

patient_ID = 37;

%Hgb, WBC count, platelet count, ANC, eosinophil, retic, RBC

baseline_patient_data = [mydata(patient_ID, 35) mydata(patient_ID, 36) mydata(patient_ID, 38) mydata(patient_ID, 39) mydata(patient_ID, 40) mydata(patient_ID, 37) mydata(patient_ID, 43)];

%Hgb, WBC, platelet, ANC, eosinophil, retic, RBC

crisis_patient_data = [mydata(patient_ID, 20) mydata(patient_ID, 21) mydata(patient_ID, 22) mydata(patient_ID, 23) mydata(patient_ID, 24) mydata(patient_ID, 43) mydata(patient_ID, 43)];

bench_data = [mydata(:, 35) mydata(:, 36) mydata(:, 38) mydata(:, 39) mydata(:, 40) mydata(:, 37) mydata(:, 43)];

neut = nanmean(bench_data(:, 4)); %baseline ANC

plat = nanmean(bench_data(:, 3)); %baseline platelets

heme = nanmean(bench_data(:, 7)) * 100 * mu_s; %rbc count

M2m = 0.04 * nanmean(bench_data(:, 2)); %.04*WBC

M1m = 0.04 * nanmean(bench_data(:, 2)); % .04*WBC

neutT = nanmean(bench_data(:, 4)) / 5; %Baseline ANC / 5
apneut = nanmean(bench_data(:,4)) / 3;  \textit{\% Baseline ANC / 3}
endoA = nanmean(bench_data(:,3)) / 3;  \textit{\% Baseline platlets / 3}
endoD = .2*nanmean(bench_data(:,3)) / 3;  \textit{\% .2 \times EA0}
adhes = heme;

benchmark_patient_data = [neut, plat, heme, M2m, M1m, neutT, apneut, endoA, endoD, adhes];

% Calculate benchmark values as the average of columns a, b, c, d, and e
%benchmark_patient_data = mean(bench_variables, 1);

% Column titles
column_titles = {
  'NA', 'PA', 'AS', 'M2', 'M1', 'NT', 'NP', 'EA', 'ED
  ', 'AM'};  \textit{\% Example column titles}

%number of RBCs from the bone marrow; 1.7 \times 10^{-11} renew a day
%o2 = .60; \textit{\% oxygen level}

[num, txt, raw] = xlsread('voc_lhs_params.xlsx');

% Extract numerical data from raw cell array
data = cell2mat(raw);

% Ensure data is imported as double precision
voc_lhs_params = double(data);
%%
scd_type = .45; .65; % determines percentage of RBCs turned sickle ??
%
BM = 1000000;
po = .10; % transition rate based on hypoxic conditions (could make 1-02)
ps = scd_type; % transition rate based on genetic mutation (polymerization rate)
% %ps = 0;
% %ps = pm + po; % transition rate from red blood cells to sickle cells
%
%
% % decay rates
mu_r = .0083; % natural cell death rate of Red Blood Cells [last 120 days] (1/120)
% mu_s = 0;
mu_s = .05; % natural cell death rate of Sickle Red Blood Cells [last 10-20 days so x10]
% mu_mar = .0025; % decay of rbc heme eaten by macrophages
mu_mas = 6.11; % decay of srbc heme eaten by macrophages x.002
mu_m = 6.956; % decay of resting macrophages x.02
%
%
% endothelial cells
beha = .3; \%transition rate from healthy EC to activated EC upon sRBC adhesion to E
bead = .19; \%.25; \%.215; \%.39 for p40; \%.35; \%.215; \%transition rate from activated EC to damaged EC based on time passed
kne = .01; \%actiivation rate of EC from NETs releasing tissue factor
kase = .333; \%activation rate of ECs from sRBC heme
bedh = .0001; \%healing rate of endothelial cells back to healthy cells from bone marrow EPCs
km1ea = .010; \%activation of ECs from pro inflammatory macrophage cytokines
mu_ec = .0027; \%decay rate of endothelial cells

\%platelets
bpha = .02; \%transition rate of platelets from inactive circulating to active platelets due to sRBCs
kep = 3.025; \%platelet activation from damaged ECs releasing TF
<<<<<<<<<<< changed to 3.025 from 3025
kasp = 3.0; \%platelet activation from hemin released by dead sRBCs
mu_p = .05; \%.09; \%.18 for p40; \%.2; \%.22; \%.5; \%natural cell death rate of Platelets

\%neutrophils
bnha = .607; \%transition rate from resting and circulating to activated N from inflammation caused by sRBC adhesion; 15.889 /N -units/day
kpn = .1; \%activation of neutrophils from activated platelets
kasn = 3.703;
\[ \mu_n = 3.978; \% \mu_n = 4.773; \% 3.978; \%\text{natural cell death rate of Neutrophils: 3.978/day ; 7.109} \]

\[ \mu_{nas} = 3.025; \%\text{death rate from eating heme sRBCs (produced quicker)} \]

\[ b_{nat} = 0.1667; \%\text{transition rate from activated N to NET formation caused by activated platelet} \]

\[ k_{m1np} = 2.898; \%\text{dying neutrophils eaten by M1 macrophages} \]

\[ k_{m2np} = 87.080; \%\text{dying neutrophils eaten by M2 macrophages} \]

\[ k_{npna} = 3.01; \%\text{dying neutrophils removed by active neutrophils} \]

\[ \mu_{np} = 1.309; \%\text{secondary apoptosis} \]

\[ \%\text{adhesion molecules} \]

\[ k_{sam} = 0.2; \%\text{activation of adhesion molecules from [sRBCs activating the endothelium] (HIGH NATURALLY)} \]

\[ k_{nam} = 0.1; \%\text{activation of adhesion molecules from [neutrophil recruitment] for the immune response} \]

\[ k_{pam} = 0.1; \%\text{activation of adhesion molecules from [neutrophil recruitment] for the immune response} \]

\[ k_{eam} = 0.16; \%\text{activation of endothelial adhesion molecules} \]

\[ \mu_{am} = 2.1; \% 2.7; \% 2.230; \% 1.025; \%\text{natural cell death rate of Adhesion molecules} \]

\[ \%\text{adhesion rates} \]

\[ a_d = 0.3; \%\text{overall adhesion rate} \]

\[ \%\text{macrophages} \]
\[ knm2 = 0.025; \text{activation of anti inflam M2 macrophages by active neutrophils} \]

\[ knpm2 = 0.045; \text{activation of anti inflam M2 macrophages by dying neutrophils} \]

\[ km2m2 = 1.624; \text{activation of anti inflam M2 macrophages by itself and its chemokines} \]

\[ kem2 = 1.0; \text{activation of anti inflam M2 macrophages by active endothelial cells} \]

\[ bm2m1 = 8.281; \text{transition rate from anti M2 to pro M1 = function of HEME LEVEL} \]

\[ mu_m1 = 6.956; \text{death rate M1 macrophages} \]

\[ km1m1 = 0.001; \text{activation of pro inflam M1 macrophages by itself and its cytokines} \]

\[ mu_m2 = 8.271; \text{death rate of M2 macrophages} \]

%specify which columns with initial patient values

\[ NA0 = \text{baseline_patient_data(:,4)}; \text{baseline ANC} \]

\[ PA0 = \text{baseline_patient_data(:,3)}; \text{baseline platelets} \]

\[ AS0 = \text{baseline_patient_data(:,7)}*100*\mu_s; \text{rbc count} \]

\[ M20 = 0.04*\text{baseline_patient_data(:,2)}; \text{.04*WBC} \]

\[ M10 = 0.04*\text{baseline_patient_data(:,2)}; \text{.04*WBC} \]

\[ NT0 = \text{baseline_patient_data(:,4)} / 5; \text{Baseline ANC / 5} \]

\[ NP0 = \text{baseline_patient_data(:,4)} / 3; \text{Baseline ANC / 3} \]

\[ EA0 = \text{baseline_patient_data(:,3)} / 3; \text{Baseline platelets / 3} \]

\[ ED0 = 0.2*(\text{baseline_patient_data(:,3)} / 3); \text{.2* EA0} \]
AMO = AS0;
kc = .0125; %background pro-inflamm cytokines

% HEALTHY PATIENT INITIAL CONDITIONS
%EA0 = 10; %activated endothelial cells in vessel wall
%ED0 = 2.94; %damaged endothelial cells %
%PA0 = 2; %activated platelet
%NA0 = 4.22; %activated neutrophil
%NT0 = 2; % number of neutrophil extracellular traps
%AS0 = 0; % initial number of dead sRBCs
%AM0 = 1; % initial number of adhesion molecules
%NP0 = 1; % initial number of dying neutrophils
%M10 = 1.5; % initial number of resting macrophages (pro inflammatory)
%M20 = 1; % initial number of monocyte derived macrophages (anti inflammatory)
%kc = 0.125; %background pro-inflamm cytokines

INIT = [ EA0 ED0 PA0 NA0 NT0 AS0 AM0 NP0 M10 M20];
tic
% ODE SOLVER
options = odeset('RelTol',1e-5,'Stats','on','OutputFcn',@odeplot);
[t,y] = ode23s(@ode_function, tspan, INIT,options);
when is there a VOC? when there are hella aggregates in the area!
when are hella aggregates in the area? when the proportion of those vs EA is high?

