

# TOWARDS AN UNDERSTANDING OF ACUTE TRAUMATIC COAGULOPATHY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Rachel Ilana Lassen LeCover

August 2020

© 2020 Rachel Ilana Lassen LeCover  
ALL RIGHTS RESERVED

# TOWARDS AN UNDERSTANDING OF ACUTE TRAUMATIC COAGULOPATHY

Rachel Ilana Lassen LeCover, Ph.D.

Cornell University 2020

In the United States, trauma is the leading cause of death for those under the age of 45, and the fourth leading cause of death for all Americans [121]. In the world as a whole, according to the World Health Organization, more than 1.3 billion people die from road traffic related injuries. Of people who are severely injured, approximately 50% live long enough to receive medical treatment, however, about a quarter of them will go on to develop acute traumatic coagulopathy (ATC), a condition which puts these people at four-fold higher risk of death. Acute traumatic coagulopathy arises when the coagulation and intrinsic immune responses misfire. This work seeks to understand how such dysfunction can occur. To understand ATC, we have built a small model of coagulation and fibrinolysis, and then embedded it into a physiologically based pharmacokinetic model of the human body. We have also explored the prediction of ATC from a data science prospective, using two emergency room datasets. Our modeling highlighted the importance of blood as a resuscitation fluid, a possible mechanism for tranexamic acid efficacy, and addressed how clotting capacity may be affected by acidosis and hypothermia, two common complications of injury. Our work to predict ATC from emergency room data demonstrated that logistic regression performs poorly on the problem, however, a recurrent neural network joined to a support vector machine may perform well at separating non-ATC from ATC patients.

## BIOGRAPHICAL SKETCH

Rachel was born and grew up in a small town along California's central coast, where she was a frequent visitor to the library. She was inspired to enter the field of chemical and biomolecular engineering by her AP Biology class, which included an experiment in which *E. coli* were transfected to produce green fluorescent protein. She graduated with honors from Johns Hopkins in 2014, earning a Bachelor's in Chemical and Biomolecular Engineering, with a minor in Applied Mathematics and Statistics. When Rachel isn't working on her thesis, she is likely to be found reading, baking, or circusing.



This document is dedicated to everyone who encouraged me to ask questions.

## ACKNOWLEDGEMENTS

A very large thanks to all of the Varner lab members, past and present who supported me during my one year at Purdue and my four years at Cornell. I would have not gotten through my first year graduate school without the support of my cohort at Purdue, so I thank you all for all of your help in teaching me thermodynamics, fluids, and math.

A huge thanks to all of the lovely people at Circus Culture, who gave me a creative outlet throughout my PhD.

The chapter related to machine learning and ATC began as a class project with Kevin Boehm for CS 6780, and without him, this project would not have gotten off the ground.

## TABLE OF CONTENTS

Biographical Sketch . . . . .	iii
Dedication . . . . .	iv
Acknowledgements . . . . .	v
Table of Contents . . . . .	vi
List of Tables . . . . .	ix
List of Figures . . . . .	xi
<b>1 Introduction</b>	<b>1</b>
1.1 Coagulation and Fibrinolysis . . . . .	2
1.1.1 Coagulation . . . . .	2
1.1.2 Fibrinolysis . . . . .	4
1.2 Complement . . . . .	6
1.3 Acute Traumatic Coagulopathy . . . . .	11
1.4 Assays . . . . .	18
1.4.1 PTT and PT . . . . .	18
1.4.2 INR . . . . .	19
1.4.3 ROTEM and TEG . . . . .	19
<b>2 Parameter Estimation via Dynamic Optimization with Particle Swarms</b>	<b>22</b>
2.1 Background . . . . .	22
2.2 Results . . . . .	25
2.2.1 DOPS explores parameter space using a combination of global methods. . . . .	25
2.2.2 DOPS minimized benchmark problems using fewer func- tion evaluations. . . . .	27
2.2.3 DOPS estimated the parameters of a human coagulation model. . . . .	31
2.2.4 Phase switching was critical to DOPS performance. . . . .	36
2.3 Discussion . . . . .	40
2.4 Conclusions . . . . .	42
2.5 Methods . . . . .	44
2.5.1 Optimization problem formulation. . . . .	44
2.5.2 Dynamic optimization with particle swarms (DOPS). . . . .	45
2.5.3 Multiswitch DOPS . . . . .	50
2.5.4 Comparison Techniques. . . . .	51
2.6 Tables . . . . .	52
<b>3 Reduced order modeling and analysis of the human complement sys-     tem</b>	<b>54</b>
3.1 Introduction . . . . .	55
3.2 Results . . . . .	60
3.2.1 Reduced order complement network. . . . .	61

3.2.2	Estimating an ensemble of reduced order complement models. . . . .	64
3.2.3	Global analysis of the reduced order complement model. .	66
3.3	Discussion . . . . .	72
3.4	Materials and Methods . . . . .	77
3.4.1	Formulation and solution of the complement model equations. . . . .	77
3.4.2	Estimating complement model parameters. . . . .	78
3.4.3	Complement model analysis. . . . .	81
<b>4</b>	<b>Kinetic Modeling of Coagulation and Fibrinolysis in Plasma</b>	<b>86</b>
4.1	Abstract . . . . .	86
4.2	Introduction . . . . .	87
4.3	Materials and methods . . . . .	92
4.3.1	Materials . . . . .	92
4.3.2	Assays . . . . .	93
4.3.3	Viscoelastometry . . . . .	93
4.3.4	Formulation and solution of model equations. . . . .	94
4.3.5	Estimating model parameters . . . . .	96
4.3.6	Global sensitivity analysis of model parameters . . . . .	99
4.3.7	Clustering of initial condition dependence . . . . .	99
4.3.8	Relearning Initial Conditions . . . . .	100
4.4	Results . . . . .	101
4.5	Discussion . . . . .	109
4.6	Conclusion . . . . .	111
<b>5</b>	<b>Kinetic Modeling of Coagulation and Fibrinolysis in Pregnancy</b>	<b>116</b>
5.1	Pregnancy and Coagulation . . . . .	116
5.2	Pre-Eclampsia . . . . .	118
5.3	Modeling Methodology . . . . .	121
5.4	Results . . . . .	123
5.4.1	Machine learning can classify pre-eclampsia . . . . .	131
5.5	Discussion . . . . .	132
5.6	Conclusions . . . . .	135
<b>6</b>	<b>Kinetic Modeling of Coagulation and Fibrinolysis in Whole Blood</b>	<b>136</b>
6.1	Introduction . . . . .	137
6.2	Materials and Methods . . . . .	141
6.2.1	Viscoelastometry Materials and Methods . . . . .	141
6.2.2	Formulation and solution of model equations . . . . .	142
6.2.3	Estimation of model parameters . . . . .	145
6.2.4	Identifiability Analysis of Initial Conditions . . . . .	146
6.3	Results . . . . .	147
6.4	Discussion . . . . .	152

6.5	Conclusions . . . . .	155
<b>7</b>	<b>Physiologically Based Pharmacokinetic Modeling of Trauma</b>	<b>156</b>
7.1	A Brief History of Physiologically Based Pharmacokinetic Modeling . . . . .	156
7.2	Mathematical Formulation . . . . .	157
7.3	Physiologically Based Pharmacokinetic Modeling of Trauma . . .	157
7.4	Modeling different treatments using Seheult bleeding model . . .	159
7.5	Simulation of Fibrinogen Supplementation . . . . .	165
7.6	Simulation of Cryoprecipitate Dosing . . . . .	171
7.7	Comparison of Whole Blood and Hextend Resuscitation in Monkeys . . . . .	172
7.8	Dosing with Tranexamic Acid (TXA) . . . . .	176
7.9	Comparison of Massive Transfusion Protocols . . . . .	183
7.10	Conclusions . . . . .	185
<b>8</b>	<b>Prediction of Acute Traumatic Coagulopathy via Machine Learning</b>	<b>186</b>
8.1	Machine Learning in Health Care . . . . .	186
8.2	Machine Learning in Trauma . . . . .	189
8.3	Data . . . . .	190
8.3.1	MIMIC-III . . . . .	190
8.3.2	Activation of Coagulation and Inflammation in Trauma Study Dataset . . . . .	192
8.4	Classification via Logistic Regression . . . . .	194
8.5	Classification via Support Vector Machine . . . . .	196
8.6	Classification via Recurrent Neural Network . . . . .	198
8.6.1	Prediction of Acute Traumatic Coagulopathy with Recurrent Neural Network . . . . .	200
8.7	Sequence Prediction . . . . .	205
<b>9</b>	<b>Conclusions</b>	<b>211</b>
<b>10</b>	<b>Appendix</b>	<b>217</b>
	<b>Bibliography</b>	<b>226</b>

## LIST OF TABLES

2.1	Table with optimization settings and results for the coagulation problem, the benchmarks and test functions using DOPS. For each problem the bounds on the parameter vector, the total number of function evaluations, the best initial objective value and the best final objective value are specified. Here <i>pnom</i> indicates the nominal or true parameter vector of the model. Nominal objective value represents the objective value using the true parameter vector or the nominal parameter vector. The CPU time is the time taken for the problem on a 2.4GHz Intel Xeon Architecture running Matlab 2014b. . . . .	52
2.2	. . . . .	53
4.1	. . . . .	114
7.1	Vascular resistances used in the body model. . . . .	167
7.2	A comparison of estimated human and rhesus macaque acute phase multipliers and lag times. . . . .	174
8.1	Accuracy of logistic regression for predicting ATC in the MIMIC-III dataset as a function of bin width. While the accuracy of the training and test sets are far above the base rate, one can observe they are very close to 1-base rate, which means if we simply predicted a patient's likelihood of having ATC at the first time point at the rate 1-base rate, we would perform nearly as well as logistic regression. . . . .	195
8.2	The calculated $F_1$ scores using a SVM on a sample of 500 MIMIC-III patients or ACIT patients at the first timepoint with various kernel types. A * denotes an ill-defined $F_1$ score as no patients were predicted to be in one of the classes. The $F_1$ scores reported are an average over five fold cross validation. . . . .	198
8.3	The estimated importance of the metaparameters in our chained LSTM-SVM model, $\pm$ one standard deviation, over 1000 metaparameter combinations. . . . .	204
8.4	The best metaparameters found via SMAC for binary classification of MIMIC-III patients using a chained LSTM-SVM via minimizing $1-F_1$ score, resulting in an average $F_1$ score of .894 over five fold cross validation. . . . .	204
8.5	Best metaparameters found after running SMAC for 908 configurations. . . . .	207
8.6	Single metaparameter importance as estimated by fAONVA via CAVE for sequence prediction after running SMAC on 908 model configurations with a mean squared error of .26. . . . .	207

8.7	Pairwise metaparameter importance as estimated by fAONVA for sequence prediction via CAVE after running SMAC. . . . .	208
-----	--	-----

## LIST OF FIGURES

1.1	A schematic of the coagulation cascade. Red lines denote inhibition, dashed lines indicate catalysis, and black lines represent reactions occurring. Inactive zymogens are shown in gray. Image taken from [283]. . . . .	3
1.2	A schematic of the fibrinolysis. Red lines denote inhibition. . . .	5
1.3	Locations of the genes that code for proteins involved in coagulation, fibrinolysis, and complement. The colored bands mark location, and are much wider than the true length of the genes for illustrative purposes. Chromosome 23 represents the X chromosome. The list of genes included was taken from [166], and their locations were taken from [79]. . . . .	9
1.4	A hive plot showing the interactions between transcription factors and coagulation related proteins. The list of genes included was taken from [166], and their associated transcription factors were taken from [79] by assembling a list of possible cis-regulatory elements for each gene, and then looking at the associated transcription factors for each cis-regulatory element, which led to a list of 328 transcription factors involved with 85 genes. . . . .	10
1.5	Typical (a) TEG and (b) ROTEM tracings. The differences in nomenclature are circled in red. Images adapted from [302]. . . .	19
2.1	Schematic of the dynamic optimization with particle swarms (DOPS) approach. <b>A:</b> Each particle represents an N dimensional parameter vector. Particles are given randomly generated initial solutions and grouped into different sub-swarms. Within each swarm the magnitude and direction of the movement a particle is influenced by the position of the best particle and also by its own experience. After every <b>g</b> number of function evaluations the particles are mixed and randomly assigned to different swarms. When the error due to the global best particle (best particle amongst all the sub-swarms) does not drop over a certain number of function evaluations, the swarm search is stopped and the search switches to a Dynamically Dimensioned Search with global best particle as the initial solution vector or candidate vector. <b>B:</b> The candidate vector performs a greedy global search for the remaining number of function evaluations. The search neighborhood is dynamically adjusted by varying the number of dimensions that are perturbed (in black) in each evaluation step. The probability that a dimension is perturbed decreases as the number of function evaluations increase. . . . .	26



2.2	Performance of DOPS and other meta-heuristics for the Ackley and Rastrigin functions. <b>A:</b> Mean scaled error versus the number of function evaluations for the 10-dimensional Ackley function. DOPS, DDS and ESS find optimal or near optimal solutions within the specified number of function evaluations. <b>B:</b> Mean scaled error versus the number of function evaluations for the 10-dimensional Rastrigin function. Nearly all the techniques find optimal or near optimal solutions within the specified number of function evaluations. <b>C:</b> Mean scaled error versus the number of function evaluations for the 300-dimensional Rastrigin function. DOPS is the only algorithm that finds an optimal or near optimal solution within the specified number of function evaluations. In all cases, the maximum number of function evaluations was $N = 4000$ . Mean and standard deviation were calculated over $\mathcal{T} = 25$ trials. A star denotes that the average value was less than $1\text{E-}6$ . . . . .	30
2.3	Schematic of the extrinsic and intrinsic coagulation cascade. Inactive zymogens upstream (grey) are activated by exposure to tissue factor (TF) following vessel injury. Tissue factor and activated factor VIIa (FVIIa) form a complex that activates factor X (fX) and IX (fIX). FXa activates downstream factors including factor VIII (fVIII) and fIX. Factor V (fV) is primarily activated by thrombin (FIIa). In addition, we included a secondary fV activation route involving FXa. FXa and FVa form a complex (prothrombinase) on activated platelets that converts prothrombin (fII) to FIIa. FIXa and FVIIIa can also form a complex (tenase) on activated platelets which catalyzes FXa formation. Thrombin also activates upstream coagulation factors, forming a strong positive feedback ensuring rapid activation. Tissue factor pathway inhibitor (TFPI) downregulates FXa formation and activity by sequestering free FXa and TF-FVIIa in a FXa-dependent manner. Antithrombin III (ATIII) inhibits all proteases. Thrombin inhibits itself binding the surface protein thrombomodulin (TM). The IIa-TM complex catalyzes the conversion of protein C (PC) to activated protein C (APC), which attenuates the coagulation response by the proteolytic cleavage of fV/FVa and fVIII/FVIIIa. . . . .	32
2.4	Error convergence rates of the nine different algorithms on the coagulation model. The objective error is the mean over $\mathcal{T} = 25$ trials. DOPS, SA, PSO and DOPS-PSO have the steepest drop in error during first 300 function evaluations. Thereafter the error drop in DDS and SA remains nearly constant whereas DOPS continues to drops further. In the allotted budget of function evaluations ESS produces a modest reduction in error. At the end of 4000 function evaluations DOPS attains the lowest error. . . . .	34

2.5	Model fits and predictions on experimental data using DOPS. (a) The model parameters were estimated using DOPS. Solid black lines indicate the simulated mean thrombin concentration using parameter vectors from 25 trials. The grey shaded region represents the 99% confidence estimate of the mean simulated thrombin concentration. The experimental data is reproduced from the synthetic plasma assays of Mann and co-workers. Thrombin generation is initiated by adding Factor TF/VIIa (5nM (blue) and 5pM (red)) to synthetic plasma containing 200 $\mu\text{mol/L}$ of phospholipid vesicles (PCPS) and a mixture of coagulation factors (II,V,VII,VIII,IX,X and XI) at their mean plasma concentrations. (b) The parameter estimates that were obtained using DOPS were tested against data that was not used in the model training. Solid black lines indicate the simulated mean thrombin concentration using parameter vectors from $\mathcal{T} = 25$ trials. The grey shaded region represents the 99% confidence estimate of the mean simulated thrombin concentration. The experimental data is reproduced from the synthetic plasma assays of Mann and co-workers. Thrombin generation is initiated by adding Factor VIIa-TF (500pM - Blue, 50pM - Pink and 10pM - purple, respectively) to synthetic plasma containing 200 $\mu\text{mol/L}$ of phospholipid vesicles (PCPS) and a mixture of coagulation factors (II,V,VII,VIII,IX,X and XI) at their mean plasma concentrations.	35
2.6	Influence of the switching strategy and sub-swarms on DOPS performance for the coagulation model. DOPS begins by using a particle swarm search and then dynamically switches (switch region), using an adaptive switching criteria, to the DDS search phase. We compared the performance of DOPS with and without DDS for different sub-swarm searches to quantify the effect of number of sub-swarms and DDS. We used one, two, four, five and eight sub-swarms, with a total of 40 particles divided evenly amongst the swarms. The results presented are the average of $\mathcal{T} = 25$ trials with $N = 4000$ function evaluations each. The convergence rates with higher swarm numbers is typically higher but there is no pronounced difference amongst four, five and eight. The multi-swarm with without DDS saturates while DOPS shows a rapid drop due to a switch to the DDS phase.	37

2.7	Comparison of DOPS and Multiswitch DOPS Performance of DOPS and Multiswitch DOPS on the CHO metabolism problem (a), the Eggholder function (b), the 100 dimensional Styblinski-Tang function (c) and the coagulation problem (d). Both methods have the same initial decrease in error, but as the number of function evaluations increases, multiswitch DOPS produces a larger decrease in error. The results presented are the average of $\mathcal{T} = 250$ trials with for the CHO metabolism problem and $\mathcal{T} = 250$ trials on the Eggholder and Styblinski-Tang functions with $\mathcal{N} = 250$ function evaluations each, and $\mathcal{T} = 25$ trials for the coagulation problem. . . . .	38
3.1	Reduced order complement model training. An ensemble of model parameters were estimated using multiobjective optimization from C3a and C5a measurements with and without zymosan [227]. The model was trained using C3a and C5a data generated from the alternative pathway (A–B) and lectin pathway initiated with 1 mg/ml zymosan (C–D). The solid black lines show the simulated mean value of C3a or C5a for the ensemble, while the dark shaded region denotes the 99% confidence interval of mean. The light shaded region denotes the 99% confidence interval of the simulated C3a and C5a concentration. All initial conditions were assumed to be at their physiological serum levels unless otherwise noted. . . . .	60
3.2	Simplified schematic of the human complement system. The complement cascade is activated through three pathways: the classical, the lectin, and the alternate pathways. Complement initiation results in the formation of classical or alternative C3 convertases, which amplify the initial complement response and signal to the adaptive immune system by cleaving C3 into C3a and C3b. C3 convertases further react to form C5 convertases which catalyze the cleavage of the C5 complement protein to C5a and C5b. C5b is critical to the formation of the membrane attack complex (MAC), while C5a recruits an adaptive immune response. . . . .	63

3.3	Global sensitivity analysis of the reduced order complement model. Sensitivity analysis was conducted on the two objectives used for model training. <b>A:</b> Sensitivity of the C3a and C5a residual w/o zymosan. <b>B:</b> Sensitivity of the C3a and C5a residual with 1 mg/ml zymosan. The bars denote the mean total sensitivity index for each parameter, while the error bars denote the 95% confidence interval. <b>C:</b> Pathways controlled by the sensitivity parameters. Bold black lines indicate the pathway involves one or more sensitive parameters, while the red lines show current therapeutics targets. Current complement therapeutics were taken from the review of Morgan and Harris [229]. . . . .	67
3.4	Reduced order complement model predictions. Simulations of C3a and C5a generated in the lectin pathway using 0.1 mg/ml, 0.01 mg/ml, and 0.001 mg/ml zymosan were compared with the corresponding experimental measurements. The solid black lines show the simulated mean value of C3a or C5a for the ensemble, while the dark shaded region denotes the 99% confidence interval of mean. The light shaded region denotes the 99% confidence interval of the simulated C3a and C5a concentration. All initial conditions were assumed to be at their physiological serum levels unless otherwise noted. . . . .	79
3.5	Pairwise sensitivity and clustering of complement model parameters in the presence of 1 mg/ml zymosan. The response of the complement model was calculated for each parameter combination following a 10% increase in parameter combinations in the presence of 1 mg/ml zymosan. The model parameters were clustered into high (blue), medium (red) and low (green) response clusters based upon the euclidian distance between the perturbed and nominal system state. . . . .	82
3.6	Robustness analysis of the complement model. Robustness coefficients were calculated for a 50%, 90% and 99% reduction in C3, C5, or C3 and C5 initial conditions. <b>A:</b> Mean robustness index for C3a and C5a generated from the alternate pathway (w/o zymosan). <b>B:</b> Mean robustness index for C3a and C5a generated from the lectin and alternate pathway (1 mg/ml zymosan). The color describes the degree of reduction of C3a or C5a following the network perturbation. Robustness coefficients were calculated using all parameter sets with Pareto rank less than five (N = 65). Mean robustness values were reported. . . . .	83

4.1	Schematic of the fibrinolysis model. Thrombin, generated by the presence of tissue factor in the coagulation model converts fibrinogen into its active form, fibrin. When tPA is present, it activates plasminogen from its inactive form to plasmin, which is capable of degrading plasmin. PAI-1 inhibits tPA and thus inhibits fibrin degradation. . . . .	91
4.2	Fibrinolysis model training and prediction. <b>A</b> Model performance on test and training data. A redder block indicates that the simulations came closer to matching the experiential measurements. <b>B</b> Top: Performance of the thrombin generation portion of the model. Experimental measurements are shown in the black dots, the thin black curve represents the mean model performance and the shaded gray represents the 95% confidence interval for $N=266$ simulations. Bottom: Performance of the fibrin generation and fibrinolysis portion of the model. The red curves represent ROTEM curves measured in two different patients at $tPA = 2$ nM and $tPA = 4$ nM, the black lines represent the simulated ROTEM curves for these patients in those cases, and the shaded dark gray represents the 99% confidence interval for the $tPA = 2$ nM case, and the shaded light gray represents the 99% confidence interval for the $tPA = 4$ nM case for $N=266$ simulations.	98
4.3	Model predictions on the left out plasma. The experimentally measured trajectory is shown in red, and the mean of $N=64$ parameter sets found based on the other plasmas is shown in black, with the 95% confidence interval shaded. The $tPA = 0$ case is shown in blue, the $tPA = 2$ case is shown in light gray, the $tPA = 4$ case is shown in dark gray, and the $tPA = 8$ case is shown in slate gray. . . . .	104
4.4	Global sensitivity analysis of the fibrinolysis model. Sensitivity analysis was performed by the model of Sobol with respect to the area under the ROTEM curve for $N = 23,200$ parameter sets. 0% dilution means all factors were at physiological level, 80% dilution means that the factors were diluted to 80% of physiological level. . . . .	106
4.5	Simulated ROTEM curve for repeated solving of the same inverse problem. The black curve shows the true ROTEM curve, where the blue curve is the average over $N=100$ attempts to find the initial conditions. The shaded blue area is the 95% confidence interval of the mean. The gray curves are the other ROTEM curves generated by the estimated initial conditions. . . . .	107
4.6	Average difference in target and found initial and experimental conditions from solving the inverse problem. The thin lines represent the standard error of the mean over $N = 50$ repetitions of solving the inverse problem. . . . .	108

4.7	Clustergram analysis of fibrinolysis model response to changes in initial conditions. Initial conditions were increased to two times their nominal values in a pairwise fashion. For diagonal entries, the the factor level was perturbed to two times its nominal level. . . . .	113
4.8	Morris analysis of fibrinolysis model response. We performed global sensitivity analysis to determine which parameters controlled the shape of the ROTEM curve output by the model. We generated parameters using the Morris method, with the parameters bounded by the minimum and maximum values they took in the rank one parameter sets generated by JuPOETS [230]. A total of $N(p + 1)$ parameter sets were generated, with $N = 1000$ and $p = 47$ in our case, for a total of 48,000 model evaluations. We ran the model with the generated parameter sets, and then transformed the ROTEM curves using fPCA (functional principle component analysis), as implemented by the R package fdapace, to extract the principle component scores. This technique allowed us to remove the time dimension and to observe how the changes in parameters affected the shape of the ROTEM curve [322]. We then used the principle component scores to calculate $\mu^*$ , the absolute value of the mean of the elementary effect, and $\sigma$ , the variance of the elementary effect. We opted to use the absolute values of the means of the elementary effects so that effects with opposite signs would not cancel each other out and artificially decrease the estimate of the mean of the elementary effect [48]. We then scaled $\mu^*$ and $\sigma$ so that the smallest value became zero and the largest value became one, using the same scaling as used for the scaled experimental measurements. This method sensitivity analysis was conducted using the SALib module encoded in the Python programming language [135]. We then used k-means clustering to group the parameters into four clusters, and colored the graph by cluster color. We used four clusters as it resulted in the best average silhouette score [281]. . . . .	115
5.1	The evolution of of the state of the coagulation system throughout pregnancy in one patient. The more the solid the lines, the further along the woman is in her pregnancy. (a) Thrombin generation curves, in plasma, as initiated with 5 pM trigger, in plasma (blue) and in plasma supplemented with thrombomodulin (green) (b) ROTEM trajectory as initiated with 5 pM trigger, in plasma (orange) and in plasma spiked with 4 nM tPA (pink).	118

5.2	Thrombin generation predictions for the test, train, and validation sets. The experimentally measured thrombin trajectories are shown in the heavy lines, the average trajectory over 20 estimated parameters sets is shown in the thin line with a shaded 95% confidence interval. The dashed lines represent the trajectories from each parameter set. Thrombin generation trajectories without additional thrombomodulin are shown in blue, thrombin generation trajectories with supplemental thrombomodulin are shown in green. (a) and (c) represent the training data (5 patients, timepoints 1 and 3. (c) represents the test set, the same five patients used for training, at timepoint 2. (d) Validation of thrombin generation portion of the model at timepoint 1 in 5 different patients (e) Validation of thrombin generation portion of the model at timepoint 2 in 5 different patients (f) of thrombin generation portion of the model at timepoint 3 in 5 different patients. . . . .	126
5.3	ROTEM curves on the training set. The first and third time points were used to estimate parameters, and the second time point served as a test. The heavy curves are the experimental measurement, the thin lines are the average prediction over $N = 13$ parameter sets, and the shaded area denotes the 95% confidence interval. ROTEM trajectories with no tPA are orange, ROTEM trajectories with 4 nM tPA are in yellow. The first column represents the first time point, the second column the second time point, the third column the third timepoint. . . . .	127
5.4	ROTEM curves on the validation set. The heavy curves are the experimental measurement, the thin lines are the average prediction over $N = 13$ parameter sets, and the shaded area denotes the 95% confidence interval. ROTEM trajectories with no tPA are orange, ROTEM trajectories with 4 nM tPA are in yellow. The first column represents the first time point, the second column the second time point, the third column the third timepoint. . . . .	128
5.5	Preliminary sensitivity analysis using Sobol total order sensitivity coefficients using $N = 100$ , and $p = 27$ resulting in the evaluation of $2N(p + 1) = 5600$ initial condition combinations evaluated. The blue bars represent the total order sensitivity coefficients, the error bars represent the confidence interval. Concentrations, when nominally non-zero were perturbed between 50% and 150% of nominal, with the exception of progesterone, which was allowed to vary between the minimum and maximum amounts measured. When the nominal concentration of a species was 0, it was perturbed between 0 and 1 nM. . . . .	129
5.6	Preliminary sensitivity analysis using Sobol total order sensitivity coefficients using $N = 500$ , and $p = 54$ resulting in the evaluation of $2N(p + 1) = 55000$ initial condition combinations evaluated. The blue bars represent the total order sensitivity coefficients, the error bars represent the confidence interval. Parameters were perturbed between 50% and 150% of the average value found in the best 13 parameter families. . . . .	130
5.7	Predictions made by consensus clustering on PCA data . . . . .	132
6.1	Calculated error of the model during leave one out cross validation. The dots show the average error on the 3 patients used in the training set, the crosses show the average performance on the left out fourth patient. . . . .	145

6.2	Simulation and analysis of the reduced order model. <b>A:</b> Prediction performance of the reduced order model on four unseen patients. The thick black line is the true ROTEM measurement, the thin black line represents the model prediction, and the shaded gray area represents the 95% confidence interval for the $N = 24$ best parameter sets. <b>B:</b> Total order sensitivity coefficients with respect to area under the ROETM curve. The gray whiskers show the 95% confidence interval for the calculated coefficients. . . . .	148
6.3	Prediction performance of the reduced order model on four unseen patients after adjustment of initial conditions. The thick black line is the true ROTEM measurement, the thin black line represents the model prediction, and the shaded gray area represents the 95% confidence interval for the $N = 24$ best parameter sets. . . . .	149
6.4	Model performance for reverse engineered initial conditions. <b>A:</b> The curves generated from minimizing the objective in 6.13. The black line shows the original generated curve, the gray curves are the curves found from optimization, the blue curve shows the average of the estimated curves, and the shaded area represents the 95% confidence interval for $N = 25$ simulations. <b>B:</b> The differences in initial conditions found as we attempted to relearn initial conditions based on ROTEM curve metrics. The bars represent the average difference, and the error bars represent the standard error of the mean for $N = 10$ simulations. . . . .	150
7.1	(a) Bleeding rate as a function of time under the Sheult et al schedule. (b) Hemostatic factor as a function of blood pressure. As blood pressure decreases, so does the hemostatic factor. . . .	160
7.2	The simplified body used in the simulations. . . . .	161
7.3	(a) The resulting maximum clot amplitude measurement from drawing blood from the simulated patient's veins throughout injury, pre-hospital, emergency room, surgical and recovery phases. The dots represent the mean maximum clot amplitude, and the shaded area represents the 95% confidence interval over fifteen selected parameter families. (b) The resulting maximum clot amplitude measurement from drawing blood from the simulated patient's veins throughout injury, pre-hospital, emergency room, surgical and recovery phases while the patient suffers from acidosis and/or hypothermia. The dots represent the mean maximum clot amplitude, and the shaded area represents the 95% confidence interval over fifteen selected parameter families.	162



7.4	Experimental measurements are shown as dots, with the IQR represented as error bars, and the simulated mean levels are shown as lines, with the 95% confidence interval shaded over a family of 24 kinetic parameter sets. (a) Fibrinogen concentrations during the first 48 hours following injury. (b) Platelet count during the first 48 hours following injury. (c) ATIII concentration during the first 48 hours following injury. . . . .	170
7.5	Fibrinogen concentrations over the first day following injury with and without cryoprecipitate supplementation. The reported concentrations from [74] are shown in dots with error bars, the simulated concentrations are shown as lines. . . . .	172
7.6	The pink corresponds to initial resuscitation with blood, and the blue corresponds to initial resuscitation with Hextend. Experimental data shown as dots, simulations shown as lines with one standard deviation shaded. (a) Fibrinogen (b) ATIII (c) FII (d) FX (e) FV . . . . .	175
7.7	TXA concentration in the first day after trauma. The 75 kg patient receives 1 gram of TXA at 90 minutes, and then another 1 gram in the following 8 hours. . . . .	177
7.8	Fibrin concentration within the wound compartment. . . . .	178
7.9	The effects of dosing a trauma patient with TXA over the first 25 hours following injury at time zero. The dashed blue lines indicate a patient that did not receive TXA, and the solid pink lines indicate a patient who received TXA. (a) Fibrin concentration within the wound compartment. (b)Fibrinogen concentration within the vein compartment. . . . .	178
7.10	The effects of dosing a trauma patient with TXA over the first 25 hours following injury at time zero with the initial tPA concentration set to ten times the nominal concentration. The dashed blue lines indicate a patient that did not receive TXA, and the solid pink lines indicate a patient who received TXA. (a) Fibrin concentration within the wound compartment. (b)Fibrinogen concentration within the vein compartment. . . . .	179
7.11	The effects of dosing a trauma patient with TXA over the first 25 hours following injury at time zero both with the initial tPA concentration set to ten times the nominal concentration. The dashed blue lines indicate a patient that received his first dose of TXA three hours and half hours after injury, and the solid orange line indicates a patient who received his first dose of TXA 90 minutes following injury. (a) Fibrin concentration within wound compartment (b) Fibrinogen concentration within vein compartment (c) plasmin concentration within vein compartment (d) Plasminogen concentration in the vein compartment . . . . .	181

7.12	Simulated maximum ROTEM amplitude under Stanford (red) or Parkland (yellow) resuscitation schedules. The mean amplitude is shown as a point every 30 minutes, and the shaded area represents one standard deviation over $N = 6$ parameter families. . .	184
8.1	The details of the patient selection process. Beginning from all the adult patient admissions in MIMIC-III, we selected only patients who were in the ICU for unplanned reasons and had at least one INR measurement. . . . .	191
8.2	A graphical representation of how patient feature vectors were constructed. The Xs denote measurements that were available at time point $t$ , and the empty boxes represent that that measurement was not available at a time point for a patient. This missingness pattern was used to construct masks to describe the missingness of data. . . . .	191
8.3	A low dimensional representation of the weights of the LSTM via tSNEs which was fairly successful at separating the patients into two groups. The red dots represent patients who developed ATC at some point within their stay, the blue ones did not. . . .	203
8.4	A generic schematic of an encoder-decoder model, with the low dimensional representation shown in blue. The inputs are $x_i$ , and the outputs are $y_i$ for timepoints $1, 2, \dots, T$ where $T$ is the length of the sequence. . . . .	206
10.1	<b>(Data fits for CHO metabolism problem)</b> Pseudo-experimental data (red x) vs. optimal solution obtained using DOPS (solid blue lines) for the 44 observed states. X axis: time [s]; Y axis: metabolite concentrations [mM]. . . . .	217
10.2	<b>(Data fits for S.cerevisiae metabolism problem)</b> Pseudo-experimental data (red x) vs. optimal solution obtained using DOPS (solid blue lines) for the 13 observed states. X axis: time [s]; Y axis: metabolite concentrations [mM]. . . . .	218
10.3	<b>(A)</b> Difference between nominal and optimal parameters for genome wide kinetic model of <i>S.cerevisiae</i> with 1759 unknown parameters using DOPS. <b>(B)</b> Difference between experimental (measured) data and data simulated with optimal parameters for genome wide kinetic model of <i>S.cerevisiae</i> with 1759 unknown parameters using DOPS. <b>(C)</b> Difference between nominal and optimal parameters for metabolic model of Chinese Hamster Ovary Cells (CHO) cells with 117 parameters using DOPS. <b>(D)</b> Difference between experimental (measured) data and data simulated with optimal parameters for metabolic model of Chinese Hamster Ovary Cells (CHO) cells with 117 parameters using DOPS. . . . .	219

10.4	Comparison of the runtime of the different optimization methods used for comparison with $\mathcal{T} = 25$ trials per method. All methods used take about the same amount of time to perform 4,000 function evaluations on the coagulation problem, as this problem is very stiff, so the majority of the time is spent solving the system of differential equations. . . . .	220
10.5	Mean convergence curves for different metaheuristics for (a) Ackley 300 dimensional and (b) Rastrigin 300 dimensional with $\mathcal{T} = 25$ trials per method. DOPS not only finds a better solution than any other technique, it finds it with fewer function evaluations . . . . .	222
10.6	Mean convergence curves for DOPS and ESS for the (a) CHO model and (b) the coagulation model with $\mathcal{T} = 25$ trials per method.	223
10.7	Variability analysis in best objective value for $\mathcal{T} = 25$ trials. (a) Ackley 300 dimensional (b) Rastrigin 300 dimensional (c) CHO model (d) coagulation. . . . .	224
10.8	Dispersion curves for DOPS on (a) CHO model (b) coagulation with $\mathcal{T} = 25$ trials per problem . . . . .	225

## CHAPTER 1

### INTRODUCTION

In the United States, trauma is the leading cause of death for those under the age of 45, and the fourth leading cause of death for all Americans [121]. The costs of trauma are high, costing approximately \$355 billion dollars, between the costs of treatment and lost productivity [110]. The number of deaths due to traumatic injury is rising in the United States with the number of trauma deaths increasing 22.8% from 2000 to 2010 (during a period which the US population only increased by 9.7%) [261]. Furthermore, traumatic injury is the leading cause of years of life lost in the United States with an estimated 3109 years of life lost per 100,000 members of the population [126]. On a global scale, the World Health Organization predicts that the global burden of deaths due to conflict, road injury, and homicide will increase at least until 2060 [205].

Of those suffering traumatic injury, 50-60% of those die immediately after injury, but the remaining 40-50% of deaths occur following arrival at a hospital with potentially treatable injuries [313]. Of these patients who survive long enough to be transported to a hospital, about a third deaths were due to hemorrhage, a potentially treatable and survivable condition [335]. Numerous studies have shown that patients who present with acute traumatic coagulopathy (ATC), a dysfunctional coagulation system, upon admission, are less likely to survive [196, 38]. While a universal test ATC has yet to be developed, it can be recognized from a prolonged prothrombin time or activated partial thromboplastin time upon hospital admission [49].

## 1.1 Coagulation and Fibrinolysis

### 1.1.1 Coagulation

The coagulation system, when functioning properly, holds two processes in a delicate balance-the production of enzymes that build clots and those that break them down. Following traumatic injury, the endothelium is damaged, exposing tissue factor, a transmembrane protein that is mainly expressed within the hemostatic envelope. When tissue factor becomes exposed to circulating blood it is activated [106]. Tissue factor serves as the initiator to the coagulation cascade, which produces thrombin, which converts fibrinogen to its active form, fibrin, which can polymerize into a clot [198]. This clot is then broken down through a process called fibrinolysis, in which tPA (tissue plasminogen activator) and uPA (urokinase plasminogen activator) convert plasminogen into its active form, plasmin, which degrades the fibrin containing clot [53]. In greater detail, tissue factor binds to FVIIa, forming the “trigger” complex, which then converts FX into its active form, FXa. FXa can then convert prothrombin (FII) to its active form (FIIa). Thrombin is a key enzyme in the coagulation cascade, as it can activate FVIII, FXI, FX, platelets, as well as itself, leading to a positive feedback loop where a very small amount of initial thrombin can lead to a large amount of thrombin generation. The thrombin generation process has a built-in negative feedback-thrombin complexes with thrombomodulin, and this complex activates protein C, which then inhibits the activities of FVa and FVIIIa [101]. Antithrombin (ATIII) inhibits the coagulation cascade at several points, including the conversion of prothrombin to thrombin, as well as the formation of the prothrombinase complex (FV-FX)a [87]. The overall coagulation cascade

is quite complex, with multiple positive and negative feedback loops, some of which are shown in Figure 1.1.

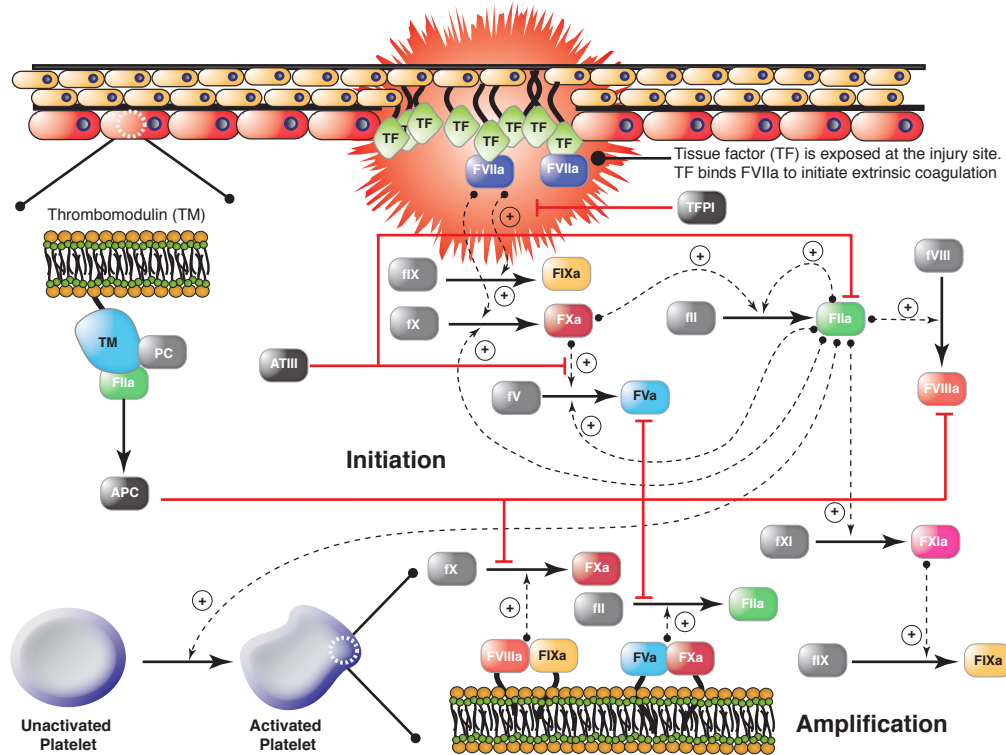


Figure 1.1: A schematic of the coagulation cascade. Red lines denote inhibition, dashed lines indicate catalysis, and black lines represent reactions occurring. Inactive zymogens are shown in gray. Image taken from [283].