% tipping point VOC
% Assign single values to a, b, and c

% S_f = (ps*(BM/(ps+mu_r)));
% N_f = y(end,4) + y(end,5);
% P_f = y(end,3);
% AM_f = y(end,7);
% EA_f = y(end,1);
% ED_f = y(end,2);

%Hgb, WBC, platelet, ANC, eosinophil, retic, RBC

% crisis__data = [mydata(:,20) mydata(:,21) mydata(:,22) mydata(:,23)
                   mydata(:,24) mydata(:,25) mydata(:,43)];

S_f = crisis__data(:,7)*100;
N_f = crisis__data(:,4) + (crisis__data(:,4) / 5);
P_f = crisis__data(:,3);
AM_f = crisis__data(:,7)*100*mu_s;
EA_f = crisis__data(:,3) / 3;
ED_f = .2*(crisis__data(:,3)) / 3;

% outputs
\[ S_{\text{sim}} = (p_s \times (BM/(p_s+\mu_r))) \]
\[ N_{\text{sim}} = y(\text{end}, 4) + y(\text{end}, 5) \]
\[ P_{\text{sim}} = y(\text{end}, 3) \]
\[ AM_{\text{sim}} = y(\text{end}, 7) \]
\[ EA_{\text{sim}} = y(\text{end}, 1) \]
\[ ED_{\text{sim}} = y(\text{end}, 2) \]

\%
\%
Calculate the thresholds based the mean of each population's data
(adjustable depending on data distribution)
\%
threshold_S = nanmean(S_f);
threshold_N = nanmean(N_f);
threshold_P = nanmean(P_f);
threshold_AM = nanmean(AM_f);
threshold_EA = nanmean(EA_f);
threshold_ED = nanmean(ED_f);

threshold_sum = threshold_S + threshold_N + threshold_P +
threshold_AM + threshold_EA + threshold_ED;

%
% Initialize the damage index
sVOC_likely = 0;

if S_sim > (threshold_S/2) && N_sim > (threshold_N/2) && P_sim > (threshold_P/2) && AM_sim > (threshold_AM/2) && EA_sim >
threshold_EA/2 && ED_sim > threshold_ED/2
    % Phase 2: Set damage index to 0.5

sVOC_likeLy = 0.8;

elseif S_sim > threshold_S/3 && N_sim > threshold_N/3 && P_sim >
threshold_P/3 && AM_sim > threshold_AM/3 && EA_sim >
threshold_EA/3 && ED_sim > threshold_ED/3

% Phase 3: Set damage index to 0.2
sVOC_likely = 0.5;

else

% No damage: Keep damage index at 0
sVOC_likely = 0;

end

% Display the result
disp(['The likelihood of sVOC is: ', num2str(sVOC_likely)]);

% plotting each cell density over time individually
figure('rend','painters','pos',[10 10 800 800])
% subplot(5,2,1)
% plot(t,R) %1 - round blood cells
% ylabel('R - RBCs')
% xlabel('time')

% subplot(5,2,2)
% plot(t,S) %2 - sickle blood cells
% ylabel('S - sRBCs')
% xlabel('time')
```matlab
subplot(7,2,1)
plot(t,y(:,1), 'LineWidth',3) % 1 - activated ECs
ylabel('EA - activated ECs')
xlabel('time')
title("Patient #' + patient_ID + ' Over ' + lengthof + ' days at
sVOC likelihood ' + sVOC_likely);

% subplot(7,2,2)
% plot(t,y(:,2), 'LineWidth',3) % 2 - damaged ECs
% ylabel('ED - damaged ECs')
% xlabel('time')

subplot(7,2,2)
plot(t,y(:,2), 'LineWidth',3) % 2 - damaged ECs
ylabel('ED - damaged ECs')
xlabel('time')

subplot(7,2,3)
plot(t,y(:,3), 'LineWidth',3) % 3 - activated platelets
ylabel('PA')
xlabel('time')

subplot(7,2,3)
plot(t,y(:,3), 'LineWidth',3) % 3 - activated platelets
ylabel('PA')
xlabel('time')

subplot(7,2,4)
plot(t,y(:,4), 'LineWidth',3) % 4 - activated neutrophils
ylabel('NA')
xlabel('time')

subplot(7,2,4)
plot(t,y(:,5), 'LineWidth',3) % 5 - neutrophil traps
ylabel('NT')
xlabel('time')
```
```
subplot(7,2,5)
plot(t,y(:,6),'LineWidth',3)  \%7 - AS heme
ylabel('AS - dead sRBCs')
xlabel('time')

subplot(7,2,6)
plot(t,y(:,7),'LineWidth',3)  \%8 - adhesion molecules
ylabel('AM')
xlabel('time')

subplot(7,2,7)
plot(t,y(:,8),'LineWidth',3)  \%9 - apoptotic neutrophils
ylabel('NP')
xlabel('time')

subplot(7,2,8)
plot(t,y(:,9),'LineWidth',3)  \%10 - M1 macrophages
ylabel('M1 - pro.inf.')
xlabel('time')

subplot(7,2,9)
plot(t,y(:,10),'LineWidth',3)  \%11 - M2 macrophages
ylabel('M2 - anti.inf.')
xlabel('time')

hold off
```

% PLOT VECTORS
% plotting all cell densities over time together
figure('render','painters','pos',[10 10 500 700])
% drawnow

% [ EA0 ED0 PA0 NA0 NT0 AR0 AS0 AM0 NP0 M10 M20];

% EA - brown
plot(t,y(:,1),"Color","#442C27","LineStyle","--","LineWidth",5)
% ylim([100,200]);

hold on % PA - green
plot(t,y(:,3),"Color","#2EBF32","LineStyle","--","LineWidth",5)
% ylim([100,200]);

hold on % NA - purple
plot(t,y(:,4),"Color","#582EBF","LineStyle",":","LineWidth",5)
% ylim([100,200]);

hold on % AS - pink
plot(t,y(:,6),"Color","#EB3EE7","LineStyle","-","LineWidth",5)
% ylim([100,200]);

hold on % AS - pink
%plot(t,y(:,7),"Color","#EB3EE7","LineStyle","-","LineWidth",2)
% ylim([100,200]);
hold on  %Adhesion molecules - grey
plot(t,y(:,7),"Color","#7E7663","LineStyle','--','LineWidth',5)
%ylim([100,200]);

hold on  %M1 macrophages - grey
plot(t,y(:,9),"Color","#7E7663","LineStyle','-','LineWidth',5)
%ylim([100,200]);

hold on  %M2 macrophages - grey
plot(t,y(:,10),"Color","#7E7663","LineStyle','--','LineWidth',5)
%ylim([100,200]);

hold on

% Extract output vectors from your ODE system solutions
output_vectors = [y(:,4) y(:,3) y(:,6) y(:,10) y(:,9) y(:,5) y(:,8)  
                  y(:,1) y(:,2) y(:,7)];  % Assuming these vectors exist

% Customize the plot
xlabel('Time');
ylabel('Value');
title('ODE System Solutions with Benchmark Values');
% Add legend for solutions if needed
grid on;
hold off;
title("Patient #" + patient_ID + " All Cell Counts Over " + lengthof + " days");
legend('EA','PA','NA','AS','AM','M1','M2')

%EA = x(1); ED = x(2); PA = x(3); NA = x(4); NT = x(5);
%AS = x(6); AM = x(7); NP = x(8); M1 = x(9); M2 = x(10);

% bench_variables = [neut, plat, heme, M2m, M1m, neutT, apneut, endoA, endoD, adhes ];

% Initialize confusion matrix
confusion_matrix = zeros(2, 2); % Rows: actual class (0 or 1),
    % Columns: predicted class (0 or 1)

% Check if benchmark values are within output vectors
for i = 1:min(size(output_vectors, 1), numel(benchmark_patient_data))

    % Determine if benchmark value falls within output vector range
    is_within_range = any(output_vectors(i, :) ==
        benchmark_patient_data(i));

    % Update confusion matrix based on the comparison
    if is_within_range
        % True positive
        confusion_matrix(2, 2) = confusion_matrix(2, 2) + 1;
    else
        % False negative

174
confusion_matrix(2, 1) = confusion_matrix(2, 1) + 1;

end

% Calculate true negatives (TN) and false positives (FP)
num_true_negatives = sum(all(output_vectors(i, :) ~= benchmark_patient_data(i)));
confusion_matrix(1, 1) = confusion_matrix(1, 1) + num_true_negatives;
confusion_matrix(1, 2) = confusion_matrix(1, 2) + sum(output_vectors(i, :) == benchmark_patient_data(i));
end

% Display confusion matrix % TN FP ; FN TP
disp('Confusion Matrix:');
disp(confusion_matrix);

TN = confusion_matrix(1, 1);
FP = confusion_matrix(1, 2);
FN = confusion_matrix(2, 1);
TP = confusion_matrix(2, 2);

% NEW PLOTT %%%%%
% EA PA NA AM
% EA , PA , NA , AM
endoA\(_c\) = mydata(patient\_ID,22) / 3;

plat\(_c\) = mydata(patient\_ID,22);

neut\(_c\) = mydata(patient\_ID,23);

adhes\(_c\) = mydata(patient\_ID,43) \* 100 \* \(\mu_s\);

% adhes\(_c\) = ad * (mydata(patient\_ID,43) \* 100 \* \(\mu_s\) + (mydata(patient\_ID,22) / 3) + mydata(patient\_ID,22) + mydata(patient\_ID,23));

crisis\_patient\_data2 = [endoA\(_c\) plat\(_c\) neut\(_c\) adhes\(_c\)];

% Hgb, WBC, platelet, ANC, eosinophil, retic, RBC

crisis\_bench\_data = [mydata(:,20) mydata(:,21) mydata(:,22) mydata(:,23) mydata(:,24) mydata(:,25) mydata(:,43)];