In the absence of injury, endothelial cells in healthy vasculature oppose clotting through the secretion of plasminogen activators and through the production of thrombomodulin, a glycoprotein which promotes the activation of Protein C, an a protein which impedes coagulation [367].

Coagulation is a complex biochemical process, which mathematical models can help us to understand and to identify key therapeutic targets. A number of kinetic models exist that describe the thrombin generation and degradation process. One of the first models considered coagulation as a set of first and second

order reactions, each of which amplified the previous one, and completely neglected any negative feedback [180]. Martonrana and Moro improved upon this model by adding negative feedback terms, but this model almost completely neglects the biochemistry of coagulation and is completely linear [203]. Khanin and Senov developed the first non linear model of coagulation, containing five differential equations, and examined its fixed points to conclude that if the system was not sufficiently stimulated, no fibrin would be produced [170]. One of the best known models, the the Hockin-Mann model consists of 34 differential equations and 42 rate constants [143]. One of the largest and most complex models of coagulation was developed by the Varner lab, containing 92 differential equations [193]. However, if we wish to embed a thrombin generation model inside a model that captures the physiology of traumatic injury, it must be small enough to solve quickly. To that aim, the Varner lab has developed a reduced order model of thrombin generation that only contains five differential equations which mostly captured the dynamics of thrombin generation [285].

### **1.1.2 Fibrinolysis**

Fibrinolysis refers to the process of breaking down a clot, of which a key component is a protein named fibrin. The thrombin produced by the coagulation cascade converts fibrinogen into fibrin which then polymerizes into a clot with the assistance of FXIIIa. The clot is later broken down by plasmin, which can be activated by either tPA or uPA, enzymes with a short half-lives that can be inhibited by plasminogen activator inhibitor-1 (PAI-1) [54]. Thrombin can help stabilize the clot through its activation of thrombin activated fibrinolysis inhibitor (TAFI), which removes a C-terminal lysine and arginines from fibrin, reducing

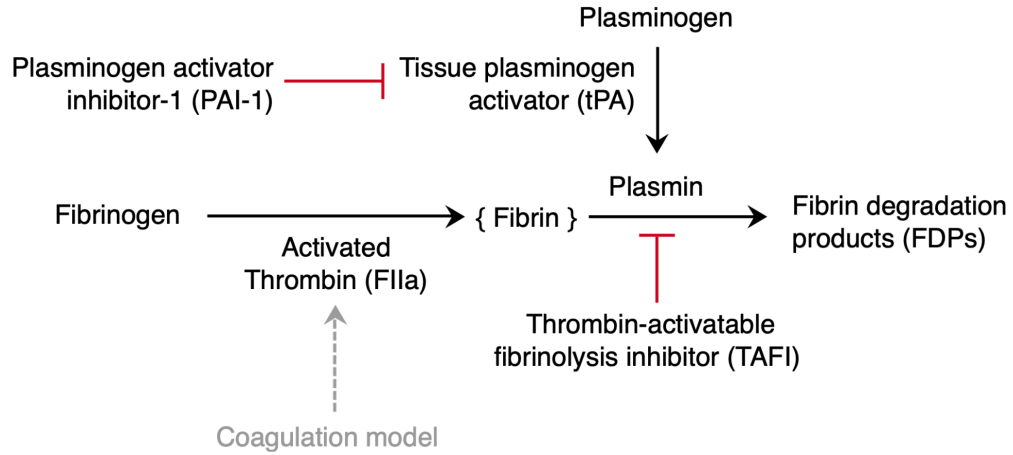


Figure 1.2: A schematic of the fibrinolysis. Red lines denote inhibition.

the possible number of plasmin binding sites.

Fibrinolysis and fibrin generation has been modeled at a variety of scales. Longstaff and Thelwell proposed a very simple model for fibrinolysis, in which the process was represented as two steps, with plasminogen being converted to plasmin by tPA, and fibrin degraded by plasmin [191]. While computationally easy to evaluate, this model greatly oversimplifies the process, and completely neglects the role of PAI-1. Fibrinolysis has been previous modeled in great detail through a 3D stochastic multi-scale model which predicted difference in lysis speeds based on clot morphology, but this model focuses only clot lysis rather than formation and degradation [15]. It has also been modeled as a mass-spring system to explore the effect of tension on lysis, revealing that with higher tension, fewer fibers (a bundle of polymerized fibrinogen) need to be cleaved to remove fibrin from a region [73]. Reifman et al modeled fibrin generation with mass action kinetics using 80 ODEs to investigate the efficacy of different prothrombin complex concentrates [218]. They used an interesting method to



model the complex kinetics of TAFI inhibiting fibrinolysis: they empirically fit a curve to the clot lysis time as a function of TAFI concentration, and then correspondingly reduced the t-PA effective concentration to account for the effects of TAFI. While this model decently predicted the final fibrin levels, it did a poor job of capturing the shape of the fibrin generation curves.

## 1.2 Complement

Complement, which takes its name from the fact it appears to “complement” the activity of antibodies in killing bacteria, is a component of the innate immune response, which can be initiated in one of three known ways [351]. The classical pathway initiates complement through the binding of the C1 complex to antibody bound to antigens on the surface of bacteria, which then cleaves C4 and C2 so that their fragments can form a C3 convertase, C4bC2a [288]. The alternative pathway does not require the binding of antibodies to pathogens to be triggered, rather, it is triggered through a process called “tickover”, where the thioester inside of C3 slowly undergoes hydrolysis, resulting in a functionally active C3 molecule [232]. The lectin pathway initiates when a mannose binding lectin binds to lectin on the surface of a bacterial cell, along with mannose-binding lectin-associated proteases 1 and 2, and this assemble acts as a C3 convertase, cleaving C3 into its active fragments [351]. Once C3b has been formed by any of these pathways, it acts as a C5 convertase, cleaving C5 into a and b fragments, of which C5b will assemble along with other complement associated proteins (C6, C7, C8, and C9) into the membrane attack complex, the final output of the complement system [231]. The membrane attack complex is believed to form a large  $\beta$  barrel pore, leading to cell death through osmotic flux [23].

The kinetic modeling of complement has not explored in as much detail as that of thrombin generation owing to the dearth of time series data [282]. Zewde et al developed a fairly large model of complement, consisting of 107 differential equations [370], however, they did not validate their model against experimental data. Korotaevskiy and coworkers developed a large model of two of the classical and alternative pathways of complement and validated it against the lysis rates of *B. burgdorferi* through the assembly of the membrane attack complex [174]. The model developed by the Varner lab includes initiation of all three pathways in only 18 differential equations with 28 parameters [282]. This model was trained against time series data of C5a and C3a and validated on unseen time series C5a and C3a data with varying levels of zymosan A, a complement pathway activator.

Although the interplay between enzymes in the coagulation cascade and the cross talk between coagulation and complement are exquisitely choreographed, it is interesting to note that physically, the genes that encode the proteins involved in these systems are scattered throughout the human genome, as seen in Figure 1.3. This dispersed layout raises questions as to how all of these genes are activated in the correct manner to produce more or less protein as needed to respond to injury. Despite the physical dispersed layout of the coagulation and complement genes throughout the human chromosomes, many of them are thought to share common transcription factors, which can be seen upon careful examination of Figure 1.4. For example, the transcription factor ATF3 (a member of the activating transcription factor family), which is hypothesized to be part of the cellular stress response, serves as a transcription factor for tissue factor, FXIII, FVIII, FVII, FV, prothrombin, von Willebrand factor, components of fibrinogen, to name just a few of the proteins in which it may play a role

in transcription [129]. The large number of edges in Figure 1.4 is indicative of the high connectivity between these proteins and their associated transcription factors, with the average degree of the graph exceeding 19. The exact mechanisms of how the liver knows how to produce more coagulation factors and fibrinogen after trauma have not yet been precisely defined, but there is some evidence that the presence of fibrin degradation products may stimulate fibrinogen production [113] and that fibrinogen production can be increased and decreased by microRNAs pulled from a library of human microRNAs[111]. The term "genomic storm" has been used to describe the changes that occur after traumatic injury, with more than 5,000 genes in leukocytes whose expression changed at least two fold after injury, a staggering number [366].<sup>1</sup>

---

<sup>1</sup>To put this number in perspective, *E. coli* is presently thought to have 4,401 genes [300].

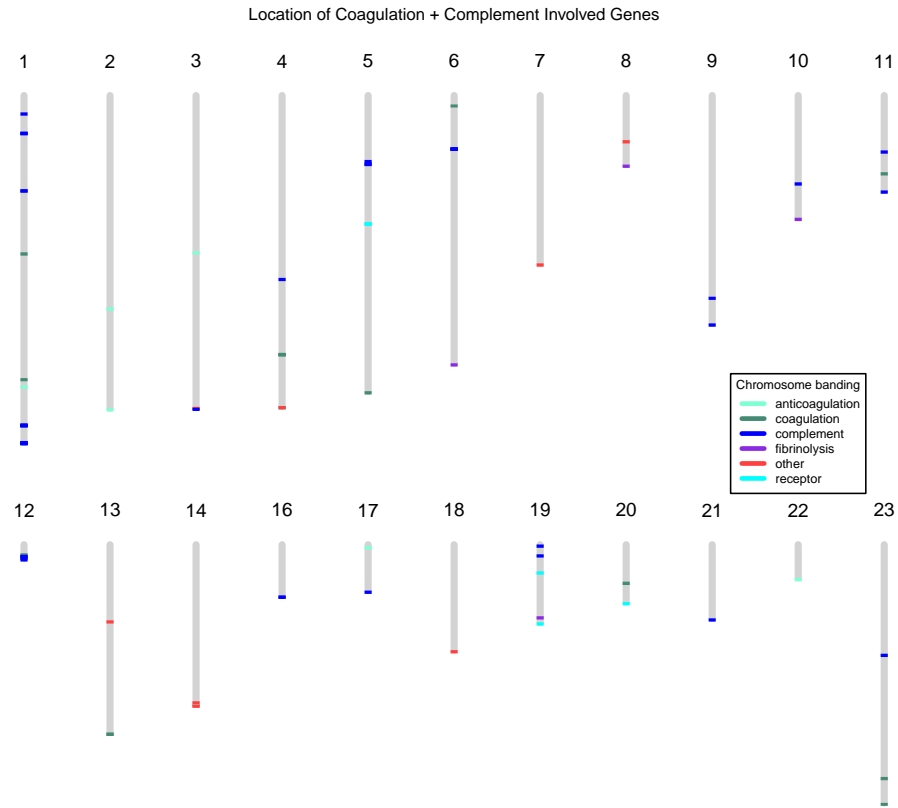


Figure 1.3: Locations of the genes that code for proteins involved in coagulation, fibrinolysis, and complement. The colored bands mark location, and are much wider than than the true length of the genes for illustrative purposes. Chromosome 23 represents the X chromosome. The list of genes included was taken from [166], and their locations were taken from [79].

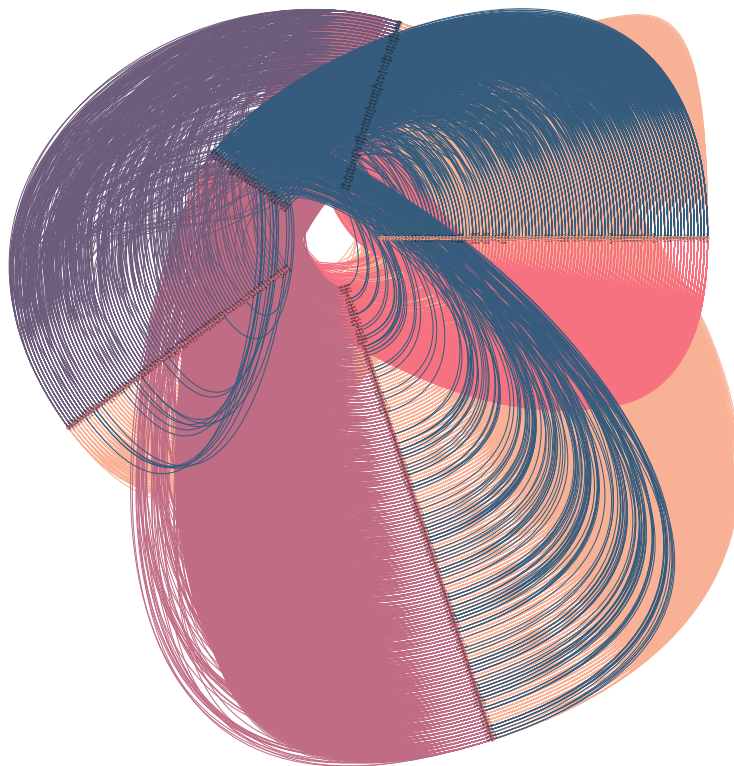


Figure 1.4: A hive plot showing the interactions between transcription factors and coagulation related proteins. The list of genes included was taken from [166], and their associated transcription factors were taken from [79] by assembling a list of possible cis-regulatory elements for each gene, and then looking at the associated transcription factors for each cis-regulatory element, which led to a list of 328 transcription factors involved with 85 genes.

### 1.3 Acute Traumatic Coagulopathy

Acute traumatic coagulopathy (ATC), otherwise known as trauma induced coagulopathy (TIC), is loosely defined as a disorder of the homeostasis following traumatic injury not attributable to any external or environmental factors. Patients that develop ATC have four fold higher mortality than other patients, higher transfusion requirements and worse organ failure rates [112]. The exact causes of ATC are an active subject of investigation, and there exist several competing theories that aim to explain ATC. One theory focuses on protein C, a protein that works to inhibit the coagulation cascade once it has been activated by thrombin. In combination with hypoperfusion that may occur due to blood loss, there may be formation of more thrombin-thrombomodulin complexes, which activate protein C, leading to increased inhibition of the coagulation system, and in this study, low protein C and high thrombomodulin levels were both significantly associated with increased mortality [36]. However, the levels of active protein C were not assessed. Contradicting this narrative, a separate study demonstrated that even when the concentrations of activated protein C were two orders higher than physiological levels, fibrinolysis occurred at normal levels [47]. A small study of severely injury patients showed that upon admission, activated protein C levels were higher in patients with disseminated intravascular coagulation (DIC) than those without, however, the criteria they used to separate the patients into DIC and non-DIC groups is unclear [181]. Another theory posits that ATC arises when fibrin is broken down at accelerated rates, due to the release of tPA from the endothelium. This theory is supported by studies showing that coagulopathic patients have lower levels of fibrinogen than non-coagulopathic patients upon hospital admission and that patient who sur-

vived had higher fibrinogen levels than patients who died from their wounds [280]. Additionally, this theory is supported by measurements of D-dimer, a fragment of degraded thrombin in trauma patients with and without ATC, with the study finding higher levels of D-dimer in the patients with ATC, both upon the day of injury and three days post injury [117]. A third theory states that ATC arises from the shedding of the glycocalyx, the negatively charged glycoprotein and glycolipid layer that lines the endothelium [90]. When this shedding occurs, it can induce auto-heparinization, reducing the tendency of the patient's blood to coagulate [62].

ATC is a fairly general term, and there are number of ways in which a patient's coagulation and fibrinolytic systems can misfire, resulting in ATC. One way to stratify trauma patients is by how fast their fibrinolytic system is running, into fibrinolysis shutdown (clots are barely being broken down,  $LY30 \leq .8\%$ ), hyperfibrinolysis (clots are being broken down faster than normal,  $LY30 \geq 3\%$ ), and physiological (clots are being broken down at a normal rate,  $.9 < LY30 < 2.9$ ) [226]. Patients who are either in fibrinolysis shutdown or hyperfibrinolysis have higher mortality than those exhibiting normal fibrinolysis [225]. The extent of clot lysis tells part of the story about clot structure, however, it appears that trauma patients, even those within the the physiological fibrinolysis group, develop clots with structure different than that of healthy controls [92].

More recent evidence has highlighted the role of platelets in ATC. When platelets taken from trauma patients were compared to platelets taken from healthy controls, the platelets from the trauma patients had impaired aggregation in response to ADP and collagen as well as increased PFA-100 closure

times [317].<sup>2</sup> This platelet dysfunction is also characterized by impaired adhesion and contractile force generation as well as a diminished response to platelet antagonists [162].

Other studies have indicated that the interplay between coagulation and immune response may play a role in the development of ATC. Thrombin, factors XIa, Xa, and IXa, in addition to plasmin have been shown to effectively cleave C3 and C5, two of the proteins associated with the complement cascade, a component of the human immune response [7]. The complement system has been shown to be very active in trauma patients, and the levels of C5b-9, which can be used as a proxy for the levels of membrane attack complex, the final product of complement, are significantly higher in more severely injured patients. Complement and coagulation may positively feedback on each other, as a study has shown that following several hours of incubation with C5a, endothelial cells express much higher levels of tissue factor mRNA [155], which in turn would stimulate the production of thrombin. A strong piece of evidence that the inflammatory response plays an important role in the development in ATC comes from a retrospective study in which the medications people had taken before injury was examined, which found that pre-hospital use of nonsteroidal anti-inflammatory drugs (NSAIDs) was associated with a 72% lower risk of ATC [237].

One group proposes that ATC arises from systems failure along the brain-cardiovascular coupling, which in turn reduces blood flow to the gut, which can then become ischemic and leaky, allowing gut-resident bacteria and fungi and their products to enter the general circulation, coupled with shedding of

---

<sup>2</sup>The PFA-100 or the Platelet Function Assay measures how long it takes for a platelet plug to form in a microscopic aperture at the end of a capillary while whole blood flows through at high shear rates.



the glycocalyx and mitochondrial dysfunction due to the unavailability of oxygen and insufficient glycogen stores [91]. This mitochondrial dysfunction then leads to the release of mitochondrial microparticles, which are derived from the inner membranes of mitochondria, which have been shown to have procoagulant properties in mice [372]. While statistically significant differences between trauma patients who develop ATC and those who do not in terms of the microparticle counts [206], exactly what leads to these changes in microparticle distributions remains unclear.

If we zoom out and examine the entire patient, a severely injured patient will probably be experiencing a degree of shock, or a lack of blood flow due to bleeding. This decrease in blood volume leads to hypotension, an increase in heart rate, and metabolic acidosis, as the body attempts to compensate for the lost blood volume and reduction in oxygenation capacity [293]. However, if the injury is severe, and significant blood has been lost, the patient's microvasculature becomes hypoperfused, resulting in increases of plasma thrombomodulin and a reduction in the measured protein C concentrations as well as an increase in tPA and D-dimer levels [37]. Under current Advanced Trauma Life Support guidelines, to increase the circulating blood volume within a patient with significant bleeding, first responders would give 1-2 liters of isotonic crystalloids<sup>3</sup>, followed by red blood cells to increase oxygen carrying capacity [316]. While this approach solves the immediate problem of the loss of circulating volume, it dilutes the coagulation capacity of the patient, since the fluid they received contains none of the proteins necessary for clot formation. More recently, the US military has begun to adopt the practice of damage control resuscitation, a set

---

<sup>3</sup>The class of isotonic crystalloids includes normal saline and lactated ringers solution, with a concentration of salts as to not to cause a volume shift between interstitial fluids, intracellular fluids, and plasma.

of resuscitation practices where the patient receives either blood products, such as red blood cells, plasma, and platelets or whole blood and is resuscitated to a lower blood pressure, while compressive devices are used to reduce the amount of bleeding, until it can be controlled surgically [50]. Although this approach results in a longer period of hypotension and hypoperfusion, resuscitating patients to a target systolic blood pressure of 70 mmHg instead of 110 mmHg did not result in a statistically significant difference in mortality [95]. This approach preserves far more of the patient's coagulation capacity, however, unless fresh whole blood is given, it still dilutes the patients blood, as a 570 mL unit of stored whole blood contains 70 mL of citrate-phosphate-dextrose solution, and unit of red blood cells is nearly a quarter additive solutions, which increase the lifespan of the red blood cells in storage, but have no coagulation potential [299].

Hypothermia (a core body temperature less than 35 °C) is another potential concern for trauma patients, as the normal hemostatic processes are greatly disturbed by injury, as traumatic injury appears to alter normal central thermoregulation and blocks the shivering response, resulting in a drop in core body temperature [320]. Additionally, trauma patients usually have their clothing removed, leading to increased surface area contact with the cool environment, and are receiving large quantities of resuscitative fluids, which are only rarely warmed to body temperature. One study of trauma patients reported that the mortality of hypothermic patients was significantly higher than those who remained warm [164]. As the proteins involved in coagulation are enzymes, the rates of the reactions they catalyze are temperature dependent, however, it is unclear if this small decrease in body temperature is large enough to significantly effect the clot formation process. One study reported that the activated partial thromoplastin time showed no significant difference between 33° and

39°C, nor was there a significant difference between 33° and 37°C in a thrombin generation assay in either pooled plasma or synthetic reconstituted plasma, however, these authors did report changes in platelet aggregation and adhesion over this temperature range [364]. Another study reports significant changes in prothrombin time and partial thromboplastin time between 37° and 37°C, in plasma samples [278]. In a porcine model, where animals were cooled to 34° C and rewarmed, prothrombin time and bleeding time were prolonged both during the cooling period and during rewarming [318].

A lack of oxygen due to reduced blood volume and reduced oxygen carrying capacity causes cells within the body to switch their metabolism from glycolysis to substrate level phosphorylation, in which pyruvate is reduced by lactate dehydrogenase to produce lactate and  $\text{NAD}^+$ , where the  $\text{NAD}^+$  is used to produce ATP, and lactate is exported from the cells and into the blood [341]. Then this lactate, in solution, tends to give up a proton, making it an Arrhenius acid, and if produced in amounts larger than it is degraded, it can lower blood pH, a condition called acidosis when pH drops below 7.35. Normal blood pH ranges between 7.35 and 7.45, but within trauma patients, it can drop below 6.6 [279]. This drop in pH can effect the activity of the proteins involved in clot formation. For example, the activity of the FVIIa/TF (trigger) complex was reduced by greater than 90% at a pH of 7.0 compared to at a pH of 7.4 in synthetic plasma [212]. Reducing pH has been shown to increase prothrombin time, activated partial thromboplastin time and platelet aggregation, showing that the coagulation cascade is susceptible to pH changes over the range observed in trauma patients. Some groups are investigating the use of vasopressors, a class of chemicals that cause an increase in blood pressure and coronary artery perfusion (and therefore increase oxygen supply to reduce the amount acidification that occurs), such as

norepinephrine and vasopressin, however, these drugs must be used cautiously, since there is conflicting evidence that these drugs boost survival [264].

A very recent paper investigated the combined effects of acidification and hypothermia on coagulation in platelet poor plasma. They performed thrombin generation assays at differing dilutions, temperatures (in the range of 31-37°C) and pHs (6.9, 7.1, and 7.4), and then recorded 5 quantitative parameters to describe the thrombin trajectory: the lag time (the time to  $\frac{1}{6}$  of the thrombin peak height), the time to peak thrombin, the velocity index (the slope of the curve to the peak), and the endogenous thrombin potential (the area under the thrombin generation curve), and then performed linear regression on these parameters. Their regressions showed that within this range, pH had little effect on their measured parameters of thrombin generation, but that hypothermia increased the area under the thrombin generation curve, in some cases [217]. As thorough as this paper is, it does not provide a definite answer to how much acidification and hypothermia effect coagulation within a trauma patient, as platelet poor plasma was used for all of the assays, not whole blood.

It is worth noting that many of the proteins involved in coagulation require carboxylation in order to function properly, and as such, are Vitamin K dependant [75], so a severe dietary lack of Vitamin K could predispose a person to coagulation system dysfunction. Calcium plays a key role in many steps of the coagulation cascade [114], so if a patient has received many citrated blood products, they may be at risk of hypocalcemia, as citrate binds to calcium and makes it unavailable <sup>4</sup> which would further inhibit their ability to clot.

While this work focuses on ATC, coagulopathy can arise whenever the body

---

<sup>4</sup>This is why blood products are citrated: to prevent them from clotting while they are in storage.

is severely perturbed. In a study of patients undergoing highly invasive surgery, such as a hepatectomy or pancreaticoduodenectomy, approximately 84% developed coagulopathy following surgery [289]. Coagulopathy can also occur due to severe infection because of the cross talk between the coagulation and immune system, and has been reported to occur in patients with Rotavirus, Varicella, Rubella, Rubeola and Influenza, as well as viral hemorrhagic fevers, caused by viruses such as Dengue and Ebola [179].

## **1.4 Assays**

### **1.4.1 PTT and PT**

The PTT or partial thromboplastin time was developed in 1953, originally to serve as a test for diagnosing hemophilia [177]. In this test, a patient's blood is citrated to prevent coagulation, then centrifuged to produce platelet poor plasma and a thromboplastin-kaolin agent is added along with calcium to start coagulation [256]. The time until a clot forms is then measured, and compared to a standard to determine if the patient is clotting more slowly or rapidly than normal. This assay provides a global view of the coagulation system, but does not identify which factors are lacking or in excess, nor does it provide any information about the state of the patient's fibrinolytic system.

The PT or prothrombin time is performed in a similar manner to the PTT, but a different set of chemicals are added to initiate coagulation, so that it measures the extrinsic pathway of the coagulation cascade, which consists of factors II, VII, and X. For a PT measurement, phospholipids, tissue factor and thromboplastin

extracts are added to the platelet poor plasma sample from the patient to kick off coagulation [269]. A PT greater than 1.5 times normal can be used to diagnose ATC [38].

### 1.4.2 INR

The international normalized ratio, or INR was developed to standardize PT times, as the contents of thromboplastin extracts are not uniform between different laboratories [269]. The INR is defined as:

$$\text{INR} = \left( \frac{\text{Patient PT}}{\text{Mean Normal PT}} \right)^{\text{ISI}} \quad (1.1)$$

where the Mean Normal PT is calculated from the plasma of at least twenty healthy individuals [255]. The ISI, or international sensitivity index is used to correct for the differences between various thromboplastin extracts. INR is used to diagnose ATC upon admission, with an  $\text{INR} \geq 1.5$  being used as the cutoff [243].

### 1.4.3 ROTEM and TEG

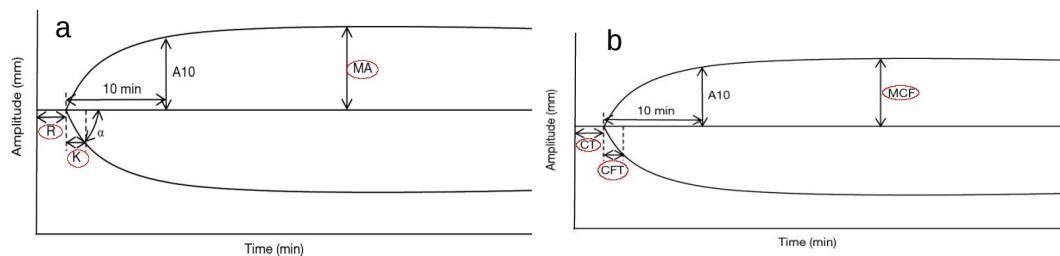


Figure 1.5: Typical (a) TEG and (b) ROTEM tracings. The differences in nomenclature are circled in red. Images adapted from [302].

ROTEM and TEG are both techniques that measure the viscoelasticity of whole blood to assess the state of the coagulation and fibrinolytic systems of a patient. In a TEG system, a small amount of whole blood is added to a cup along with reagents to initiate coagulation, to which a pin with a torsion wire are attached. The cup is oscillated around the pin and as a clot forms, the changes in the viscoelastic properties of the blood cause an increase in the rotation transmitted to the torsion wire, which is detected by a transducer [360]. A ROTEM system operates in a similar manner, however, the main difference between the two systems is that in a ROTEM system, the cup containing the blood is stationary and the pin attached to the torsion wire rotates [241]. Both systems produce a tracing of clot amplitude as a function of time, however, different nomenclature is used to describe features of this curve depending on the test system used, as highlighted in Figure 1.5. The time (in minutes) to reach an amplitude of 2mm is called the CT in ROTEM and the R in TEG. The time (in minutes) for a clot to reach an amplitude of 20mm after it has reached an amplitude of 2mm is called CFT in ROTEM and K in TEG. The maximum amplitude a clot reaches is called MA in TEG and MCF in ROTEM. ROTEM can be used diagnose ATC, with a clot amplitude at 5 minutes less than 35 mm used as a cut-off [76]. Although clot amplitude is reported in units of mm, these units do not directly correspond to clot diameter or radius, rather they refer to mm of torsional amplitude based on the rotation of either the cup or the pin [130]. This amplitude can be converted to a shear modulus,  $G$ , the ratio of shear stress to shear strain using the formula  $G = \frac{5000A}{100-A}$  where  $A$  is the measured clot amplitude [142]. These viscoelastic tests can be used to guide fluid resuscitation, where an elevated CT/R value can be treated with FFP or PCC, and a decreased  $\alpha$  angle can be treated with cryoprecipitate or fibrinogen concentrate, and a lower

than normal MCF/MA value can be treated with platelets or cryoprecipitate or fibrinogen concentrate [131].



## CHAPTER 2

# PARAMETER ESTIMATION VIA DYNAMIC OPTIMIZATION WITH PARTICLE SWARMS

### 2.1 Background

<sup>1</sup> Mathematical modeling has evolved as a powerful paradigm to analyze, and ultimately design complex biochemical networks [11, 338, 160, 172, 145]. Mathematical modeling of biochemical networks is often an iterative process. First, models are formulated from existing biochemical knowledge, and then model parameters are estimated using experimental data [6, 14, 10]. Parameter estimation is typically framed as a non-linear optimization problem wherein the residual (or objective function) between experimental measurements and model simulations is minimized using an optimization strategy [222]. Optimal parameter estimates are then used to predict unseen experimental data. If the validation studies fail, model construction and calibration are repeated iteratively until satisfactory results are obtained. As our biological knowledge increases, model formulation may not be as significant a challenge, but parameter estimation will likely remain difficult.

Parameter estimation is a major challenge to the development of biochemical models. Parameter estimation has been a well studied engineering problem for decades [242, 25, 369, 24]. However, the complex dynamics of large biological systems and noisy, often incomplete experimental data sets pose a unique estimation challenge. Often optimization problems involving biological systems are non-linear and multi-modal i.e., typical models have multiple local min-

---

<sup>1</sup>This work has been previously published in [283]

ima or maxima [222, 14]. Non-linearity coupled with multi-modality renders local optimization techniques such as pattern search [146], Nelder-Mead simplex methods [238], steepest descent or Levenberg-Marquardt [228] incapable of reliably obtaining globally optimal solutions as these methods often terminate at local minimum. Though deterministic global optimization techniques (for example algorithms based on branch and bound) can handle non-linearity and multi-modality [103, 148], the absence of derivative information, discontinuous objective functions, non-smooth regions or the lack of knowledge about the objective function hampers these techniques.

Meta-heuristics like Genetic Algorithms (GAs) [119], Simulated Annealing (SA) [171], Evolutionary Programming [107] and Differential Evolution (DE) [321, 336, 352, 244] have all shown promise on non-linear multi-modal problems [323]. These techniques do not make any assumptions, nor do they require, *a priori* information about the structure of the objective function. Meta-heuristics are often very effective at finding globally optimal or near optimal solutions. For example, Mendes et al. used SA to estimate rate constants for the inhibition of HIV proteinase [211], while Modchang et al. used a GA to estimate parameters for a model of G-protein-coupled receptor (GPCR) activity [219]. Parameter estimates obtained using the GA stratified the effectiveness of two G-protein agonists, N6-cyclopentyladenosine (CPA) and 5'-N-ethylcarboxamidoadenosine (NECA). Tashkova et al. compared different meta-heuristics for parameter estimation on a dynamic model of endocytosis; DE was the most effective of the approaches tested [328]. Banga and co-workers have also successfully applied scatter-search to estimate model parameters [343, 276, 96]. Hybrid approaches, which combine meta-heuristics with local optimization techniques, have also become popular. For example, Egea et al. developed the enhanced scatter search

(eSS) method [96], which combined scatter and local search methods, for parameter estimation in biological models [344]. However, despite these successes, a major drawback of most meta-heuristics remains the large number of function evaluations required to explore parameter space. Performing numerous potentially expensive function evaluations is not desirable (and perhaps not feasible) for many types of biochemical models. Alternatively, Tolson and Shoemaker found, using high-dimensional watershed models, that perturbing only a subset of parameters was an effective strategy for estimating parameters in expensive models [332]. Their approach, called Dynamically Dimensioned Search (DDS), is a simple stochastic single-solution heuristic that estimates nearly optimal solutions within a specified maximum number of function (or model) evaluations. Thus, while meta-heuristics are often effective at estimating globally optimal or nearly optimal solutions, they require a large number of function evaluations to converge to a solution.

In this study, we developed Dynamic Optimization with Particle Swarms (DOPS), a novel hybrid meta-heuristic that combines the global search capability of multi-swarm particle swarm optimization with the greedy refinement of dynamically dimensioned search (DDS). The objective of DOPS is to obtain near optimal parameter estimates for large biochemical models within a relatively few function evaluations. DOPS uses multi-swarm particle swarm optimization to generate nearly optimal candidate solutions, which are then greedily updated using dynamically dimensioned search. While particle swarm techniques are effective, they have the tendency to become stuck in small local regions and lose swarm diversity, so we combined multi-swarm particle optimization with DDS to escape these local regions and continue towards better solutions [61]. We tested DOPS using a combination of classic optimization test functions, bio-

chemical benchmark problems and real-world biochemical models. First, we tested the performance of DOPS on the Ackley and Rosenbrock functions, and published biochemical benchmark problems. Next, we used DOPS to estimate the parameters of a model of the human coagulation cascade. On average, DOPS outperformed other common meta-heuristics like differential evolution, a genetic algorithm, CMA-ES (Covariance Matrix Adaptation Evolution Strategy), simulated annealing, single-swarm particle swarm optimization, and dynamically dimensioned search on the optimization test functions, benchmark problems and the coagulation model. For example, DOPS recovered the nominal parameters for the benchmark problems using an order of magnitude fewer function evaluations than eSS in all cases. It also produced parameter estimates for the coagulation model that predicted unseen coagulation data sets. Thus, DOPS is a promising hybrid meta-heuristic for the estimation of biochemical model parameters in relatively few function evaluations.

## **2.2 Results**

### **2.2.1 DOPS explores parameter space using a combination of global methods.**

DOPS combines a multi-swarm particle swarm method with the dynamically dimensioned search approach of Shoemaker and colleagues (Fig. 2.1). The goal of DOPS is to estimate optimal or near optimal parameter vectors for high-dimensional biological models within a specified number of function evaluations. Toward this objective, DOPS begins by using a multi-swarm parti-

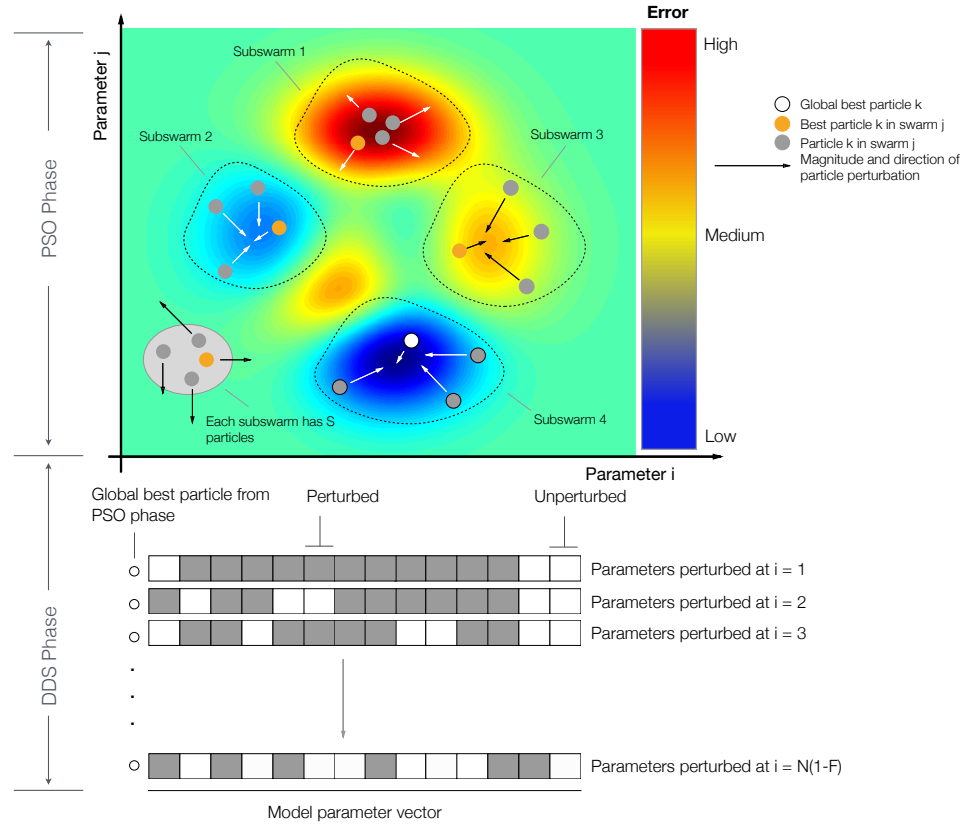


Figure 2.1: Schematic of the dynamic optimization with particle swarms (DOPS) approach. **A:** Each particle represents an  $N$  dimensional parameter vector. Particles are given randomly generated initial solutions and grouped into different sub-swarms. Within each swarm the magnitude and direction of the movement a particle is influenced by the position of the best particle and also by its own experience. After every  $g$  number of function evaluations the particles are mixed and randomly assigned to different swarms. When the error due to the global best particle (best particle amongst all the sub-swarms) does not drop over a certain number of function evaluations, the swarm search is stopped and the search switches to a Dynamically Dimensioned Search with global best particle as the initial solution vector or candidate vector. **B:** The candidate vector performs a greedy global search for the remaining number of function evaluations. The search neighborhood is dynamically adjusted by varying the number of dimensions that are perturbed (in black) in each evaluation step. The probability that a dimension is perturbed decreases as the number of function evaluations increase.

cle swarm search and then dynamically switches, using an adaptive switching criteria, to the DDS approach. The particle swarm search uses multiple sub-swarms wherein the update to each particle (corresponding to a parameter vec-

tor estimate) is influenced by the best particle amongst the sub-swarm, and the current globally best particle. Particle updates occur within sub-swarms for a certain number of function evaluations, after which the sub-swarms are reorganized. This sub-swarm mixing is similar to the regrouping strategy described by Zhao et al. [371]. DOPS switches out of the particle swarm phase based upon an adaptive switching criteria that is a function of the rate of error convergence. If the error represented by the best particle does not decrease for a threshold number of function evaluations, DOPS switches automatically to the DDS search phase. The DDS search is initialized with the globally best particle from the particle swarm phase, thereafter, the particle is greedily updated by perturbing a subset of dimensions for the remaining number of function evaluations. The identity of the parameters perturbed is chosen randomly, with fewer parameters perturbed the higher the number of function evaluations.

### **2.2.2 DOPS minimized benchmark problems using fewer function evaluations.**

On average, DOPS performed similarly or outperformed four other meta-heuristics for the Ackley and Rastrigin test functions (Fig. 2.2). The Ackley and Rastrigin functions both have multiple local extrema and attain a global minimum value of zero. In each case, the maximum number of function evaluations was fixed at  $N = 4000$ , and  $\mathcal{T} = 25$  independent experiments were run with different initial parameter vectors. DOPS found optimal or near optimal solutions for both the 10-dimensional Ackley (Fig. 2.2A) and Rastrigin (Fig. 2.2B) functions within the budget of function evaluations. In each of the 10-dimensional

cases, other meta-heuristics such as DDS and DE also performed well. However, DOPS consistently outperformed all other approaches tested. This performance difference was more pronounced as the dimension of the search problem increased; for a 300-dimensional Rastrigin function, DOPS was the only approach to find an optimal or near optimal solution within the function evaluation budget (Fig. 2.2B). Taken together, DOPS performed at least as well as other meta-heuristics on small dimensional test problems, but was especially suited to large dimensional search spaces. Next, we tested DOPS on benchmark biochemical models of varying complexity.