% clean the data by removing NaN values

endoA\_val\(_c\) = crisis\_bench\_data(:,3); % Baseline platlets / 3

neut\_val\(_c\) = crisis\_bench\_data(:,4); % baseline ANC

plat\_val\(_c\) = crisis\_bench\_data(:,3); % baseline platelets

adhes\_val\(_c\) = crisis\_bench\_data(:,7);

% parameters for the standard deviation calculation

endoAc\_for\_std = std(endoA\(_c\)(~any(isnan(endoA\(_c\)), 2), :) / 3; % Baseline platlets / 3

neutc\_for\_std = std(neut\(_c\)(~any(isnan(neut\(_c\)), 2), :)); % baseline ANC

platc\_for\_std = std(plat\(_c\)(~any(isnan(plat\(_c\)), 2), :)); % baseline platelets

adhesc\_for\_std = std(adhes\(_c\)(~any(isnan(adhes\(_c\)), 2), :)) \* 100 \* \(\mu_s\);
crisis_benchmark_std = [endoAc_for_std platc_for_std neutc_for_std
       adhesc_for_std];

title("Patient #" + patient_ID + " Main Cell Counts Over " +
       lengthof + " days at sVOC likelihood " + sVOC_likely);

legend('EA', 'PA', 'NA', 'AM', '', 'pop. avg')

<title>plotting sRBCs with platelets, neutrophils, plasma AGEs</title>

figure('rend', 'painters', 'pos', [50 50 700 900])

hold on %sickle cells
plot(tspan, S, "Color", "#900C3F", 'LineStyle', '-', 'LineWidth', 2.5) % sRBCS

%hold on %activated ECs
plot(t, y(:, 1), "Color", "#442C27", 'LineStyle', '--', 'LineWidth', 3.5)

hold on %platelets
plot(t, y(:, 3), "Color", "#2EBF32", 'LineStyle', '--', 'LineWidth', 3.5)

hold on %neutrophils
plot(t, y(:, 4), "Color", "#582EBF", 'LineStyle', '--', 'LineWidth', 3.5)

hold on %Adhesion molecules
plot(t,y(:,7),"Color","#7E7663","LineStyle","--","LineWidth",3.5)

hold on

column_titles2 = {'EA', 'PA', 'NA', 'AM'}; % Example column titles

for i = 1:numel(crisis_patient_data2)
    % Plot benchmark value point
    plot(t(end), crisis_patient_data2(i), 'rx', 'MarkerSize', 17);

    % Draw a line through the benchmark value point
    line([t(1) , t(end)], [crisis_patient_data2(i), crisis_patient_data2(i)], 'Color', 'r', 'LineStyle', ':', 'LineWidth', 1.25);

    % Add label for benchmark value point
    text(t(end), crisis_patient_data2(i), [ column_titles2{i} ], 'VerticalAlignment', 'bottom', 'HorizontalAlignment', 'right');

    % Calculate upper and lower bounds for shaded regions
    y_upper = crisis_patient_data2(i) + crisis_benchmark_std(i);
    y_lower = crisis_patient_data2(i) - crisis_benchmark_std(i);

    % Define x-coordinates for shaded regions (time points)
    x_fill = [t(1), t(end), t(end), t(1)];
% Define y-coordinates for upper and lower shaded regions
y_fill_upper = [y_upper, y_upper, y_lower, y_lower];

% Plot shaded regions for standard deviation
fill(x_fill, y_fill_upper, [0.8, 0.8, 0.8], 'EdgeColor', 'r', 'FaceAlpha', 0.2); % red = [1,0,0]
end

% Customize the plot
xlabel('Time');
ylabel('Value');
title('ODE System Solutions with Crisis Benchmark Values');

% Add legend for solutions if needed
grid on;

hold off;

title("Patient #" + patient_ID + " Main Cell Counts Over " + lengthof + " days at sVOC likelihood " + sVOC_likely);

legend('EA','PA','NA','AM','', 'crisis data')

toc
%-----------------------------------
%% SIMULATION %%
%MAKE FUNCTION FOR SYSTEM(to put into solver)

function dy = ode_function(t,x)
    global BM ps mu_r mu_p mu_n mu_nar mu_m mu_nas mu_e
        mu_am bedh beha bead bpha bnha bnat knm2 knpm2 km2m2 kem2 bm2m1
        mu_m2 mu_m1 km1m1 kc
        global kps kse kpe kne karp kasp kep kns kpn ken karn kasn ad ksam
        keam kpam knam km1ea mu_mar mu_mas km1np km2np knpna mu_np mu_m
        mu_ec kare kase

    EA = x(1); ED = x(2); PA = x(3); NA = x(4); NT = x(5);
    AS = x(6); AM = x(7); NP = x(8); M1 = x(9); M2 = x(10);

% sources:
    s_E = 55; %39; %25 for p40; %30; %44; %100; %static starting amount
        of endothelial in the body ( 10 ,100 for SCD s.s)
    s_P = 48; %41; %30 for p40; %75; %number inactively circulating the
        blood (30 for reg, 75 for SCD s.s.)
    s_N = 40; %50; %number of neutrophils resting in the blood ( 10 for
        healthy , 50 for s.s )
    s_AM = 30; %number of adhesion molecules in the area bc SCD ( 10 for
        healthy, 30 for s.s)
    s_M = 30; %number of macrophages on standy in the area ( 10 for
        healthy, 30 for s.s)
\%(mu_m + km1m1\cdot M1 + kc + bm2m1\cdot M2 + knm2\cdot NA\cdot N2 + knpm2\cdot NP\cdot N2 + km2m2\cdot M2)\%

\%R = \frac{BM}{(ps+\mu_r)};

\%S = (ps\cdot R) = (ps\cdot \frac{BM}{(ps+\mu_r)})

% activated endothelial cells over time
\quad dEA = s\_E\cdot (kne\cdot (NT+NA) + kase\cdot AS + km1ea\cdot M1\cdot EA)/(bead\cdot EA + (kne\cdot (NT+NA) + kase\cdot AS + km1ea\cdot M1\cdot EA)) - bead\cdot EA

% damaged endothelial cells over time
\quad dED = bead\cdot EA - bedh\cdot ED

% activated platelets over time
\quad dPA = s\_P\cdot (kasp\cdot AS + kep\cdot EA)/(\mu_p + (kasp\cdot AS + kep\cdot EA)) - \mu_p\cdot PA

% activated neutrophils over time
\quad dNA = s\_N\cdot (kasn\cdot AS + kpn\cdot PA)/(\mu_n + (kasn\cdot AS + kpn\cdot PA) + bnat\cdot NA) - bnat\cdot NA - \mu_n\cdot NA

% neutrophil traps over time
\quad dNT = bnat\cdot NA - \mu_n\cdot NT

% dead sRBCs over time
\quad dAS = \mu_s\cdot (ps\cdot (BM/(ps+\mu_r))) - mu\_nas\cdot AS\cdot NA - mu\_mas\cdot AS\cdot (M1+M2)
% adhesion molecules over time

\[ dAM = s_{AM} \left( k_{sam} \left( p_{s} \left( B_{M} / (p_{s} + \mu_r) \right) \right) + k_{pam} \cdot PA + k_{nam} \cdot (N_{A} + N_{T}) + k_{eam} \cdot EA \right) / \left( \mu_{am} + k_{sam} \left( p_{s} \left( B_{M} / (p_{s} - \mu_r) \right) \right) + k_{pam} \cdot PA + k_{nam} \cdot N_{A} + k_{eam} \cdot EA \right) \] - \mu_{am} \cdot AM

% apoptosis of neutrophils over time

\[ dNP = \mu_{n} \cdot N_{A} - k_{m1np} \cdot N_{P} \cdot M_{1} - k_{m2np} \cdot N_{P} \cdot M_{2} - k_{npna} \cdot N_{A} - \mu_{np} \cdot N_{P} \]

% M1 macrophages over time

\[ dM1 = \left( s_{M} \left( k_{m1m1} \cdot M_{1} + k_{c} + b_{m2m1} \cdot M_{2} \right) / \left( \mu_{m} + k_{m1m1} \cdot M_{1} + k_{c} + b_{m2m1} \cdot M_{2} + k_{nmm2} \cdot N_{A} \cdot M_{2} + k_{npnm2} \cdot N_{P} \cdot M_{2} + k_{m2m2} \cdot M_{2} + k_{em2} \cdot E_{A} \cdot M_{2} \right) \right) + b_{m2m1} \cdot k_{npnm2} \cdot M_{2} \cdot A_{S} - b_{m2m1} \cdot M_{1} - \mu_{m1} \cdot M_{1} \]

% M2 macrophages over time

\[ dM2 = \left( s_{M} \left( k_{nm2} \cdot N_{A} \cdot M_{2} + k_{npnm2} \cdot N_{P} \cdot M_{2} + k_{m2m2} \cdot M_{2} + k_{em2} \cdot E_{A} \cdot M_{2} + k_{m1m1} \cdot M_{1} + k_{c} + b_{m2m1} \cdot M_{2} \right) \right) - b_{m2m1} \cdot k_{npnm2} \cdot M_{2} \cdot A_{S} + b_{m2m1} \cdot M_{1} - \mu_{m2} \cdot M_{2} \]

\[ dy = [dE_{A}; dE_{D}; dPA; dN_{A}; dN_{T}; dA_{S}; dA_{M}; dN_{P}; dM_{1}; dM_{2}]; \]

end

% -------------------------------------------------------------------
6.3. LHS Sampling Code for the sVOC Model from [61]:

```matlab
function s=LHS_Call(xmin,xmean,xmax,xsd,nsample,distrib,threshold)
% s= latin_hs ( xmean , xsd , nsample , nvar )
% LHS from normal distribution, no correlation
% method of Stein
% Stein , M. 1987. Large Sample Properties of Simulations Using Latin
% Hypercube Sampling.
% Technometrics 29:143-151

if nsample ==1
    s=xmean;
    return
end

if nargin <7
    threshold =1e20;
end

[sample,nvar]=size(xmean);

if distrib == 'norm' % you only need to specify xmean & xsd
    ran=rand(nsample,nvar);
    s=zeros(nsample,nvar);
    % method of Stein
    for j=1: nvar
        idx=randperm(nsample);
        P=(idx.'-ran(:,j))/nsample; % probability of the cdf
```
\[ s(:,j) = \text{xmean}(j) + \text{ltqnorm}(P) \times \text{xsd}(j); \] % this can be replaced by any inverse distribution function