Villaverde and co-workers published a set of benchmark biochemical problems to evaluate parameter estimation methods [344]. They ranked the example problems by computational cost from most to least expensive. We evaluated the performance of DOPS on problems from the least and most expensive categories. The least expensive problem was a metabolic model of Chinese Hamster Ovary (CHO) with 35 metabolites, 32 reactions and 117 parameters [342]. The biochemical reactions were modeled using modular rate laws and generalized Michaelis–Menten kinetics. On the other hand, the expensive problem was a genome scale kinetic model of *Saccharomyces cerevisiae* with 261 reactions, 262 variables and 1759 parameters [311]. In both cases, synthetic time series data generated with known parameter values, was used as training data to estimate the model parameters. For the *Saccharomyces cerevisiae* model, the time series data consisted of 44 observables, while for the CHO metabolism problem the data corresponded to 13 different metabolite measurement sets. The number of function evaluations was fixed at  $N = 4000$ , and we trained both models against the synthetic experimental data. DOPS produced good fits to the synthetic data (Fig. 10.1 and Fig. 10.2), and recapitulated the nominal parameter

values using only  $N \leq 4000$  function evaluations (Fig. 10.3). On the other hand, the enhanced scatter search (eSS) with a local optimizer method, took on order  $10^5$  function evaluations for the same problems. DOPS required a comparable amount of time (Fig. 10.4), faster convergence (Fig. 10.5 and Fig. 10.6), and also had lower variability in the best value obtained (Fig. 10.7) across multiple runs when compared to other meta-heuristics. Thus, DOPS estimated the parameters in benchmark biochemical models, and recovered the original parameters from the synthetic data, using fewer function evaluations. Next, we compared the performance of DOPS with four other meta-heuristics for a model of the human coagulation cascade.



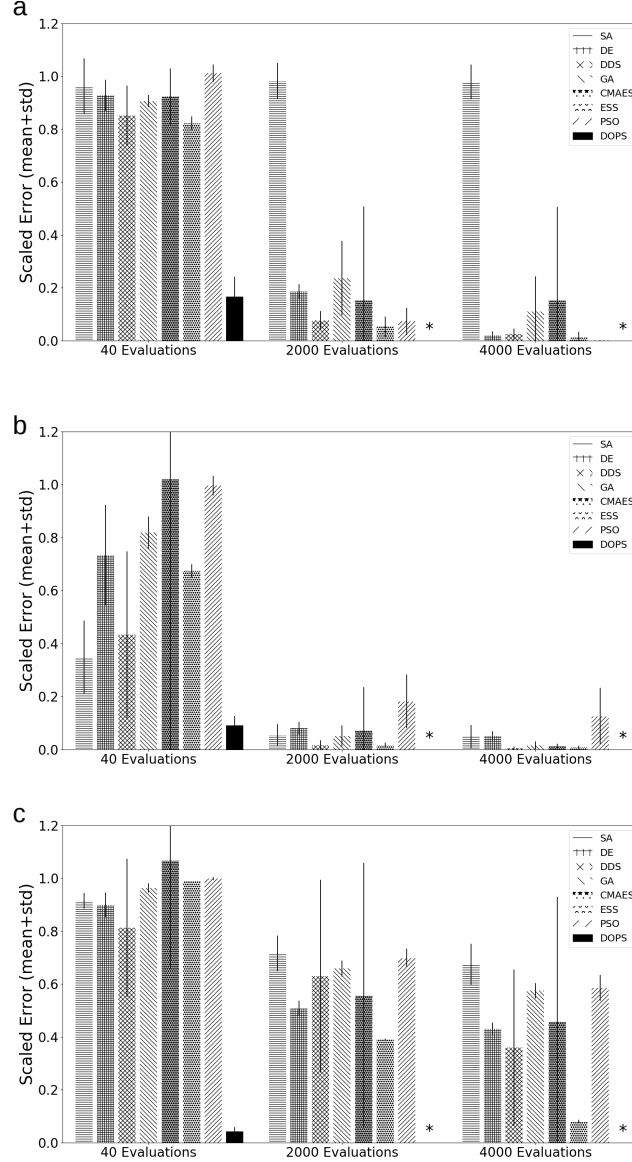


Figure 2.2: Performance of DOPS and other meta-heuristics for the Ackley and Rastrigin functions. **A:** Mean scaled error versus the number of function evaluations for the 10-dimensional Ackley function. DOPS, DDS and ESS find optimal or near optimal solutions within the specified number of function evaluations. **B:** Mean scaled error versus the number of function evaluations for the 10-dimensional Rastrigin function. Nearly all the techniques find optimal or near optimal solutions within the specified number of function evaluations. **C:** Mean scaled error versus the number of function evaluations for the 300-dimensional Rastrigin function. DOPS is the only algorithm that finds an optimal or near optimal solution within the specified number of function evaluations. In all cases, the maximum number of function evaluations was  $N = 4000$ . Mean and standard deviation were calculated over  $\mathcal{T} = 25$  trials. A star denotes that the average value was less than  $1\text{E-}6$ .

### 2.2.3 DOPS estimated the parameters of a human coagulation model.

Coagulation is an archetype biochemical network that is highly interconnected, containing both negative and positive feedback (Fig. 2.3). The biochemistry of coagulation, though complex, has been well studied [200, 201, 197, 346, 87, 108, 8], and reliable experimental protocols have been developed to interrogate the system [143, 55, 199, 193]. Coagulation is mediated by a family of proteases in the circulation, called factors and a key group of blood cells, called platelets. The central process in coagulation is the conversion of prothrombin (FII), an inactive coagulation factor, to the master protease thrombin (FIIa). Thrombin generation involves three phases, initiation, amplification and termination. Initiation requires a trigger event, for example a vessel injury which exposes tissue factor (TF), which leads to the activation of factor VII (FVIIa) and the formation of the TF/FVIIa complex. Two converging pathways, the extrinsic and intrinsic cascades, then process and amplify this initial coagulation signal. There are several control points in the cascade that inhibit thrombin formation, and eventually terminate thrombin generation. Tissue Factor Pathway Inhibitor (TFPI) inhibits upstream activation events, while antithrombin III (ATIII) neutralizes several of the proteases generated during coagulation, including thrombin. Thrombin itself also inadvertently plays a role in its own inhibition; thrombin, through interaction with thrombomodulin, protein C and endothelial cell protein C receptor (EPCR), converts protein C to activated protein C (APC) which attenuates the coagulation response by proteolytic cleavage of amplification complexes. Termination occurs after either prothrombin is consumed, or thrombin formation is neutralized by inhibitors such as APC or ATIII. Thus, the human coagu-

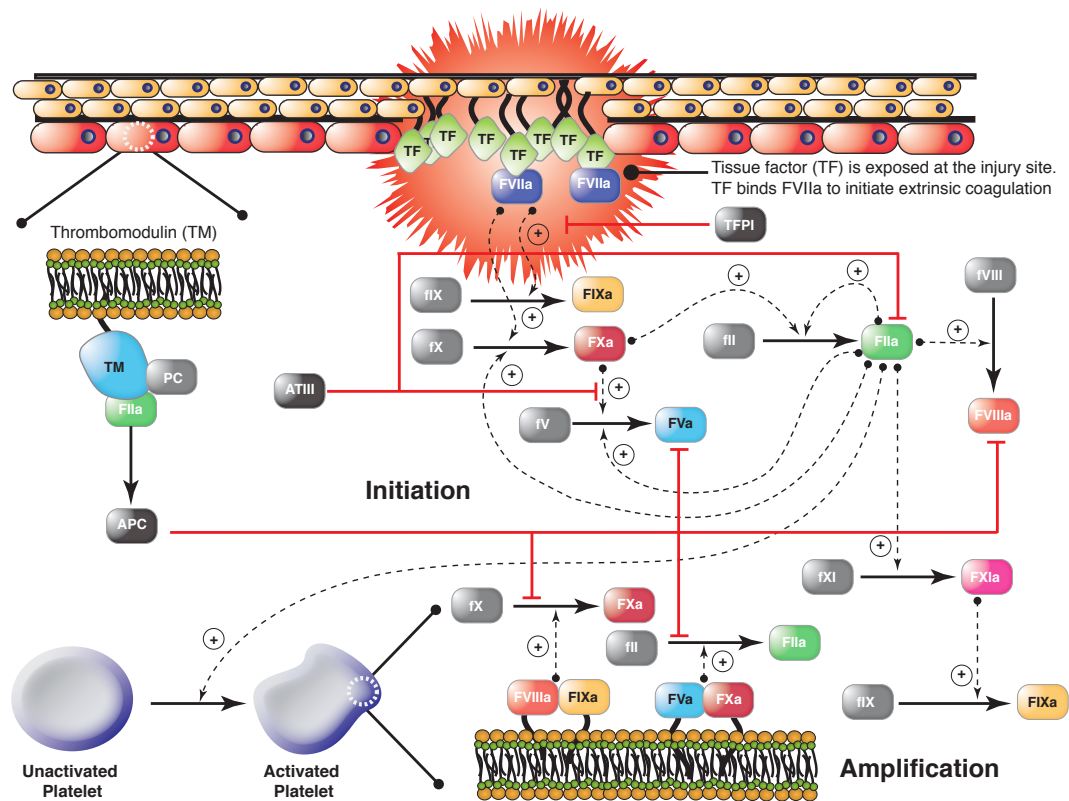


Figure 2.3: Schematic of the extrinsic and intrinsic coagulation cascade. Inactive zymogens upstream (grey) are activated by exposure to tissue factor (TF) following vessel injury. Tissue factor and activated factor VIIa (FVIIa) form a complex that activates factor X (fX) and IX (fIX). FXa activates downstream factors including factor VIII (fVIII) and fX. Factor V (fV) is primarily activated by thrombin (FIIa). In addition, we included a secondary fV activation route involving FXa. FXa and FVa form a complex (prothrombinase) on activated platelets that converts prothrombin (fII) to FIIa. FIXa and FVIIIa can also form a complex (tenase) on activated platelets which catalyzes FXa formation. Thrombin also activates upstream coagulation factors, forming a strong positive feedback ensuring rapid activation. Tissue factor pathway inhibitor (TFPI) down-regulates FXa formation and activity by sequestering free FXa and TF-FVIIa in a FXa-dependent manner. Antithrombin III (ATIII) inhibits all proteases. Thrombin inhibits itself binding the surface protein thrombomodulin (TM). The IIa-TM complex catalyzes the conversion of protein C (PC) to activated protein C (APC), which attenuates the coagulation response by the proteolytic cleavage of fV/FVa and fVIII/FVIIIa.

lation cascade is an ideal test case; coagulation is challenging because it contains both fast and slow dynamics, but also accessible because of the availability of comprehensive data sets for model identification and validation. In this study, we used the coagulation model of Luan et al. [193], which is a coupled system of non-linear ordinary differential equations where biochemical interactions were modeled using mass action kinetics. The Luan model contained 148 parameters and 92 species and has been validated using 21 published experimental datasets.

DOPS estimated the parameters of a human coagulation model for TF/VIIa initiated coagulation without anticoagulants (Fig. 2.5a). The objective function was an unweighted linear combination of two error functions, representing coagulation initiated with different concentrations of TF/FVIIa (5pM, 5nM) [143]. The number of function evaluations was restricted to  $N = 4000$  for each algorithm we tested, and we performed  $\mathcal{T} = 25$  trials of each experiment to collect average performance data (Table 2.1). DOPS converged faster and had a lower final error compared to the other algorithms (Fig. 2.4). Within the first 25% of function evaluations, DOPS produced a rapid drop in error followed by a slower but steady decline (Fig. 10.8b). Approximately between 500-1000 function evaluations DOPS switched to the dynamically dimensioned search phase, however this transition varied from trial to trial since the switch was based upon the local convergence rate. On average, DOPS minimized the coagulation model error to a greater extent than the other meta-heuristics. However, it was unclear if the parameters estimated by DOPS had predictive power on unseen data. To address this question, we used the final parameters estimated by DOPS to simulate data that was not used for training (coagulation initiated with 500pM, 50pM, and 10pM TF/VIIa). The optimal or near optimal parameters obtained by

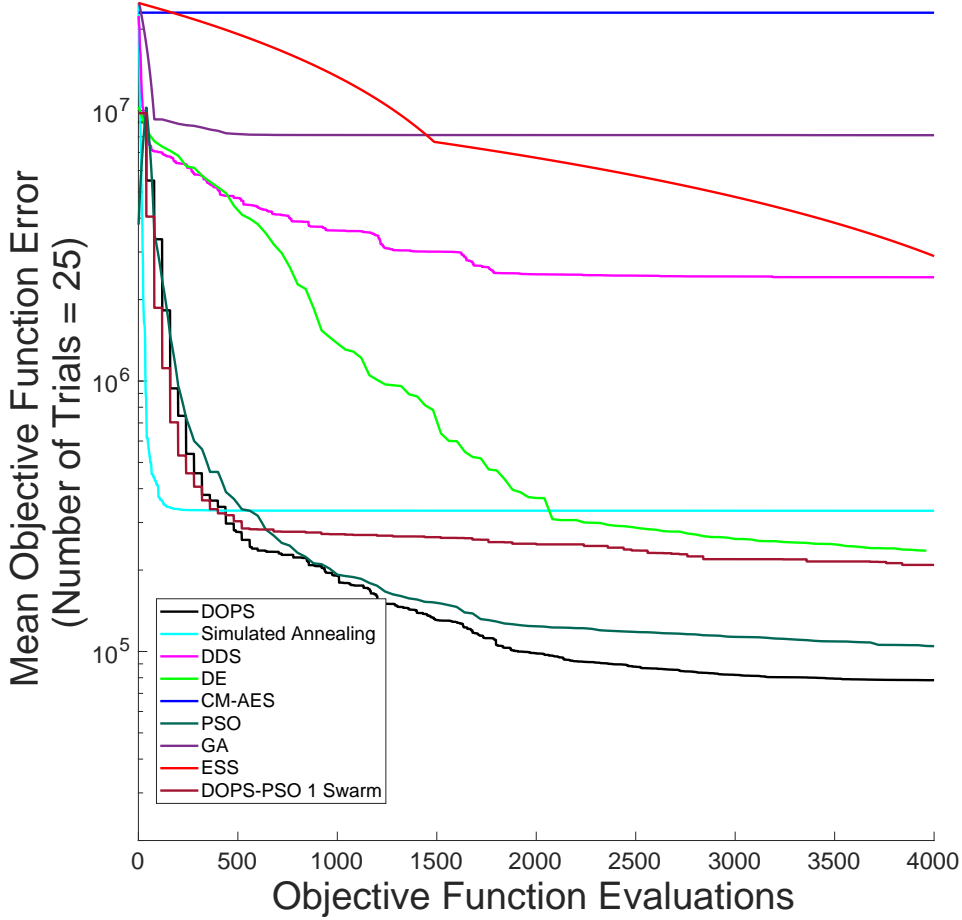


Figure 2.4: Error convergence rates of the nine different algorithms on the coagulation model. The objective error is the mean over  $\mathcal{T} = 25$  trials. DOPS, SA, PSO and DOPS-PSO have the steepest drop in error during first 300 function evaluations. Thereafter the error drop in DDS and SA remains nearly constant whereas DOPS continues to drops further. In the allotted budget of function evaluations ESS produces a modest reduction in error. At the end of 4000 function evaluations DOPS attains the lowest error.

DOPS predicted unseen coagulation datasets (Fig. 2.5b). The normalized standard error for the coagulation predictions was consistent with the training error, with the exception of the 50pM TF/VIIa case which was a factor 2.65 worse (Table 2.2). However, this might be expected as coagulation initiation with 50pM

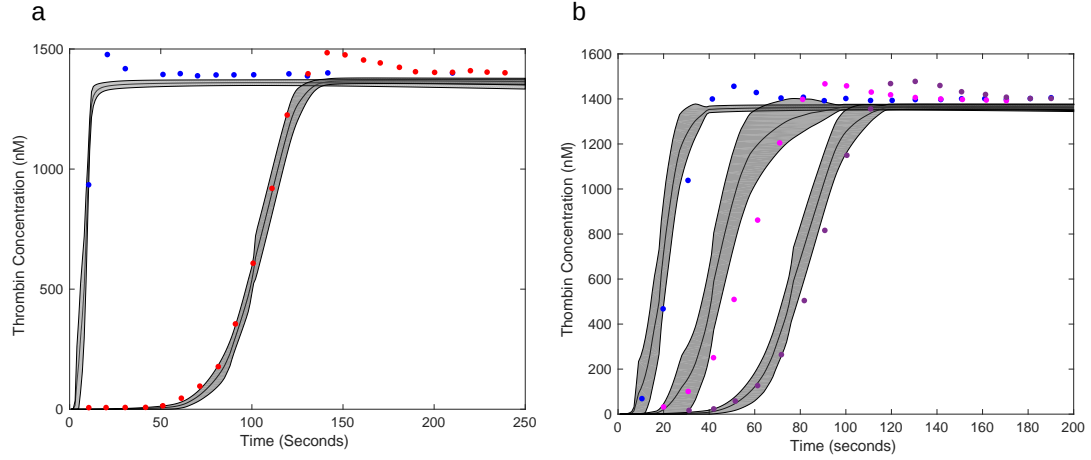


Figure 2.5: Model fits and predictions on experimental data using DOPS. (a) The model parameters were estimated using DOPS. Solid black lines indicate the simulated mean thrombin concentration using parameter vectors from 25 trials. The grey shaded region represents the 99% confidence estimate of the mean simulated thrombin concentration. The experimental data is reproduced from the synthetic plasma assays of Mann and co-workers. Thrombin generation is initiated by adding Factor TF/VIIa (5nM (blue) and 5pM (red)) to synthetic plasma containing 200  $\mu\text{mol/L}$  of phospholipid vesicles (PCPS) and a mixture of coagulation factors (II,V,VII,VIII,IX,X and XI) at their mean plasma concentrations. (b) The parameter estimates that were obtained using DOPS were tested against data that was not used in the model training. Solid black lines indicate the simulated mean thrombin concentration using parameter vectors from  $\mathcal{T} = 25$  trials. The grey shaded region represents the 99% confidence estimate of the mean simulated thrombin concentration. The experimental data is reproduced from the synthetic plasma assays of Mann and co-workers. Thrombin generation is initiated by adding Factor VIIa-TF (500pM - Blue, 50pM - Pink and 10pM - purple, respectively) to synthetic plasma containing 200  $\mu\text{mol/L}$  of phospholipid vesicles (PCPS) and a mixture of coagulation factors (II,V,VII,VIII,IX,X and XI) at their mean plasma concentrations.

TF/FVIIa was the farthest away from the training conditions. Taken together, DOPS estimated parameter sets with predictive power on unseen coagulation data using fewer function iterations than other meta-heuristics. Next, we explored how the number of sub-swarms and the switch to DDS influenced the

performance of the approach.

#### **2.2.4 Phase switching was critical to DOPS performance.**

A differentiating feature of DOPS is the switch to dynamically dimensioned search following stagnation of the initial particle swarm phase. We quantified the influence of the number of sub-swarms and the switch to DDS on error convergence by comparing DOPS with and without DDS for different numbers of sub-swarms (Fig. 2.6). We considered multi swarm particle swarm optimization with and without the DDS phase for  $N = 4000$  function evaluations and  $\mathcal{T} = 25$  trials on the coagulation model. We used one, two, four, five and eight sub-swarms, with a total of 40 particles divided evenly amongst the swarms. Hence, we did not consider swarm numbers of three and seven. All other algorithm parameters remained the same for all cases. Generally, the higher sub-swarm numbers converged in fewer function evaluations, where the optimum particle partitioning was in the neighborhood of five sub-swarms. However, the difference in convergence rate was qualitatively similar for four, five and eight sub-swarms, suggesting there was an optimal number of particles per swarm beyond which there was no significant advantage. The multi-swarm particle swarm optimization stagnated after 25% of the available function evaluations irrespective of the number of sub-swarms. However, DOPS (with five sub-swarms) switched to DDS after detecting the stagnation. The DDS phase refined the globally best particle to produce significantly lower error on average when compared to multi-swarm particle swarm optimization alone. Thus, the automated switching strategy was critical to the overall performance of DOPS. However, it was unclear if multiple strategy switches could further improve

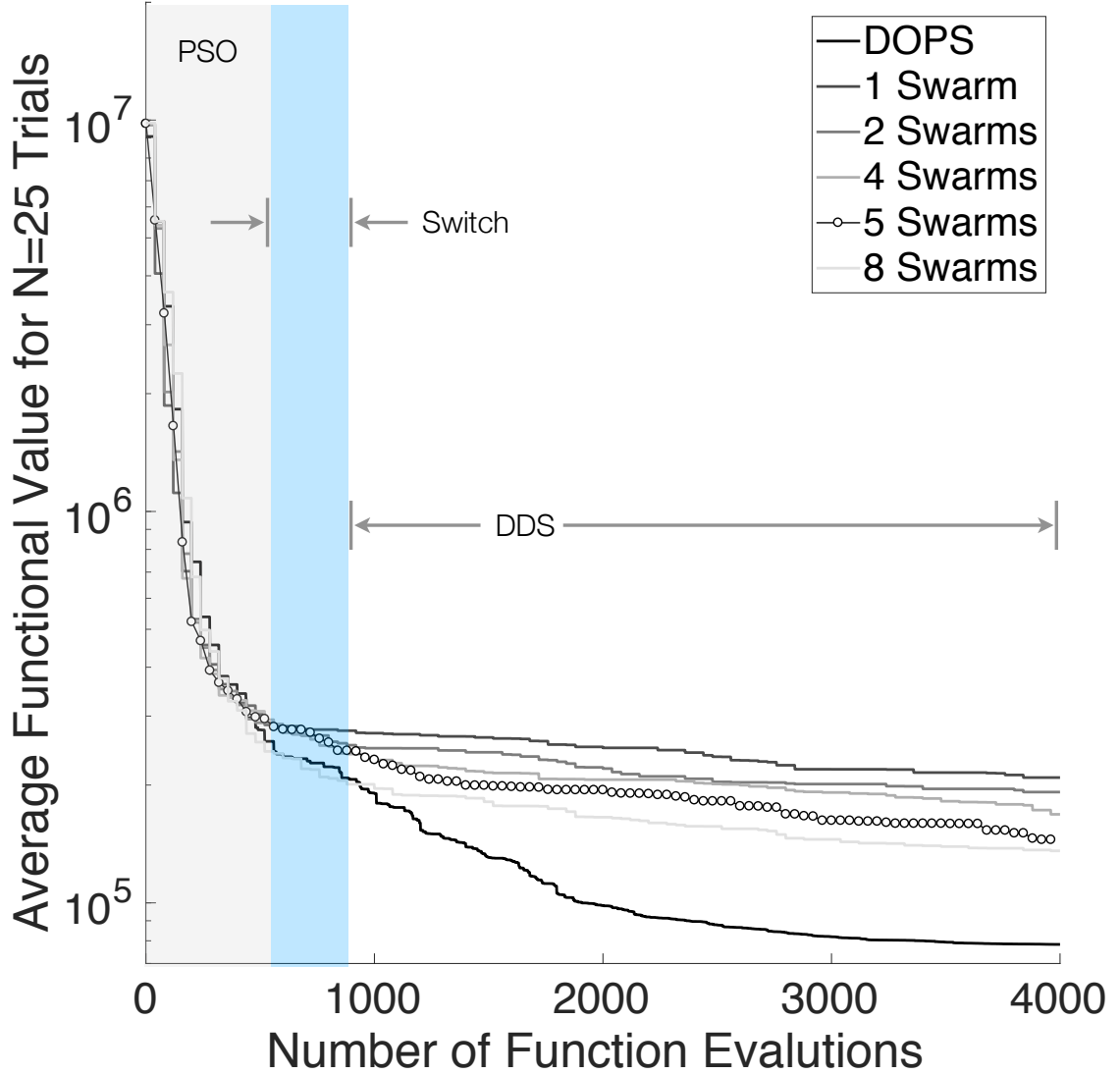


Figure 2.6: Influence of the switching strategy and sub-swarms on DOPS performance for the coagulation model. DOPS begins by using a particle swarm search and then dynamically switches (switch region), using an adaptive switching criteria, to the DDS search phase. We compared the performance of DOPS with and without DDS for different sub-swarm searches to quantify the effect of number of sub-swarms and DDS. We used one, two, four, five and eight sub-swarms, with a total of 40 particles divided evenly amongst the swarms. The results presented are the average of  $\mathcal{T} = 25$  trials with  $N = 4000$  function evaluations each. The convergence rates with higher swarm numbers is typically higher but there is no pronounced difference amongst four, five and eight. The multi-swarm with without DDS saturates while DOPS shows a rapid drop due to a switch to the DDS phase.



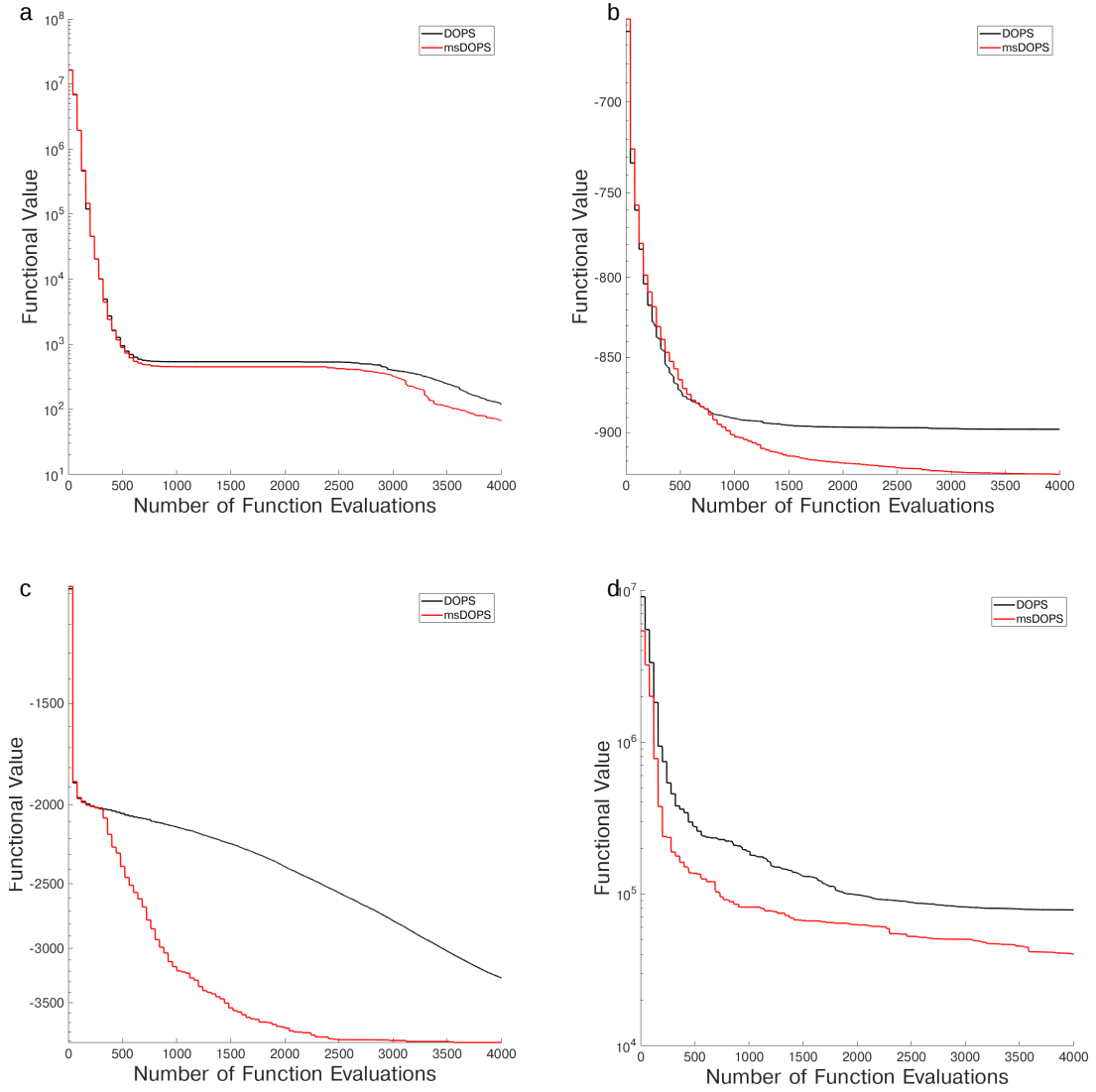


Figure 2.7: Comparison of DOPS and Multiswitch DOPS Performance of DOPS and Multiswitch DOPS on the CHO metabolism problem (a), the Eggholder function (b), the 100 dimensional Styblinski-Tang function (c) and the coagulation problem (d). Both methods have the same initial decrease in error, but as the number of function evaluations increases, multiswitch DOPS produces a larger decrease in error. The results presented are the average of  $\mathcal{T} = 250$  trials with for the CHO metabolism problem and  $\mathcal{T} = 250$  trials on the Eggholder and Styblinski-Tang functions with  $\mathcal{N} = 250$  function evaluations each, and  $\mathcal{T} = 25$  trials for the coagulation problem.

performance.

We explored the performance of DOPS if it was permitted to switch between the PSO (Particle Swarm Optimization) and DDS modes multiple times. This mode ( $\text{msDOPS}$ ) had comparable performance to DOPS on 10-d Ackley and Rastrigin functions, as well as on the 300-dimensional Rastrigin function. However,  $\text{msDOPS}$  performed better than DOPS on the CHO metabolism problem (Fig. 2.7a), with the average functional value being nearly half that of DOPS. To further distinguish DOPS from  $\text{msDOPS}$ , we compared the performance of each algorithm on the Eggholder function, a difficult function to optimize given its multiple minima [158].  $\text{msDOPS}$  outperformed DOPS on the Eggholder function, however, neither version reached the true minimum at -959.6407 on any trial with a budget of  $N = 4000$  function evaluations (Fig. 2.7b). We also explored the performance of  $\text{msDOPS}$  and DOPS on the 100 dimensional Styblinski-Tang function [159] (Fig. 2.7c). In this comparison,  $\text{msDOPS}$  significantly outperformed DOPS, finding the true minimum before exhausting its function evaluation budget, while DOPS does not reach the minimum. Since the performance of  $\text{msDOPS}$  was promising on these problems, we measured its performance on the coagulation problem. Surprisingly, DOPS performed similarly to  $\text{msDOPS}$  on the coagulation problem (Fig. 2.7d); the final average objective value for DOPS reached 0.9413% of the initial functional value, compared to 0.9428% for  $\text{msDOPS}$ . Taken together, these results indicate that switching plays a key role in DOPS's performance and that for some classes of problems, multiple switching between modes produces a faster drop in objective value. However, the coagulation model results suggested the advantage of  $\text{msDOPS}$  was problem specific.

## 2.3 Discussion

In this study, we developed dynamic optimization with particle swarms (DOPS), a novel meta-heuristic for parameter estimation. DOPS combined multi-swarm particle swarm optimization, a global search approach, with the greedy strategy of dynamically dimensioned search to estimate optimal or nearly optimal solutions in a fixed number of function evaluations. We tested the performance of DOPS and four widely used meta-heuristics on the Ackley and Rastrigin test functions, a set of biochemical benchmark problems and a model of the human coagulation cascade. We also compared the performance of DOPS to enhanced Scatter Search (eSS), another widely used meta-heuristic approach. As the number of parameters increased, DOPS outperformed the other meta-heuristics, generating optimal or nearly optimal solutions using significantly fewer function evaluations compared with the other methods. We tested the solutions generated by DOPS by comparing the estimated and true parameters in the benchmark studies, and by using the coagulation model to predict unseen experimental data. For both benchmark problems, DOPS retrieved the true parameters in significantly fewer function evaluations than other meta-heuristics. For the coagulation model, we used experimental coagulation measurements under two different conditions to estimate optimal or nearly optimal parameters. These parameters were then used to predict unseen coagulation data; the coagulation model parameters estimated by DOPS predicted the correct thrombin dynamics following TF/FVIIa induced coagulation without anticoagulants. Lastly, we showed the average performance of DOPS improved when combined with dynamically dimensioned search phase, compared to an identical multi-swarm approach alone, and that multiple mode switches could

improve performance for some classes of problems. Taken together, DOPS is a promising meta-heuristic for the estimation of parameters in large biochemical models.

Meta-heuristics can be effective tools to estimate optimal or nearly optimal solutions for complex, multi-modal functions. However, meta-heuristics typically require a large number of function evaluations to converge to a solution compared with techniques that use derivative information. DOPS is a combination of particle swarm optimization, which is a global search method, and dynamically dimensioned search, which is a greedy evolutionary technique. Particle swarm optimization uses collective information shared amongst swarms of computational particles to search for global extrema. Several particle swarm variants have been proposed to improve the search ability and rate of convergence. These variations involve different neighborhood structures, multi-swarms or adaptive parameters. Multi-swarm particle swarm optimization with small particle neighborhoods has been shown to be better in searching on complex multi-modal solutions [371]. Multi-swarm methods generate diverse solutions, and avoid rapid convergence to local optima. However, at least for the coagulation problem used in this study, multi-swarm methods stagnated after approximately 25% of the available function evaluations; only the introduction of dynamically dimensioned search improved the rate of error convergence. Dynamically dimensioned search, which greedily perturbs only a subset of parameter dimensions in high dimensional parameter spaces, refined the globally best particle and produced significantly lower error on average when compared to multi-swarm particle swarm optimization alone. However, dynamically dimensioned search, starting from a initial random parameter guess, was not as effective on average as DOPS. The initial solutions generated by the

multi swarm search had a higher propensity to produce good parameter estimates when refined by dynamically dimensioned search. Thus, our hybrid combination of two meta-heuristics produced better results than either constituent approach, and better results than other meta-heuristic approaches on average. This was true of not only the convergence rate on the coagulation problem, but also the biochemical benchmark problems; DOPS required two-orders of magnitude fewer function evaluations compared with enhanced Scatter Search (eSS) to estimate the biochemical benchmark model parameters. What remains to be explored is the performance of DOPS compared to techniques that utilize derivative information, either on their own or in combination with other meta-heuristics, and the performance of DOPS in real-world applications compared with other meta-heuristics such as hybrid genetic algorithms e.g., see [223]. Gradient methods perform well on smooth convex problems which have either a closed form of the gradient of the function being minimized, or a form that can be inexpensively estimated numerically. While the biological problems DOPS is intended for often do not have this form, perhaps the solutions could be further improved by following (or potentially replacing) the DDS phase with a gradient based technique when applicable. Taken together, the combination of particle swarm optimization and dynamically dimensioned search performed better than either of these constituent approaches alone, and required fewer function evaluations compared with other common meta-heuristics.

## 2.4 Conclusions

DOPS performed well on many different systems with no pre-optimization of algorithm parameters, however there are many research questions that should

be pursued further. DOPS comfortably outperformed existing, widely used meta-heuristics for high dimensional global optimization functions, biochemical benchmark models and a model of the human coagulation system. However, it is possible that highly optimized versions of common meta-heuristics could surpass DOPS; we should compare the performance of DOPS with optimized versions of the other common meta-heuristics on both test and real-world problems to determine if a performance advantage exists in practice. Next, DOPS has a hybrid architecture, thus the particle swarm phase could be combined with other search strategies such as local derivative based approaches to improve convergence rates. We could also consider multiple phases beyond particle swarm and dynamically dimensioned search, for example switching to a gradient based search following the dynamically dimensioned search phase. Lastly, we should update DOPS to treat multi-objective problems. The identification of large biochemical models sometimes requires training using qualitative, conflicting or even contradictory data sets. One strategy to address this challenge is to estimate experimentally constrained model ensembles using multiobjective optimization. Previously, we developed Pareto Optimal Ensemble Techniques (POETs) which integrates simulated annealing with Pareto optimality to identify models near the optimal tradeoff surface between competing training objectives [21]. Since DOPS consistently outperformed simulated annealing on both test and real-world problems, we expect a multi-objective form of DOPS would more quickly estimate solutions which lie along high dimensional trade-off surfaces.

## 2.5 Methods

### 2.5.1 Optimization problem formulation.

Model parameters were estimated by minimizing the difference between model simulations and  $\mathcal{E}$  experimental measurements. Simulation error is quantified by an objective function  $K(\mathbf{p})$  (typically the Euclidean norm of the difference between simulations and measurements) subject to problem and parameter constraints:

$$\begin{aligned}
 \min_{\mathbf{p}} K(\mathbf{p}) &= \sum_{i=1}^{\mathcal{E}} (g_i(t_i, \mathbf{x}, \mathbf{p}, \mathbf{u}) - y_i)^2 \\
 \text{subject to } \dot{\mathbf{x}} &= \mathbf{f}(t, \mathbf{x}(t), \mathbf{p}), \mathbf{u}(t), \mathbf{p}) \\
 \mathbf{x}(t_0) &= \mathbf{x}_0 \\
 \mathbf{c}(t, \mathbf{x}, \mathbf{p}, \mathbf{u}) &\geq \mathbf{0} \\
 \mathbf{p}^L &\leq \mathbf{p} \leq \mathbf{p}^U
 \end{aligned} \tag{2.1}$$

The term  $K(\mathbf{p})$  denotes the objective function (sum of squared error),  $t$  denotes time,  $g_i(t_i, \mathbf{x}, \mathbf{p}, \mathbf{u})$  is the model output for experiment  $i$ , while  $y_i$  denotes the measured value for experiment  $i$ . The quantity  $\mathbf{x}(t, \mathbf{p})$  denotes the state variable vector with an initial state  $\mathbf{x}_0$ ,  $\mathbf{u}(t)$  is a model input vector,  $\mathbf{f}(t, \mathbf{x}(t), \mathbf{p}), \mathbf{u}(t), \mathbf{p})$  is the system of model equations (e.g., differential equations or algebraic constraints) and  $\mathbf{p}$  denotes the model parameter vector (quantity to be estimated). The parameter search (or model simulations) can be subject to  $\mathbf{c}(t, \mathbf{x}, \mathbf{p}, \mathbf{u})$  linear or non-linear constraints, and parameter bound constraints where  $\mathbf{p}^L$  and  $\mathbf{p}^U$  denote the lower and upper parameter bounds, respectively. Optimal model parame-

ters are then given by:

$$\mathbf{p}^* = \arg \min_{\mathbf{p}} K(\mathbf{p}) \quad (2.2)$$

In this study, we considered only parameter bound constraints, and did not include the  $\mathbf{c}(t, \mathbf{x}, \mathbf{p}, \mathbf{u})$  linear or non-linear problem constraints. However, additional these constraints can be handled, without changing the approach, using a penalty function method.

### 2.5.2 Dynamic optimization with particle swarms (DOPS).

DOPS combines multi-swarm particle swarm optimization with dynamically dimensioned search (Fig. 2.1) and (Algo. 1). The goal of DOPS is to estimate optimal or near optimal parameter vectors for high-dimensional biological models within a specified number of function evaluations. Toward this objective, DOPS begins by using a particle swarm search and then dynamically switches, using an adaptive switching criteria, to a DDS search phase.

#### Phase 1: Particle swarm phase.

Particle swarm optimization is an evolutionary algorithm that uses a population of particles (solutions) to find an optimal solution [68, 2]. Each particle is updated based on its experience (particle best) and the experience of all other particles within the swarm (global best). The particle swarm phase of DOPS begins by randomly initializing a swarm of  $\mathcal{K}$ -dimensional particles (represented as  $z_i$ ), wherein each particle corresponded to a  $\mathcal{K}$ -dimensional parameter vector. After initialization, particles were randomly partitioned into  $k$  equally sized



**input** : A randomized swarm of particles of size  $NP \times K$  and fixed number of function evaluations  $N$   
**output**: Optimized parameter vector of size  $1 \times K$

```

1 Initialize the particles randomly and assign particles randomly to various
  sub-swarms;
2 while  $j \leq N$  do
3   if  $\text{mod}(j,G)=0$  then
4     | Reassign particles to different sub-swarms;
5   end
6   for  $i \leftarrow 1$  to  $NS$  do
7     | Update particles within sub-swarms according to equation 3;
8   end
9   Find best particle  $\mathcal{G}$  amongst all sub-swarms;
10  if  $\text{besterror}(j) \geq 0.99 * \text{besterror}(j + 1)$  then
11    | failurecounter  $\leftarrow$  failurecounter + 1;
12  else
13    | failurecounter  $\leftarrow$  0;
14  end
15  if failurecounter  $\geq$  threshold then
16    |  $\mathcal{G} \leftarrow \text{DDS}(\mathcal{G}, N - j)$ ;
17    | return  $\mathcal{G}$ 
18  else
19    |  $j \leftarrow j + 1$ ;
20  end
21  return  $\mathcal{G}$ 
22 end

```

**Algorithm 1:** Pseudo code for the dynamic optimization with particle swarms (DOPS) method.

sub-swarms  $\mathcal{S}_1, \dots, \mathcal{S}_k$ . Particles within each sub-swarm  $\mathcal{S}_k$  were updated according to the rule:

$$z_{i,j} = \theta_{1,j-1} z_{i,j-1} + \theta_2 r_1 (\mathcal{L}_i - z_{i,j-1}) + \theta_3 r_2 (\mathcal{G}_k - z_{i,j-1}) \quad (2.3)$$

where  $(\theta_1, \theta_2, \theta_3)$  were adjustable parameters,  $\mathcal{L}_i$  denotes the best solution found by particle  $i$  within sub-swarm  $\mathcal{S}_k$  for function evaluation  $1 \rightarrow j - 1$ , and  $\mathcal{G}_k$  denotes the best solution found over all particles within sub-swarm  $\mathcal{S}_k$ . The quantities  $r_1$  and  $r_2$  denote uniform random vectors with the same dimension as the number of unknown model parameters ( $\mathcal{K} \times 1$ ). Equation (2.3) is similar to the

general particle swarm update rule, however, it does not contain velocity terms. In DOPS, the parameter  $\theta_{1,j-1}$  is similar to the inertia weight parameter for the velocity term described by Shi and Eberhart [305]; Shi and Eberhart proposed a linearly decreasing inertia weight to improve convergence properties of particle swarm optimization. Our implementation of  $\theta_{1,j-1}$  is inspired by this and the decreasing perturbation probability proposed by Tolson and Shoemaker [332]. It is an analogous equivalent to inertia weight on velocity. However  $\theta_{1,j-1}$  places inertia on the position rather than velocity and uses the same rule described by Shi and Eberhart to adaptively change with the number of function evaluations:

$$\theta_{1,j} = \frac{(\mathcal{N} - j) * (w_{max} - w_{min}))}{(\mathcal{N} - 1)} + w_{min} \quad (2.4)$$

where  $\mathcal{N}$  represents the total number of function evaluations,  $w_{max}$  and  $w_{min}$  are the maximum and minimum inertia weights, respectively. In this study, we used  $w_{max} = 0.9$  and  $w_{min} = 0.4$ , however, these values are user configurable and could be changed depending upon the problem being explored. Similarly,  $\theta_2$  and  $\theta_3$  were treated as constants, where  $\theta_2 = \theta_3 = 1.5$ . While updating the particles, parameter bounds were enforced using reflection boundary conditions (Algo. 2).

```

1 if  $z_{i,j}^{old} < z_i^{min}$  then
2    $z_{i,j}^{new} = z_{i,j}^{old} + (z_i^{min} - z_{i,j}^{old})$  if  $z_{i,j}^{new} > z_i^{max}$  then
3      $z_{i,j}^{new} = z_i^{max}$ 
4   end
5 end
6 if  $z_{i,j}^{old} > z_i^{max}$  then
7    $z_{i,j}^{new} = z_{i,j}^{old} + (z_i^{old} - z_{i,j}^{max})$  if  $z_{i,j}^{new} < z_i^{min}$  then
8      $z_{i,j}^{new} = z_i^{min}$ 
9   end
10 end
```

**Algorithm 2:** Pseudo code for the reflective boundary conditions used by the dynamic optimization with particle swarms (DOPS) method.