\text{end}
\text{end}

\text{if distrib == 'unif'} % you only need to specify xmin & xmax
    \text{if xmin==0}
        xmin=1e-300;
    \text{end}
    nvar=length(xmin);
    ran=rand(nsample,nvar);
    s=zeros(nsample,nvar);
    \text{for j=1: nvar}
        idx=randperm(nsample);
        P=(idx'-ran(:,j))/nsample;
        xmax(j);
        xmin(j);
        xmax(j)/xmin(j);
        \text{if (xmax(j)<1 & xmin(j)<1) || (xmax(j)>1 & xmin(j)>1)}
            'SAME RANGE';
            \text{if (xmax(j)/xmin(j))<threshold} % It uses the log scale if the order of magnitude of [xmax-xmin] is bigger than threshold
                '<1e3: LINEAR SCALE';
                s(:,j) = xmin(j) + P.* (xmax(j)-xmin(j));
            \text{else}
                '>=1e3: LOG SCALE';
        \text{end}
    \text{end}
\text{end}
\begin{verbatim}
function dydt=sVOCmodel(t,y,LHSmatrix,x,runs)
function dydt=sVOCmodel(t,y,LHSmatrix,x,runs)
    %% PARAMETERS %%
    Parameter_settings_LHS;
    % s=LHSmatrix(x,1);
    % muT=LHSmatrix(x,2);

    s(:,j) = log(xmin(j)) + P.*abs(abs(log(xmax(j))) - abs(log(xmin(j))));
    s(:,j) = exp(s(:,j));

    else
        'e- to e+';
        if (xmax(j)/xmin(j))<threshold  \% It uses the log scale
            if the order of magnitude of [xmax-xmin] is bigger than threshold
                '<1e3: LINEAR SCALE';
                s(:,j) = xmin(j) + P.* (xmax(j)-xmin(j));
            else
                '>=1e3: LOG SCALE';
                s(:,j) = log(xmin(j)) + P.*abs(log(xmax(j))-log(xmin(j)));
                s(:,j) = exp(s(:,j));
        end
    end
end

% hist(s) \% plots the histogram of the pdf
\end{verbatim}
% r=LHSmatrix(x,3);
% k1=LHSmatrix(x,4);
% k2=LHSmatrix(x,5);
% mub=LHSmatrix(x,6);
% N=LHSmatrix(x,7);
% muV=LHSmatrix(x,8);
% dummy_LHS=LHSmatrix(x,9);
% dummy_LHS=LHSmatrix(x,10);

scd_type = LHSmatrix(x,1);

BM = LHSmatrix(x,2);
%po = LHSmatrix(x,3);
ps = LHSmatrix(x,3);

% decay rates
mu_r = LHSmatrix(x,4);
mu_s = LHSmatrix(x,5);
mu_mas = LHSmatrix(x,8);
u_m = LHSmatrix(x,9);

% endothelial cells
beha = LHSmatrix(x,13);
bead = LHSmatrix(x,14);
kne = LHSmatrix(x,29);
kase = LHSmatrix(x,27);
bedh = LHSmatrix(x,12);
km1ea = LHSmatrix(x,38);
mu_ec = LHSmatrix(x,43);

\textit{Platelets}

bpha = LHSmatrix(x,15);
kep = LHSmatrix(x,28);
kasp = LHSmatrix(x,30);
mu_p = LHSmatrix(x,6);

\textit{Neutrophils}

bnha = LHSmatrix(x,16);
kpn = LHSmatrix(x,31);
kasn = LHSmatrix(x,32);
mu_n = LHSmatrix(x,7);
mu_nas = LHSmatrix(x,10);
bnat = LHSmatrix(x,17);
k1np = LHSmatrix(x,39);
k2np = LHSmatrix(x,40);
knpna = LHSmatrix(x,41);
mu_np = LHSmatrix(x,42);

\textit{Adhesion molecules}

ksam = LHSmatrix(x,34);
knam = LHSmatrix(x,37);
kpam = LHSmatrix(x,36);
keam = LHSmatrix(x,35);
mu_am = LHSmatrix(x,11);
% adhesion rates
ad = LHSmatrix(x,33);

% macrophages
knm2 = LHSmatrix(x,18);
knpm2 = LHSmatrix(x,19);
km2m2 = LHSmatrix(x,20);
kem2 = LHSmatrix(x,21);
bm2m1 = LHSmatrix(x,22);
mu_m1 = LHSmatrix(x,24);
km1m1 = LHSmatrix(x,25);
mu_m2 = LHSmatrix(x,23);

% sources:
s_E = LHSmatrix(x,44);
s_P = LHSmatrix(x,45);
s_N = LHSmatrix(x,46);
s_AM = LHSmatrix(x,47);
s_M = LHSmatrix(x,48);
kc = LHSmatrix(x,26);

EA = y(1); ED = y(2); PA = y(3); NA = y(4); NT = y(5);
AS = y(6); AM = y(7); NP = y(8); M1 = y(9); M2 = y(10);

%(mu_m + km1m1*M1 + kc + bm2m1*M2 + knm2*NA*M2 + knpm2*NP*M2 + km2m2 * M2)
%R = (BM/(ps+mu_r));
%S = (ps*R); = (ps*(BM/(ps+mu_r)));

% activated endothelial cells over time
dEA = s_E*(kne*(NT+NA) + kase*AS + km1ea*M1*EA)/(bead*EA + (kne*(NT+NA) + kase*AS + km1ea*M1*EA)) - bead*EA;

% Check if a certain amount of time has passed
% Equation 2 activation condition
%if t >= .5*length(t) %&& EA > % Activate Equation 2 after half of the time and if x(1) >
  % dED = bead*EA - bedh*ED ; % Activate Equation 2
%else
  % dED = 0; % Keep Equation 2 inactive
%end

%damaged endothelial cells over time
dED = bead*EA - bedh*ED;
%dED = bead*EA/(mu_ec + bead*EA) - bedh*ED

% healthy platelets over time
%dPH = PH - bpha*PH
% source - activated platelets

% activated platelets over time
\[ d_P A = s_P \times (k_{ap} \times \text{AS} + k_{ep} \times \text{EA}) / (\mu_p + (k_{ap} \times \text{AS} + k_{ep} \times \text{EA})) - \mu_p \times \text{PA}; \]

% healthy neutrophils over time
% \[ d_{NH} = N - b_{nha} \times \text{NH} \]
% source - activated neutrophils
% activated neutrophils over time
% \[ d_{NA} = s_N \times (k_{asn} \times \text{AS} + k_{pna} \times \text{PA}) / (\mu_n + (k_{asn} \times \text{AS} + k_{pna} \times \text{PA}) + b_{nat} \times \text{NA}) - b_{nat} \times \text{NA} - \mu_n \times \text{NA}; \]
% neutrophil traps over time
% \[ d_{NT} = b_{nat} \times \text{NA} - \mu_n \times \text{NT}; \]
% dead RBCs over time
% \[ d_{AR} = \mu_r \times (BM / (p_s + \mu_r)) - \mu_{nar} \times \text{AR} \times \text{NA} - \mu_{mar} \times \text{AR} \times (M1 + M2) \]
% dead sRBCs over time
% \[ d_{AS} = \mu_s \times (p_s \times (BM / (p_s + \mu_r))) - \mu_{nas} \times \text{AS} \times \text{NA} - \mu_{mas} \times \text{AS} \times (M1 + M2); \]
% adhesion molecules over time
% \[ d_{AM} = s_{AM} \times (k_{sam} \times (p_s \times (BM / (p_s + \mu_r))) + k_{pam} \times \text{PA} + k_{nam} \times \text{NA} + k_{eam} \times \text{EA}) / (\mu_{am} + (k_{sam} \times (p_s \times (BM / (p_s - \mu_r))) + k_{pam} \times \text{PA} + k_{nam} \times \text{NA} + k_{eam} \times \text{EA})) - \mu_{am} \times \text{AM}; \]
% adhesion molecules over time
% \[ k_{se} \times (p_s \times (BM / (p_s - \mu_r))) \times E + k_{pe} \times P \times E + k_{ne} \times N \times E - \mu_{am} \times \text{AM} \]
% apoptotic neutrophils over time
dNP = \mu_n \cdot NA - k_{m1np} \cdot NP \cdot M1 - k_{m2np} \cdot NP \cdot M2 - k_{npna} \cdot NA - \mu_{np} \cdot NP;

%M1 macrophages over time
dM1 = (s_M \cdot (k_{m1m1} \cdot M1 + k_c + b_{m2m1} \cdot M2) + b_{m2m1} \cdot k_{npm2} \cdot M2 \cdot NP - \mu_{m1} \cdot M1)
        = \frac{s_M \cdot (k_{m1m1} \cdot M1 + k_c + b_{m2m1} \cdot M2)}{(\mu_m + k_{m1m1} \cdot M1 + k_c + b_{m2m1} \cdot M2 + k_{nm2} \cdot NA \cdot M2 + k_{npm2} \cdot NP \cdot M2 + k_{km2} \cdot M2 + k_{kem2} \cdot EA \cdot M2)} + b_{m2m1} \cdot k_{npm2} \cdot M2 \cdot AS + b_{m2m1} \cdot M2 - \mu_{m1} \cdot M1;