After every  $M$  function evaluations, particles were randomly redistributed to a new sub-swarm, and updated according to Eqn. (2.3). This process continued for a maximum of  $\mathcal{F}N$  functions evaluations, where  $\mathcal{F}$  denotes the fraction of function evaluations used during the particle swarm phase of DOPS:

$$\mathcal{F} = \left( \frac{NP}{N} \right) j \quad (2.5)$$

The quantity  $NP$  denotes the total number of particles in the swarm,  $N$  denotes the total possible number of function evaluations, while the counter  $j$  denotes the number of successful particle swarm iterations (each costing  $NP$  function evaluations). If the simulation error stagnated e.g., did not change by more than 1% for a specified number of evaluations (default value of 4), the swarm phase was terminated and DOPS switched to exploring parameter space using the DDS approach using the remaining  $(1 - \mathcal{F})N$  function evaluations.

## Phase 2: DDS phase.

**input** : Candidate vector  $\mathcal{G}$  from swarm search and  $(1 - \mathcal{F})N$  evaluations

**output**: Optimized parameter vector of size  $1 \times K$

```

1 while  $j \leq (1 - \mathcal{F})N$  do
2   | Assign probability of perturbation to each dimension  $\mathcal{P}_i$  according to
   | equation 7;
3   | Select a subset of dimensions based on a threshold value for perturbation;
4   | Update candidate solution  $\mathcal{G}(J)$  according to equation 5;
5   | Ensure updated solution  $\mathcal{G}_{new}(J)$  is within bounds using Algorithm 2;
6 end
```

**Algorithm 3:** Pseudo code for the Dynamically Dimensioned Search (DDS) method.

Dynamically Dimensioned Search (DDS) is a single solution based search algorithm. DDS is used to obtain good solutions to high-dimensional search problems within a fixed number of function evaluations. DDS starts as a global

search algorithm by perturbing all the dimensions. Later the number of dimensions that are perturbed is decreased with a certain probability. The probability that a certain dimension is perturbed reduces (a minimum of one dimension is always perturbed) as the iterations increase. This causes the algorithm to behave as a local search algorithm as the number of iterations increase. The perturbation magnitude of each dimension is from normal distribution with zero mean. The standard deviation that was used in the original DDS paper and the current study is 0.2. DDS performs a greedy search where the solution is updated only if it is better than the previous solution. The combination of perturbing a subset of dimensions along with greedy search indirectly relies on model sensitivity to a specific parameter combination. The reader is requested to refer to the original paper by Tolson and Shoemaker for further detail [332].

At the conclusion of the swarm phase, the overall best particle,  $\mathcal{G}_k$ , over the  $k$  sub-swarms was used to initialize the DDS phase. DOPS takes at least  $(1 - \mathcal{F})N$  function evaluations during the DDS phase and then terminates the search. For the DDS phase, the best parameter estimate was updated using the rule:

$$\mathcal{G}_{new}(J) = \begin{cases} \mathcal{G}(\mathbf{J}) + \mathbf{r}_{normal}(\mathbf{J})\sigma(\mathbf{J}), & \text{if } \mathcal{G}_{new}(\mathbf{J}) < \mathcal{G}(\mathbf{J}). \\ \mathcal{G}(\mathbf{J}), & \text{otherwise.} \end{cases} \quad (2.6)$$

where  $\mathbf{J}$  is a vector representing the subset of dimensions that are being perturbed,  $\mathbf{r}_{normal}$  denotes a normal random vector of the same dimensions as  $\mathcal{G}$ , and  $\sigma$  denotes the perturbation amplitude:

$$\sigma = R(\mathbf{p}^U - \mathbf{p}^L) \quad (2.7)$$

where  $R$  is the scalar perturbation size parameter,  $\mathbf{p}^U$  and  $\mathbf{p}^L$  are  $(\mathcal{K} \times 1)$  vectors that represent the maximum and minimum bounds on each dimension. The

set  $\mathbf{J}$  was constructed using a probability function  $\mathcal{P}_i$  that represents a threshold for determining whether a specific dimension  $j$  was perturbed or not;  $\mathcal{P}_i$  is monotonically decreasing function of function evaluations:

$$\mathcal{P}_i = 1 - \log \left[ \frac{i}{(1 - \mathcal{F})\mathcal{N}} \right] \quad (2.8)$$

where  $i$  is the current iteration. After  $\mathcal{P}_i$  is determined, we drew  $\mathcal{P}_j$  from a uniform distribution for each dimension  $j$ . If  $\mathcal{P}_j < \mathcal{P}_i$  was included in  $\mathbf{J}$ . Thus, the probability that a dimension  $j$  was perturbed was inversely proportional to the number of function evaluations. DDS updates are greedy;  $\mathcal{G}_{new}$  becomes the new solution vector only if it is better than  $\mathcal{G}$ .

### 2.5.3 Multiswitch DOPS

We investigated whether switching search methods more than once would result in better performance; this DOPS variant is referred to as multiswitch DOPS or msDOPS. msDOPS begins with the PSO phase and uses the same criteria as DOPS to switch to the DDS phase. However, msDOPS can switch back to a PSO search when the DDS phase has reduced the functional value to 90% of its initial value. Should the DDS phase fail to improve the functional value sufficiently, this version is identical to DOPS. When the switch from DDS to PSO occurs, we use the best solution from DDS to seed the particle swarm. DOPS and msDOPS source code is available for download under a MIT license at <http://www.varnerlab.org>.

## 2.5.4 Comparison Techniques.

The implementations of particle swarm optimization, simulated annealing, and genetic algorithms are the ones given in Matlab R2017A (`particleswarm`, `simulannealbnd` and `ga`). The implementation of DE used was developed by R.Storn and available at <http://www1.icsi.berkeley.edu/~storn/code.html>. The version of eSS used was Release 2014B - AMIGO2014bench VERSION WITH eSS MAY-2014-BUGS FIXED - JRB, released by the Process Engineering Group IIM-CSIC. The genetic algorithm, particle swarm, and differential evolution algorithms were run with a 40 particles to be directly comparable to the number of particles used in the PSO phase of DOPS. For comparison, the version of CM-AES used was `cmaes.m`, Version 3.61.beta from [https://www.lri.fr/~hansen/cmaes\\_inmatlab.html](https://www.lri.fr/~hansen/cmaes_inmatlab.html). The scripts used to run the comparison methods are also available at <http://www.varnerlab.org>.

## 2.6 Tables

Table 2.1: Table with optimization settings and results for the coagulation problem, the benchmarks and test functions using DOPS. For each problem the bounds on the parameter vector, the total number of function evaluations, the best initial objective value and the best final objective value are specified. Here *pnom* indicates the nominal or true parameter vector of the model. Nominal objective value represents the objective value using the true parameter vector or the nominal parameter vector. The CPU time is the time taken for the problem on a 2.4GHz Intel Xeon Architecture running Matlab 2014b.

	Coagulation	B1	B4	Ackley	Rastrigin
Evaluations	4000	4000	4000	4000	4000
Lower Bound	0.001.pnom	0.2.pnom	0.2.pnom	-15	-5.12
Upper Bound	1000.pnom	5.pnom	5.pnom	30	5.12
CPU Time	10.1 hrs	38.3 hrs	6.2 min	2.8 s	2.6 s
Scaled initial error	1.0	1.0	1.0	1.0	1.0
Scaled final error	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
Scaled nominal error	0.42	0.1	< 0.01	0	0

Table 2.2: Error analysis for the human coagulation model. The coagulation model was trained on coagulation initiated with TF/FVIIa at 5 nM and the 5 pM to obtain the optimal parameters. Using these optimal parameters, coagulation dynamics were predicted for varying initiator concentrations (500 pM, 50 pM and 10 pM). Model agreement with measurements was quantified using normalized squared error. The normalized squared error is defined as  $N.S.E. = (1/\max(\mathbf{X})) * (\|\mathbf{Y}, \mathbf{X}\|/\sqrt{N})$  where  $\mathbf{X}$  is the experimental data,  $\mathbf{Y}$  is the model simulation data interpolated onto the experimental time scale and  $N$  is the total number of experimental time points.

TF/FVIIa concentration	Normalized S.E.	Category
5 nM	0.1336	Training
500 pM	0.2242	Prediction
50 pM	0.3109	Prediction
10 pM	0.2023	Prediction
5 pM	0.1170	Training



# CHAPTER 3

## REDUCED ORDER MODELING AND ANALYSIS OF THE HUMAN COMPLEMENT SYSTEM

### Abstract

<sup>1</sup> Complement is an important pathway in innate immunity, inflammation, and many disease processes. However, despite its importance, there are few validated mathematical models of complement activation. In this study, we developed an ensemble of experimentally validated reduced order complement models. We combined ordinary differential equations with logical rules to produce a compact yet predictive model of complement activation. The model, which described the lectin and alternative pathways, was an order of magnitude smaller than comparable models in the literature. We estimated an ensemble of model parameters from *in vitro* dynamic measurements of the C3a and C5a complement proteins. Subsequently, we validated the model on unseen C3a and C5a measurements not used for model training. Despite its small size, the model was surprisingly predictive. Global sensitivity and robustness analysis suggested complement was robust to any single therapeutic intervention. Only the simultaneous knockdown of both C3 and C5 consistently reduced C3a and C5a formation from all pathways. Taken together, we developed a validated mathematical model of complement activation that was computationally inexpensive, and could easily be incorporated into pre-existing or new pharmacokinetic models of immune system function. The model described experimental data, and predicted the need for multiple points of therapeutic intervention

---

<sup>1</sup>This chapter has previously been published as [282].

to fully disrupt complement activation.

### 3.1 Introduction

Complement is an important pathway in innate immunity. It plays a significant role in inflammation, host defense as well as many disease processes. Complement was discovered in the late 1880s where it was found to 'complement' the bactericidal activity of natural antibodies [246]. However, research over the past decade has suggested the importance of complement extends beyond innate immunity. For example, complement contributes to tissue homeostasis [265]. It has also been linked with several diseases including Alzheimers, Parkinson's, multiple sclerosis, schizophrenia, rheumatoid arthritis and sepsis [266, 272]. Complement also plays positive and negative roles in cancer; attacking tumor cells with altered surface proteins in some cases, while potentially contributing to tumor growth in others [288, 267]. Lastly, several other important biochemical systems are integrated with complement including the coagulation cascade, the autonomous nervous system and inflammation [267]. Thus, complement is important in a variety of beneficial and potentially harmful functions in the body. Despite its importance, there have been few approved complement specific therapeutics, largely because of safety concerns and challenging pharmacokinetic constraints, however, progress is being made [268].

The complement cascade involves many soluble and cell surface proteins, receptors and regulators [349, 350]. The outputs of complement are the Membrane Attack Complex (MAC), and the inflammatory mediator proteins C3a and C5a. The membrane attack complex, generated during the terminal phase of the re-

sponse, forms transmembrane channels which disrupt the membrane integrity of targeted cells, leading to cell lysis and death. On the other hand, the C3a and C5a proteins act as a bridge between innate and adaptive immunity, and play an important role in regulating inflammation [288]. Complement activation takes place through three pathways: the classical, the lectin and the alternate pathways. The classical pathway is triggered by antibody recognition of foreign antigens or other pathogens. A multimeric protein complex C1 binds antibody-antigen complexes and undergoes a conformational change, leading to an activated form with proteolytic activity. The activated C1-complex cleaves soluble complement proteins C4 and C2 into C4a, C4b, C2a and C2b, respectively. The C4a and C2b fragments bind to form the C4bC2a protease, also known as the classical pathway C3 convertase (CP C3 convertase). The lectin pathway is initiated through the binding of L-ficolin or Mannose Binding Lectin (MBL) to carbohydrates on the surfaces of bacterial pathogens. These complexes, in combination with mannose-associated serine proteases 1 and 2 (MASP-1/2), also cleave C4 and C2, leading to additional CP C3 convertase. Thus, the classical and lectin pathways, initiated by different cues on foreign surfaces, converge at the CP C3 convertase. On the other hand, the alternate pathway is activated by a 'tickover' mechanism in which complement protein C3 is spontaneously hydrolyzed to form an activated intermediate C3w; C3w recruits factor B and factor D, leading to the formation of C3wBb. C3wBb cleaves C3 into C3a and C3b, where the C3b fragment further recruits additional factor B and factor D to form C3bBb, the alternate C3 convertase (AP C3 convertase) [249]. The role of classical and alternate C3 convertases is varied. First, AP C3 convertases mediate signal amplification. AP C3 convertases cleave C3 into C3a and C3b; the C3b fragment is then free to form additional alternate C3 convertases, thereby form-

ing a positive feedback loop. Next, AP/CP C3 convertases link complement initiation with the terminal phase of the cascade through the formation of C5 convertases. Both classical and alternate C3 convertases can recruit C3b subunits to form the classical pathway C5 convertase (C4bC2aC3b, CP C5 convertase), and the alternate pathway C5 convertase (C3bBbC3b, AP C5 convertase), respectively. Both C5 convertases cleave C5 into the C5a and C5b fragments. The C5b fragment, along with the complement proteins C6, C7, C8 and multiple C9s, form the membrane attack complex. On the other hand, both C3a and C5a are important inflammatory signals involved in several responses [349, 350]. Thus, the complement cascade attacks invading pathogens, while acting as a beacon for adaptive immunity.

The complement cascade is regulated by plasma and host cell surface proteins which balance host safety with effectiveness. The initiation of the classical pathway via complement protein C1 is controlled by the C1 Inhibitor (C1-Inh); C1-Inh irreversibly binds to and deactivates the active subunits of C1, preventing chronic complement activation [348]. Regulation of upstream processes in the lectin and alternate pathways also occurs through the interaction of the C4 binding protein (C4BP) with C4b, and factor H with C3b [31]. Interestingly, both factor H and C4BP are capable of binding their respective targets while in convertase complexes as well. At the host cell surface, membrane co-factor protein (MCP or CD46) can interact with C4b and C3b, which protects the host cell from complement self-activation [270]. Delay accelerating factor (DAF or CD55) also recognizes and dissociates both C3 and C5 convertases on host cell surfaces [195]. More generally the well known inflammation regulator Carboxypeptidase-N has broad activity against the complement proteins C3a, C4a, and C5a, rendering them inactive by cleavage of carboxyl-terminal argi-

nine and lysine residues [186]. Although Carboxypeptidase-N does not directly influence complement activation, it silences the important inflammatory signals produced by complement. Lastly, assembly of the MAC complex itself can be inhibited by vitronectin and clusterin in the plasma, and CD59 at the host surface [56, 370]. Thus, there are many points of control which influence complement across the three activation pathways.

Developing quantitative mathematical models of complement could be crucial to fully understanding its role in the body. Traditionally, complement models have been formulated as systems of linear or non-linear ordinary differential equations (ODEs). For example, Hirayama et al., modeled the classical complement pathway as a system of linear ODEs [139], while Korotaevskiy and co-workers modeled the classical, lectin and alternate pathways as a system of non-linear ODEs [174]. More recently, large mechanistic models of sections of complement have also been proposed. For example, Liu et al., analyzed the formation of the classical and lectin C3 convertases, and the regulatory role of C4BP using a system of 45 non-linear ODEs with 85 parameters [187]. Zewde and co-workers constructed a detailed mechanistic model of the alternative pathway which consisted of 107 ODEs and 74 kinetic parameters and delineated between the fluid, host and pathogen surfaces [370]. However, these previous studies involved large models with little experimental validation. Thus, while these models are undoubtedly important theoretical tools, it is unclear if they can describe or quantitatively predict complement measurements. The central challenge of complement model identification is the estimation of model parameters from experimental measurements. Unlike other important cascades, such as coagulation where there are well developed experimental tools and publicly available data sets, the data for complement is relatively sparse. Data sets with

missing or incomplete data, and limited dynamic data also make the identification of large mechanistic complement models difficult. Thus, reduced order approaches which describe the biology of complement using a limited number of species and parameters could be important for pharmacokinetic model development, and for our understanding of the varied role of complement in the body.

## 3.2 Results

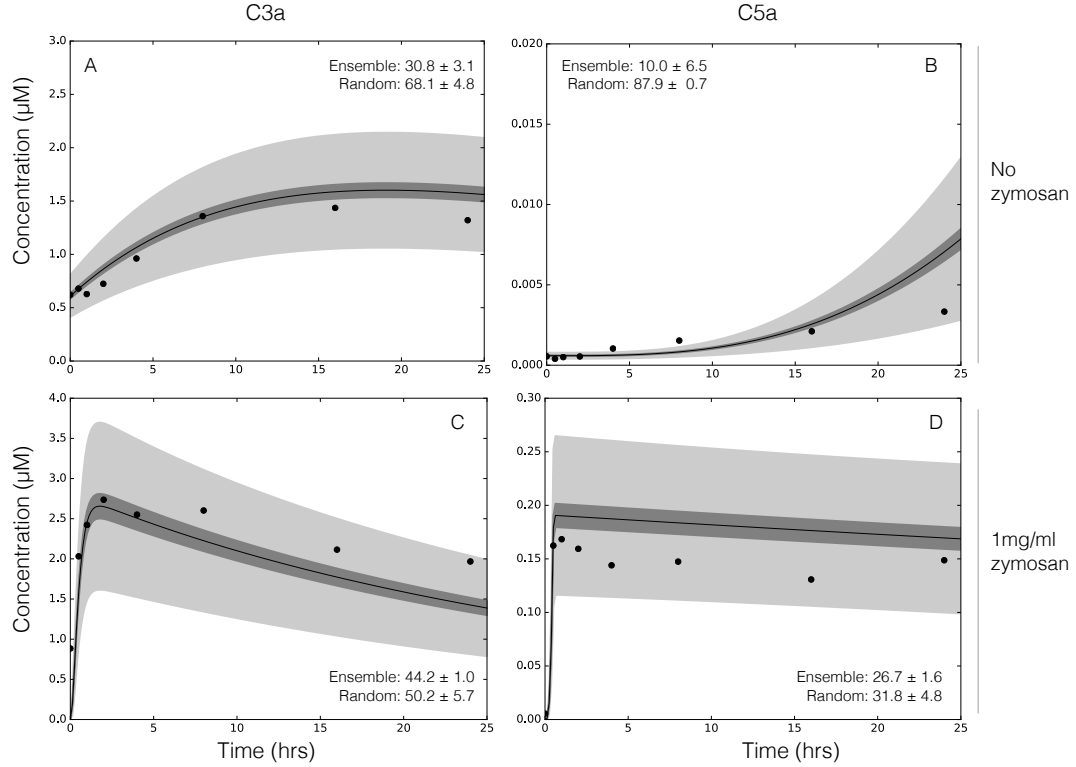


Figure 3.1: Reduced order complement model training. An ensemble of model parameters were estimated using multiobjective optimization from C3a and C5a measurements with and without zymosan [227]. The model was trained using C3a and C5a data generated from the alternative pathway (A–B) and lectin pathway initiated with 1 mg/ml zymosan (C–D). The solid black lines show the simulated mean value of C3a or C5a for the ensemble, while the dark shaded region denotes the 99% confidence interval of mean. The light shaded region denotes the 99% confidence interval of the simulated C3a and C5a concentration. All initial conditions were assumed to be at their physiological serum levels unless otherwise noted.

In this study, we estimated an ensemble of experimentally validated reduced order complement models using multiobjective optimization. The modeling approach combined ordinary differential equations with logical rules to produce a complement model with a limited number of equations and parameters. The reduced order model, which described the lectin and alternative pathways, con-

sisted of 18 differential equations with 28 parameters. Thus, the model was an order of magnitude smaller and included more pathways than comparable models in the literature. We estimated an ensemble of model parameters from *in vitro* time series measurements of the C3a and C5a complement proteins. Subsequently, we validated the model on unseen C3a and C5a measurements not used for model training. Despite its size, the model was surprisingly predictive. After validation, we performed global sensitivity and robustness analysis to estimate which parameters and species controlled model performance. Sensitivity analysis suggested CP C3 and C5 convertase parameters were critical, while robustness analyses suggested complement was robust to any single therapeutic intervention; only the knockdown of both C3 and C5 consistently reduced C3a and C5a formation for all cases. Taken together, we developed a reduced order complement model that was computationally inexpensive, and could easily be incorporated into pre-existing or new pharmacokinetic models of immune system function. The model described experimental data, and predicted the need for multiple points of intervention to disrupt complement activation.

### **3.2.1 Reduced order complement network.**

The complement model described the alternate and lectin pathways (Fig. 3.2). A trigger event initiated the lectin pathway (encoded as a logical rule), which activated the cleavage of C2 and C4 into C2a, C2b, C4a and C4b, respectively. Classical Pathway (CP) C3 convertase (C4aC2b) then catalyzed the cleavage of C3 into C3a and C3b. The alternate pathway was initiated through the spontaneous hydrolysis of C3 into C3a and C3b. The C3b fragments generated by hydrolysis (or by CP C3 convertase) could then form the alternate pathway (AP)



C3 convertase (C3bBb). We did not consider C3w, nor the formation of the initial alternate C3 convertase (C3wBb). Rather, we assumed C3w was equivalent to C3b and only modeled the formation of the main AP C3 convertase. Both the CP and AP C3 convertases catalyzed the cleavage of C3 into C3a and C3b. A second C3b fragment could then bind with either the CP or AP C3 convertase to form the CP or AP C5 convertase (C4bC2aC3b or C3bBbC3b). Both C5 convertases catalyzed the cleavage of C5 into the C5a and C5b fragments. In this study, we simplified the model by assuming both factor B and factor D were in excess. However, we did explicitly account for the action of two other control proteins, factor H and C4BP. Lastly, we did not consider MAC formation, instead we stopped at C5a and C5b. Lectin pathway activation, and C3/C5 convertase activity were modeled using a combination of saturation kinetics and non-linear transfer functions, which resulted in a significant size reduction of the model, while maintaining performance. Binding interactions were modeled using mass-action kinetics, where we assumed all binding was irreversible. Thus, while the reduced order complement model encoded significant biology, it was highly compact consisting of only 18 differential equations and 28 model parameters. Next, we estimated an ensemble of model parameters from time series measurements of the C3a and C5a complement proteins.

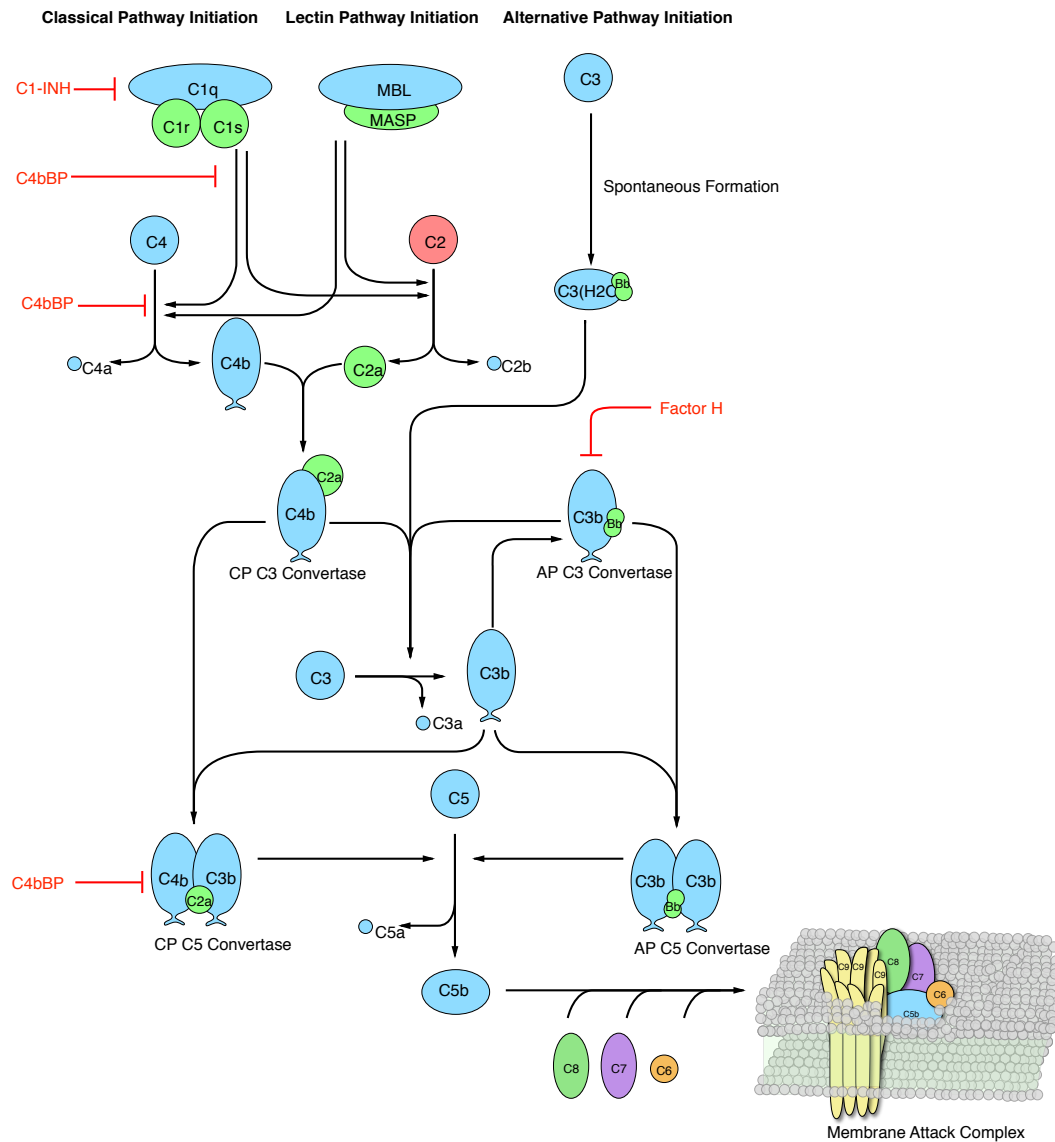


Figure 3.2: Simplified schematic of the human complement system. The complement cascade is activated through three pathways: the classical, the lectin, and the alternate pathways. Complement initiation results in the formation of classical or alternative C3 convertases, which amplify the initial complement response and signal to the adaptive immune system by cleaving C3 into C3a and C3b. C3 convertases further react to form C5 convertases which catalyze the cleavage of the C5 complement protein to C5a and C5b. C5b is critical to the formation of the membrane attack complex (MAC), while C5a recruits an adaptive immune response.

### 3.2.2 Estimating an ensemble of reduced order complement models.

A critical challenge for the development of any dynamic model is the estimation of model parameters. We estimated an ensemble of complement model parameters using *in vitro* time-series data sets generated with and without zymosan, a lectin pathway activator [227]. The residual between model simulations and experimental measurements was minimized using the Pareto Optimal Ensemble Technique (JuPOETs) [20] starting from a initial guess generated by the dynamic optimization with particle swarms (DOPS) routine. Unless otherwise specified, all initial conditions were assumed to be at their mean physiological values. While we had significant training data, the parameter estimation problem was underdetermined (we were not able to uniquely determine model parameters). Thus, instead of using the best-fit yet uncertain parameter set, we estimated an ensemble of probable parameter sets to quantify model uncertainty ( $N = 2100$ , see materials and methods). The complement model ensemble captured the behavior of both the alternate and lectin pathways (Fig. 3.1). To estimate alternate pathway model parameters, we used C3a and C5a measurements in the absence of zymosan (Fig. 3.1A and B). On the other hand, lectin pathway parameters were estimated from C3a and C5a measurements in the presence of 1mg/ml zymosan (Fig. 3.1C and D). The reduced order model reproduced a panel of alternate and lectin pathway data sets in the neighborhood of physiological factor and inhibitor concentrations. The model fit for parameter sets estimated by JuPOETs, quantified by the Akaike information criterion (AIC), was statistically significantly different than a random parameter control for each case at a 95% confidence level. However, it was unclear whether the reduced order model

could predict new data, without updating the model parameters. To address this question, we fixed the model parameters and simulated data sets not used for model training.

We tested the predictive power of the reduced order complement model with data not used during model training (Fig. 3.4). Six validation cases were considered, three for C3a and C5a each, respectively. Similar to model training, we compared the AIC for each prediction case to a randomized parameter family. All model parameters and initial conditions were fixed for the validation simulations (with the exception of zymosan, and other experimentally mandated changes). The ensemble of reduced order models predicted the qualitative dynamics of C3a formation (Fig. 3.4, top), and C5a formation (Fig. 3.4, bottom) at three inducer concentrations. For each training case, the AIC was statistically significantly different than the random parameter control for a 95% confidence level. The rate of C3a formation and C3a peak time were directly proportional to initiator dose. Similarly, the C5a plateau and rate of formation were also directly proportional to initiator dose, with the lag time being indirectly proportional to initiator exposure for both C3a and C5a. However, there were shortcomings with model performance. First, while the overall C3a trend was captured (within the 99% confidence interval), the C3a dynamics were too fast with the exception of the low dose case. We believe the C3a time scale was related to our choice of training data, how we modeled the tickover mechanism, and factor B and D limitation. We trained the model using either no or 1 mg/ml zymosan, but predicted cases in a different initiator range; comparing training to prediction, the model performance e.g., the shape of the C3a trajectory was biased towards either high or very low initiator doses. Next, tickover was modeled as a first-order generation processes where C3wBb formation and

activity was lumped into the AP C3 convertase. Thus, we skipped an important upstream step which could influence AP C3 convertase formation by attenuating the rate C3 cleavage into C3a and C3b. We also assumed both factor B and factor D were not limiting, thereby artificially accelerating the rate of AP C3 convertase formation. The C5a predictions followed a similar trend as C3a; we captured the long-time C5a behavior but over predicted the time scale of C5 cleavage. However, because the C5a time scale depends strongly upon C3 convertase formation, we can likely correct the C5 issues by fixing the rate of C3 cleavage. Despite these shortcomings, we qualitatively predicted experimental measurements not used for model training typically within the 99% confidence of the ensemble, for three inducer levels. Next, we used global sensitivity and robustness analysis to determine which parameters and species controlled the performance of the complement model.

### **3.2.3 Global analysis of the reduced order complement model.**

We conducted sensitivity analysis to estimate which parameters controlled the performance of the reduced order complement model. We calculated the total sensitivity of the C3a and C5a residual to changes in model parameters with and without zymosan (Fig. 3.3). In the absence of zymosan (where only the alternative pathway is active), the most sensitive parameter was the rate constant governing the assembly of the AP C3 convertase, as well as the rate constant controlling basal C3b formation via the tickover mechanism. The C5a trajectory was sensitive to the AP C5 convertase kinetic parameters (Fig. 3.3A). Interestingly, neither the rate nor the saturation constant governing AP C3 convertase activity were sensitive in the absence of zymosan. Thus, C3a formation in the al-

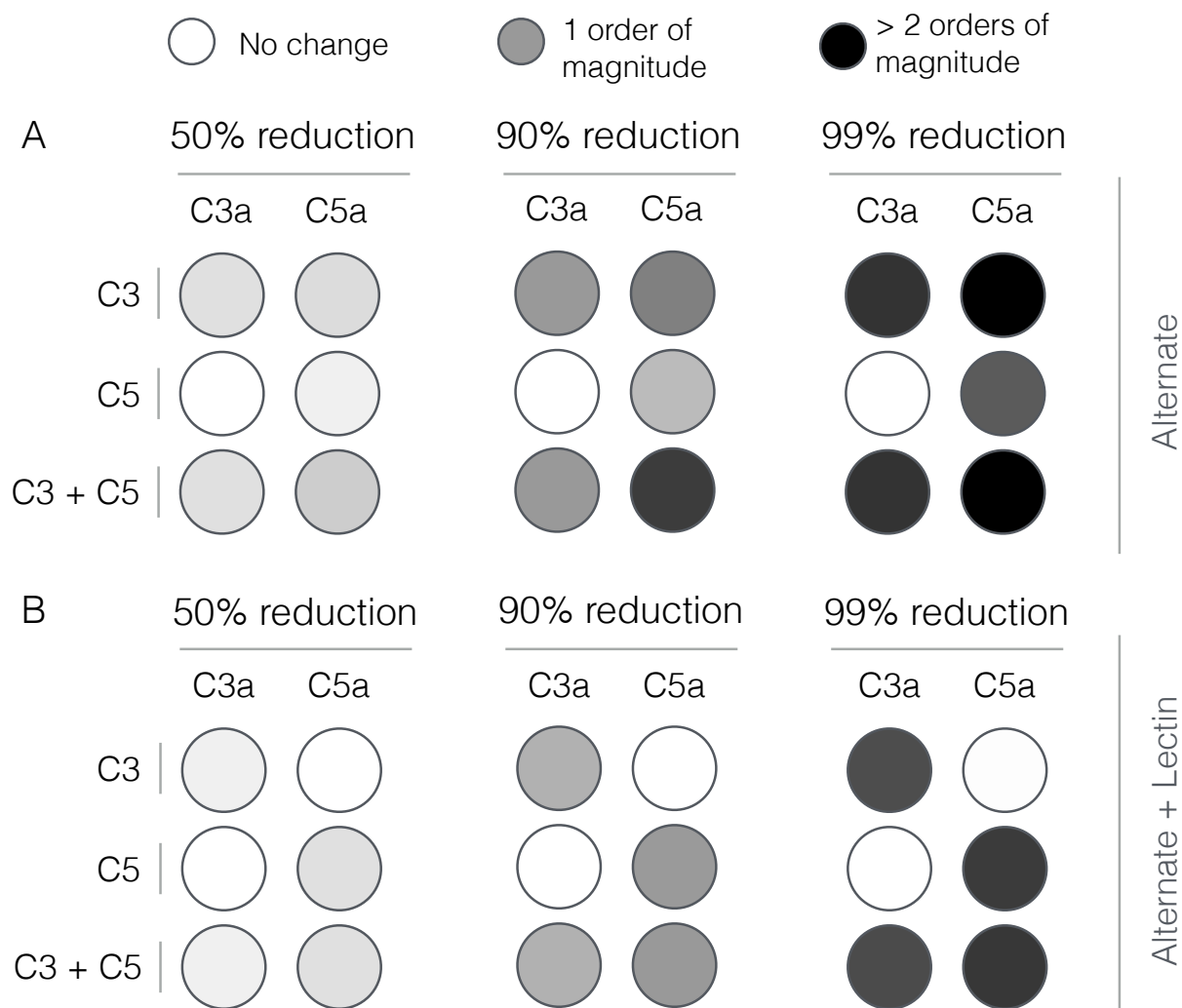


Figure 3.3: Global sensitivity analysis of the reduced order complement model. Sensitivity analysis was conducted on the two objectives used for model training. **A:** Sensitivity of the C3a and C5a residual w/o zymosan. **B:** Sensitivity of the C3a and C5a residual with 1 mg/ml zymosan. The bars denote the mean total sensitivity index for each parameter, while the error bars denote the 95% confidence interval. **C:** Pathways controlled by the sensitivity parameters. Bold black lines indicate the pathway involves one or more sensitive parameters, while the red lines show current therapeutics targets. Current complement therapeutics were taken from the review of Morgan and Harris [229].

ternative pathway was more heavily influenced by the spontaneous hydrolysis of C3, rather than AP C3 convertase activity, in the absence of zymosan. In the presence of zymosan, the C3a residual was controlled by the formation and ac-

tivity of the CP C3 convertase, as well as tickover and degradation parameters. On the other hand, the C5a residual was controlled by the formation and activity of CP C5 convertase, and tickover C3b formation in the presence of zymosan (Fig. 3.3B). The lectin initiation parameters were sensitive, but to a lesser extent than CP convertase kinetic parameters and tickover C3b formation. Thus, sensitivity analysis suggested that CP C3/C5 convertase formation and activity dominated in the presence of zymosan, but tickover parameters and AP C5 convertase were more important without initiator. AP C3 convertase assembly was important, but its activity was not. Next, we compared the sensitivity results to current therapeutic approaches; pathways involving sensitive parameters have been targeted for clinical intervention (Fig. 3.3C). In particular, the sensitivity analysis suggested AP/CP C5 convertase inhibitors, or interventions aimed at attenuating C3 or C5 would most strongly influence complement performance. Thus, there was at least a qualitative overlap between sensitivity and the potential of biochemical efficacy. However, total sensitivity coefficients quantify how simultaneous changes in many parameters e.g., rate or saturation constants affect model performance (in this case model fit). To better understand the role of each parameter, and parameter combination, we explored how finite changes in parameter combinations influenced model performance.

Pairwise parameter perturbations identified crosstalk within the complement model (Fig. 3.5). We perturbed each pairwise parameter combination by 10%, and calculated the distance between the perturbed and nominal state for each parameter set in the ensemble. We then clustered the mean response of each parameter combination based upon the euclidian distance between the perturbed and nominal states into low (green), medium (red) and high (blue) response clusters. A low response (white) meant the parameter perturbations did

not significantly change the system state compared with the nominal case. Four of the 28 parameters (or approximately 14% of the overall model parameters) were in the high response cluster (Fig. 3.5, blue cluster). These parameters included the rate constant controlling the basal formation of C3b (#12), C3a degradation (#26) as well as the catalytic rate constant governing CP C3 convertase activity (#22). The only C5 related parameter in the high response group was the rate constant controlling the formation of CP C5 convertase (#15). Approximately, 36%, or 10 of the 28 model parameters, were clustered in the medium impact cluster (Fig. 3.5, red cluster). Three parameters (#10, #1, #27) were especially important in this cluster; The reaction order governing CP C3 convertase activity was important (#10), along with the rate constant controlling C4a and C4b formation from C4 in the lectin initiation pathway (#1), and the constant controlling the inhibitory action of C4BP (#27). Lastly, 50% of the model parameters were clustered in the low response cluster (Fig. 3.5, green cluster). Many of these parameters influenced complement activation; for example, parameter #23 (the CP C3 convertase saturation constant) was important, just not to the extent of other model parameters. Pairwise synergistic interactions between parameters were also identified. For example, in the high impact cluster, three synergistic relationships were identified, a single positive and two negative cases. Parameters #12 (rate constant governing basal C3b formation) and #15 (formation of CP C5 convertase) acted synergistically to increase the system response. On the other hand, simultaneously changing parameters #12 and #22 or #15 and #26 decreased the system response relative to a single perturbation. However, the most striking examples of synergy occurred in the medium impact cluster; for example, simultaneously increasing parameters #13 (rate constant governing AP C3 convertase formation) and #19 (saturation constant governing AP



C5 convertase activity) significantly changed the model state. Changes in parameter #3 (rate constant governing C2a and C2b formation from C2) showed both positive and negative synergistic effects depending upon the other parameter that was perturbed. Taken together, sensitivity coefficients quantified how changes in parameters or parameter combinations affected model performance. However, individual parameters e.g., rate or saturation constants are not easily druggable. To more closely simulate a clinical intervention e.g., administration of anti-complement inhibitors, we performed knock-down analysis on the initial values of C3 and C5 in the absence and presence of flow.