%M2 macrophages over time
dM2 = (s_M \cdot (k_{nm2} \cdot NA \cdot M2 + k_{npm2} \cdot NP \cdot M2 + k_{km2} \cdot M2 + k_{kem2} \cdot EA \cdot M2) - b_{m2m1} \cdot k_{npm2} \cdot NP - b_{m2m1} \cdot M2 - \mu_{m2} \cdot M2)
        = \frac{s_M \cdot (k_{nm2} \cdot NA \cdot M2 + k_{npm2} \cdot NP \cdot M2 + k_{km2} \cdot M2 + k_{kem2} \cdot EA \cdot M2)}{(\mu_m + k_{nm2} \cdot NA \cdot M2 + k_{npm2} \cdot NP \cdot M2 + k_{km2} \cdot M2 + k_{kem2} \cdot EA \cdot M2 + k_{km1} \cdot M1 + k_c + b_{m2m1} \cdot M2)} - b_{m2m1} \cdot k_{npm2} \cdot M2 \cdot AS - b_{m2m1} \cdot M2 - \mu_{m2} \cdot M2;

dydt = [dEA; dED; dPA; dNA; dNT; dAS; dAM; dNP; dM1; dM2];

%-------------------------------------------------------------------
% Tinf - T1death - T1inf;
% T1inf - T2death;
% Vrelease - Tinf - Vdeath];
The results should be compared to the PRCC results section in Supplementary Material D and Table D.1 for different N (specified by "runs" in the script below)

clear all;
close all;

% Sample size N
runs=250;
%runs = 25;

%alpha = .05;

%tspan = 0:.01:14;%days!
%t_end = 14; % length of the simulations
%tspan = (1:.01:t_end);

% LHS MATRIX %
Parameter_settings_LHS;

%LHS_CALL = (xmin,xmean,xmax,xsd,nsample,distrib,threshold)
scd_type_LHS = LHS_Call(.25, scd_type, .9, 0, runs, 'unif'); % baseline = .45
BM_LHS = LHS_Call(10e4, BM, 10e8, 0, runs, 'unif'); % baseline = 10e6
ps_LHS = LHS_Call(.01, ps, .9, 0, runs, 'unif'); % baseline = .45
mu_r_LHS = LHS_Call(1e-4, mu_r, .9, 0, runs, 'unif'); % baseline =
mu_s_LHS = LHS_Call(1e-4, mu_s, .9, 0, runs, 'unif'); % baseline = .05
mu_p_LHS = LHS_Call(1e-2, mu_p, .9, 0, runs, 'unif'); % baseline = .25
mu_n_LHS = LHS_Call(1, mu_n, 10, 0, runs, 'unif'); % baseline = .03
mu_mas_LHS = LHS_Call(1, mu_mas, 10, 0, runs, 'unif'); % baseline = .002
mu_m_LHS = LHS_Call(1, mu_m, 10, 0, runs, 'unif'); % baseline = .02
mu_nas_LHS = LHS_Call(1, mu_nas, 10, 0, runs, 'unif'); % baseline = .025
mu_am_LHS = LHS_Call(.1, mu_am, 10, 0, runs, 'unif'); % baseline = .025

deh_LHS = LHS_Call(1e-5, deh, .79, 0, runs, 'unif'); % baseline = .5
deha_LHS = LHS_Call(.1, deha, .5, 0, runs, 'unif'); % baseline = .3
bead_LHS = LHS_Call(1e-5, bead, .9, 0, runs, 'unif'); % baseline = .25
bpha_LHS = LHS_Call(1e-5, bpha, .9, 0, runs, 'unif'); % baseline = .3
bnha_LHS = LHS_Call(1e-5, bnha, .9, 0, runs, 'unif'); % baseline = .3
bnat_LHS = LHS_Call(1e-5, bnat, .9, 0, runs, 'unif'); % baseline = .002

knm2_LHS = LHS_Call(1e-3, knm2, .9, 0, runs, 'unif'); % baseline = .25
knpm2_LHS = LHS_Call(1e-3, knpm2, .9, 0, runs, 'unif'); % baseline = .45
km2m2_LHS = LHS_Call(1e-3, km2m2, .9, 0, runs, 'unif'); % baseline = .1
kem2_LHS = LHS_Call(1e-3, kem2, .9, 0, runs, 'unif'); % baseline = .02
bm2m1_LHS = LHS_Call(1, bm2m1, 10, 0, runs, 'unif'); % baseline = .5
mu_m2_LHS = LHS_Call(1, mu_m2, 10, 0, runs, 'unif'); % baseline = 3.2
mu_m1_LHS = LHS_Call(1, mu_m1, 10, 0, runs, 'unif'); % baseline = 2.2
km1m1_LHS = LHS_Call(1e-4, km1m1, .9, 0, runs, 'unif'); % baseline = .001
kc_LHS = LHS_Call(1e-3, kc, .5, 0, runs, 'unif'); % baseline = .0125
kase_LHS = LHS_Call(1e-2, kase, 10, 0, runs, 'unif'); % baseline = .3
kep_LHS = LHS_Call(1, kep, 10, 0, runs, 'unif'); % baseline = .01
kne_LHS = LHS_Call(1e-3, kne, 10, 0, runs, 'unif'); % baseline = .12
kasp_LHS = LHS_Call(1e-2, kasp, 10, 0, runs, 'unif'); % baseline = .05
kpn_LHS = LHS_Call(1e-2, kpn, 10, 0, runs, 'unif'); % baseline = .05
kasn_LHS = LHS_Call(1, kasn, 10, 0, runs, 'unif'); % baseline = .002
ad_LHS = LHS_Call(1e-2, ad, 5, 0, runs, 'unif'); % baseline = .03
ksam_LHS = LHS_Call(1e-2, ksam, 5, 0, runs, 'unif'); % baseline = .6
keam_LHS = LHS_Call(1e-2, keam, 5, 0, runs, 'unif'); % baseline = .3
kpam_LHS = LHS_Call(1e-2, kpam, 5, 0, runs, 'unif'); % baseline = .5
knam_LHS = LHS_Call(1e-2, knam, 5, 0, runs, 'unif'); % baseline = .5
km1ea_LHS = LHS_Call(1e-4, km1ea, .3, 0, runs, 'unif'); % baseline = .10
km1np_LHS = LHS_Call(1e-1, km1np, 5, 0, runs, 'unif'); % baseline = .021
km2np_LHS = LHS_Call(1e-1, km2np, 90, 0, runs, 'unif'); % baseline = .021
knpna_LHS = LHS_Call(1e-1, knpna, 5, 0, runs, 'unif'); % baseline = .01
\texttt{mu\_np\_LHS} = LHS\_Call(1e\-2, mu\_np, 3, 0, runs, 'unif'); \texttt{\% baseline = 0.025}

\texttt{mu\_ec\_LHS} = LHS\_Call(1e\-4, mu\_ec, 1, 0, runs, 'unif'); \texttt{\% baseline = 0.002}

\texttt{s\_E\_LHS} = LHS\_Call(10, s\_E, 125, 0, runs, 'unif'); \texttt{\% baseline = 100}

\texttt{s\_P\_LHS} = LHS\_Call(10, s\_P, 100, 0, runs, 'unif'); \texttt{\% baseline = 75}

\texttt{s\_N\_LHS} = LHS\_Call(10, s\_N, 75, 0, runs, 'unif'); \texttt{\% baseline = 50}

\texttt{s\_AM\_LHS} = LHS\_Call(10, s\_AM, 50, 0, runs, 'unif'); \texttt{\% baseline = 30}

\texttt{s\_M\_LHS} = LHS\_Call(10, s\_M, 50, 0, runs, 'unif'); \texttt{\% baseline = 30}

%% LHS MATRIX and PARAMETER LABELS

\texttt{LHS\_matrix = [scd\_type\_LHS, BM\_LHS, ps\_LHS, mu\_r\_LHS, mu\_s\_LHS, mu\_p\_LHS,}
\texttt{   mu\_n\_LHS, mu\_mas\_LHS, mu\_m\_LHS, mu\_nas\_LHS, mu\_am\_LHS, bedh\_LHS,}
\texttt{   beha\_LHS, bead\_LHS, bpha\_LHS, bnha\_LHS, bnat\_LHS,}
\texttt{   knm2\_LHS, knpm2\_LHS, km2m2\_LHS, kem2\_LHS, bm2m1\_LHS, mu\_m2\_LHS,}
\texttt{   mu\_m1\_LHS, km1m1\_LHS, kc\_LHS, kase\_LHS, kep\_LHS, kne\_LHS,}
\texttt{   kasp\_LHS,}
\texttt{   kpn\_LHS, kasn\_LHS, ad\_LHS, ksam\_LHS, keam\_LHS, kpam\_LHS, knam\_LHS,}
\texttt{   km1ea\_LHS, kminp\_LHS, km2np\_LHS, knpna\_LHS,}
\texttt{   mu\_np\_LHS, mu\_ec\_LHS, s\_E\_LHS, s\_P\_LHS, s\_N\_LHS, s\_AM\_LHS, s\_M\_LHS];}

\texttt{for \textbf{x} = 1: runs \texttt{\% Run solution x times choosing different values}}
\texttt{\hspace{3em} f=@ODE\_LHS;}
\texttt{\hspace{3em} \texttt{\%f=@ODE\_LHS;}}
\texttt{\hspace{3em} \texttt{x}}
\texttt{\hspace{3em} LHS\_matrix(x,:);}
\texttt{\hspace{3em} [t,y]=ode23s(@(t,y)f(t,y,LHS\_matrix,x,runs),\texttt{tspan},y0,[]);}
% [t, y] = ode15s (@(t,y)f(t,y,LHSmatrix,x,runs),tspan,y0,[]); 

A = [t y]; % [time y] 
%% Save the outputs at ALL time points [tspan] 
% T_lhs(:,x) = Anew(:,1);  
% CD4_lhs(:,x) = Anew(:,2);  
% T1_lhs(:,x) = Anew(:,3);  
% T2_lhs(:,x) = Anew(:,4);  
% V_lhs(:,x) = Anew(:,5);  

EA_lhs(:,x) = A(:,2);  
ED_lhs(:,x) = A(:,3);  
PA_lhs(:,x) = A(:,4);  
NA_lhs(:,x) = A(:,5);  
NT_lhs(:,x) = A(:,6);  
AS_lhs(:,x) = A(:,7);  
AM_lhs(:,x) = A(:,8);  
NP_lhs(:,x) = A(:,9);  
M1_lhs(:,x) = A(:,10);  
M2_lhs(:,x) = A(:,11);  

% EA0 ED0 PA0 NA0 NT0 AS0 AM0 (8) NP0 M10 M20 
%% Save only the outputs at the time points of interest [time_points]:  
%% MORE EFFICIENT 
% EA_lhs(:,x) = A(time_points+1,1);
ED_lhs(:,x)=A(time_points+1,2);
PA_lhs(:,x)=A(time_points+1,3);
NA_lhs(:,x)=A(time_points+1,4);
NT_lhs(:,x)=A(time_points+1,5);
AS_lhs(:,x)=A(time_points+1,6);
AM_lhs(:,x)=A(time_points+1,7);
NP_lhs(:,x)=A(time_points+1,8);
M1_lhs(:,x)=A(time_points+1,9);
M2_lhs(:,x)=A(time_points+1,10);

end

%% Save the workspace
save Model_LHS.mat;

% tipping point VOC
% Assign single values to a, b, and c (replace these with your actual values)
S_f = (ps*(BM/(ps+mu_r)));
N_f = y(end,4) + y(end,5);
P_f = y(end,3);
AM_f = y(end,7);
EA_f = y(end,1);
ED_f = y(end,2);

% Calculate the thresholds based on the 75th percentile of each population's data (adjustable depending on data distribution)
threshold_S = quantile(S_f, 0.75);
threshold_N = quantile(N_f, 0.75);
threshold_P = quantile(P_f, 0.75);
threshold_AM = quantile(AM_f, 0.75);
threshold_EA = quantile(EA_f, 0.75);
threshold_ED = quantile(ED_f, 0.75);

% Initialize the damage index
sVOC_likely = 0;

% Check conditions for each population and update the damage index accordingly
if S_f > threshold_S && N_f > threshold_N && P_f > threshold_P &&
   AM_f > threshold_AM && EA_f > threshold_EA && ED_f > threshold_ED
   % Phase 1: Set sVOC likelihood to 0.2
   sVOC_likely = 0.2;
elseif S_f > (threshold_S/2) && N_f > (threshold_N/2) && P_f > (thres
   % Phase 2: Set damage index to 0.5
   sVOC_likely = 0.5;
elseif S_f > threshold_S/3 && N_f > threshold_N/3 && P_f > thres
   % Phase 3: Set damage index to 0.8
   sVOC_likely = 0.8;
else
   % No damage: Keep damage index at 0

sVOC_likely = 0;
end

% Display the result
disp(['The likelihood of sVOC is: ', num2str(sVOC_likely)]);

% plotting each cell density over time individually
figure('rend','painters','pos',[10 10 800 800])
%s subplot(5,2,1)
%plot(t,R) %1 - round blood cells
%ylabel('R - RBCs')
%xlabel('time')

% subplot(5,2,2)
%plot(t,S)% 2 - sickle blood cells
%ylabel('S - sRBCs')
%xlabel('time')

 subplot(7,2,1)
plot(t,y(:,1),'LineWidth',3) % 1 - activated ECs
ylabel('EA - activated ECs')
xlabel('time')

% subplot(7,2,2)
%plot(t,y(:,2),'LineWidth',3) % 2 - damaged ECs
subplot(7,2,2)
plot(t,y(:,3),'LineWidth',3) % 3 - activated platelets
ylabel('PA - activated platelets')
xlabel('time')

subplot(7,2,3)
plot(t,y(:,4),'LineWidth',3) % 4 - activated neutrophils
ylabel('NA - activated neutrophils')
xlabel('time')

subplot(7,2,4)
plot(t,y(:,5),'LineWidth',3) % 5 - neutrophil traps
ylabel('NT - neutrophil traps')
xlabel('time')

% subplot(7,2,6)
% plot(t,y(:,6)) % 6 - AR heme
% ylabel('AR - dead RBCs')
% xlabel('time')

subplot(7,2,5)
plot(t,y(:,6),'LineWidth',3) % 7 - AS heme
ylabel('AS - dead sRBCs')
xlabel('time')
subplot(7,2,6)
plot(t,y(:,7),'LineWidth',3) %8 - adhesion molecules
ylabel('AM - adhesion molecules')
xlabel('time')

subplot(7,2,7)
plot(t,y(:,8),'LineWidth',3) %9 - apoptotic neutrophils
ylabel('NP - apoptotic N')
xlabel('time')

subplot(7,2,8)
plot(t,y(:,9),'LineWidth',3) %10 - M1 macrophages
ylabel('M1 - pro.inf.')
xlabel('time')

subplot(7,2,9)
plot(t,y(:,10),'LineWidth',3) %11 - M2 macrophages
ylabel('M2 - anti.inf. macrophages')
xlabel('time')

hold off

%PLOT VECTORS
%plotting all cell densities over time together
figure('rend','painters','pos',[10 10 500 700])
%drawnow

% [ EAO EDO PAO NTO ARO ASO AMO NPO M10 M20];

%EA - brown
plot(t,y(:,1),"Color","#442C27","LineStyle","-",'LineWidth',5)
ylim([100,200]);

% PA - green
plot(t,y(:,3),"Color","#2EBF32","LineStyle","-",'LineWidth',5)
ylim([100,200]);

% NA - purple
plot(t,y(:,4),"Color","#582EBF","LineStyle",":\',"LineWidth',5)
ylim([100,200]);

%AS - pink
plot(t,y(:,6),"Color","#EB3EE7","LineStyle","-",'LineWidth',5)
ylim([100,200]);

% Adhesion molecules - grey
plot(t,y(:,7),"Color","#7E7663","LineStyle","-",'LineWidth',5)
% ylim ([100,200]);

hold on % M1 macrophages - grey
plot (t,y(:,9),"Color","#7E7663","LineStyle","-","LineWidth",5)

% ylim ([100,200]);

hold on % M2 macrophages - grey
plot (t,y(:,10),"Color","#7E7663","LineStyle","--","LineWidth",5)

% ylim ([100,200]);

hold off

legend ('EA', 'PA', 'NA', 'AS', 'AM', 'M1', 'M2')

% plotting sRBCs with platelets, neutrophils, plasma AGEs

figure ('rend', 'painters', 'pos', [10 10 500 700])
%drawnow %sickle cells
%plot (tspan, S, "Color", "#900C3F", 'LineStyle', '-', 'LineWidth', 2.5) % sRBCS

% hold on % activated ECs
plot (t,y(:,1),"Color"," #442C27","LineStyle","--","LineWidth",3.5)

hold on % platelets
plot (t,y(:,3),"Color","#2EBF32","LineStyle","--","LineWidth",3.5)

hold on % neutrophils
plot (t,y(:,4),"Color","#582EBF","LineStyle","--","LineWidth",3.5)
hold on  % Adhesion molecules
plot(t,y(:,7),"Color","#7E7663","LineStyle",'--','LineWidth',3.5)
hold off

legend('EA','PA','NA','AM')

figure('rend','painters','pos',[10 10 500 800])
plot(t,VOC_sev)
title(" Estimated VOC severity over time")

for i = 1:25

% tipping point VOC
% Assign single values to a, b, and c (replace these with your actual values)
S_f = (ps*(BM/(ps+mu_r)));
N_f = NA_lhs(end,i) + NT_lhs(end,i);  % y(end,4) + y(end,5);
P_f = PA_lhs(end,i);
AM_f = AM_lhs(end,i);
EA_f = EA_lhs(end,i);
ED_f = ED_lhs(end,i);

% Calculate the thresholds based on the 75th percentile of each population's data (adjustable depending on data distribution)
threshold_S = quantile(S_f, 0.75);
threshold_N = quantile(N_f, 0.75);
threshold_P = quantile(P_f, 0.75);
threshold_AM = quantile(AM_f, 0.75);
threshold_EA = quantile(EA_f, 0.75);
threshold_ED = quantile(ED_f, 0.75);

% Initialize the damage index
sVOC_likely = 0;

% Check conditions for each population and update the damage index accordingly
if S_f > threshold_S && N_f > threshold_N && P_f > threshold_P &&
   AM_f > threshold_AM && EA_f > threshold_EA && ED_f >
   threshold_ED
   % Phase 1: Set sVOC likelihood to 0.2
   sVOC_likely = 0.2;
elseif S_f > (threshold_S/2) && N_f > (threshold_N/2) && P_f >
   (threshold_P/2) && AM_f > (threshold_AM/2) && EA_f > threshold_EA
   /2 && ED_f > threshold_ED/2
   % Phase 2: Set damage index to 0.5
   ...
sVOC_likely = 0.5;

elseif S_f > threshold_S /3 && N_f > threshold_N /3 && P_f >
threshold_P /3 && AM_f > threshold_AM /3 && EA_f > threshold_EA /3
&& ED_f > threshold_ED /3