Knock-down analysis in the absence of flow suggested there was no single intervention that inhibited complement activation in the presence of both initiation pathways (Fig. 3.6). Robustness coefficients quantify the response of a protein to a macroscopic structural or operational perturbation to a biochemical network. Here, we computed how the C3a and C5a trajectories responded to a decrease in the initial abundance of C3 and/or C5 with and without lectin initiator. We simulated the addition of different doses of anti-complement inhibitor cocktails by decreasing the initial concentration of C3, C5 or the combination of C3 and C5 by 50%, 90% and 99%. This would be conceptually analogous to the administration of a C3 inhibitor e.g., Compstatin alone or combination with Eculizumab (Fig. 3.3C). The response of the complement model to different knock-down magnitudes was non-linear; a 90% knock-down had an order of magnitude more impact than a 50% knock-down. As expected, a C5 knock-down had no effect on C3a formation for either the alternate (Fig. 3.6A) or lectin pathways (Fig. 3.6B). However, C3a and to a greater extent C5a abundance decreased with decreasing C3 concentration in the alternate pathway (Fig. 3.6A). This agreed with the sensitivity results; changes in AP C3-convertase formation

affected the downstream dynamics of C5a formation. Thus, if we only considered the alternate pathway, C3 alone could be a reasonable target, especially given that C5a formation was surprisingly robust to C5 levels in the alternate pathway. Yet, when both pathways were activated, C5a levels were robust to the initial C3 concentration (Fig. 3.6B); even 1% of the nominal C3 was able to generate enough AP/CP C5 convertase to maintain C5a formation. Thus, the only reliable intervention that consistently reduced both C3a and C5a formation for all cases was a knockdown of both C3 and C5. For example, a 90% decrease of both C3 and C5 reduced the formation of C5a by an order of magnitude, while C3a was reduced to a lesser extent (Fig. 3.6B).

### 3.3 Discussion

In this study, we estimated an ensemble of experimentally validated reduced order complement models using multiobjective optimization. The modeling approach combined ordinary differential equations with logical rules to produce a complement model with a limited number of equations and parameters. The reduced order model, which described the lectin and alternative pathways, consisted of 18 differential equations with 28 parameters. Thus, the model was an order of magnitude smaller and included more pathways than comparable mathematical models in the literature. We estimated an ensemble of model parameters from *in vitro* time series measurements of the C3a and C5a complement proteins. Subsequently, we validated the model on unseen C3a and C5a measurements that were not used for model training. Despite its small size, the model was surprisingly predictive. After validation, we performed global sensitivity and robustness analysis to estimate which parameters and species controlled model performance. These analyses suggested complement was robust to any single therapeutic intervention. The only intervention that consistently reduced C3a and C5a formation for all cases was a knockdown of both C3 and C5. Taken together, we developed a reduced order complement model that was computationally inexpensive, and could easily be incorporated into pre-existing or new pharmacokinetic models of immune system function. The model described experimental data, and predicted the need for multiple points of intervention to disrupt complement activation.

There has been a paucity of validated mathematical models of complement pathway activation. To our knowledge, this study is one of the first complement models that combined multiple initiation pathways with experimental valida-

tion of important complement products like C5a. However, there have been several theoretical models of components of the cascade in the literature. Liu and co-workers modeled the formation of C3a through the classical pathway using 45 non-linear ODEs [187]. In contrast, in this study we modeled lectin mediated C3a formation using only five ODEs. Though we did not model all the initiation interactions in detail, especially the cross-talk between the lectin and classical pathways, we successfully captured C3a dynamics with respect to different concentrations of lectin initiators. The model also captured the dynamics of C3a and C5a formed from the alternate pathway using only seven ODEs. The reduced order model predictions of C5a were qualitatively similar to the theoretical complement model of Zewde et al., which involved over 100 ODEs [370]. However, we found that the C3a produced in the alternate pathway was nearly three orders of magnitude greater than the C5a generated. While this was in agreement with the experimental data [227], it differed from the theoretical predictions made by Zewde et al., who showed C3a was eight orders of magnitude greater than the C5a concentration [370]. In our model, the time profile of both C3a and C5a generated changed with respect to the quantity of zymosan (the lectin pathway initiator). In particular, the C3a peak time was directly proportional to initiator, while the lag phase for generation was inversely proportional to the initiator concentration. Korotaevskiy et al. showed a similar trend using a theoretical model of complement, albeit for much shorter time scales [174]. Thus, the reduced order complement model performed at least as well as existing larger mechanistic models, despite being significantly smaller.

Global analysis of the complement model suggested potentially important therapeutic targets. Complement malfunctions are implicated in a spectrum of diseases, however the development of complement specific therapeutics has

been challenging [266, 229]. Previously, we have shown that mathematical modeling and analysis can be useful tools to estimate therapeutically important mechanisms [194, 236, 329, 262]. In this study, we analyzed a validated ensemble of reduced order complement models to better understand the strengths and weaknesses of the cascade. In the presence of an initiator, C3a and C5a formation was sensitive to CP C3/C5 convertase assembly and activity, and to a lesser extent lectin initiation parameters. Formation of the CP convertases can be inhibited by targeting upstream protease complexes like MASP-1,2 from the lectin pathway (or C1r, C1s from classical pathway). For example, Omeros, a protease inhibitor that targets the MASP-2 complex, has been shown to inhibit the formation of downstream convertases [298]. Lampalizumab and Bikacimab, which target factor B and factor D respectively, or naturally occurring proteins such as Cobra Venom Factor (CVF), an analogue of C3b, could also attenuate AP convertase formation [345, 168, 152]. Removing supporting molecules could also destabilize the convertases. For example, Novemed Therapeutics developed the antibody, NM9401 against propedien, a small protein that stabilizes alternate C3 convertase [16]. Lastly, convertase catalytic activity could be attenuated using small molecule protease inhibitors. All of these approaches are consistent with the results of the sensitivity analysis. On the other hand, robustness analysis suggested C3a and C5a generation could only be significantly attenuated by modulating the free levels of C3 and C5. The most commonly used anti-complement drug Eculizumab, targets the C5 protein [229]. Several other antibodies targeting C5 are also being developed; for example, LFG316 targets C5 in Age-Related Macular Degeneration [277], while Mubodina is used to treat Atypical Hemolytic-Uremic Syndrome (aHUS) [210]. Other agents such as Coversin [357] or the aptamer Zimura [97] could also be used to knockdown C5. The

peptide inhibitor Compstatin and its derivatives are promising approaches for the inhibition of C3 [204]. However, while the knockdown of C3 and C5 affect C3a and C5a levels downstream, the abundance, turnover rate and population variation of these proteins make them difficult targets [310, 325]. For example, the eculizumab dosage must be significantly adjusted during the course of treatment for aHUS [245]. A validated complement model, in combination with personalized pharmacokinetic models of immune system function, could be an important development for the field.

The performance of the complement model was impressive given its limited size. However, there are several questions that should be explored further. A logical progression for this work would be to expand the network to include the classical pathway and the formation of the membrane attack complex (MAC). However, time course measurements of MAC abundance (and MAC formation dynamics) are scarce, making the inclusion of MAC challenging. On the other hand, inclusion of classical pathway activation is straightforward. Liu et al., have shown cross-talk between the activation of the classical and lectin pathways through C reactive proteins (CRP) and L-ficolin (LF) under inflammation conditions [187]. Thus, inclusion of these species, in addition to a lumped activation term for the classical pathway should allow us to capture classical activation. Next, we should address the C3a time scale issue. We believe the C3a time scale was related to our choice of training data, how we modeled the tickover mechanism, and factor B and D limitation. Tickover was modeled as a first-order generation processes where C3wBb formation and activity was lumped into the AP C3 convertase. Thus, we skipped an important step which could strongly influence AP C3 convertase formation by slowing down the rate C3 cleavage into C3a and C3b. The model should be expanded to in-

clude the C3wBb intermediate, where C3wBb catalyzes C3 cleavage at a slow rate compared to normal AP or CP C3 convertases. We also assumed both factor B and factor D were not limiting, thereby artificially accelerating the rate of AP C3 convertase formation. This shortcoming could be addressed by including balances around factor B and D, and including these species in the appropriate kinetic rates. The C5a predictions also had an accelerated time scale. However, because the C5a time scale depended strongly upon C3 convertase formation, we can likely correct the C5 issues by fixing the rate of C3 cleavage. Lastly, we should also consider including the C2-bypass pathway, which was not included in the model. The C2-bypass mediates lectin pathway activation, without the involvement of MASP-1/2. Thus, this pathway could be important for understanding the role of MASP-1/2 inhibitors on complement activation.

## 3.4 Materials and Methods

### 3.4.1 Formulation and solution of the complement model equations.

We used ordinary differential equations (ODEs) to model the time evolution of complement proteins ( $x_i$ ) in the reduced order model:

$$\frac{1}{\tau_i} \frac{dx_i}{dt} = \sum_{j=1}^{\mathcal{R}} \sigma_{ij} r_j(\mathbf{x}, \epsilon, \mathbf{k}) \quad i = 1, 2, \dots, \mathcal{M} \quad (3.1)$$

where  $\mathcal{R}$  denotes the number of reactions and  $\mathcal{M}$  denotes the number of proteins in the model. The quantity  $\tau_i$  denotes a time scale parameter for species  $i$  which captures unmodeled effects. For the current study,  $\tau$  scaled with the level of initiator ( $z$ ) for C5a and C5b;  $\tau_i = z/z^*$  for  $i = \text{C5a, C5b}$  where  $z^*$  was 1mg/ml,  $\tau_i = 1$  for all other species. The quantity  $r_j(\mathbf{x}, \epsilon, \mathbf{k})$  denotes the rate of reaction  $j$ . Typically, reaction  $j$  is a non-linear function of biochemical and enzyme species abundance, as well as unknown model parameters  $\mathbf{k}$  ( $\mathcal{K} \times 1$ ). The quantity  $\sigma_{ij}$  denotes the stoichiometric coefficient for species  $i$  in reaction  $j$ . If  $\sigma_{ij} > 0$ , species  $i$  is produced by reaction  $j$ . Conversely, if  $\sigma_{ij} < 0$ , species  $i$  is consumed by reaction  $j$ , while  $\sigma_{ij} = 0$  indicates species  $i$  is not connected with reaction  $j$ . Species balances were subject to the initial conditions  $\mathbf{x}(t_o) = \mathbf{x}_o$ .

Rate processes were written as the product of a kinetic term ( $\bar{r}_j$ ) and a control term ( $v_j$ ) in the complement model. The kinetic term for the formation of C4a, C4b, C2a and C2b, lectin pathway activation, and C3 and C5 convertase activity was given by:

$$\bar{r}_j = k_j^{max} \epsilon_i \left( \frac{x_s^\eta}{K_{js}^\eta + x_s^\eta} \right) \quad (3.2)$$



where  $k_j^{max}$  denotes the maximum rate for reaction  $j$ ,  $\epsilon_i$  denotes the abundance of the enzyme catalyzing reaction  $j$ ,  $\eta$  denotes a cooperativity parameter, and  $K_{js}$  denotes the saturation constant for species  $s$  in reaction  $j$ . We used mass action kinetics to model protein-protein binding interactions within the network:

$$\bar{r}_j = k_j^{max} \prod_{s \in m_j^-} x_s^{-\sigma_{sj}} \quad (3.3)$$

where  $k_j^{max}$  denotes the maximum rate for reaction  $j$ ,  $\sigma_{sj}$  denotes the stoichiometric coefficient for species  $s$  in reaction  $j$ , and  $s \in m_j^-$  denotes the set of *reactants* for reaction  $j$ . We assumed all binding interactions were irreversible.

The control terms  $0 \leq v_j \leq 1$  depended upon the combination of factors which influenced rate process  $j$ . For each rate, we used a rule-based approach to select from competing control factors. If rate  $j$  was influenced by  $1, \dots, m$  factors, we modeled this relationship as  $v_j = \mathcal{I}_j(f_{1j}(\cdot), \dots, f_{mj}(\cdot))$  where  $0 \leq f_{ij}(\cdot) \leq 1$  denotes a regulatory transfer function quantifying the influence of factor  $i$  on rate  $j$ . The function  $\mathcal{I}_j(\cdot)$  is an integration rule which maps the output of regulatory transfer functions into a control variable. Each regulatory transfer function was modeled using a Hill function. In this study, we used  $\mathcal{I}_j \in \{min, max\}$  [286]. If a process has no modifying factors,  $v_j = 1$ . The model equations were implemented in Julia and solved using the CVODE routine of the Sundials package [28, 138]. The model code and parameter ensemble is freely available under an MIT software license and can be downloaded from the Varnerlab website [339].

### 3.4.2 Estimating complement model parameters.

We estimated a single initial parameter set using the Dynamic Optimization with Particle Swarms (DOPS) technique [284]. DOPS is a novel hybrid meta-

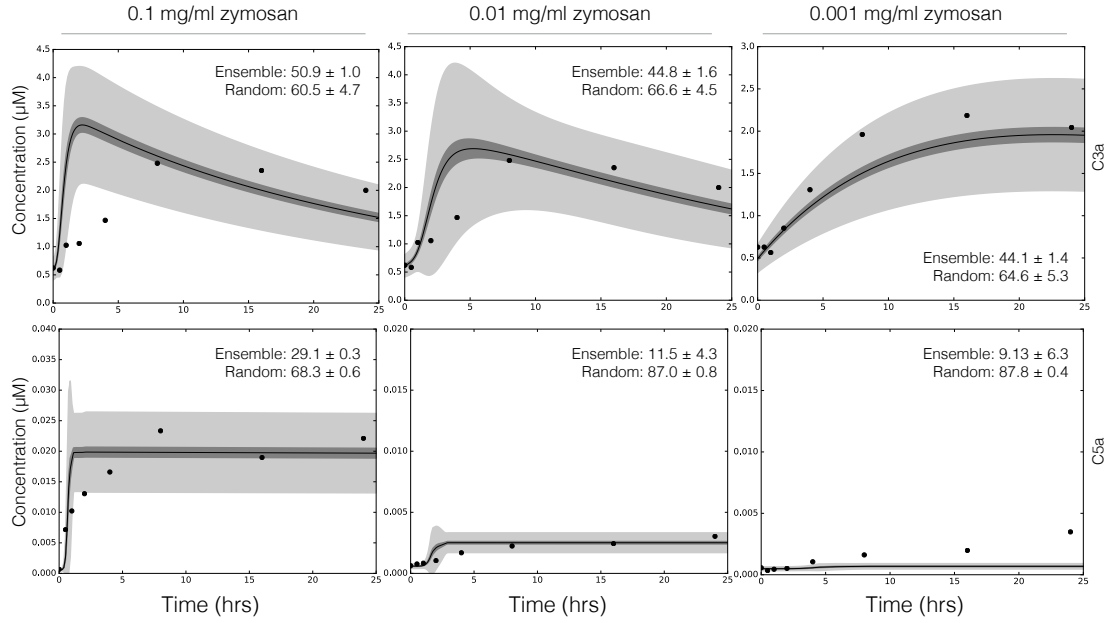


Figure 3.4: Reduced order complement model predictions. Simulations of C3a and C5a generated in the lectin pathway using 0.1 mg/ml, 0.01 mg/ml, and 0.001 mg/ml zymosan were compared with the corresponding experimental measurements. The solid black lines show the simulated mean value of C3a or C5a for the ensemble, while the dark shaded region denotes the 99% confidence interval of mean. The light shaded region denotes the 99% confidence interval of the simulated C3a and C5a concentration. All initial conditions were assumed to be at their physiological serum levels unless otherwise noted.

heuristic which combines a multi-swarm particle swarm method with the dynamically dimensioned search approach of Shoemaker and colleagues [333]. DOPS minimized the squared residual between simulated and C3a and C5a measurements with and without zymosan as a single objective. The best fit set estimated by DOPS served as the starting point for multiobjective ensemble generation using Pareto Optimal Ensemble Technique in the Julia programming language (JuPOETs) [20]. JuPOETs is a multiobjective approach which inte-

grates simulated annealing with Pareto optimality to estimate model ensembles on or near the optimal tradeoff surface between competing training objectives. JuPOETs minimized training objectives of the form:

$$O_j(\mathbf{k}) = \sum_{i=1}^{\mathcal{T}_j} \left( \hat{\mathcal{M}}_{ij} - \hat{y}_{ij}(\mathbf{k}) \right)^2 + \left( \frac{\mathcal{M}'_{ij} - \max y_{ij}}{\mathcal{M}'_{ij}} \right)^2 \quad (3.4)$$

subject to the model equations, initial conditions and parameter bounds  $\mathcal{L} \leq \mathbf{k} \leq \mathcal{U}$ . The first term in the objective function measured the shape difference between the simulations and measurements. The symbol  $\hat{\mathcal{M}}_{ij}$  denotes a scaled experimental observation (from training set  $j$ ) while the symbol  $\hat{y}_{ij}$  denotes the scaled simulation output (from training set  $j$ ). The quantity  $i$  denotes the sampled time-index and  $\mathcal{T}_j$  denotes the number of time points for experiment  $j$ . The scaled measurement is given by:

$$\hat{\mathcal{M}}_{ij} = \frac{\mathcal{M}_{ij} - \min_i \mathcal{M}_{ij}}{\max_i \mathcal{M}_{ij} - \min_i \mathcal{M}_{ij}} \quad (3.5)$$

Under this scaling, the lowest measured concentration become zero while the highest equaled one, where a similar scaling was defined for the simulation output. The second-term in the objective function quantified the absolute error in the estimated concentration scale, where the absolute measured concentration (denoted by  $\mathcal{M}'_{ij}$ ) was compared with the largest simulated value. In this study, we minimized two training objectives, the total C3a and C5a residual w/o zymosan ( $O_1$ ) and the total C3a and C5a residual for 1 mg/ml zymosan ( $O_2$ ). JuPOETs identified an ensemble of  $N = 2100$  parameter sets which were used for model simulations and uncertainty quantification subsequently. JuPOETs is open source, available under an MIT software license. The JuPOETs source code is freely available from the JuPOETs GitHub repository [340]. The objective functions used in this study are available in the GitHub model repository [339].

The simulation and prediction performance of the complement model was measured using the Akaike information criterion (AIC) [5]. In this study, we implemented the AIC as:

$$AIC = 2N_p + N_m \ln \left( \frac{1}{M} \sum_{\tau} (x_{\tau} - y_{\tau})^2 \right) \quad (3.6)$$

where  $N_p, N_m$  denotes the number of parameters, and the number of experimental measurements, respectively. The summation term in Eq. (3.6) denotes the residual between the model simulation ( $x$ ) and experimental measurements ( $y$ ), where the residual is normalized by the scale of the experimental data ( $M$ ). We compared the AIC for the model parameters estimated in this study, with a random parameter control generated to have a similar order of magnitude. The mean and standard deviation of the AIC was calculated over the parameter ensemble and the random parameter control were reported in this study.

### 3.4.3 Complement model analysis.

#### Global sensitivity analysis.

We conducted global sensitivity analysis to estimate which parameters and species controlled the performance of the reduced order model. We computed the total variance-based sensitivity index of each parameter relative to the training residual for the C3a/C5a alternate and C3a/C5a lectin objectives using the Sobol method [312]. The sampling bounds for each parameter were established from the minimum and maximum value for that parameter in the parameter ensemble. We used the sampling method of Saltelli *et al.* to compute a family of  $N(2d + 2)$  parameter sets which obeyed our parameter ranges, where  $N$  was the number of trials per parameters, and  $d$  was the number of parameters in

the model [287]. In our case,  $N = 400$  and  $d = 28$ , so the total sensitivity indices were computed using 23,200 model evaluations. The variance-based sensitivity analysis was conducted using the SALib module encoded in the Python programming language [135].

### Pairwise sensitivity analysis and clustering.

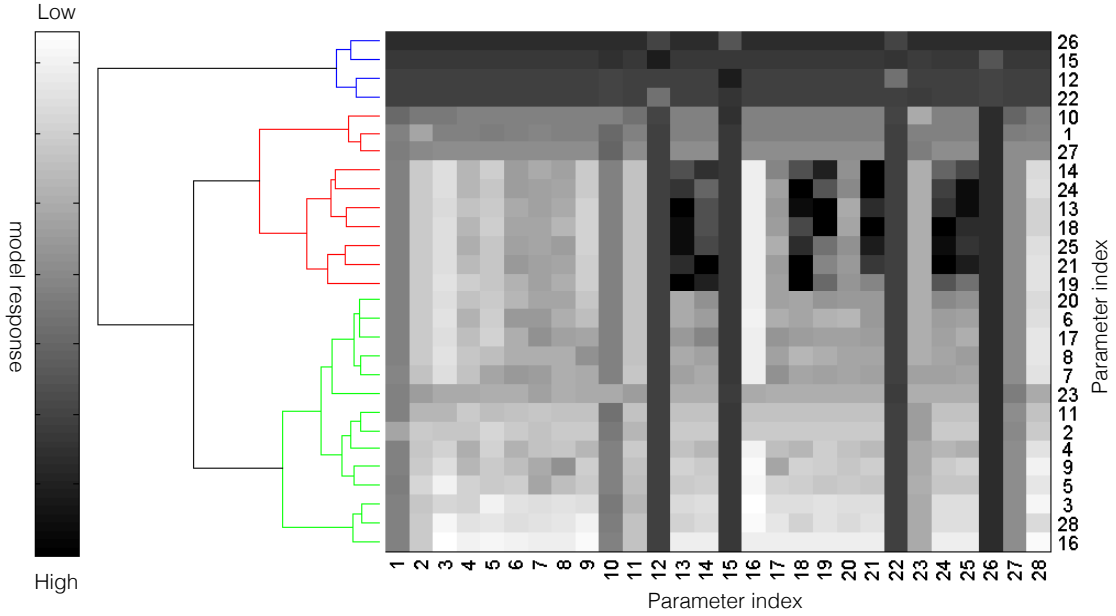


Figure 3.5: Pairwise sensitivity and clustering of complement model parameters in the presence of 1 mg/ml zymosan. The response of the complement model was calculated for each parameter combination following a 10% increase in parameter combinations in the presence of 1 mg/ml zymosan. The model parameters were clustered into high (blue), medium (red) and low (green) response clusters based upon the euclidian distance between the perturbed and nominal system state.

We perturbed each pair of model parameters by 10% of their nominal value, and then calculated the euclidian distance between the perturbed and nominal system states for physiological conditions. We repeated this calculation for each member of the parameter ensemble, and calculated the mean differences between the perturbed and nominal states. We then clustered the resulting  $\log_{10}$

transformed mean distances using the `Clustergram` routine in MATLAB (The Mathworks, Natick MA). We considered three clusters, high, medium and low displacement.

### Robustness analysis.

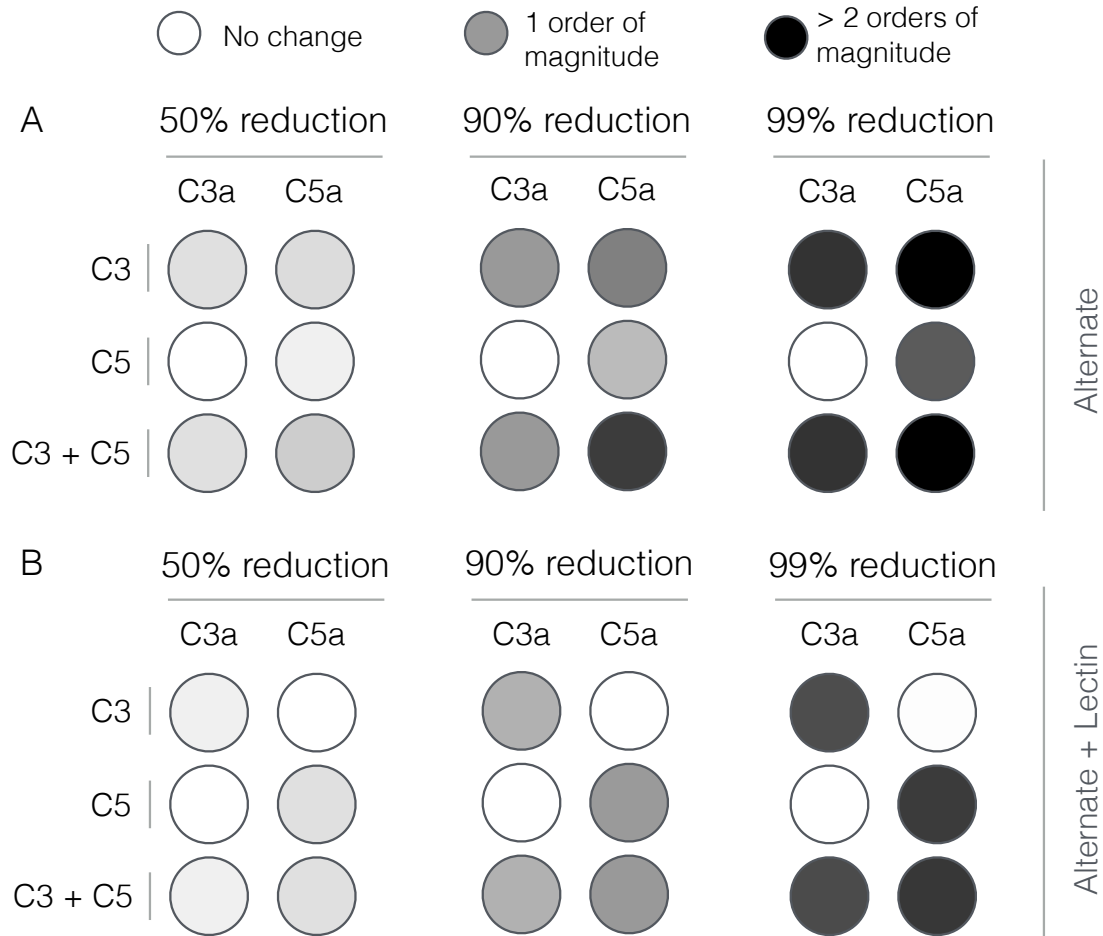


Figure 3.6: Robustness analysis of the complement model. Robustness coefficients were calculated for a 50%, 90% and 99% reduction in C3, C5, or C3 and C5 initial conditions. **A:** Mean robustness index for C3a and C5a generated from the alternate pathway (w/o zymosan). **B:** Mean robustness index for C3a and C5a generated from the lectin and alternate pathway (1 mg/ml zymosan). The color describes the degree of reduction of C3a or C5a following the network perturbation. Robustness coefficients were calculated using all parameter sets with Pareto rank less than five ( $N = 65$ ). Mean robustness values were reported.

Robustness coefficients quantify the response of a marker to a structural or operational perturbation to the network architecture. Robustness coefficients were calculated as shown previously [330]. Log-transformed robustness coefficients denoted by  $\hat{\alpha}(i, j, t_o, t_f)$  were defined as:

$$\hat{\alpha}(i, j, t_o, t_f) = \log_{10} \left[ \left( \int_{t_o}^{t_f} x_i(t) dt \right)^{-1} \left( \int_{t_o}^{t_f} x_i^{(j)}(t) dt \right) \right] \quad (3.7)$$

Here,  $t_o$  and  $t_f$  denote the initial and final simulation time, while  $i$  and  $j$  denote the indices for the marker and the perturbation, respectively. A value of  $\hat{\alpha}(i, j, t_o, t_f) > 0$ , indicates increased marker abundance, while  $\hat{\alpha}(i, j, t_o, t_f) < 0$  indicates decreased marker abundance following perturbation  $j$ . If  $\hat{\alpha}(i, j, t_o, t_f) \sim 0$ , perturbation  $j$  did not influence the abundance of marker  $i$ . In this study, we perturbed the initial condition of C3 or C5 or a combination of C3 and C5 by 50%, 90% and 99% and measured the area under the curve (AUC) of C3a or C5a with and without lectin initiator. We computed the robustness coefficients for a subset of the parameter ensemble ( $N = 65$ ) and reported the mean robustness value.

## **Funding**

This material is based upon work supported by, or in part by, the U. S. Army Research Laboratory and the U. S. Army Research Office under contract/grant number W911NF1020114.



## CHAPTER 4

### KINETIC MODELING OF COAGULATION AND FIBRINOLYSIS IN PLASMA

#### 4.1 Abstract

<sup>1</sup> Rotational thromboelastometry (ROTEM) provides an assessment of a patient's coagulation and fibrinolytic systems. ROTEM has become one of the most important and cost effective point-of-care techniques to rapidly test for hyper or hypo fibrinolysis. Given the prevalence of ROTEM, it is of tremendous clinical importance to develop mathematical models that integrate ROTEM measurements with the underlying biochemistry of coagulation. However, the development of such models has been slow, and plagued by challenges due to unavailability and complexity of fibrin degradation products coupled with an incomplete understanding of the degradation mechanism, missing patient data, patient variability and a lack of consensus about the mechanisms that contribute to hyperfibrinolysis or hypofibrinolysis. In this study, we developed an effective mathematical model which integrates ROTEM measurements with the biochemistry of coagulation and clot formation. The effective model consists of 24 ordinary differential equations combined with logical rules describing unmodeled regulatory mechanisms. The effective model describes the three phases of coagulation, the influence of key regulatory species, as well as thrombin mediated formation of cross-linked fibrin. Model parameters were estimated using thrombin generation curves and ROTEM measurements from platelet poor plasma isolated from blood taken from healthy individuals. After validating

---

<sup>1</sup>Under preparation as "Effective modeling of the human coagulation and fibrinolytic pathways" A. Sagar, R. LeCover, M. Bravo, T. Orfeo, K.E. Brummel-Ziedins, A. Pusateri, J. Varner

that our model captured the thrombin generation curves, we used them to predict ROTEM measurements for a separate set of patients. Our model captured the trends of ROTEM trajectories change shape with increasing concentrations of tPA, and had the lowest error at high concentrations of tPA. Variance based sensitivity analysis emphasized the importance of tPA in the area under the ROTEM curve, and through posing the inverse problem, we found that we could, on average, recover the protein concentrations within 30% of their true values.

## 4.2 Introduction

Trauma is the leading cause of death and disability for persons 36 years old and younger, surpassing all other causes combined [175]. In addition to its cost in lives, trauma has a large economic cost; it accounts for approximately \$671 billion per year, in health care costs and lost productivity [109]. Hemorrhage accounts for 40% of all trauma deaths, where the control of bleeding is especially challenging in the presence of blood coagulation disorders, collectively known as coagulopathy [291]. Trauma induced coagulopathy can follow from several mechanisms, for example, coagulation factor depletion, dysregulation of the protein C pathway, shedding of the glycocalyx, and hyperfibrinolysis [78]. However, adverse outcomes associated with coagulopathy are not limited to death from acute blood loss. Too much clotting can also be associated with a poor prognosis. Organ dysfunction, multiple organ failure and increased susceptibility to sepsis [102] are all potential consequences of prolonged shock resulting from coagulopathy [290]. For example, fatal multiple organ failure in severely injured trauma patients is often characterized by a hypofibrinolytic

state (fibrinolytic shutdown), in which clot lysis is impaired [208]. In trauma patients, rapid assessment of the coagulation and fibrinolytic systems could allow for early detection of coagulopathy.

Rotational thromboelastometry, the family of tests to which ROTEM belongs, has been shown capable of assessing coagulation abnormalities and rapidly predicting which patients would require massive transfusion, at a cost comparable to the cost of assessing one aspect of the coagulation system [169]. ROTEM has been used to guide treatment of trauma patients to determine the appropriate doses of fibrinogen concentrate and prothrombin complex concentrate, and this ROTEM guided treatment resulted in lower mortality than predicted by both the trauma injury severity score and the revised injury severity classification score [296]. Furthermore, when ROTEM is used in combination with a conventional clotting assay, such as the international normalized ratio, these assays in combination increase the detection of patients with coagulopathy as well as the sensitivity in predicting which patients will need massive transfusion [70]. Owing to the importance and prevalence ROTEM in hospitals around the world, we sought to build a mathematical model that could connect the underlying biochemistry to a ROTEM trajectory.

The immediate response of the body to injury is to activate the coagulation system, which generates a barrier (hemostatic plug) to arrest blood loss. Damage to the vasculature results in the exposure of circulating blood to tissue factor (TF) expressed by extravascular cells. FVIIa present in plasma in subnanomolar concentrations binds to TF and the resulting complex (extrinsic tenase) triggers a series of enzymatic events leading to thrombin formation [40]. Thrombin catalyzes the conversion of fibrinogen to fibrin, the activation of

FXIII to FXIIIa, and the activation of platelets. The resulting hemostatic plug is composed of activated platelets bound within and to an insoluble protein scaffolding constructed of self-assembling, polymerized fibrin molecules covalently crosslinked by FXIIIa [306]. The clot building response is downregulated by stoichiometric inhibitors present in plasma including antithrombin, which targets a number of coagulation proteases which propagate the reaction, and TF pathway inhibitor (TFPI) which inhibits the initiating complex (extrinsic tenase) [87]. A dynamic feedback pathway of inhibition is also engaged, through the action of thrombomodulin. Thrombin bound to thrombomodulin, a protein expressed on the surface of endothelial cells, activates protein C to activated protein C (APC). APC downregulates coagulation by inactivating the cofactors FVa and FVIIIa and their precursors FV and FVIII, thereby shutting down the two enzymatic complexes (prothrombinase and intrinsic tenase) involved in propagating the coagulant process [101].

Ultimately, repair of the injured vessel requires removal of the platelet-fibrin structure by fibrinolysis. Fibrinolysis involves the activation of the protease plasmin from its circulating precursor plasminogen by tissue type plasminogen activator (tPA), and once activated, plasmin catalyzes the proteolysis of the fibrin matrix. The conversion of plasminogen to plasmin at physiologically relevant rates proceeds through the assembly of a complex between tPA, plasminogen and fibrin, thus localizing plasmin formation to the injury site [116]. The trigger for the initiation of fibrinolysis is the formation of fibrin. As lytic processes initiate as the clot forms, stabilization of the forming clot requires a robust system capable of suppressing lysis until a sufficient degree of tissue repair and remodeling is achieved. This is accomplished by two mechanisms: directly via the action of the stoichiometric inhibitors PAI-1 (plasminogen acti-

vator inhibitor 1), the primary plasma inhibitor of tPA, and  $\alpha 2$ -antiplasmin, the primary plasma inhibitor of plasmin; and indirectly through activation of TAFI, a carboxypeptidase that modifies fibrin, thereby suppressing lysis by impairing the binding of plasminogen and tPA to fibrin [234]. Thrombin dependent activation of TAFI appears to be the physiologically relevant pathway and thus represents the route by which thrombin contributes to clot stabilization [240]. In TIC, depletion of PAI-1, overexpression of tPA and fibrinogen depletion have all been hypothesized to contribute to hyperfibrinolysis [37] while elevated levels of PAI-1 have been associated with fibrinolytic shutdown [99].

We and many others have worked to understand coagulation and clot formation, using both mechanistic models e.g., [143, 194, 192] and reduced order modeling approaches [286]. Longstaff and Thelwell proposed a simple two step model for fibrinolysis in which plasminogen is converted to plasmin by tPA, and fibrin is then degraded by plasmin [191]. While computationally easy to evaluate, this model neglected the role of key proteins such as PAI-1. On the other hand, fibrinolysis has also been modeled in great detail using a three-dimensional stochastic multiscale approach [15]. The complex model successfully predicted the difference in lysis speeds based on clot morphology, but neglected clot formation and degradation. Reifman *et al* modeled fibrin generation with mass action kinetics using 80 ODEs to investigate the efficacy of different prothrombin complex concentrates [218]. They used an interesting method to model the complex kinetics of TAFI inhibiting fibrinolysis: they empirically fit a curve to the clot lysis time as a function of TAFI concentration, and then correspondingly reduced the tPA effective concentration to account for the effects of TAFI. While this model decently predicted the final fibrin levels, it did a poor job of capturing the shape of the fibrin generation curves. Taken together, this il-

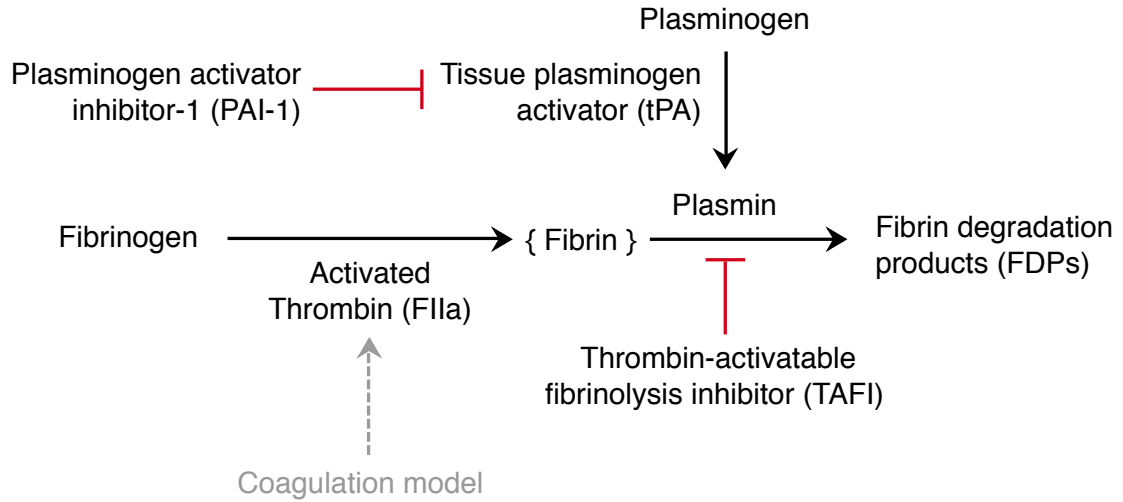


Figure 4.1: Schematic of the fibrinolysis model. Thrombin, generated by the presence of tissue factor in the coagulation model converts fibrinogen into its active form, fibrin. When tPA is present, it activates plasminogen from its inactive form to plasmin, which is capable of degrading plasmin. PAI-1 inhibits tPA and thus inhibits fibrin degradation.

illuminates the need for a combined model of coagulation and fibrinolysis which can accurately describe clot formation as well as the clot degradation process.

In this study, we formulated an effective model of coagulation and fibrinolysis. A set of individual platelet poor plasmas were supplemented with varying concentrations of tPA and subjected to a constant TF stimulus, yielding time course data reflecting the fibrin forming and lysing capacities of each plasma. We formulated a logical rules/ordinary differential equation model, trained the model on ROTEM data collected at various concentrations of tPA, and then successfully predicted previously unseen data, with model accuracy increasing as the concentration of tPA in the plasma increased. Through the use of leave one out cross validation, we confirmed that our parameter estimation process had not overfit the data. We used a variance based method to identify the most sensitive parameters, which were found to be related to tPA kinetics. We believe that this is the first study to connect a kinetic model with ROTEM curves, a key

development, as ROTEM is playing an increasingly large role both in surgery and in assessing the coagulation state of a patient. Our model could be especially relevant through solving the inverse problem, namely, estimating the concentrations of proteins in the coagulation system from a ROTEM trajectory.

## **4.3 Materials and methods**

### **4.3.1 Materials**

Full-length (residues 1-263) recombinant tissue factor (Tf) was purchased from Haematologic Technologies (Essex Junction VT, USA). 1,2-Dioleoyl-sn-Glycero-3-Phospho-L-Serine (PS) and 1,2-Dioleoyl-sn-Glycero-3-Phosphocholine (PC) were purchased from Avanti Polar Lipids, Inc. (Alabaster AL, USA). Preparations of phospholipid vesicles (PCPS) composed of 75% PC and 25% PS were made as described [137] as was the TF/PCPS reagent [51]. Corn trypsin inhibitor (CTI) was purified in-house [51] as was D-Phe-Pro-Arg-CH<sub>2</sub>Cl (FPR-ck) [361]. Tissue plasminogen activator (tPA) (two chain) was purchased from Molecular Innovations, Inc., MI, USA).  $\alpha$ 2-antiplasmin was purchased from Enzyme Research Labs Inc. (IN, USA) and plasmin from Haematologic Technologies Inc. (Essex Junction VT, USA). Aprotinin was purchased from Sekisui Diagnostics (Stamford CT, USA). Units of transfusion plasma were obtained from the University of Vermont Medical Center Blood Bank Plasma from each unit was distributed into aliquots and immediately frozen at -80°C until used.