% Phase 3: Set damage index to 0.8
sVOC_likely = 0.8;

else
% No damage: Keep damage index at 0
sVOC_likely = 0;
end

%plotting each cell density over time individually
figure('rend','painters','pos',[10 10 800 800])

subplot(5,2,1)
plot(t,R) %1 - round blood cells
ylabel('R - RBCs')
xlabel('time')

subplot(5,2,2)
plot(t,S) %2 - sickle blood cells
ylabel('S - sRBCs')
xlabel('time')

subplot(7,2,1)
plot(t,EA lhs(:,i),'LineWidth',3) % 1 - activated ECs
ylabel('EA - activated ECs')
xlabel('time')
title("RUN #" + i + " Solo Cell Counts at sVOC likelihood ")
sVOC_likely);

subplot(7,2,2)
plot(t,ED_lhs(:,i),'LineWidth',3) % 2 - damaged ECs
ylabel('ED - damaged ECs')
xlabel('time')

subplot(7,2,3)
plot(t,PA_lhs(:,i),'LineWidth',3) % 3 - activated platelets
ylabel('PA')
xlabel('time')

subplot(7,2,4)
plot(t,NA_lhs(:,i),'LineWidth',3) % 4 - activated neutrophils
ylabel('NA')
xlabel('time')

subplot(7,2,5)
plot(t,NT_lhs(:,i),'LineWidth',3) % 5 - neutrophil traps
ylabel('NT')
xlabel('time')

% subplot(7,2,6)
% plot(t,y(:,6)) % 6 - AR heme
% ylabel('AR - dead RBCs')
% xlabel('time')

subplot(7,2,6)
plot(t, AS_lhs(:,i), 'LineWidth', 3) % 7 - AS heme
ylabel('AS - dead sRBCs')
xlabel('time')

subplot(7,2,7)
plot(t, AM_lhs(:,i), 'LineWidth', 3) % 8 - adhesion molecules
ylabel('AM')
xlabel('time')

subplot(7,2,8)
plot(t, NP_lhs(:,i), 'LineWidth', 3) % 9 - apoptotic neutrophils
ylabel('NP')
xlabel('time')

subplot(7,2,9)
plot(t, M1_lhs(:,i), 'LineWidth', 3) % 10 - M1 macrophages
ylabel('M1 - pro.inf.')
xlabel('time')

subplot(7,2,10)
plot(t, M2_lhs(:,i), 'LineWidth', 3) % 11 - M2 macrophages
ylabel('M2 - anti.inf.')
xlabel('time')

hold off
runnumbera = 0;
runnumber = runnumbera + i;

% Define the filename for saving the plot
filename = sprintf('LHS_Transient_plots_%s.png',runnumber);

%filename = sprintf('PRCC_plots%.png', PRCC_var);

% Specify the folder where you want to save the images
folder = '/Users/macuser/Desktop/PhD/Ch 4 VOC ODE/CHAPTER 4 analysis/LHS + PRCC';

% Save the plot as an image in the specified folder
saveas(gcf, fullfile(folder, filename));

% Close the current figure to avoid cluttering
close(gcf);

end

PRCC_PLOT(LHSmatrix, EA_lhs, t_end, PRCC_var, y_var_label(1,1))
PRCC_PLOT(LHSmatrix, ED_lhs, t_end, PRCC_var, y_var_label(1,2));
PRCC_PLOT(LHSmatrix, PA_lhs, t_end, PRCC_var, y_var_label(1,3))
PRCC_PLOT(LHSmatrix, NA_lhs, t_end, PRCC_var, y_var_label(1,4))
PRCC_PLOT(LHSmatrix, NT_lhs, t_end, PRCC_var, y_var_label(1,5))
PRCC_PLOT(LHSmatrix, AS_lhs, t_end, PRCC_var, y_var_label(1,6))
PRCC_PLOT(LHSmatrix, AM_lhs, t_end, PRCC_var, y_var_label(1,7))
PRCC_PLOT(LHSmatrix, NP_lhs, t_end, PRCC_var, y_var_label(1,8))
PRCC_PLOT(LHSmatrix, M1_lhs, t_end, PRCC_var, y_var_label(1,9))
PRCC_PLOT(LHSmatrix, M2_lhs, t_end, PRCC_var, y_var_label(1,10))

%PRCC_PLOT(LHSmatrix, EA_lhs, .5*t_end, PRCC_var, y_var_label)
```matlab
% PRCC_PLOT(LHSmatrix, EA_lhs, t_end, PRCC_var, y_var_label)

% [prcc, sign, sign_label] = PRCC(LHSmatrix, y, t_end, PRCC_var, alpha);
[prcc, sign, sign_label] = PRCC(LHSmatrix, EA_lhs, t_end, PRCC_var, .05);
```

```
% PARAMETER BASELINE VALUES

scd_type = .65; % determines percentage of RBCs turned sickle

BM = 1000000; % RBC x 10e6

po = .10; % transition rate based on hypoxic conditions (could make 1 - O2)

ps = scd_type; % transition rate based on genetic mutation (polymerization rate)

% decay rates

mu_r = .0083; % natural cell death rate of Red Blood Cells [last 120 days] (1/120)

mu_s = .05; % natural cell death rate of Sickle Red Blood Cells [last 10-20 days so x10]

mu_mas = 6.11; % decay of srbc heme eaten by macrophages .03

mu_m = 6.956; % decay of resting macrophages .02

% endothelial cells

beha = .3; % transition rate from healthy EC to activated EC upon srBC adhesion to E
```
bead = .0027; \( \% \) transition rate from activated EC to damaged EC based on time passed

kne = .01; \( \% \) activation rate of EC from NETs releasing tissue factor .12

kase = .33; \( \% \) activation rate of ECs from sRBC heme

bedh = .0001; \( \% \) healing rate of endothelial cells back to healthy cells from bone marrow EPCs

km1ea = .010; \( \% \) activation of ECs from pro inflammatory macrophage cytokines

mu_ec = .0027; \( \% \) decay rate of endothelial cells

Beads

bpha = .02; \( \% \) transition rate of platelets from inactive circulating to active platelets due to sRBCs

kep = 4; \( \% \) platelet activation from damaged ECs releasing TF

kasp = 3; \( \% \) platelet activation from hemin released by dead sRBCs

mu_p = .25; \( \% \) natural cell death rate of Platelets

Neutrophils

bnha = .607; \( \% \) transition rate from resting and circulating to activated N from inflammation caused by sRBC adhesion; 15.889 \( /N \) units/day

kpn = .1; \( \% \) activation of neutrophils from activated platelets

kasn = 3.703;

mu_n = 7.109; \( \% \) natural cell death rate of Neutrophils: 3.978/day

mu nas = 3.025; \( \% \) death rate from eating heme sRBCs (produced quicker)

bnat = .1667; \( \% \) transition rate from activated N to NET formation
caused by activated platelet

\[ k_{1np} = 2.898; \text{% dying neutrophils eaten by M1 macrophages} \]
\[ k_{2np} = 87.08; \text{% dying neutrophils eaten by M2 macrophages} \]
\[ k_{npa} = 3.01; \text{% dying neutrophils removed by active neutrophils} \]
\[ \mu_{np} = 1.309; \text{% secondary apoptosis} \]

% adhesion molecules

\[ k_{sam} = 1.2; \text{% activation of adhesion molecules from [sRBCs activating the endothelium] (HIGH NATURALLY)} \]
\[ k_{nam} = 1.1; \text{% activation of adhesion molecules from [neutrophil recruitment] for the immune response} \]
\[ k_{pam} = 1.1; \text{% activation of adhesion molecules from [neutrophil recruitment] for the immune response} \]
\[ k_{eam} = 1.16; \text{% activation of endothelial adhesion molecules} \]
\[ \mu_{am} = 1.025; \text{% natural cell death rate of Adhesion molecules} \]

% adhesion rates

\[ a_d = .3; \text{% overall adhesion rate} \]

% macrophages

\[ k_{nm2} = .025; \text{% activation of anti inflam M2 macrophages by active neutrophils} \]
\[ k_{npm2} = .045; \text{% activation of anti inflam M2 macrophages by dying neutrophils} \]
\[ k_{2m2} = 1.624; \text{% activation of anti inflam M2 macrophages by itself and its chemokines} \]
60 \text{kem2} = 1.0; \% activation of anti inflam M2 macrophages by active endothelial cells
61 \text{bm2m1} = 8.281; \% transition rate from anti M2 to pro M1 = function of HEME LEVEL
62 \text{mu_m1} = 6.956; \% death rate M1 macrophages
63 \text{km1m1} = .001; \% activation of pro inflam M1 macrophages by itself and its cytokines
64 \text{mu_m2} = 8.271; \% death rate of M2 macrophages

66 \% sources:
67 \text{s_E} = 100; \% static starting amount of endothelial in the body (10,100 for SCD s.s)
68 \text{s_P} = 75; \% number inactively circulating the blood (30 for reg, 75 for SCD s.s.)
69 \text{s_N} = 50; \% number of neutrophils resting in the blood (10 for healthy, 50 for s.s)
70 \text{s_AM} = 30; \% number of adhesion molecules in the area bc SCD (10 for healthy, 30 for s.s)
71 \text{s_M} = 30; \% number of macrophages on standy in the area (10 for healthy, 30 for s.s)

74 \% Parameter Labels
75 \text{PRCC_var} = \{'scd_type', 'BM', 'ps', 'mu_r', 'mu_s', 'mu_p', 'mu_n', 'mu_mas', 'mu_m', 'mu_nas', 'mu_am', 'bedh', 'beha', 'bead', 'bpha', 'bnha', 'bnat', ...
76 'knm2', 'knpm2', 'km2m2', 'kem2', 'bm2m1', 'mu_m2', 'mu_m1', 'km1m1', 'kc', 'kase', 'kep', 'kne', 'kasp', ...