### 4.3.2 Assays

Assays assessing the functional level of fibrinogen were performed by the University of Vermont Medical Center Clinical Laboratory. ELISA methods were used to assess protein antigen levels: plasminogen (PG-EIA, Affinity Biologicals ON, Canada); thrombin activatable fibrinolysis inhibitor (TAFI) (Zymutest (TAFI) Total Ag, Hyphen Biomedical, Neville sur Olse, FR); plasminogen activator inhibitor 1 (PAI-1) (R&D Systems Inc., MN, USA).  $\alpha$ -thrombin-antithrombin complex ( $\alpha$ -TAT) levels were determined via in house ELISA as described previously [359]. Plasmin-antiplasmin (PAP) complex levels were determined using the Immunoclone PAP ELISA (Sekisui Diagnostics CT, USA) with the following modification. A stock of 6  $\mu$ M PAP complex was generated in house by reacting 9  $\mu$ M  $\alpha$ 2 antiplasmin (Enzyme Research Labs Inc., IN, USA) with 6  $\mu$ M plasmin (Haematologic Technologies Inc., VT, USA). Quantitative formation of PAP was confirmed via SDS-PAGE analysis. This stock was used in place of the kit calibrator to generate a standard curve.

### 4.3.3 Viscoelastometry

Citrate plasmas were recalcified (15 mM  $\text{CaCl}_2$  final, volume change of 1.5%) for 3 min at 37°C. The tissue factor reagent (20  $\mu$ L, 5 pM final) and tPA (20  $\mu$ L, 2 nM, 4 nM or 8 nM final) or 20  $\mu$ L buffer (no tPA control) were prealiquoted as separate drops into 8 ROTEM cups. Each reaction was run in duplicate. 300  $\mu$ L of recalcified plasma was then added to each cup and data collection initiated. The contents of ROTEM cups were collected as described previously [359, 273].



#### 4.3.4 Formulation and solution of model equations.

The effective model equations described the time evolution of fibrinolysis and coagulation proteins ( $x_i$ ) using equations of the form:

$$\frac{dx_i}{dt} = \sum_{j=1}^{\mathcal{R}} \sigma_{ij} r_j(\mathbf{x}, \epsilon, \mathbf{k}) \quad i = 1, 2, \dots, \mathcal{M} \quad (4.1)$$

$$\mathcal{A}(\mathbf{x}) = \mathcal{A}_o + \mathcal{A}_1(\mathbf{x}) \left[ \frac{\mathcal{F}(\mathbf{x})^2}{K(\mathbf{x})^2 + \mathcal{F}(\mathbf{x})^2} \right] \quad (4.2)$$

The quantity  $\mathcal{R}$  denotes the number of reactions,  $\mathcal{M}$  denotes the number of proteins in the model,  $x_i$  denotes the concentration (nM) of species  $i$ , and  $\mathcal{A}$  denotes an output function which transforms the biochemical model species concentrations into a ROTEM signal. The quantity  $r_j(\mathbf{x}, \epsilon, \mathbf{k})$  denotes the rate of reaction  $j$ . Typically, reaction  $j$  is a non-linear function of biochemical ( $\mathbf{x}$ ) and enzyme ( $\epsilon$ ) species abundance, as well as unknown model parameters  $\mathbf{k}$  ( $\mathcal{K} \times 1$ ). The quantity  $\sigma_{ij}$  denotes the stoichiometric coefficient for species  $i$  in reaction  $j$ . If  $\sigma_{ij} > 0$ , species  $i$  is produced by reaction  $j$ . Conversely, if  $\sigma_{ij} < 0$ , species  $i$  is consumed by reaction  $j$ , while  $\sigma_{ij} = 0$  indicates species  $i$  is not connected with reaction  $j$ . Species balances were subject to the initial conditions  $\mathbf{x}(t_o) = \mathbf{x}_o$ .

Rate processes were written as the product of a kinetic term ( $\bar{r}_j$ ) and a control term ( $v_j$ ). Enzyme catalyzed rates were modeled using multiple saturation kinetics:

$$\bar{r}_j = k_j^{\text{cat}} \epsilon_i \prod_{s \in m_j^-} \left( \frac{x_s}{K_{js} + x_s} \right) \quad (4.3)$$

where  $k_j^{\text{cat}}$  denotes the catalytic rate constant for reaction  $j$ ,  $\epsilon_i$  denotes the abundance of the enzyme catalyzing reaction  $j$ ,  $K_{js}$  denotes the saturation constant for species  $s$  in reaction  $j$ , and where the product is carried out over  $s \in m_j^-$ , the set of *reactants* for reaction  $j$ . We used mass action kinetics to model protein-

protein binding interactions within the network:

$$\bar{r}_j = k_j^+ \prod_{s \in m_j^-} x_s^{-\sigma_{sj}} \quad (4.4)$$

where  $k_j^+$  denotes the on-rate constant for reaction  $j$ ,  $\sigma_{sj}$  denotes the stoichiometric coefficient for species  $s$  in reaction  $j$ , and  $s \in m_j^-$  denotes the set of *reactants* for reaction  $j$ . We assumed all binding interactions were irreversible.

The control terms  $0 \leq v_j \leq 1$  depended upon the combination of factors which influenced rate process  $j$ . For each rate, we used a rule-based approach to select from competing control factors. If rate  $j$  was influenced by  $1, \dots, m$  factors, we modeled this relationship as  $v_j = \mathcal{I}_j(f_{1j}(\cdot), \dots, f_{mj}(\cdot))$  where  $0 \leq f_{ij}(\cdot) \leq 1$  denotes a regulatory transfer function quantifying the influence of factor  $i$  on rate  $j$ . The function  $\mathcal{I}_j(\cdot)$  is an integration rule which maps the output of regulatory transfer functions into a control variable. In this study, we used  $\mathcal{I}_j \in \{\min, \max\}$  and Hill-like transfer functions [286]. If a rate process had no modifying factors,  $v_j = 1$ .

The output function  $\mathcal{A}(\mathbf{x})$  transforms the biochemical species abundance predicted by the model into a ROTEM signal. Modeling the connection between blood rheology and the local fibrin concentration is complex and beyond the scope of this study. Instead, we developed an empirical model which consisted of a scale function, which captured the amplitude of the ROTEM signal, and a shape function which captured the signal shape. The quantity  $\mathcal{A}_o$  denotes the baseline ROTEM measurement which we assumed was constant across plasma samples ( $\mathcal{A}_o \simeq 1.5$ ). On the other hand, the shape function:

$$0 \leq \frac{\mathcal{F}(\mathbf{x})^2}{K(\mathbf{x})^2 + \mathcal{F}(\mathbf{x})^2} \leq 1 \quad (4.5)$$

was modified by a scale function. The quantity  $\mathcal{F}(\mathbf{x})$  denotes the unweighted

sum of fibrin species in the model, while the saturation function  $K(\mathbf{x})$  was given by:

$$K(\text{tPA}) = 5100 - 615 \times (\text{tPA}) \quad (4.6)$$

where these values were estimated from the ROTEM curves. The scale function  $\mathcal{A}_1(\mathbf{x})$  was given by:

$$\mathcal{A}_1(\mathbf{x}) = \begin{cases} t \leq t_d & 0 \\ t > t_d & S \times [1 - \exp(-\tau(t - t_d))] \end{cases} \quad (4.7)$$

where

$$\tau = 0.0035 \times \left[ 1 - \frac{\text{FIIa}}{\max(\text{FII})} \right] \quad (4.8)$$

The material model delay parameter,  $t_d = \mathcal{N}(\mu_d, \sigma_d)$ , was identified from the training data. The scale parameter  $S$ , modeled as a linear function of tPA abundance, was defined as:

$$S(\text{tPA}) = \mathcal{N}(\mu_1, \sigma_1) - \mathcal{N}(\mu_2, \sigma_2) \times (\text{tPA}) \quad (4.9)$$

The intercept and slope distributions were modeled as Gaussian distributions, where the mean and standard deviation for each distribution  $(\mu_i, \sigma_i)$  were estimated directly from the training data. The model and output equations were implemented in the Julia programming language and solved using the `ODE23s` routine of the `ODE` package [28]. The model code and parameter ensemble is freely available under an MIT software license and can be downloaded from <http://www.varnerlab.org>.

### 4.3.5 Estimating model parameters

We estimated an ensemble of model parameters using Pareto Optimal Ensemble Technique in the Julia programming language (JuPOETs) [21]. JuPOETs is a

multiobjective approach which integrates simulated annealing with Pareto optimality to estimate model ensembles on or near the optimal tradeoff surface between competing training objectives. JuPOETs minimized training objectives of the form:

$$O_j(\mathbf{k}) = \sum_{i=1}^{\mathcal{T}_j} \left( \hat{\mathcal{M}}_{ij} - \hat{y}_{ij}(\mathbf{k}) \right)^2 + \left( \frac{\mathcal{M}'_{ij} - \max y_{ij}}{\mathcal{M}'_{ij}} \right)^2 \quad (4.10)$$

subject to the model equations, initial conditions and parameter bounds  $\mathcal{L} \leq \mathbf{k} \leq \mathcal{U}$ . The first term in the objective function measured the shape difference between the simulations and measurements. The symbol  $\hat{\mathcal{M}}_{ij}$  denotes a scaled experimental observation (from training set  $j$ ) while the symbol  $\hat{y}_{ij}$  denotes the scaled simulation output (from training set  $j$ ). The quantity  $i$  denotes the sampled time-index and  $\mathcal{T}_j$  denotes the number of time points for experiment  $j$ . The scaled measurement is given by:

$$\hat{\mathcal{M}}_{ij} = \frac{\mathcal{M}_{ij} - \min_i \mathcal{M}_{ij}}{\max_i \mathcal{M}_{ij} - \min_i \mathcal{M}_{ij}} \quad (4.11)$$

Under this scaling, the lowest measured concentration became zero while the highest equaled one, where a similar scaling was defined for the simulation output. The second term in the objective function quantified the absolute error in the estimated concentration scale, where the absolute measured concentration (denoted by  $\mathcal{M}'_{ij}$ ) was compared with the largest simulated value. In this study, we minimized three training objectives, the total error estimated for tPA = 0, 2, 4, 8 nM for three plasmas from our plasma data set. JuPOETs identified an ensemble of  $N \approx 2500$  parameter sets which were used for model simulations and uncertainty quantification subsequently. JuPOETs is open source, available under an MIT software license from the JuPOETs GitHub repository at <https://github.com/varnerlab/POETs.jl>. The objective functions used in this study are available in the GitHub model repository available from <http://varnerlab.org>.

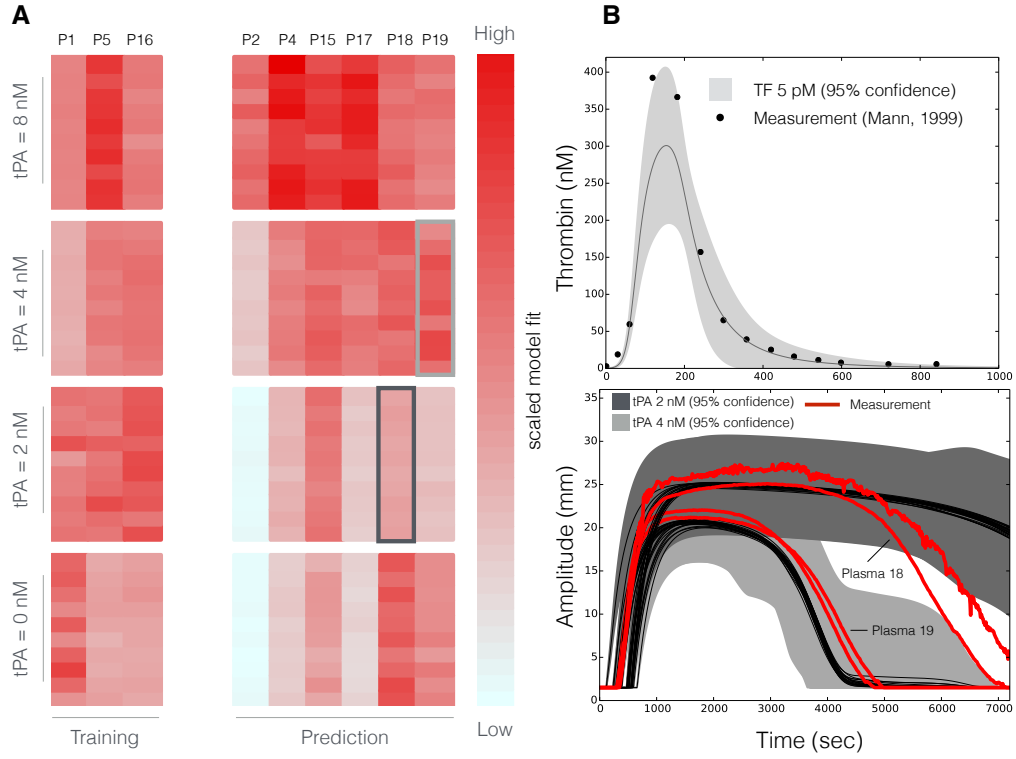


Figure 4.2: Fibrinolysis model training and prediction. **A** Model performance on test and training data. A redder block indicates that the simulations came closer to matching the experiential measurements. **B** Top: Performance of the thrombin generation portion of the model. Experimental measurements are shown in the black dots, the thin black curve represents the mean model performance and the shaded gray represents the 95% confidence interval for  $N=266$  simulations. Bottom: Performance of the fibrin generation and fibrinolysis portion of the model. The red curves represent ROTEM curves measured in two different patients at  $tPA = 2$  nM and  $tPA = 4$  nM, the black lines represent the simulated ROTEM curves for these patients in those cases, and the shaded dark gray represents the 99% confidence interval for the  $tPA = 2$  nM case, and the shaded light gray represents the 99% confidence interval for the  $tPA = 4$  nM case for  $N=266$  simulations.

To confirm that we were not over-fitting the experiment data, we performed leave one out cross validation. In this cross validation step, we minimized the objectives as defined in Eqn 4.10, over eight of the nine plasmas, and then calculated the mean squared error over the remaining plasma.

### 4.3.6 Global sensitivity analysis of model parameters

Global sensitivity analysis was used to estimate which parameters controlled the performance of the effective model. We computed the total variance-based sensitivity index of each parameter relative to the area under the ROTEM curve using the Sobol method [312]. The sampling bounds for each parameter were established from the minimum and maximum value for that parameter in the parameter ensemble. We used the sampling method of Saltelli *et al.* to compute a family of  $N(2d + 2)$  parameter sets which obeyed our parameter ranges, where  $N$  was the number of trials per parameters, and  $d$  was the number of parameters in the model [287]. In the case of fibrinolysis without platelets,  $N = 400$  and  $d = 28$ , so the total sensitivity indices were computed using 23,200 model evaluations. The variance-based sensitivity analysis was conducted using the SALib module encoded in the Python programming language [135].

### 4.3.7 Clustering of initial condition dependence

We sought to understand how ROTEM trajectories would change as concentrations of proteins associated with coagulation, fibrin generation, and fibrinolysis were altered. We perturbed the initial conditions to two times their nominal values in a pair wise fashion and calculated the distances between the nominal and perturbed ROTEM curves. We then clustered the  $\log_{10}$  transformed distances between the ROTEM trajectories using the `Clustergram` routine in Matlab (The Mathworks, Natick MA). We clustered the  $\log_{10}$  distances with a linear shift to make all of the values positive.

### 4.3.8 Relearning Initial Conditions

As ROTEM curves are often summarized by a few descriptive metrics, we decided to investigate if we could estimate the initial protein concentrations that generated the ROTEM curve from these metrics. To see if we could estimate the initial conditions, we first generated a simulated ROTEM curve generated by parameters within the range of a half to two times nominal initial conditions, and then calculated the CT (coagulation time, the time it takes for the curve to reach an amplitude of 2 mm, in seconds), CFT (clot formation time, the time it takes for the curve to reach an amplitude of 20 mm, after it has reached an amplitude of 2mm), the  $\alpha$ -Angle (the slope of the line between the CT and CFT on the ROTEM curve, in degrees), the MCF (the maximum clot firmness reached, in mm), ML (the amount of lysis that occurs by 120 minutes, described as a percent of the MCF), and area under the ROTEM curve-the target parameters. Our objective function for the minimization was defined as:

$$\sum_{j=1}^6 \beta_j \left( \frac{y_{t_j} - y_{e_j}}{w_j} \right)^2 \quad (4.12)$$

where  $w_j$  is the scale associated with metric  $j$ ,  $y_{t_j}$  is the target value for that metric and  $y_{e_j}$  is the current value of that metric based on the present initial condition estimate, and  $\beta_j$  is the weight applied to metric  $j$ . We weighted the area under the curve more heavily than the other metrics to better recover curves of the desired shape.

We then sought to relearn the initial conditions. To do so, we minimized the difference between the target parameters and the same metrics calculated for a new curve by altering the initial conditions. We used a particle swarm technique to perform this optimization to avoid getting stuck in local minima, as implemented in [221], with a swarm of  $N = 50$  particles for 500 iterations.

We approached this process in two ways: (i) we repetitively solved for initial conditions for the same ROTEM curve and; (ii) we generated a family of nominal ROTEM curves (with initial conditions between 50% and 150% of nominal values), and estimated the initial conditions for these curves one time for each curve. In the first approach, we were able to estimate initial conditions for one ROTEM curve, while under the second approach, we confirmed that this technique was generalizable to more than one specific ROTEM curve.

## 4.4 Results

In the model, activated thrombin (FIIa) converted fibrinogen into a series of fibrin intermediates which are then degraded by plasmin into fibrin degradation products (Fig. 4.1). Thrombin generation is modeled as described in [286]. Briefly, a trigger compound (FVIIa-TF) activates thrombin, which then can self-activate (thrombin converts prothrombin to thrombin). Antithrombin-III (ATIII), tissue factor pathway inhibitor (TFPI), and activated protein C (APC) oppose thrombin generation, and the effects of factors V, X, VIII, and IX are modeled via logical control. The generated thrombin then converts fibrinogen into fibrin monomers, which polymerize through a series of intermediates into fibrin. However, this model does not account for the full spectrum of possible fibrin polymers and monomers, rather, it assumes the presence of two pseudo-intermediates, one of which can be degraded by plasmin. While the model does contain equations that account for the action of uPA, the concentration of uPA in the system was assumed to be zero, so the kinetic parameters for these equations should not be considered to be accurate. Since this is a simplified model, it only considers one type of reaction between thrombin and fibrino-



gen, and accounts for the actions of FXIIIa through a control term. This model further simplifies the biology by assuming that the concentration of TAFI is constant, and through the use of only one splicing variant of fibrinogen, although multiple splice variants exist and interact differently with thrombin [82].

The model was trained using data from three patients and tested using measurements from six patients not used in the model training (Fig. 4.2). The model consists of two modules, a coagulation model which calculate the activated thrombin abundance, and a fibrinolysis module describing clot formation. The ROTEM data sets did not include thrombin trajectories, thus we used historical data from Mann and coworkers [42] to estimate coagulation model parameters. The coagulation model captured tissue factor initiated thrombin generation in the absence of platelets, although it under predicted the peak thrombin concentration (Fig. 4.2B). We then fixed the coagulation model parameters, and estimated the fibrinolysis model parameters by minimizing the squared difference between model simulated and experimentally measured ROTEM signals for patients P1, P5 and P16 as a function of tPA (Fig 4.2A, first column). The training performance of the model varied significantly with patient and tPA concentration. For example, the model was more consistent with P5 at higher tPA levels, but P2 was better captured at lower tPA levels. Similarly, the predicted ROTEM curves varied with patient and tPA concentration in a non obvious way (Fig. 4.2C). For example, the ROTEM signal for P2 was not captured by the simulation at  $\text{tPA} = 2 \text{ nM}$  or less; as the tPA concentration decreased, prediction performance declined for P2. However, other patients such as P18 exhibited a more complex pattern with relatively better performance at  $\text{tPA} = 0, 4$  and  $8 \text{ nM}$  and worse performance at  $\text{tPA} = 2 \text{ nM}$ . In general, the model performed best at high concentrations of tPA, and the performance declined as the amount of

tPA added decreased. For example, consider P18 and P19 at tPA = 2 nM and 4 nM, respectively (Fig. 4.2C). In the case of P18 tPA = 4 nM, the simulation captured the clotting initiation time, and the maximum amplitude of clotting but failed to describe the clot amplitude decay rate. On the other hand, all three features were captured in the P19 tPA = 4 nM case. Taken together, the combined coagulation and fibrinolysis model captured the clotting initiation time, maximum clot amplitude and the rate of clot decay for patient samples not used in model training. However, the training and prediction simulations were conducted with an arbitrary partitioning of the patient data, and the simulation performance varied significantly with patient and tPA abundance. Thus, it was unclear if the overall performance of the model could be improved with a different partitioning of the patient data. Toward this question, we used leave-one out cross validation to systematically vary the makeup of the training and validation data sets.

Leave-one out cross validation (LOOCV) confirmed that model captured the ROTEM curves, even when parameters were estimated by an alternative method (Fig. 4.3). In all cases, the parameter sets found by LOOCV capture the effects of adding increasing amounts of tPA. The simulated curves decrease in amplitude and the time at which they begin to turn over decreases as more tPA is added, and in the absence of tPA (blue), the simulations capture that no decrease in amplitude should occur. Many of the plasmas, in the absence of tPA, show an increase in clot amplitude throughout the experiment, (quickly in the beginning, and slowing down at the end), but our model predicts that after the clot reaches its final amplitude following the rapid growth phase, the clot amplitude will remain constant. With the exception of plasma 2, the parameter sets found through LOOCV perform better than those selected randomly (Table 4.1).

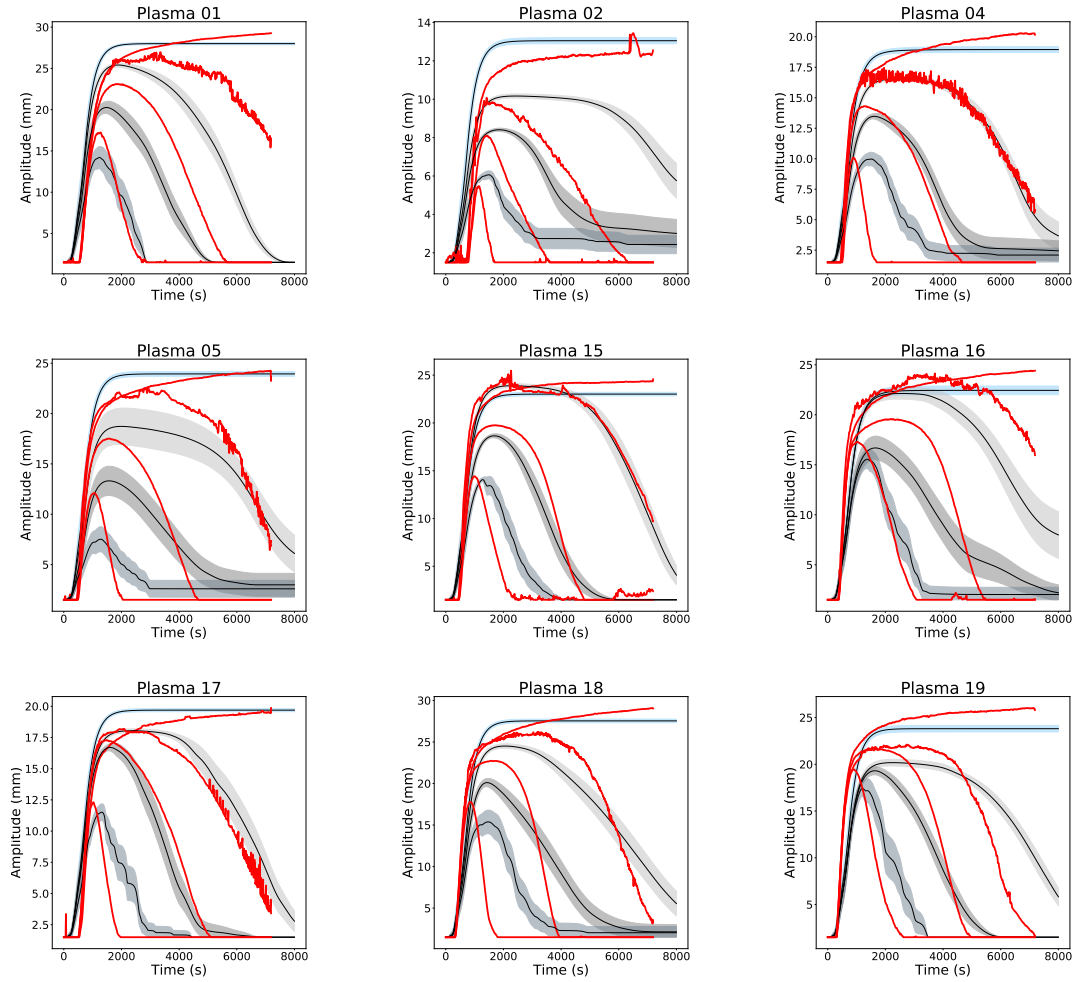


Figure 4.3: Model predictions on the left out plasma. The experimentally measured trajectory is shown in red, and the mean of  $N=64$  parameter sets found based on the other plasmas is shown in black, with the 95% confidence interval shaded. The  $tPA = 0$  case is shown in blue, the  $tPA = 2$  case is shown in light gray, the  $tPA = 4$  case is shown in dark gray, and the  $tPA = 8$  case is shown in slate gray.

However, plasma 2 had significantly lower amplitude than any other plasma at all values of tPA, so it is not unexpected that when it was left out of the training case, JuPOETs struggled to find suitable parameters, which resulted in the randomly selected parameter sets performing better. It is interesting to note that parameters estimated via LOOCV did a better job of capturing ROTEM curves at lower concentrations of tPA than the parameters estimated by our initial partitioning of the data. Taken together, LOOCV confirmed that the our estimated parameters described fibrin generation and degradation without overfitting the data. Given that we had developed a model and found parameters which captured ROTEM curves better (overall) than randomly generated parameters, we investigated which inputs altered the curves generated by the model.

Global sensitivity analysis confirmed the central role of tPA in fibrinolysis, both in diluted and undiluted plasma (Fig. 4.4). Global sensitivity coefficients were calculated for the area under the ROTEM curve in diluted and undiluted patient samples. The area under the ROTEM curve was largely controlled by tPA abundance, as well as the rate constant controlling how fast tPA can convert plasminogen into its active form. This result agreed with the large differences observed in curve shape and amplitude observed experimentally when differing amounts of tPA are added (Fig. 4.3). The contribution of the TAFI control parameters to the area under the ROTEM curve decreased when we simulated diluted plasma, as did the contribution of the initial amount of TAFI, perhaps because the decreased amount of prothrombin available in diluted plasma decreases the amount of thrombin present and available to activate TAFI. The role of tPA is additionally highlighted by the pairwise perturbations of the initial conditions (Fig 4.7), where it remained a prime contributor to the difference in the area under the curve regardless of which other initial condition was al-

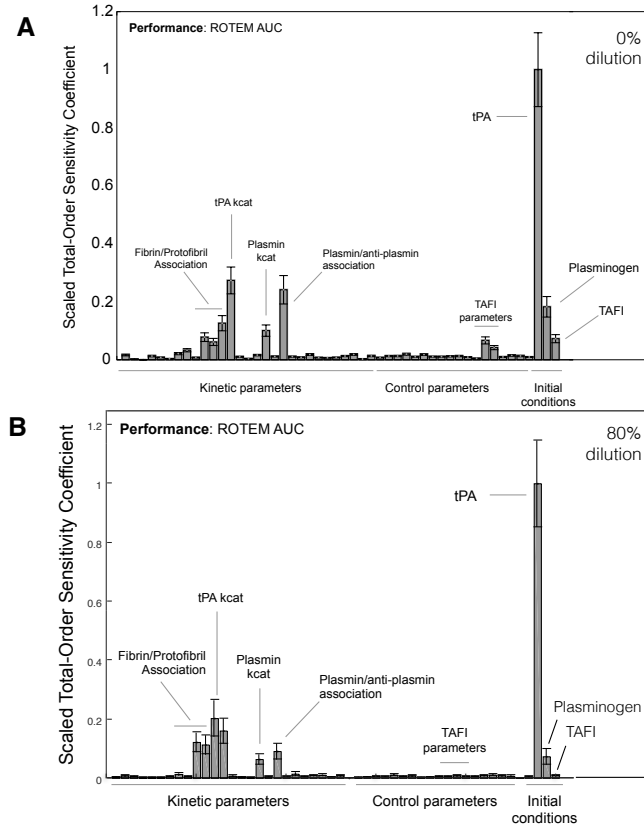


Figure 4.4: Global sensitivity analysis of the fibrinolysis model. Sensitivity analysis was performed by the model of Sobol with respect to the area under the ROTEM curve for  $N = 23,200$  parameter sets. 0% dilution means all factors were at physiological level, 80% dilution means that the factors were diluted to 80% of physiological level.

tered. The contribution of tPA is slightly diminished when the amount of fibrinogen is also perturbed. The initial conditions that alter the area under the ROTEM curve the most when pairwise perturbed (TAFI, PAI-1, fibrinogen, and tPA) are all directly related to fibrin generation, and not further upstream, related to thrombin activation. This is logical, as our metric for sensitivity is the shape of the ROTEM curve, which is directly related to fibrin generation and degradation. While this process requires the presence of thrombin, as long as enough thrombin is present to begin the fibrin generation process, a ROTEM curve of satisfactory amplitude will be generated. Taken together, these sensi-

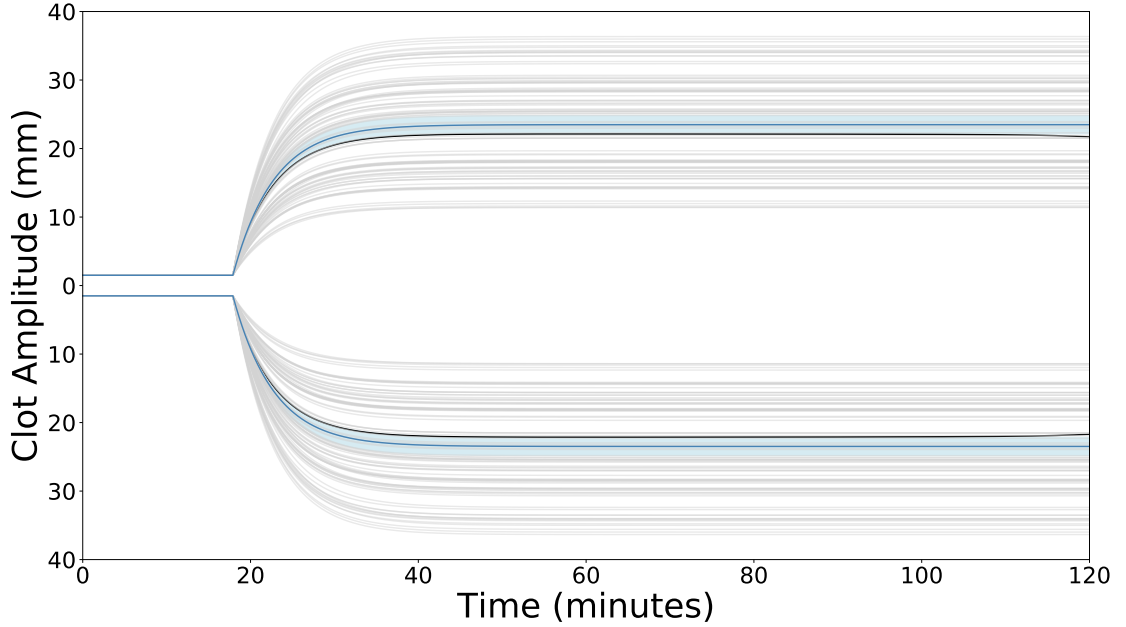


Figure 4.5: Simulated ROTEM curve for repeated solving of the same inverse problem. The black curve shows the true ROTEM curve, where the blue curve is the average over  $N=100$  attempts to find the initial conditions. The shaded blue area is the 95% confidence interval of the mean. The gray curves are the other ROTEM curves generated by the estimated initial conditions.

tivity results indicated likely targets for drugs that seek to alter the fibrin generation and degradation pathways. While this sensitivity analysis revealed which initial conditions altered the area under the ROTEM curve, we also sought to analyze the sensitivity of our model through an alternate approach, by attempting to see if we could estimate the initial conditions that created a specific curve, given the commonly reported ROTEM metrics for that curve.

Initial sample composition could be estimated from ROTEM measurements (Fig 4.5). To estimate the initial sample composition, we formulated an optimization problem which minimized the error between selected ROTEM metrics as a function of the initial protein concentrations. Through posing the inverse problem, we discovered that we could learn a possible set of initial conditions based on limited information about a ROTEM curve. However, since these ini-

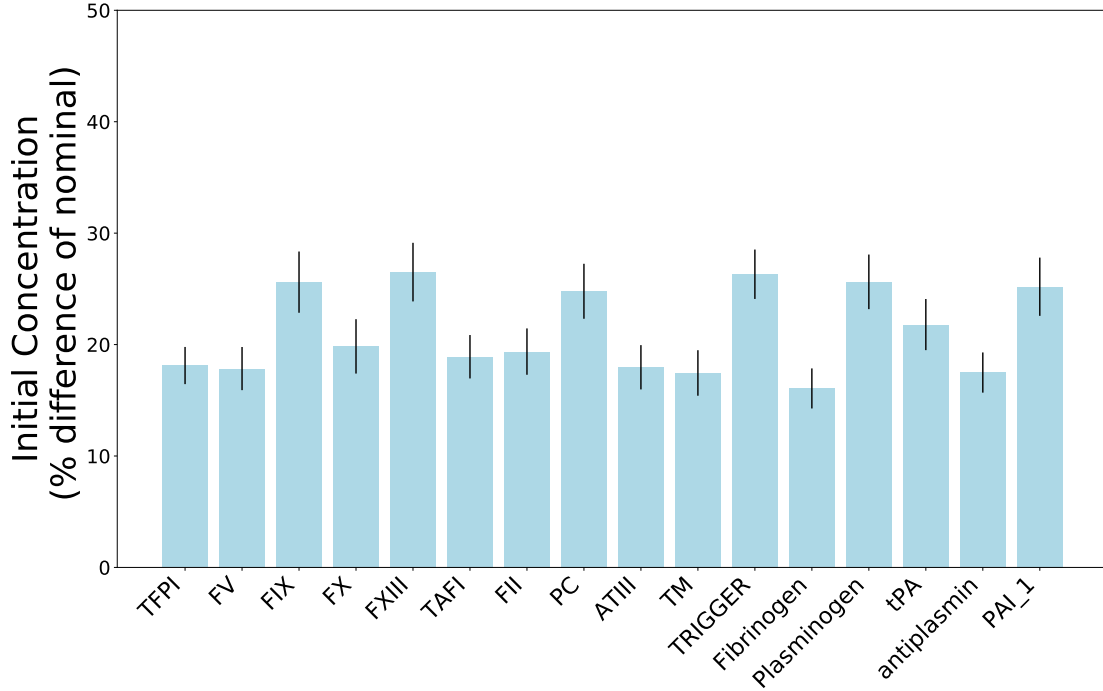


Figure 4.6: Average difference in target and found initial and experimental conditions from solving the inverse problem. The thin lines represent the standard error of the mean over  $N = 50$  repetitions of solving the inverse problem.

tial conditions are not unique, we repeated posing the inverse problem to learn which initial conditions could vary the most while still accurately describing the ROTEM curve based on these metrics. We found it was possible for the initial amount of fibrinogen to differ greatly-this is probably because as long as there was sufficient fibrinogen to reach the MCF, it does not matter if it was in excess. Additionally, the initial concentrations of thrombin and antithrombin were robust to perturbations. This is possibly because the amounts of thrombin and antithrombin can balance each other out. A high concentration of thrombin will be regulated by the presence of antithrombin, and vis-a-versa. Overall, we were fairly successful at recovering the initial conditions (Fig 4.6), coming within 30% or closer to their true values. When we focused on one specific ROTEM curve and repetitively solved for the initial conditions, we came close to recovering the

initial curve, when all of the found curves were averaged (Fig 4.5). All of the curves shown in this figure had errors on the same scale as those when we estimated initial conditions for a number of different ROTEM curves. Some of these curves underestimate the MCF while others tend to overestimate it, but when averaged, they come close to capturing the MCF of the true initial conditions. Taken together, these two methods reveal that further analysis of previously reported ROTEM metrics on trauma patients may hold further information about the state of the patient's coagulation and fibrinolytic systems.

## 4.5 Discussion

In this study, we constructed an effective model of fibrinolysis, which considered the dynamics of thrombin activation. We then created a function to transform the concentrations into ROTEM curves, and estimated kinetic and control parameters for the models using a Pareto Optimal Ensemble Technique. We then validated the model by measuring its performance on datasets that were not used in the parameter estimation. Viscoelastic tests are playing an increasingly important role in identifying the presence of aberrant fibrinolysis in trauma patients and in evaluating therapeutic responses. Empirically validated computational models linking the biochemistry of coagulation and fibrinolysis to physical changes in clot integrity should prove useful in probing the underlying mechanisms of aberrant viscoelastic profiles. In addition models that use individual-specific coagulation/fibrinolytic factor composition data may provide insight into preconditions favoring the development of trauma induced coagulopathy.



We then performed sensitivity analysis to observe which targets in the system could potentially be therapeutic targets. In the model of fibrinolysis without platelets both the total order sensitivity constants and the hierarchical clustering identify tPA's  $k_{cat}$  as sensitive, even under dilution conditions. Therefore, tPA could potentially serve as a therapeutic target. Additionally, the initial condition of tPA was very sensitive in both cases, suggesting that if a drug could render tPA inactive or increase its concentration, clot breakdown times could be significantly altered. The drug Alteplase (Cathflow Activase) does exactly that, given that it consists of recombinant version of tPA [213]. The CRASH-2 trial revealed that giving trauma patients tranexamic acid (TXA) within 3 hours of injury reduced the risk of death from haemorrhage [362]. TXA is believed to reversibly bind to plasminogen, and once bound, it blocks plasminogen and plasmin from binding to fibrin, and even if tPA converts plasminogen that is bound to TXA to plasmin, this plasmin will be unable to degrade fibrin [207]. While TXA is not directly considered in our model, the importance of tPA may partially explain why TXA is such an effective drug. Even though TXA does not bind with tPA, it mutes tPA's effects by deactivating tPA's target, plasminogen. Additionally, the fibrin control parameters appear to play a large role in determining the area under the ROTEM curve, as well as the rate constant for fibrinogen activation.

In the present model, the formation of a clot is represented by the formation of fiber, which is formed by protofibrils. The polymerization of protofibrils into oligomers, and the growth of the oligomers into a clot could potentially be included to further refine the model. Furthermore, the rate at which plasmin breaks down a clot is a function of the clot's structure [72]. Future work could include generating different types of fibers, with differing kinetic proper-

ties with respect to plasmin. Additionally, the role of platelets was completely neglected in the clot formation process, as this work focused on plasma, not whole blood. In the future, we could formulate a thrombin generation model which includes thrombin generation on activated platelets and then includes the platelet contribution to the generated clot. We investigated if the initial conditions that generated a particular ROTEM trajectory could be identified from commonly reported metrics, such as clotting time and alpha angle, and found that we were able to, on average relearn the initial conditions. In the future, we could investigate if measuring the concentration of one factor in the patient's blood in combination with a ROTEM curve could boost the accuracy of the estimation of the other factors, and which additional measurement would provide the most information. We could also investigate if another non-gradient dependant optimization method, such as simulated annealing performs better at estimating factor concentrations from ROTEM metrics.

## 4.6 Conclusion

In conclusion, we have combined our previous model of thrombin generation with an effective model of fibrin generation and fibrinolysis to describe ROTEM curves. We trained the model on three patient plasmas and then predicted ROTEM curves on six different patient plasmas, with prediction accuracies declining at higher tPA concentrations. Sensitivity analysis of the model elucidated that the area under the ROTEM curve is highly controlled by the concentration of tPA, while the shape of the ROTEM curve is highly controlled by a combination of thrombin generation and fibrin generation parameters. We found that based on six different ROTEM curve parameters, we could find a

set of initial conditions which described the curve, but that certain precursor species could vary greatly while still having the same target metrics.

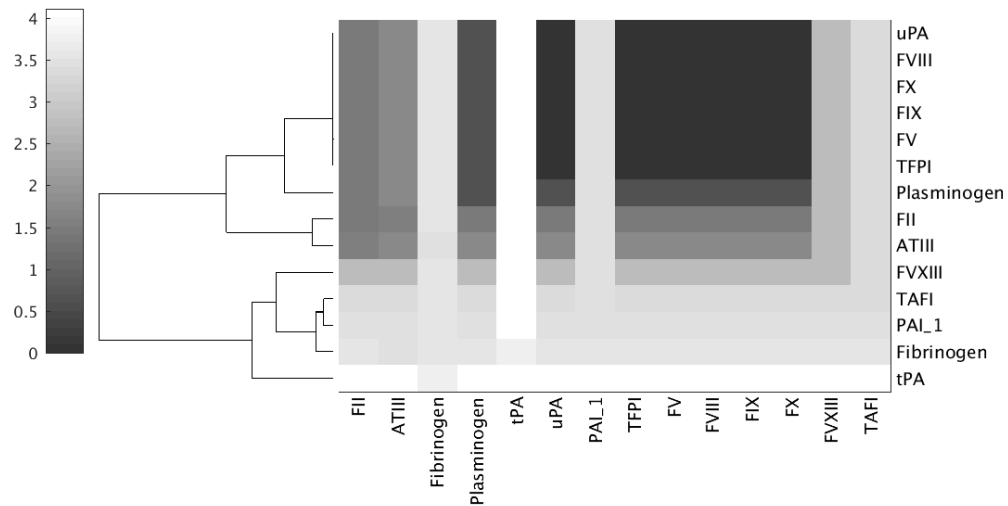


Figure 4.7: Clustergram analysis of fibrinolysis model response to changes in initial conditions. Initial conditions were increased to two times their nominal values in a pairwise fashion. For diagonal entries, the the factor level was perturbed to two times its nominal level.

Table 4.1

Plasma	Training	Prediction	Random
1	$0.0120 \pm .008$	$0.97 \pm .3$	$6.72 \pm .3$
2	$0.0091 \pm .006$	$2.22 \pm 1.4$	$1.62 \pm 1.4$
4	$0.0107 \pm .01$	$2.00 \pm 3$	$3.09 \pm 4$
5	$0.0104 \pm .007$	$2.41 \pm 6$	$4.50 \pm 6$
15	$0.0089 \pm .008$	$1.00 \pm 1.5$	$5.13 \pm 2$
16	$0.0113 \pm .009$	$1.77 \pm 5$	$5.39 \pm 6$
17	$0.0105 \pm .009$	$1.17 \pm 1.3$	$4.27 \pm 1.4$
18	$0.0126 \pm .01$	$5.04 \pm 9$	$10.21 \pm 9$
19	$0.0128 \pm .007$	$1.27 \pm 2$	$8.51 \pm 2$

Errors for the leave one out cross validation. The training column includes all parameter sets in which this plasma was used in the training set, the training column represents the results when this plasma was not trained on, and the random column represents model performance on  $N = 64$  parameter sets randomly selected from the parameters generated for the Morris sampling. Values are presented as the mean  $\pm$  one standard deviation.

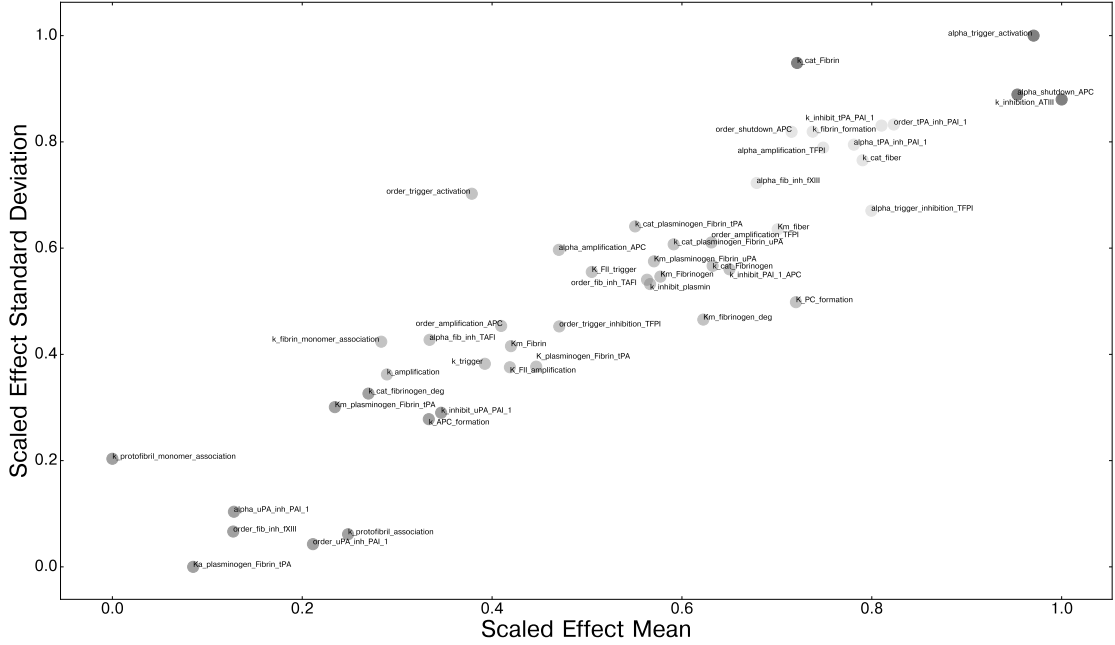


Figure 4.8: Morris analysis of fibrinolysis model response. We performed global sensitivity analysis to determine which parameters controlled the shape of the ROTEM curve output by the model. We generated parameters using the Morris method, with the parameters bounded by the minimum and maximum values they took in the rank one parameter sets generated by JuPOETS [230]. A total of  $N(p + 1)$  parameter sets were generated, with  $N = 1000$  and  $p = 47$  in our case, for a total of 48,000 model evaluations. We ran the model with the generated parameter sets, and then transformed the ROTEM curves using fPCA (functional principle component analysis), as implemented by the R package fdapace, to extract the principle component scores. This technique allowed us to remove the time dimension and to observe how the changes in parameters affected the shape of the ROTEM curve [322]. We then used the principle component scores to calculate  $\mu^*$ , the absolute value of the mean of the elementary effect, and  $\sigma$ , the variance of the elementary effect. We opted to use the absolute values of the means of the elementary effects so that effects with opposite signs would not cancel each other out and artificially decrease the estimate of the mean of the elementary effect [48]. We then scaled  $\mu^*$  and  $\sigma$  so that the smallest value became zero and the largest value became one, using the same scaling as used for the scaled experimental measurements. This method sensitivity analysis was conducted using the SALib module encoded in the Python programming language [135]. We then used k-means clustering to group the parameters into four clusters, and colored the graph by cluster color. We used four clusters as it resulted in the best average silhouette score [281].

CHAPTER 5

**KINETIC MODELING OF COAGULATION AND FIBRINOLYSIS IN  
PREGNANCY**

### **5.1 Pregnancy and Coagulation**

Pregnancy perturbs the human body in many ways. From the early stages of pregnancy, the cardiac output increases by 20% by 8 weeks of gestation, and profound vasodilation occurs, increasing renal plasma flow by more than 40%, and significant changes occur in the endocrine system, altering glucose metabolism [315]. Pregnancy also vastly alters the coagulation system, with concentrations of factors VII, VIII, X, XII, XIII and fibrinogen rising throughout pregnancy, and with the concentration of FXI falling [134]. Factor VIII levels appear to rise dramatically towards the end of pregnancy, and fall back to normal within eight weeks of delivery [133]. PAI-1 and PAI-2 (plasminogen activator inhibitor) levels increase throughout pregnancy, with the villous cells of the fetus producing PAI-2, so that the concentration of PAI-2 found in the mother's blood correlates with the birthweight [34]. Platelet count has been shown to decline slightly (around 10%) during the course of pregnancy, both in women with a normal pregnancy, and in women with pregnancy complicated by hypertension or pre-eclampsia [45]. Although platelet count decreases in pregnancy, platelet activation (as measured by platelet binding to wells that were precoated with fibrinogen) is higher in pregnant women, as well as higher levels of P-selectin, another marker of platelet activation, interestingly, the women with pregnancy induced hypertension had even higher levels of P-selectin, but their platelets had the same binding affinity as normotensive

pregnant women [167].

Through the use of thrombin generation assays (Figure 5.1a) and viscoelastic tests (Figure 5.1b), we can observe how the coagulation system evolves through the course of pregnancy. The peak amount of thrombin generated increases as the pregnancy becomes more advanced, both in plasma and plasma supplemented with thrombomodulin, in agreement with finding that overall, pregnancy is a hypercoagulable state [157]. This hypercoagulable state is also apparent in the ROTEM trajectory, where both in the presence and absence of tPA, the maximum amplitude of the clot formed increases with pregnancy duration. Additionally, the further along in the pregnancy, the faster the clot appears to form. Although it is well accepted that pregnancy results in dramatic changes to the coagulation system, to the best of our knowledge, no one has developed a mathematical model specifically focused on coagulation and fibrinolysis in pregnancy. One model appears to describe how thrombin generation differences between adult, maternal, and cord blood, however, this model was only validated against samples taken at the end of pregnancy [263].



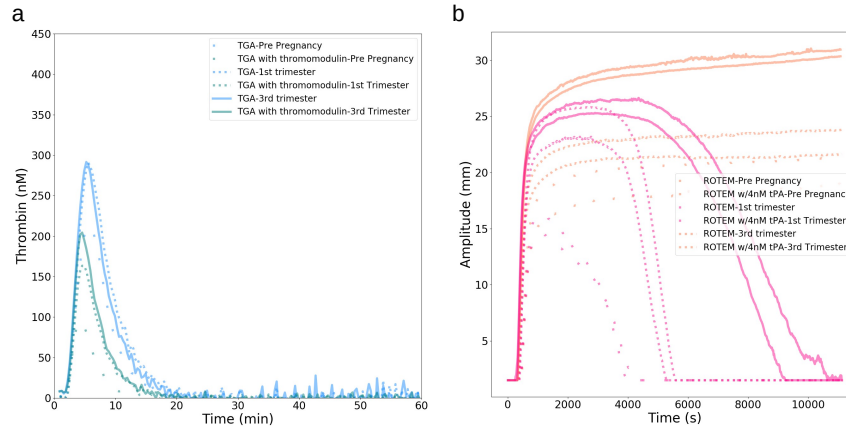


Figure 5.1: The evolution of of the state of the coagulation system throughout pregnancy in one patient. The more the solid the lines, the further along the woman is in her pregnancy. (a) Thrombin generation curves, in plasma, as initiated with 5 pM trigger, in plasma (blue) and in plasma supplemented with thrombomodulin (green) (b) ROTEM trajectory as initiated with 5 pM trigger, in plasma (orange) and in plasma spiked with 4 nM tPA (pink).

## 5.2 Pre-Eclampsia

For as long as women have been having babies and there have been writing systems extant to describe daily life, seizures and convulsions associated with pregnancy have been recorded, with one of the oldest records being from approximately 2200 BC in Egypt, suggesting that something be placed in a woman's mouth the day she gives birth to prevent her from biting her tongue [60]. Originally, the term eclampsia referred to convulsions that occurred but could be stopped once the precipitating event was removed, namely, these convulsions could be due to states other than pregnancy, but in the 19th century, a connection was made between these convulsions during birth and in the weeks following and premonitory symptoms, such as headache, temporary loss of vision, severe pain in the stomach, and edema of the hands, arms, neck, and face [26].

One of the major causes of material mortality is pre-eclampsia (estimated to contribute to 15-20% of material deaths in developed countries), in addition to causing perinatal deaths, preterm births, and intrauterine growth restrictions, it is a condition where the mother's vascular system responds abnormally to placentation, the growth of the placenta [307]. In between two to eight percent of pregnancies, pre-eclampsia develops, and women of certain ethnic groups, such as African-American and Filipino women, as well of women with low socioeconomic status, women with diabetes, chronic hypertension, and obesity appear to be at higher risk of pre-eclampsia [319]. Pre-eclampsia is diagnosed when a woman has diastolic blood pressure greater than or equal to 90 mmHg in addition to proteinuria (a concentration greater than 300 mg/L of protein in urine) on or after 20 weeks of pregnancy in a woman who was normotensive prior to pregnancy [214].

The root causes of pre-eclampsia remain an active area of investigation, but the disease appears to progress in two phases. The first phase relates to poor placentation, where the fetus doesn't properly connect to the mother's circulation system, and placental flow defects have been shown to be an early and powerful predictor of pre-eclampsia [254]. The second stage is apparent in the mother, whose endothelium becomes activated, via the intrinsic system, and changes in the surface receptors of her monocytes, granulocytes and lymphocytes mirroring that of those observed in sepsis [259]. Although the placenta plays a key role in our current understanding of pre-eclampsia, having an abnormally sized or perfused placenta is not sufficient or necessary for pre-eclampsia, rather, phase two can develop from a variety of different dysfunctions [275].

Women with pre-eclampsia have changes in their coagulation and fibri-

nolytic systems that exceed those observed in normal pregnancy. Women with severe pre-eclampsia (blood pressure exceeding 160/110 mmHg before delivery, proteinuria ranging from .5-10.0 g per day and generalized oedema) had much higher levels of fibrin degradation products in their serum as well as their urine as compared to women with mild pre-eclampsia and normal pregnancy, both preceding delivery and 1 day post delivery, interestingly, these severely pre-eclamptic women also had much higher levels of cyrofibrinogen but lower levels of plasminogen [150]. Antithrombin III (ATIII) levels in patients with pre-eclampsia have been shown to correlate with the severity of the disorder, with woman with more elevated blood pressure having lower levels of ATIII than women with mild pre-eclampsia, yet women with pre-eclampsia, regardless of severity had ATIII levels one standard deviation below the average compared to ATIII levels in normotensive pregnant women [355]. Levels of thrombomodulin, a glycoprotein which binds to thrombin and deactivates it, are elevated in women with pre-eclampsia as compared to normal pregnant women and healthy non-pregnant women [93]. Not only are levels of PAI-1 elevated in women with pre-eclampsia as compared to normal pregnant women, so are levels of TNF $\alpha$  (tumor necrosis factor) antigen, PAI-1, and tissue factor mRNA within portions of the placenta, leading the authors to propose that TNF $\alpha$  may induce production of PAI-1 and tissue factor [104]. Women with pre-eclampsia are usually hypercoagulable compared to normal pregnancy [94].

Thromboelastic measurements, such as ROTEM and TEG have begun to enter into obstetric practice, and have further documented the hypercoagulable state of pregnancy, however, the majority of studies in this area have focused on normal pregnancy, and have not yet been able to distinguish between healthy pregnant women and women with mild pre-eclampsia [153]. In one report,

women with pre-eclampsia have higher endogenous thrombin generation potential (area under a thrombin generation curve) than those with normal pregnancy, however, this does not appear to be a perfect separating criteria, as some of the women with pre-eclampsia had lower endogenous thrombin generation potentially than normotensive women [98].

### 5.3 Modeling Methodology

To further verify and refine our previously developed reduced order kinetic models of coagulation [285], we used the matched protein concentrations and thrombin generation trajectories in patients to re-estimate parameters and expand the model. As the previous iteration of the model did not consider the explicit inhibition of thrombin by activated protein C, we added this functionality to the model. The data we received consisted of thrombin generation trajectories in plasma, with and without the addition of thrombomodulin as well as thromoelastic assays, initiated by tissue factor, both without tissue plasminogen activator and with the addition of tissue plasminogen activator to the final concentration of 4nM. These dynamic assays were accompanied by a wide range of coagulation factor measurements, including prothrombin, factors V, XIII, IX,X,XIII, protein C, antithrombin-III, thrombomodulin, fibrinogen, antiplasmin, PAI-1, and TAFI. We constructed a reduced order model similar to that described in Chapter 4.6, however, we added an additional species to the model PAI-2, since PAI-2 is synthesized by the placenta during pregnancy [12]. We also added a switch like function to control the degradation of the formed clot, since we observed that PAI-1 concentrations correlated well with having an extended degradation time, as in some women, the time for the ROTEM to

complete was close to four hours, while in other women, it only took two hours for the test to complete because their clot degraded much faster. To capture this wide variation in clot degradation rates, we added a control term that reduced the rate that tPA broke down the clot if the initial concentration of PAI-1 in the sample exceeded a threshold. We also modified the control term for fibrinolysis to capture the large range in decline times. We formulated this control term as

$$C_{\text{fibrinolysis-new}} = \max(0, \tanh(t - t_{\text{delay-tPA}})) C_{\text{fibrinolysis}} \quad (5.1)$$

where  $t_{\text{delay,tPA}} = b + m * \text{Progesterone} + \epsilon$ , and  $b$  and  $m$  were estimated from linear regression, and  $\epsilon \sim N(0, \sigma^2)$ , where  $\sigma$  was also estimated from the same linear regression.

We estimated parameters for this model in a stepwise fashion. We divided the experimental data from  $N = 10$  patients into test, training and validation sets, with the training set consisting of the first and third timepoints from five patients, while the second timepoint was used as the test set, and the other five patients remaining as the validation set. We first fit the thrombin generation without additional thrombomodulin added as a multiobjective optimization problem using JuPOETs [20]. To ensure that our parameters accurately described the observed trajectories, we minimized an objective function which included the mean squared error between the experimental curves and the simulated curves and the absolute error between area under the experimental and simulated curves. Lastly, we added additional penalties if the simulated curves did not reach an amplitude of 50nM or exceeded an amplitude of 500nM or if the peak occurred after 9 minutes. After finding a family of 10 parameter sets that adequately described the thrombin generation process, we then bound

the thrombin generation parameters within the model above and below by the largest and smallest values found for these parameters found within this family of 10 parameter sets and posed another multiobjective optimization problem to estimate parameters for the fibrin related portion of the model. We minimized the mean squared error between the ROTEM trajectory with 4nM tPA, with penalties for no dynamics or for a ROTEM trajectory that didn't come down in the presence of tPA for the 5 patients in the training set and the 1st and 3rd visit, and then checked how our parameters predicted their ROTEM trajectory at their 2nd visit.

## 5.4 Results

Initial re-fitting simulations of the matched patient training data sets using the reduced-order model were consistent with the experimental measurements (Fig. 5.2). Our model successfully captured the differences in peak thrombin generation that were induced by the addition of thrombomodulin to the plasma. However, our model appears to overpredict the height of the thrombin peak more than it tends to underpredict, and in some cases, it predicts an earlier peak than is experimentally observed. It appears that our model does not quite succeed in capturing the amount of variance that is present in the thrombin generation curves with extra thromomodulin, with our model overpredicting the amount of thrombin generated. The predictions in the validation and test set look no worse than those in the training set, indicating that we have not overfit our model.

Our model succeeded in capturing the differences in shape in ROTEM

curves between the no tPA case and the 4 nM tPA case, as seen in Figure 5.3, with the curve not declining in amplitude in the absence of tPA. However, our method of drawing from a random distribution to determine the control term for tPA activation is only partially successful in capturing the wide variation in decline times observed, with the model predicting a faster decline than measured for patient 1 and 36. This model does succeed in capturing the much slower decline of patient 36 at the third timepoint. In general, our model, on average, predicts a higher clot amplitude for the  $\text{tPA} = 0$  case, but the true amplitude is captured in most cases within the 95% confidence interval. The model mostly succeeds in capturing the variation in maximum clot amplitude, but this could perhaps be improved by allowing different fibrin species to contribute varying amounts to the predicted clot amplitude.

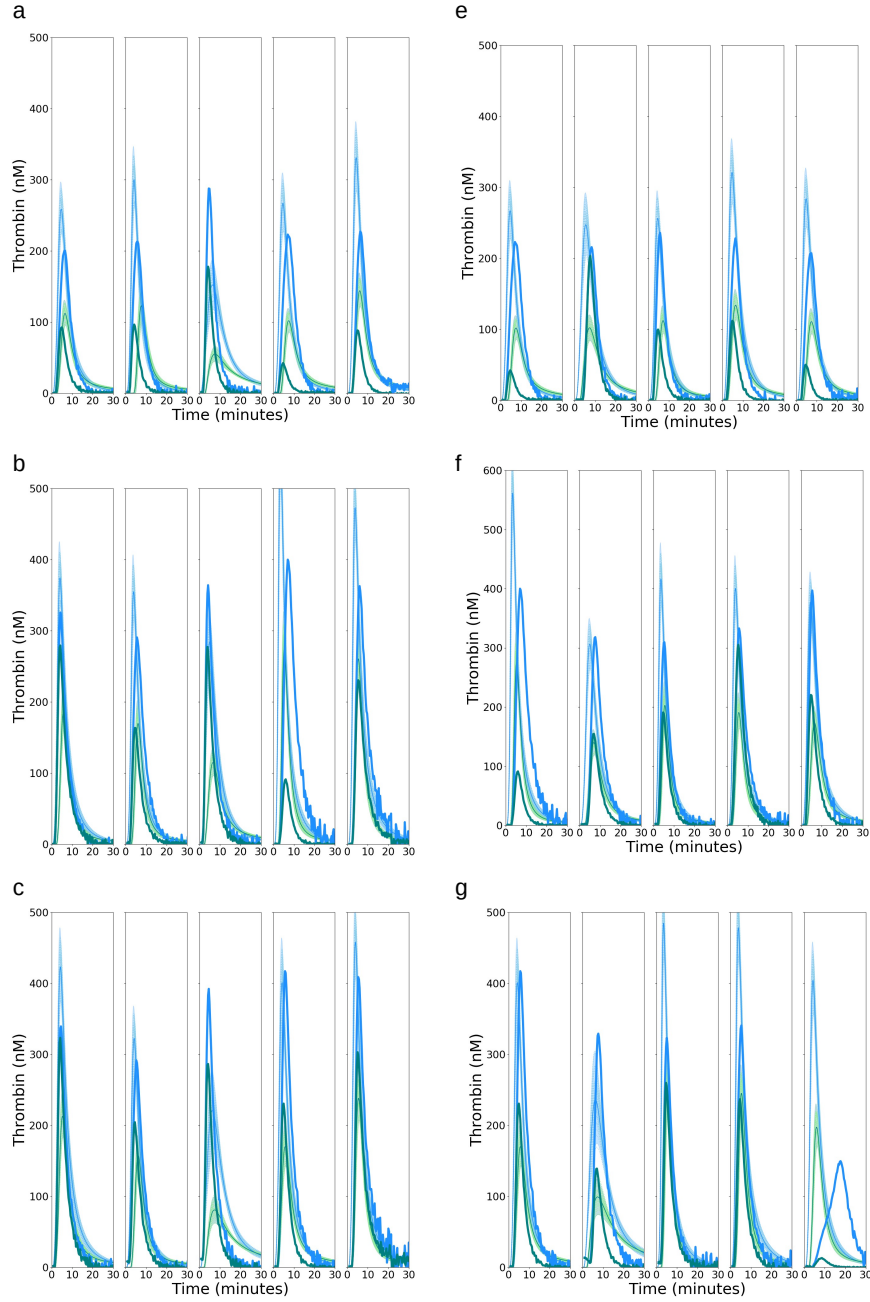
Model performance appears to be comparable on the five patients set aside as a validation set, as shown in Figure 5.4. The model does not completely capture the huge delay in decline time in patient 139, however, this is not surprising, as there were no patients within the training set that had not declined in amplitude in 180 minutes. In a few cases, the predicted clot amplitude for the  $\text{tPA} = 0$  case was higher in the model than experimental measured, but in the majority of the cases, the true amplitude fell within the 95% confidence interval of the predictions.

We used the method of Sobol as encoded in the `SALib` Python package to estimate the sensitivity of the area under the ROTEM curve of our model. As seen in Figure 5.5, the area under the curve appears to be controlled largely by the initial tPA concentration as well as the the initial fibrinogen concentration. This is a logical result, as the initial amount of tPA present contributes to the

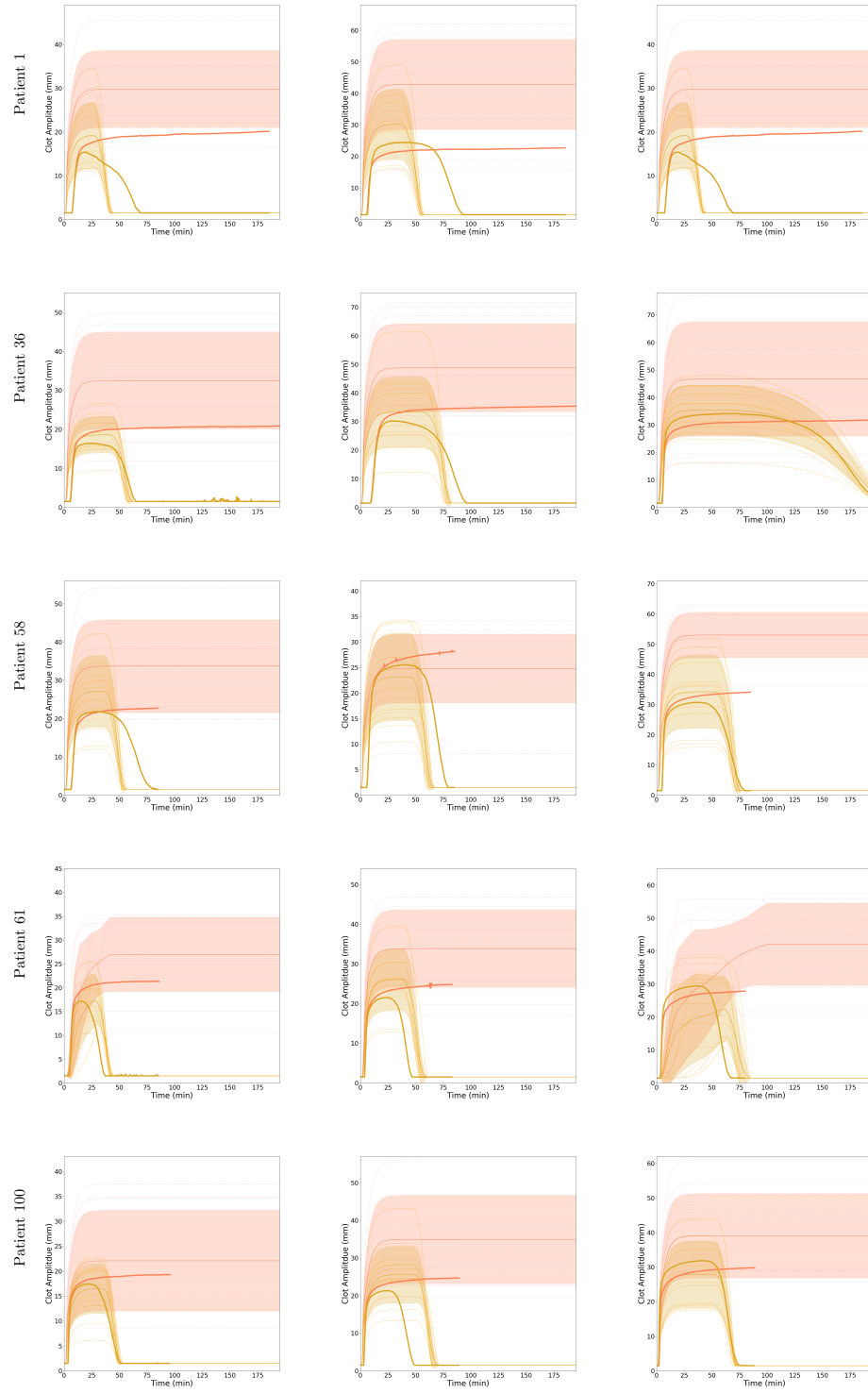
amount of plasmin that will be produced, which will effect how quickly the clot is broken down, which will alter the area under the ROTEM curve. It is also expected that the initial amount of fibrinogen present will play a large role in the area under the curve, as fibrinogen is the protein that is converted into an active form which then polymerizes to form the clot. Since we have formulated the control term which alters the rate that tPA activates to be a function of progesterone, the initial progesterone concentration also appears to play an important role in determining the area under the ROTEM curve.

When we applied the method of Sobol to the area under the ROTEM curve as a function of the parameters of the model, it revealed that one of the control parameters for TFPI plays a large role in determining the area under the curve, as seen in Figure 5.6. This control parameter plays a role in how fast trigger initiates the thrombin generation process, so in turn, this parameter continues to be important in determining the area under the ROTEM curve. In the current formulation, it appears that the TFPI control term dominates out the other factors that play into the the control term for trigger initiation. This analysis also highlighted the importance of  $K_m$  for the reaction in which plasminogen is converted to its active form. This is a parameter one would expect to play a large role in the shape of the ROTEM curve, as plasminogen is the protein responsible for the degradation of the formed clot, so the rate that this protein is produced would be expected to shape both the maximum amplitude of the ROTEM trajectory and the duration of the clot degradation process.

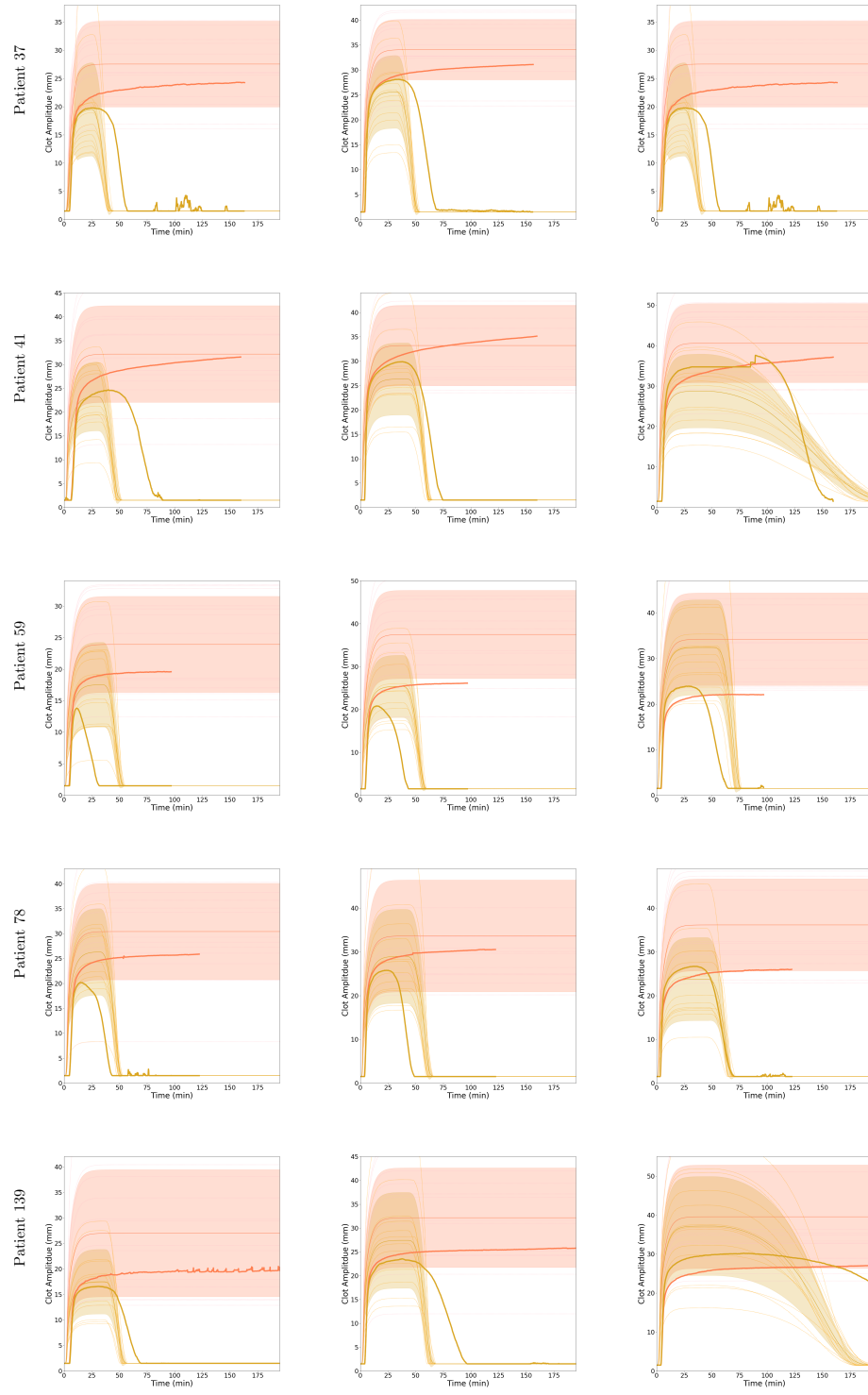




**Figure 5.2:** Thrombin generation predictions for the test, train, and validation sets. The experimentally measured thrombin trajectories are shown in the heavy lines, the average trajectory over 20 estimated parameters sets is shown in the thin line with a shaded 95% confidence interval. The dashed lines represent the trajectories from each parameter set. Thrombin generation trajectories without additional thrombomodulin are shown in blue, thrombin generation trajectories with supplemental thrombomodulin are shown in green. (a) and (c) represent the training data (5 patients, timepoints 1 and 3. (c) represents the test set, the same five patients used for training, at timepoint 2. (d) Validation of thrombin generation portion of the model at timepoint 1 in 5 different patients (e) Validation of thrombin generation portion of the model at timepoint 2 in 5 different patients (f) of thrombin generation portion of the model at timepoint 3 in 5 different patients.



**Figure 5.3:** ROTEM curves on the training set. The first and third time points were used to estimate parameters, and the second time point served as a test. The heavy curves are the experimental measurement, the thin lines are the average prediction over  $N = 13$  parameter sets, and the shaded area denotes the 95% confidence interval. ROTEM trajectories with no tPA are orange, ROTEM trajectories with 4 nM tPA are in yellow. The first column represents the first time point, the second column the second time point, the third column the third timepoint.



**Figure 5.4:** ROTEM curves on the validation set. The heavy curves are the experimental measurement, the thin lines are the average prediction over  $N = 13$  parameter sets, and the shaded area denotes the 95% confidence interval. ROTEM trajectories with no tPA are orange, ROTEM trajectories with 4 nM tPA are in yellow. The first column represents the first time point, the second column the second time point, the third column the third timepoint.

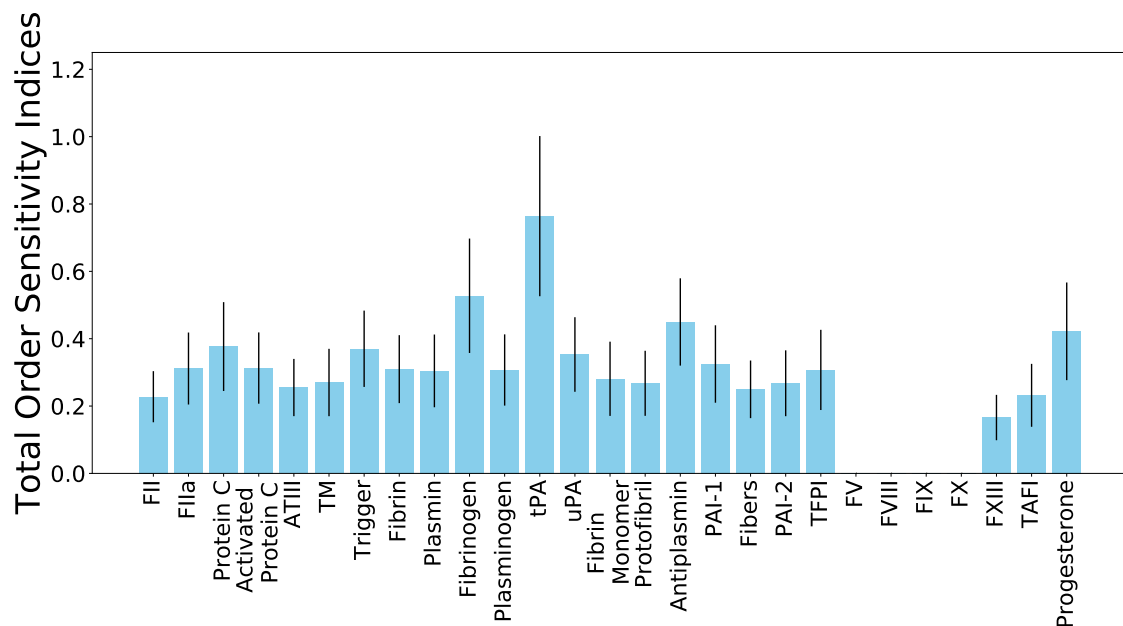


Figure 5.5: Preliminary sensitivity analysis using Sobol total order sensitivity coefficients using  $N = 100$ , and  $p = 27$  resulting in the evaluation of  $2N(p + 1) = 5600$  initial condition combinations evaluated. The blue bars represent the total order sensitivity coefficients, the error bars represent the confidence interval. Concentrations, when nominally non-zero were perturbed between 50% and 150% of nominal, with the exception of progesterone, which was allowed to vary between the minimum and maximum amounts measured. When the nominal concentration of a species was 0, it was perturbed between 0 and 1 nM.

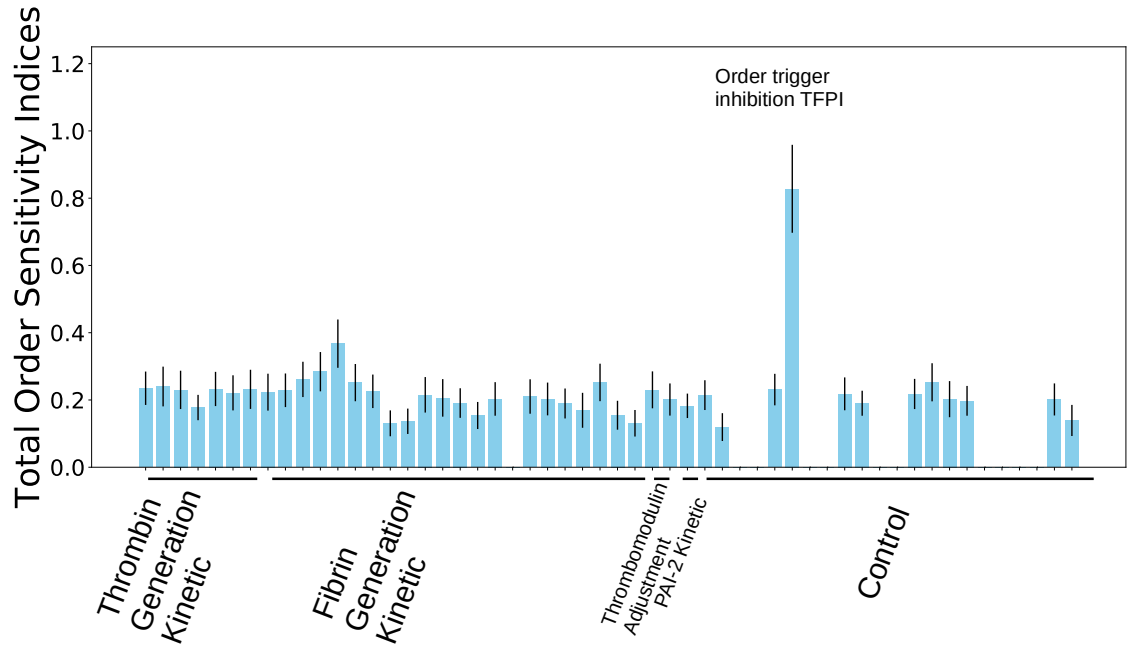


Figure 5.6: Preliminary sensitivity analysis using Sobol total order sensitivity coefficients using  $N = 500$ , and  $p = 54$  resulting in the evaluation of  $2N(p + 1) = 55000$  initial condition combinations evaluated. The blue bars represent the total order sensitivity coefficients, the error bars represent the confidence interval. Parameters were perturbed between 50% and 150% of the average value found in the best 13 parameter families.