213
t_end = 15; \% length of the simulations
tspan = (0:.01: t_end); \% time points where the output is calculated

%time_points = [1 .5*t_end]; \% time points of interest for the US analysis

% INITIAL CONDITION FOR THE ODE MODEL

% initialize with patient's baseline levels
% load file

voc_data = xlsread('VOC Data.xlsx', 1, 'A2:AU46');

save('voc_data.mat', 'voc_data');
% load(voc_data.mat)

% renaming data to use:
mydata = voc_data;

% choose rows (patient)
% gained platelets (negative number difference) :: 6, 14, 27, 35, {37}, 44
% lost platelets (positive number difference) :: 4, 11, 17, 21, 29, 30, 40

patient_ID = 37;

% Hgb, WBC count, platelet count, ANC, eosinophil, retic, RBC
baseline_patient_data = [mydata(patient_ID,35) mydata(patient_ID,36) mydata(patient_ID,38) mydata(patient_ID,39) mydata(patient_ID,40) mydata(patient_ID,37) mydata(patient_ID,43)];

% Hgb, WBC, platelet, ANC, eosinophil, retic
crisis_patient_data = [mydata(patient_ID,20) mydata(patient_ID,21) mydata(patient_ID,22) mydata(patient_ID,23) mydata(patient_ID,24) mydata(patient_ID,25)];

% specificity which columns with initial patient values

% NA0 = Baseline ANC = baseline_patient_data(:,4)
% PA0 = Baseline platlets = baseline_patient_data(:,3)
% AS0 = RBC * 100 * mu_s = baseline_patient_data(:,7)
% M20 = 0.04 * WBC = 0.04 * baseline_patient_data(:,2)
% M10 = 0.04 * WBC = 0.04 * baseline_patient_data(:,2)
% NT0 = Baseline ANC / 5 = baseline_patient_data(:,4) / 5
% NP0 = Baseline ANC / 3 = baseline_patient_data(:,4) / 3
% EA0 = Baseline platlets / 3 % (EA activates PA) = baseline_patient_data(:,3) / 3
% ED0 = 0.2 * (Baseline platlets / 3)
% AM0 = AS0

NA0 = baseline_patient_data(:,4);
PA0 = baseline_patient_data(:,3);
AS0 = baseline_patient_data(:,7)*100*mu_s;
M20 = .04*baseline_patient_data(:,2);
M10 = .04*baseline_patient_data(:,2);
NT0 = baseline_patient_data(:,4) / 5;
NP0 = baseline_patient_data(:,4) / 3;
EA0 = baseline_patient_data(:,3) / 3;
ED0 = .2*(baseline_patient_data(:,3) / 3);
AM0 = AS0;
kc = .0125; % background pro-inflamm cytokines

% STEADY STATE INITIAL CONDITIONS
% EA0 = 10000; % activated enothelial cells in vessel wall
% ED0 = 294; % damaged endothelial cells
% PA0 = 200; % activated platelets
% NA0 = 422; % activated neutrophils
% NT0 = 200; % number of neutrophil extracellular traps
% AS0 = 100; % initial number of dead sRBCs
% AM0 = 100; % initial number of adhesion molecules
% NP0 = 100; % initial number of dying neutrophils
% M10 = 150; % initial number of resting macrophages (pro inflammatory)
% M20 = 100; % initial number of monocyte derived macrophages (anti
1 inflammatory)
% kc = .0125; %background pro-inflamm cytokines

153 \( y_0 = [ \text{EAO EDO PAO NAO NTO ASO AMO NPO M10 M20}]; \)

154 y0 = [T0, T1, T2, V];

156 % Variables Labels
Y: The model outputs, $T \times N$, where $T$ is the number of time points and $N$ is the number of runs.

$s$: A single time point ordinal. If $T$ is the number of time points then $s$ is a single value in the range $[1, T]$, i.e. $1 \leq s \leq T$. For example if $T$ is 10 then $s$ is in the range $[1, 10]$, i.e. $1 \leq s \leq 10$.

PRCC_var: A cell array of string names of the $k$ varied parameters. This is from the settings file, and is in the Matlab workspace and result .mat file (Model_LHS.mat) after running Model_LHS.m.

y_var: A cell array of string names of the model outputs. This is from the settings file, and is in the Matlab workspace and result .mat file (Model_LHS.mat) after running Model_LHS.m, with name y_var_label.

For example:

PRCC_PLOT(LHSmatrix, V_lhs, 1, PRCC_var, y_var_label)
PRCC_PLOT(LHSmatrix, V_lhs, 2, PRCC_var, y_var_label)
function PRCC_PLOT(X, Y, s, PRCC_var, y_var)

Y=Y(s,:);
%Y=y;

[a k]=size(X); % Define the size of LHS matrix
Xranked=rankingN(X);
Yranked=ranking1(Y);

for i=1:k % Loop for the whole submatrices, Zi
    c1=['LHStemp=Xranked;LHStemp(:,',num2str(i),')= [];Z',num2str(i),'
        = [ones(a,1) LHStemp];LHStemp=[];';
    eval(c1);
end

for i=1:k
    c2=['[b',num2str(i),',bint',num2str(i),',r',num2str(i),']=
        regress(Yranked,Z',num2str(i),');'';
    c3=['[b',num2str(i),',bint',num2str(i),',rx',num2str(i),']=
        regress(Xranked(:,',num2str(i),'),Z',num2str(i),');'';
    eval(c2);
    eval(c3);
end

for i=1:k
    c4=['r',num2str(i)];
    c5=['rx',num2str(i)];
    [r p]=corr(eval(c4),eval(c5));
    a=['[PRCC , p-value] = ' '[', num2str(r) ', ', num2str(p) ']' ];
    % ' Time point=' num2str(s-1)];
    figure, plot((eval(c4)),(eval(c5)),'.'), title(a),...
legend(PRCC_var{i}),xlabel(PRCC_var{i}),ylabel(y_var);%

eval(c6);

% Define the filename for saving the plot
filename = sprintf('PRCC_plot_%s.png', PRCC_var{i});
%filename = sprintf('PRCC_plots%.png', PRCC_var);

% Specify the folder where you want to save the images
folder = '/Users/macuser/Desktop/PhD/Ch 4 VOC ODE/CHAPTER 4 analysis/LHS + PRCC';

% Save the plot as an image in the specified folder
saveas(gcf, fullfile(folder, filename));

% Close the current figure to avoid cluttering
close(gcf);
end
Quindel D. Jones, originally from Jackson, Mississippi, got a first hand experience of the health disparities prevalent in this country for Black Americans. In a city of food deserts, high stress policing, and underpaid labor, illness and disease is a common occurrence in the Deep South, USA. After watching her family develop disease conditions due to multiple factors, with very little explanation, understanding, or resources to combat their symptoms, she developed an initial interest in pathology, the study of the cause and effects of disease. After graduating high school, she began attending the Georgia Institute for Technology in 2015 with a major in Systems & Industrial Engineering. Though she was only there a year before her transfer to her alma mater Jackson State University, her love for systems and public health grew exponentially after her time there. Her time at Jackson State also exposed her to new opportunities and interests, learning of biostatistics and mathematical modeling through the Field of Dreams Conference and her REU at RIT. At the end of her undergraduate career she attended the Joint Math Meetings 2019, where she would meet her Ph.D. academic advisor, Dr. Rebecca Segal, through one of her presentations. Upon the realization that mathematics could be utilized for finding solutions within medicine and public health, she embarked on this journey to become a mathematical scientist and applied to Virginia Commonwealth University to pursue a Ph.D. in Systems Modeling & Analysis. This allowed her to learn more about mathematical biology and begin her research career studying disease dynamics and impact, particularly in Sickle Cell disease. She hopes to continue her post-graduate career modeling biological systems and using her skills as a mathematical scientist and public speaker for public health disparity research, with a particular interest in under-funded and under-researched diseases prevalent in the Black community.
References


[24] Prabhakar Deonikar and Mahendra Kavdia. “Low micromolar intravascular cell-free hemoglobin concentration affects vascular NO bioavailability in sickle cell disease: a computational analysis”.

224


[53] Huan Lei and George E. Karniadakis. “Probing vasoocclusion phenomena in sickle cell anemia via mesoscopic simulations”. In: Proceedings of the National Academy of Sciences 110.28 (July 2013), pp. 11326–11330. DOI: 10.1073/pnas.1221297110


[67] Navier-Stokes Equations. URL: https://www.grc.nasa.gov/www/k-12/airplane/nseqs.html


[81] RePORT. URL: https://report.nih.gov/funding/categorical-spending#/


[87] Richa Sharma et al. “Macrophage metabolic rewiring improves heme-suppressed efferocytosis and tissue damage in sickle cell disease”. In: *Blood* 141.25 (June 2023), pp. 3091–3108. ISSN: 0006-4971. DOI: [10.1182/blood.2022018026](https://doi.org/10.1182/blood.2022018026) (visited on 02/10/2024).


CPT: Pharmacometrics & Systems Pharmacology 10.7 (2021), pp. 696–708. ISSN: 2163-8306. DOI: 
10.1002/psp4.12638

lab to ward. Apr. 2023. URL: https://www.frontiersin.org/journals/oncology/articles/