### 5.4.1 Machine learning can classify pre-eclampsia

One interesting question is whether we could distinguish patients with a history of pre-eclampsia versus nulliparous women based upon biochemical and clinical data sets gathered during the initial medical visit. Toward this questions, we used patient data in combination with an unsupervised clustering approach to classify patient status (Fig. 5.7). The data provided to us for machine learning comprised of 30 rows, representing 30 patients and 61 columns, representing measurements of coagulation related proteins, two pregnancy related hormones (estradiol and progesterone), immune system related proteins, as well as metrics from a thrombin generation assay (with and without the addition of additional thromomodulin) and from a viscoelastic assay (with and without the addition of tPA) all taken at the first medical visit. The objective was to divide the data into two equal parts using unsupervised learning methods. The dataset consisted of two groups of women - nulliparous women (never having delivered a child) and women who were pre-eclamptic in a previous pregnancy resulting in birth. To get a sense of how non-linear vs linear the relationship among variables was, we investigated the co-variance of the data. This revealed that some variables were highly correlated. As to resolve the issue of highly correlated variables, we used principal component analysis (PCA) on the data. Upon examination of just two components of the PCA, roughly 60% of the variance in the data could be explained. In order to retain most information from the data we decided to keep five components which explained roughly 95% of the variance. After that, we used the consensus classifier of the k-means and a Gaussian Mixture Model. The clustered data can be separated using a linear decision boundary (Fig. 5.7). Surprisingly, we found that using the consensus classifier of k-means with the Gaussian Mixture Model correctly classified 22 out of the 30 patients (a

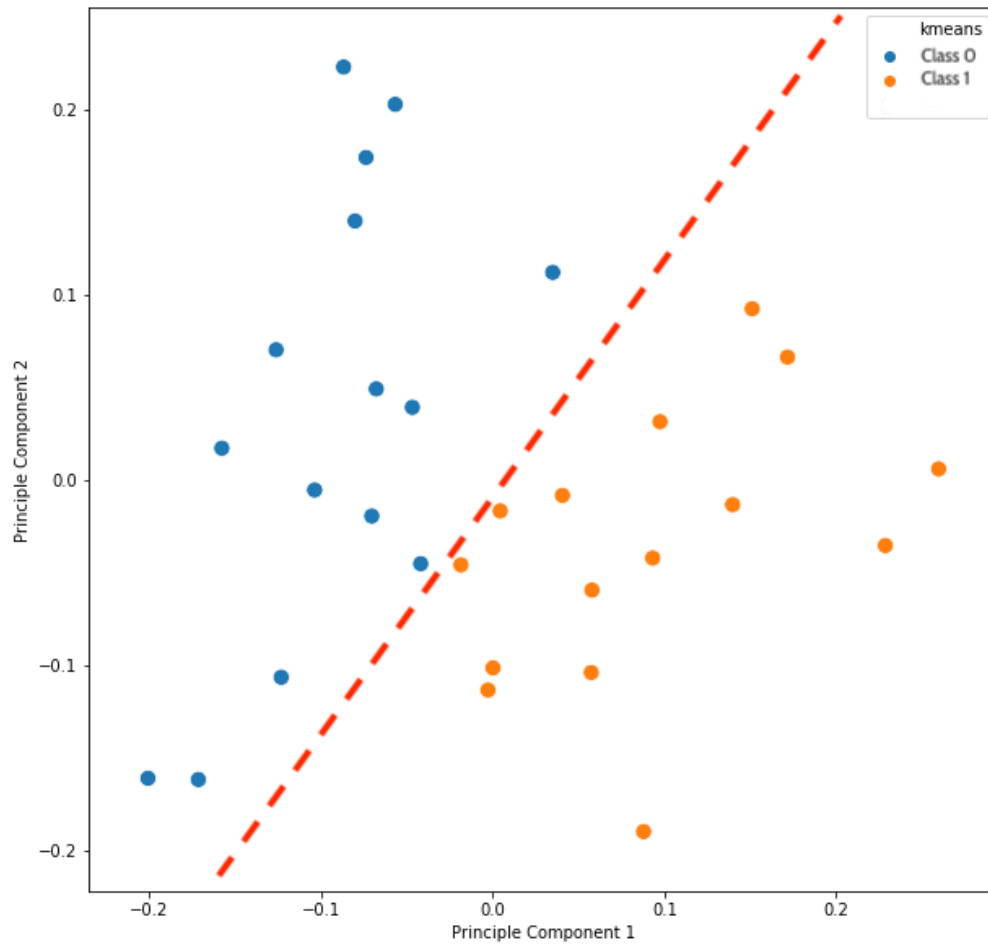


Figure 5.7: Predictions made by consensus clustering on PCA data

73% success rate). To check if this trend of highly linearly correlated variables sustains in other linearity decomposition methods we used Sparse Coding and Probabilistic Principle Component Analysis, however, neither of these methods equally divided the data among the two classes.

## 5.5 Discussion

Machine learning showed some promise in separating patients with pre-eclampsia from non-pre-eclamptic patients in a an unsupervised setting. How-

ever, when the data is reduced in dimensions via PCA, it can be seen that with the measurements presently available, there is no obvious separation along these dimensions. It is possible that we could improve the separation between the two classes of patients by constructing additional features from the time series measurements (the thrombin generation assays and the thromboelastic assays) other than the metrics that are normally reported (such as the area under the curve, or MCF). However, if this is not the case, this is also an interesting result, as it means that clinicians looking for a predictor of pre-eclampsia will need to consider other biomarkers and examine alternate theories of what leads to pre-eclampsia. It is also possible that we could see better separation between patients if we added additional covariates to the data, such as factors previously known to be associated with pre-eclampsia, and then we could examine if any of the immune and coagulation associated biomarkers improved accuracy and/or specificity of the separation.

It is interesting to note that by the third trimester, some women had developed clots that were very resistant to lysis, for example, Patient 139's ROTEM assay took more than 48 hours to run, as her clot took more than twenty hours to degrade. This stands in sharp contract to other patients, (Patients 100 and 78, for example), for which the entire assay took less than 14 hours. The time from peak amplitude to reaching an amplitude of 0 mm (duration of the decline time) does not appear to be strongly correlated with PAI1 or PAI2 levels (the absolute value of the Pearson correlation coefficient does not exceed .3 for patients in the third trimester), however, the duration of the decline time appears to be negatively correlated with FXIII levels (Pearson correlation coefficient -.51 for patients in the third trimester), which may indicate that in certain patients, these altered FXIII levels are changing how the clot is cross linked and making it more resis-



tant to fibrinolysis. One study reports that FXIII levels are statistically significantly lower during the second and third trimesters, however this difference is not large enough to rise to the level of statistical significance when compared to non pregnant women [301], so it remains unclear if changes in FXIII concentration are responsible for the large increase in the duration of the degradation time. In addition to a negative correlation with FXIII, this increase in duration of the decline of clot amplitude may also be associated with an increase in fibrinogen (Pearson correlation coefficient .69 in third trimester patients) as well  $\gamma'$  fibrinogen (Pearson correlation coefficient of .57 in third trimester patients). This prolongation of the duration of the decline time occurred in both in normal pregnancy and pre-eclampsic patients, suggesting the root cause of it is not the same as the factors that give rise to pre-eclampsia. It may prove informative to look at the fibrin structures of these clots to see if polymerization has proceeded differently between these two groups, and to see if the women with the clots that were more resistant to lysis have more tightly packed clots, or mechanically stronger clots that are less permeable increased amount of  $\gamma'$  fibrinogen present [105]. In the future, we may wish to further refine our model by including  $\gamma'$  fibrinogen as a distinct species, with a different rate constant for its degradation.

It is well known that the coagulation and complement systems are intimately linked, with proteins in the coagulation cascade in turn activating compounds in the complement system [7], and with certain components of the coagulation system deactivating portions of the complement system [19]. In the future, if we had measurements of complement related proteins, we could further refine this model to reflect this interplay. Pregnancy offers a unique opportunity to capture these interactions, because complement levels are known to change over

the course of pregnancy [13], and have been shown to be somewhat different between women with normal pregnancies and pre-eclamptic women [85]. We could further refine the control terms in our model to capture the interactions between complement and coagulation.

## 5.6 Conclusions

This work, to the best of our knowledge, presents one of the most detailed models of coagulation and fibrinolysis during pregnancy. It is strengthened by longitudinal samples, which allow us to track how the state of a woman's coagulation system evolves over the course of her pregnancy. Although our model did not perfectly capture the ROTEM trajectories for each woman and each point within her pregnancy, it did capture major trends. We discovered that it was difficult to separate patients with a normal pregnancy from those who would go on to develop pre-eclampsia in an un-supervised setting with the present set of thrombin generation, thromboelastic, and biochemical measurements, which suggests that if there is one master protein that controls the development of pre-eclampsia, it was not one measured in this study, and strengthens the argument that pre-eclampsia is a condition with a complex underlying etiology. We also discovered that some women, towards the end of their pregnancy begin to develop clots that are very resistant to lysis, however, this phenomenon does not appear to be linked to pre-eclampsia, as women within both groups developed this very lysis resistant clots. In the future, microscopy may allow us to confirm our hypothesis that coagulation proceeded differently in these patients resulting in tighter packing of the fibrin fibers which form the clot.

## CHAPTER 6

### KINETIC MODELING OF COAGULATION AND FIBRINOLYSIS IN WHOLE BLOOD

<sup>1</sup> Thromboelastic testing provides an assessment of a patient's coagulation and fibrinolytic systems. In recent years, thromboelastic testing has become an important point of care technique. However, its direct connection with the underlying biochemistry of coagulation and clot formation is not obvious. Toward this issue, we describe a validated reduced order mathematical model of coagulation and fibrinolysis, consisting of 22 ordinary differential equations, which described clot formation from initiation of the coagulation cascade through the degradation of polymerized fibrin by plasmin. We trained the model via leave one out cross validation on ROTEM measurements, a common thromboelastic test, on four patients, and then predicted ROTEM trajectories on four unseen patients, in whole blood and whole blood with the addition of 2 nM tissue plasminogen activator. Following model validation, sensitivity analysis suggested which biochemical interactions and species controlled the system response. Lastly, we investigated if we could estimate protein concentrations from commonly reported thromboelastic metrics. These estimation studies suggested we could (on average) relearn the initial fibrinogen concentration to within 20% of its true value. Taken together, this work presents a model which connects the underlying biochemistry of coagulation and clot formation in patients to a common point of care thrombelastographic test.

---

<sup>1</sup>This work has been previously published in [178].

## 6.1 Introduction

In the United States, traumatic injury is the leading cause of death for those under 45 years of age. Of those suffering traumatic injury, 50-60% die immediately after injury, while the remaining 40-50% of deaths occur following arrival at a hospital with potentially treatable injuries [313]. Of the patients who survive long enough to be transported to a hospital, about a third die from hemorrhage, a potentially treatable and survivable condition [335]. Several studies have shown that patients who present with acute traumatic coagulopathy (ATC), a dysfunctional coagulation system, upon admission, are less likely to survive [196, 38]. Acute traumatic coagulopathy is an open research topic: some researchers believe that activated protein C plays a key role in ATC by inhibiting coagulation [90], while others believe that the depletion of fibrinogen or excessive fibrinolysis may be the cause of ATC [209].

Thromboelastic testing is an important tool in assessing ATC, since it can be used to identify ATC within five minutes, before conventional laboratory results are available [77]. Rotational thromboelastometry (ROTEM) and Thromboelastography (TEG), both thrombelastographic tests, can quickly assess the viscoelastic properties of blood following traumatic injury [169]. ROTEM assesses thrombin generation, clot formation and fibrinolysis, and when used to guide transfusion, it reduces the amount of bleeding in patients [3]. Viscoelastic assays can detect hypocoagulability when other commonly performed tests such as activated partial thromboplastin time (APTT), prothrombin time (PT), and platelet count do not detect any abnormalities [161]. Thus, thromboelastic assays such as ROTEM are important tools for the detection of acute traumatic coagulopathy.

The coagulation system is a balance between the production of enzymes that build clots and those that break them down. Following traumatic injury, damage to the endothelium can result in the exposure of tissue factor, a transmembrane protein that is mainly constitutively expressed on cells outside the hemostatic envelope to circulating blood [106]. Tissue factor serves as the initiator to the coagulation process, which ultimately produces the enzyme thrombin, which in turn converts fibrinogen molecules to fibrin monomers. Fibrin monomers spontaneously associate, with the resulting insoluble polymer forming the protein scaffolding of a blood clot, a structure which includes activated platelets and red blood cells [198]. This clot is then broken down through a process called fibrinolysis, in which tissue plasminogen activator (tPA) and/or urokinase plasminogen activator (uPA) convert plasminogen into its active form, plasmin, which degrades the fibrin scaffold of the clot [53]. In greater detail, when tissue factor on the surface of extravascular cells is exposed to blood, it binds to FVIIa, an enzyme with poor catalytic potential present in blood at subnanomolar concentrations, forming the “trigger” complex (extrinsic tenase). The extrinsic tenase then efficiently converts FX and FIX into their active form, FXa and FIXa. FXa can then convert prothrombin (FII) to its active form, thrombin (FIIa). Thrombin is the key enzymatic product of the coagulation process, as it can activate the procofactors FVIII and FV to their active cofactor forms (FVIIIa, FVa), as well as FXI, and platelets. The activation of platelets provides phospholipid surface for the assembly of the two other primary enzymatic complexes of the coagulant response, the intrinsic tenase (FIXa-FVIIIa complex) which amplifies factor Xa production, and prothrombinase (FVa-FXa complex) which generates the vast majority of FIIa. The overall result is a positive feedback loop where a very small amount of initial thrombin can lead to a large

amount of thrombin generation. The thrombin generation process has a built-in negative feedback loop in vivo. Thrombin complexes with thrombomodulin, a transmembrane protein expressed on the luminal surface of endothelial cells and this complex activates protein C, which then proteolytically inactivates of FVa, FVIIIa, FV and FVIII [101]. This thrombin dependent pathway is the central dynamic anticoagulant mechanism limiting clot formation to the injury site. Negative regulation of thrombin levels is also controlled by Antithrombin, an abundant plasma protease inhibitor that forms irreversible 1:1 complexes with most of the proteinases participating in the coagulant response including thrombin, factor Xa, factor IXa, TF-FVIIa, factor XIa and the prothrombinase complex [87]. Of all the proteinase inhibitors in plasma it appears to be the primary inhibitor of thrombin. As noted, the thrombin produced by the coagulation process converts fibrinogen into fibrin which then polymerizes into a clot with the assistance of FXIIIa. The clot is later broken down by plasmin, which can be activated by either tPA or uPA, enzymes with short half-lives that can be inhibited by plasminogen activator inhibitor-1 (PAI-1) [54]. Thrombin can help stabilize the clot through its activation of thrombin activated fibrinolysis inhibitor (TAFI), which removes a C-terminal lysine and arginines from fibrin, reducing the possible number of plasmin binding sites.

Mathematical models of coagulation and clot formation can help us understand and identify key therapeutic targets implicated in ATC. Several kinetic models exist that describe the thrombin generation and degradation process. One of the best known models, the Hockin-Mann model, consists of 34 ordinary differential equations and 42 rate constants [143]. One of the most complex models of coagulation, containing 92 differential equations, was developed by Luan et al., [193]. However, if we wish to embed a thrombin generation model

inside a model that captures the whole-body physiology of traumatic injury, it must be small enough to solve quickly. Toward this objective, reduced order models have been developed which mostly captured thrombin generation dynamics [285]. Fibrinolysis and fibrin generation have also been modeled at a variety of scales. Longstaff and Thelwell proposed a very simple model for fibrinolysis, in which the process was represented as two steps, with plasminogen being converted to plasmin by tPA, and fibrin degraded by plasmin [191]. While computationally easy to evaluate, this model greatly oversimplified the process, and neglected the role of important regulators such as PAI-1. Fibrinolysis has been previously modeled in great detail through a 3D stochastic multi-scale model which predicted difference in lysis speeds based on clot morphology, but this model focused only on clot lysis rather than formation and degradation [15]. Reifman and Mitrophanov modeled fibrin generation with mass action kinetics using 80 ODEs to investigate the efficacy of different prothrombin complex concentrates [218]. They used an interesting method to model the complex kinetics of TAFI inhibition of fibrinolysis: they empirically fit a curve to the clot lysis time as a function of TAFI concentration, and then correspondingly reduced the effective tPA concentration to account for the effects of TAFI. While this model predicted the final fibrin levels, it failed to capture the shape of the fibrin generation curves.

In this work, we present a model of coagulation and fibrinolysis which uses an effective modeling approach to describe the complete coagulation process—from thrombin generation through fibrin degradation. The results generated by the model were compared to ROTEM data to test its validity. The model captured the differences in ROTEM trajectory in the absence and presence of tPA for four unseen patients. However, there were fine features of the response that

were not properly captured, potentially because of missing biochemistry in the model. Sensitivity analysis identified important biochemical components driving the system response, components that could be useful therapeutic targets. Lastly, the model could be used, along with thrombelastographic test results, to estimate some of the biochemical initial conditions in the sample. Taken together, this study presents one of the first models to connect the underlying biochemistry of coagulation and clot formation in patients to common point of care thrombelastographic test results.

## **6.2 Materials and Methods**

### **6.2.1 Viscoelastometry Materials and Methods**

Corn trypsin inhibitor (CTI), a factor XIIa inhibitor used to prevent contact pathway activation, was prepared as previously described [52]. Recombinant tissue factor (TF1-263) was purchased from Haematologic Technologies (Essex Junction, VT) and used to prepare a relipidated TF reagent as described [52]. Recombinant two chain tissue plasminogen activator (tPA) was purchased from Molecular Innovations (Novi, Michigan) with a reported specific activity of 2,516,682 IU/mg (relative to WHO International Standard for tPA, NIBSC 98/714).

Eight healthy individuals (4 males/4 females) were recruited for this study. All participants gave informed consent prior to blood collection. The participation of all individuals was approved by the University of Vermont Committees on Human Research. Blood was drawn by venipuncture (19 gauge needle). The first 3 ml were discarded and the next 10 mL collected and transferred to a 15



mL conical tube preloaded with 200  $\mu\text{L}$  of 5 mg/mL CTI (0.1 mg/mL final). CTI treated blood was divided into two 5 mL aliquots and placed in a 37°C water bath. Ten  $\mu\text{L}$  of a 1  $\mu\text{M}$  t-PA stock (final 2 nM tPA) was added to one 5 mL aliquot and 10  $\mu\text{L}$  of carrier buffer to the other. 320  $\mu\text{L}$  of CTI treated blood with or without tPA was transferred to ROTEM cups preloaded with 20  $\mu\text{L}$  of solution containing 85 pM TF reagent and data collection initiated. Reactions were run in duplicate. Complete blood cell counts were performed using the pocH-100i automated hematology analyzer (Sysmex, Mundelein, IL, USA).

## 6.2.2 Formulation and solution of model equations

The reduced order model consisted of 22 ordinary differential equations of the form:

$$\frac{1}{\tau_i} \frac{dx_i}{dt} = \sum_{j=1}^R \sigma_{ij} r_j(\mathbf{x}, \mathbf{k}) v_j(\mathbf{x}, \mathbf{k}) \quad (6.1)$$

The rate of change of species  $x_i$ , divided by its characteristic time scale,  $\tau_i$  is equal to the rate at which it is produced or consumed by the  $R$  reactions in the system. The quantity  $\sigma_{ij}$  denotes the stoichiometric coefficient for species  $i$  in reaction  $j$ . If  $\sigma_{ij} > 0$ , species  $i$  is produced by reaction  $j$ . Conversely, if  $\sigma_{ij} < 0$ , species  $i$  is consumed by reaction  $j$ , while  $\sigma_{ij} = 0$  indicates that species  $i$  does not participate in reaction  $j$ . The rate of reaction  $j$ ,  $r_j(\mathbf{x}, \mathbf{k}) v_j(\mathbf{x}, \mathbf{k})$ , is the product of the kinetics of the reaction ( $r_j$ ) and a logical control term ( $v_j$ ). The quantity  $\mathbf{k}$  denotes the (unknown) model parameters. Enzyme catalyzed rates were modeled using multiple saturation kinetics:

$$r_j = k_j^{\max} \epsilon_i \prod_{s \in m_j^-} \left( \frac{x_s}{K_{js} + x_s} \right) \quad (6.2)$$

where  $k_j^{max}$  denotes the rate constant for reaction  $j$  (1/hr),  $\epsilon_i$  denotes the abundance of the enzyme catalyzing reaction  $j$ ,  $K_{js}$  denotes the saturation constant for species  $s$  in reaction  $j$ , and  $s \in m_j^-$ , the set of *reactants* that participate in reaction  $j$ . Mass action kinetics were used to model the rate of protein-protein binding interactions:

$$r_j = k_j^{max} \prod_{s \in m_j^-} x_s^{-\sigma_{sj}} \quad (6.3)$$

where  $k_j^{max}$  denotes the maximum rate for reaction  $j$ ,  $\sigma_{sj}$  denotes the stoichiometric coefficient for species  $s$  in reaction  $j$ , and  $m_j$  denotes the set of *reactants* for reaction  $j$ .

The logical control terms  $0 \leq v_j \leq 1$  depended upon the combination of factors which influenced rate process  $j$ . For each rate, we used a rule-based approach to select from competing control factors. If rate  $j$  was influenced by  $1, \dots, l$  factors, we modeled this relationship as  $v_j = \mathcal{I}_j(f_{1j}(\cdot), \dots, f_{lj}(\cdot))$  where  $0 \leq f_{ij}(\cdot) \leq 1$  denotes a regulatory transfer function quantifying the influence of factor  $i$  on rate  $j$ . The function  $\mathcal{I}_j(\cdot)$  is an integration rule which maps the output of regulatory transfer functions into a control variable. In this study, we used  $\mathcal{I}_j \in \{min, max\}$  and Hill-like transfer functions. If a rate process had no modifying factors,  $v_j = 1$ .

Platelet activation was modeled using the method developed by Chatterjee et al, [55]. Reactions involving platelets were scaled by  $\eta \cdot \epsilon^{-1}$ , where  $\eta \leq 1$  is used to alter the rate constant for the reaction; we adjusted  $\eta$  to account for differing platelet counts. The term  $\epsilon$ , which is related to the fraction of activated platelets, is governed by:

$$\frac{d\epsilon}{dt} = k(\epsilon_{max} - \epsilon) \quad (6.4)$$

$$\epsilon_{max} = \epsilon_{max0} + (1 - \epsilon_{max0}) * f(\mathbf{S}) \quad (6.5)$$

where:

$$\aleph = \max_{t' \in [0, t]} \text{FIIa}(t') \quad (6.6)$$

and:

$$f(\aleph) = \frac{\aleph^\alpha}{\aleph^\alpha + \beta^\alpha} \quad (6.7)$$

Rate constants for processes involving platelets were then adjusted as  $k = k\eta \cdot \varepsilon^{-1}$ , and  $\alpha$  was estimated as a model parameter.

To simulate ROTEM dynamics we developed a function  $A(\dots)$  that mapped biochemical concentrations to the ROTEM signal (amplitude):

$$A = A_0 + A_1 \left( \frac{R^2}{K(\text{tPA})^2 + R^2} \right) \quad (6.8)$$

where  $A_0$  denotes the baseline ROTEM measurement,  $A_1$  denotes a maximum amplitude scaling factor (assumed constant across plasmas), and the term in parentheses denotes an empirical shape function. The shape saturation term  $K(\text{tPA})$  was given by:

$$K(\text{tPA}) = 1000 + 100 \times \text{tPA} \quad (6.9)$$

and  $R$  denotes the weighted average of the fibrin species contribution ( $\mathcal{F}$ ) and the platelet contribution ( $\mathcal{P}$ ) to the ROTEM amplitude; they were assumed to contribute equally [314, 173]:

$$R = 1/2 \times \mathcal{F} + 1/2 \times \mathcal{P} \quad (6.10)$$

The quantity  $\mathcal{F}$  is the sum of all fibrin species in the model, while  $\mathcal{P}$  is given by:

$$\mathcal{P} = \frac{\text{\#platelets}}{\text{\#Nominal platelets}} * c_{fibrin} * \varepsilon \quad (6.11)$$

where the patient's platelet count ( $\text{\#platelets}$ ; measured) was normalized by  $\text{\#Nominal platelets}$  (assumed to be 300,000 per microliter),  $c_{fibrin}$  denotes the fibrin concentration predicted by the model, and  $\varepsilon$  is governed by Eqn. (6.4).

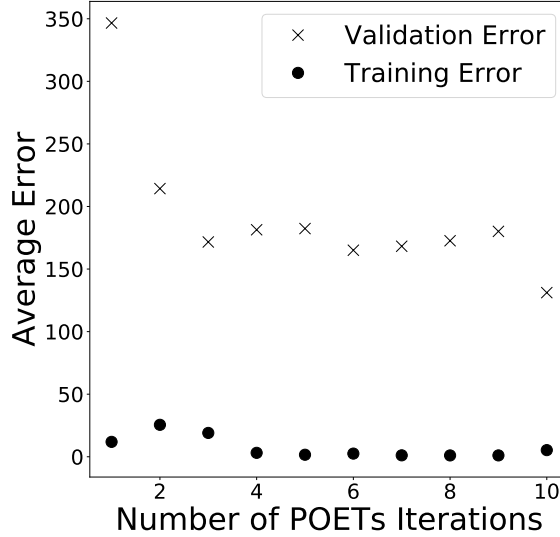


Figure 6.1: Calculated error of the model during leave one out cross validation. The dots show the average error on the 3 patients used in the training set, the crosses show the average performance on the left out fourth patient.

Model equations were encoded in Julia (v1.1), and solved using the `DifferentialEquations` package. Model equations, initial conditions and parameters are available at [https://github.com/varnerlab/Fibrinolysis\\_model\\_w\\_platelets](https://github.com/varnerlab/Fibrinolysis_model_w_platelets).

### 6.2.3 Estimation of model parameters

Model parameters for the thrombin generation portion of the model were estimated by minimizing the mean squared error between the thrombin generation predicted by the model and experimentally measured thrombin generation profiles in platelet rich plasma (data not shown). Next, we estimated the parameters for the fibrin generation and degradation based on ROTEM data using the Pareto Optimal Ensemble Technique (POETs), a multi-objective optimization technique [22]. Each objective minimized the mean squared error between

the observed ROTEM trajectories and the simulated trajectories for each patient for the  $tPA = 2$  nM case. We utilized leave one out cross validation. We divided the 8 patients into a training and a test group. We then trained the model on three patients in the training group (in the presence of tPA) and calculated the model prediction error on the left out fourth patient. The average error on the left out patient (the validation error) declined dramatically between the first and second iteration of POETs, and continued to slowly decline in the subsequent iterations, with the validation error remaining approximately one order of magnitude larger than the training error. The lowest validation error was found at 10 POETs iterations; thus, we used that parameter set to predict ROTEM curves on the four unseen patients. The final family of parameter sets consisted of  $N=24$  members; the best two parameter sets per objective were selected throughout all four cases of the leave one out cross validation.

## 6.2.4 Identifiability Analysis of Initial Conditions

We sought to estimate the maximum number of initial conditions we could identify from a ROTEM curve, as this information could be highly valuable in an emergency room setting to determine what course of treatment should be given. We numerically calculated sensitivities using central differences, so that  $s_{i,j} = \frac{dy}{d\theta_i} \approx \frac{y(t)_+ - y(t)_-}{2h}$ , where  $\theta_i$  represents initial condition  $i$ , and  $y(t)$  represents the simulated ROTEM curve at either the increase or decrease in the specified initial condition, and  $h = \theta_+ - \theta_-$ , the step size in parameter space. We collapsed the time dimension by calculating the root mean squared sensitivity,  $s_i^{msqr} = \sqrt{\frac{1}{N_D} \sum_{j=1}^{N_D} s_{i,j}^2}$ , where  $N_D$  represents the number of time points in our simulation. Parameters were noted as not sensitive if their magnitude was four

orders of magnitude smaller than the maximum root mean squared sensitivity [115]. We can then normalize the root mean squared sensitivities of the sensitive parameters  $\bar{s}_i = \frac{s_i^{msqr}}{\|s_i^{msqr}\|}$  for  $i = 1, \dots, N_{\theta_{sensitive}}$ .

From the the normalized sensitivities, we examined their collinearity, as suggested in [41]. In this approach, we know parameters are linearly independent if there exist  $k$  constants  $\alpha_i \neq 0$  such that:

$$\alpha_1 \bar{s}_{K1} + \alpha_2 \bar{s}_{K2} + \dots + \alpha_k \bar{s}_{Kk} = 0 \quad (6.12)$$

We can calculate the the degree of collinearity among the parameters via the collinearity index,  $CI_k = \frac{1}{\sqrt{\lambda_{K,min}}}$ , where  $\lambda_{K,min}$  is the smallest eigenvalue of  $\bar{S}_K^T \bar{S}_K$ , where  $\bar{S}_K$  is assembled placing each of the  $k$  normalized sensitivity vectors in the columns of  $\bar{S}_K$ . Per Brun [41], a subset of parameters are identifiable if their collinearity index is less than twenty.

We then used algorithm four from [115], to find the largest identifiable subset of parameters.

## 6.3 Results

We predicted ROTEM curves on four unseen patients with and without tPA (Fig. 6.2A). The model captured the change in curve shape with the addition of tPA, as well as the final amplitude in the absence of tPA. The model approximately captured the shape of the curves in the absence of tPA; it did not reproduce the slight decline in amplitude that occurred after reaching the maximum. Additionally, the measured ROTEM curves were more symmetric in their increase and decrease than the predictions in the case of tPA = 2 nM, which had a

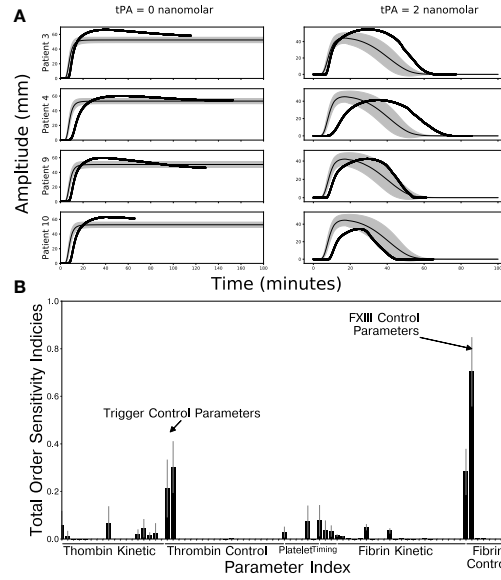


Figure 6.2: Simulation and analysis of the reduced order model. **A**: Prediction performance of the reduced order model on four unseen patients. The thick black line is the true ROTEM measurement, the thin black line represents the model prediction, and the shaded gray area represents the 95% confidence interval for the  $N = 24$  best parameter sets. **B**: Total order sensitivity coefficients with respect to area under the ROETM curve. The gray whiskers show the 95% confidence interval for the calculated coefficients.

more rapid increase than decline.

We used the global sensitivity analysis [136] to evaluate which parameters had the largest effect on the simulated ROTEM curve (Fig. 6.2B). To examine the sensitivity of the model to the  $p = 77$  parameters, we sampled 78,000 parameter families, with the values of the parameter families falling between 50% and 150% of the estimated parameter values. We evaluated the sensitivity of the area under the curve (AUC) of the ROTEM signal, as that summarizes the ROTEM curve, as well as commonly reported metrics computed from the ROTEM trajectory. Global sensitivity analysis suggested the area under the ROTEM curve was largely controlled by FXIII and trigger control parameters.

We found that we could improve the model performance by allowing the

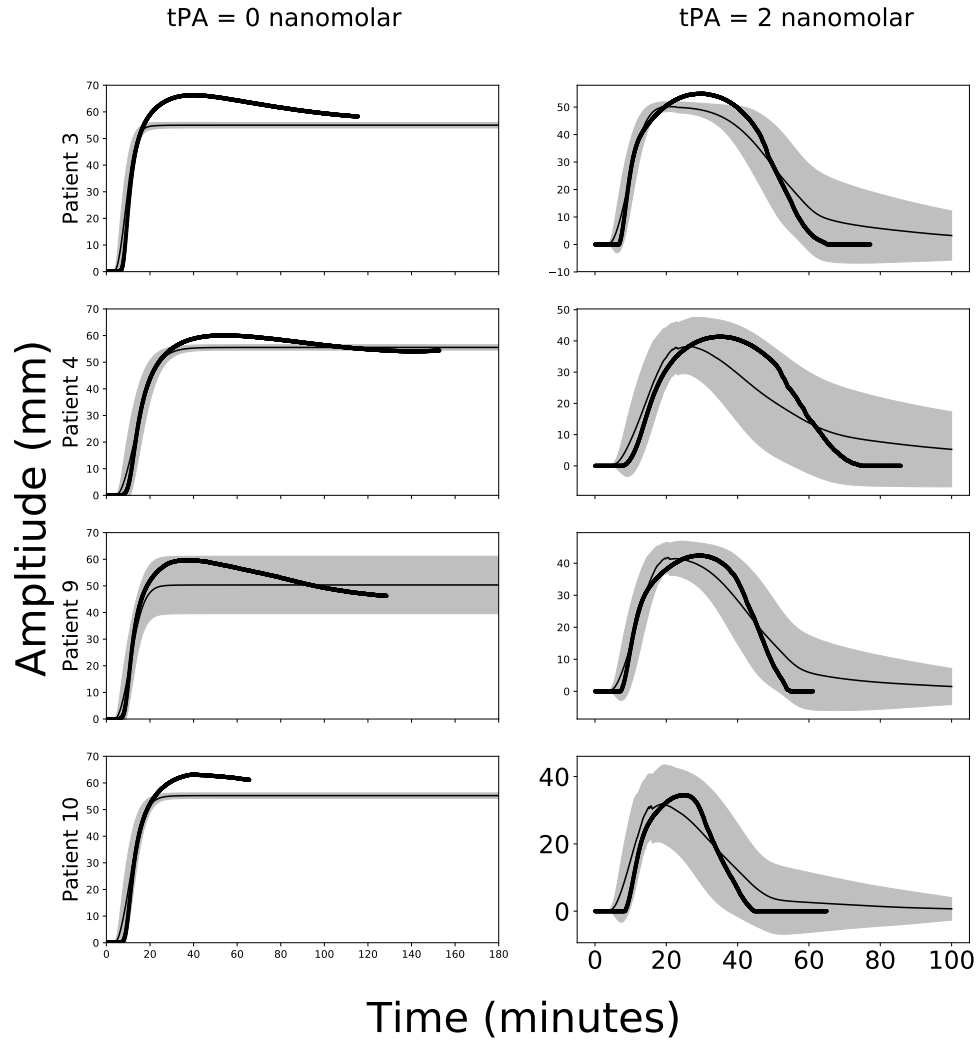


Figure 6.3: Prediction performance of the reduced order model on four unseen patients after adjustment of initial conditions. The thick black line is the true ROTEM measurement, the thin black line represents the model prediction, and the shaded gray area represents the 95% confidence interval for the  $N = 24$  best parameter sets.



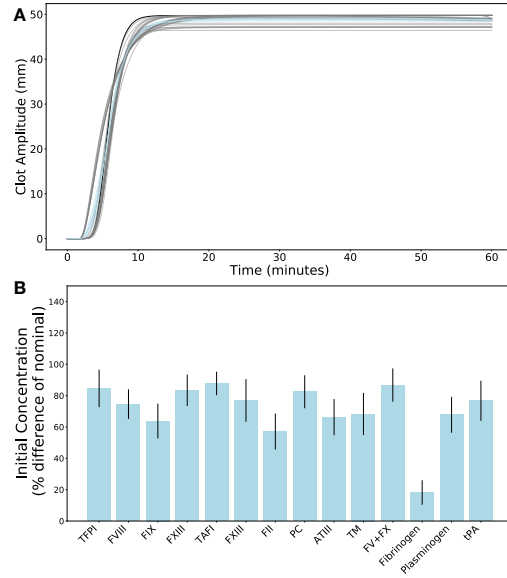


Figure 6.4: Model performance for reverse engineered initial conditions. **A:** The curves generated from minimizing the objective in 6.13. The black line shows the original generated curve, the gray curves are the curves found from optimization, the blue curve shows the average of the estimated curves, and the shaded area represents the 95% confidence interval for  $N = 25$  simulations. **B:** The differences in initial conditions found as we attempted to relearn initial conditions based on ROTEM curve metrics. The bars represent the average difference, and the error bars represent the standard error of the mean for  $N = 10$  simulations.

initial conditions of the 4 unseen patients to vary within 50-150% of normal, a range that is accepted in humans (Fig. 6.3). However, even with these adjustments, we fail to capture the slight kink observed in the  $tPA = 0$  nM case. Our model does not reproduce these dynamics because we set the concentration of uPA equal to zero, and the slight clot amplitude observed could be due to the action of uPA.

Our theoretical identifiability analysis based on a ROTEM curve as a function of its initial conditions found six (prothrombin, ATIII, FV\_FXa, prothrombinase platelets, fibrinogen, tPA) of the initial conditions to be sensitive, and all of these six conditions fell within the largest identifiable subset. However,

two of these initial conditions (FV\_FXa and prothrombinase platelets) represent intermediates which play key roles during coagulation, but do not exist in appreciable quantities outside of activation of the coagulation processes, leaving us with four useful theoretically identifiable initial conditions. One of the remaining four, tPA, is physiologically present at much lower levels than it was added during the assays conducted for the development of this model, so while large changes in tPA may be identifiable, we do not believe it will be identifiable under nominal conditions.

The initial concentration of fibrinogen could be estimated to within 20% of its true value from the ROTEM curve (Fig. 6.4). A critical question when interpreting ROTEM curves is their relationship with the underlying biochemistry of coagulation and fibrinolysis. Thus, having formulated a model, we wished to see if we could estimate the biochemical initial conditions that generated a ROTEM profile by solving an inverse problem; we treated biochemical initial conditions as unknown parameters to be estimated by repeatedly solving the model for different initial conditions and minimizing the difference between the predicted and measured ROTEM curves. However, in the literature, ROTEM curves are not summarized as trajectories over time, rather certain metrics are reported, including *CT*, clot formation time (time from beginning of the test until the clot reaches an amplitude of 2 mm), *CFT*, the time since the *CT* until the clot reaches an amplitude of 20 mm, the  $\alpha$  angle, the angle that describes the slope between the 2 mm and 20 mm points, the maximum amplitude or *MCF* (maximum clot firmness), and the maximum lysis, the extent of clot lysis at the end of the assay [176]. Thus, the objective function (to be minimized) for the

inverse problem was defined as:

$$O = \left( \sum_{j=1}^6 \left( \frac{y_{t_j} - y_{e_j}}{w_j} \right)^2 \beta_j \right)^{1/2} \quad (6.13)$$

where  $w_j$  is a scale associated with metric  $j$ , taken from the median values reported in [292],  $y_{t_j}$  is the target value for a metric, and  $y_{e_j}$  is the current value of a metric based on the present initial condition estimate. Lastly,  $\beta_j$  is the weight applied to metric  $j$ . The problem was constrained such that protein concentrations fell between 0% and 150% of their nominal values. We approached this problem in two ways: repetitively solving for the initial concentrations for one generated ROTEM curve (Fig. 6.4A), and by solving for the initial conditions for a number of different ROTEM curves simultaneously (Fig. 6.4B). The first approach allowed us to determine the accuracy that could be obtained in re-generating a ROTEM curve, while the second approach permitted us to observe how close the re-learned initial conditions were to those that generated the ROTEM curve. The inverse problem was solved using a particle swarm approach (as coded in the Julia package `Optim.jl` [220]) as the gradient is difficult and potentially expensive to calculate.

## 6.4 Discussion

The model qualitatively captured the ROTEM curves with and without tPA. However, there were fine features of the ROTEM response that were not properly captured by the model. For example, the model slightly over-predicted the clot amplitude in the case of 2 nM tPA for patient 10 and under-predicted patient 3 for the same case. This is likely because the same initial conditions, with the exception of platelet counts which were measured, were used for all pa-

tients. However, the spectrum of what is considered a normal concentration for proteins involved in the coagulation and fibrinolysis pathways is quite broad, with values ranging between 50% and 150% of the normal range [42]. Thus, if we adjusted the initial conditions within this range, we may be able to find a set of initial conditions afforded a better model description of the data without adjusting the kinetic parameters. Furthermore, while the model captured the shape of the curves in the absence of tPA, it failed to capture the asymmetry of the curves when tPA was added, nor did it capture the slight decline in amplitude occurring in the absence of tPA. This discrepancy could possibly be resolved by adding uPA to the model; uPA also activates plasmin, leading to clot degradation. Lastly, the model could be further strengthened by training with matched ROTEM curves and factor measurements; this would allow us to confirm that the model produces the observed ROTEM curves when given the true initial conditions.

Global sensitivity analysis suggested the area under the ROTEM curve was largely controlled by FXIII and trigger control parameters. FXIII, when converted to its active form, FXIIIa, plays a large role in determining how resistant fibrin clots are to fibrinolysis [43]. Thus, it follows that FXIII/FXIIIa would play a large role in determining the area under the ROTEM curve. This result concurs with the results reported by Wettstein et al, in which they reported that decreased FXIII availability correlated with unexplained intraoperative bleeding as well as a decline in MCF [358]. Trigger (tissue factor bound to FVIIa) initiates the coagulation process that catalyzes key steps in the formation of the clot. Thus, by changing the parameters that control coagulation initiation, we can dramatically change the area under a ROTEM curve. Global sensitivity analysis also revealed that the trigger control parameters played a large role in the max-

imum clot firmness, as do the kinetic parameters involved in the initial burst of thrombin generation. While ROTEM only directly measures fibrin generation and lysis, the importance of the thrombin generation process that leads to fibrin generation can not be neglected. Even if a patient has a clinically normal amount of fibrin, if they are lacking a key protein in the thrombin generation cascade, we would expect this to be evident from an altered MCF. Additionally, the trigger control parameters were important to determining the CFT, as well as the kinetic parameters that determine the rate at which fibrinogen is activated to fibrin and the rate constant that determines how quickly plasmin breaks down fibrin. Taken together, sensitivity analysis suggests that for maximum efficacy, patients should receive drugs that target the tissue-factor-FVIIa complex or FXIII. FXIII represents a particular target of interest, since while FXIIIa helps protect clots from degradation in the absence of fibrin, it will not prompt clot formation, and furthermore, through encouraging clot stability, bleeding and transfusion requirements could be reduced.

Solving the inverse problem revealed which initial conditions could be easily estimated from commonly reported ROTEM metrics. By attempting to solve for the initial conditions of one ROTEM curve a number of times, we learned that we can discover a set of initial conditions that, on average, will reproduce a ROTEM curve very close to the curve which we sought. When generalized, this process identified which initial conditions were robust (can tolerate a large change in value while still producing a comparable ROTEM curve) and which were sensitive (a small change in their value completely alters the ROTEM curve). The initial amount of fibrinogen appeared to be very sensitive, with it being recovered within 20% of its true value. Fibrinogen is the key protein that forms the majority of the clot, so it is logical that we recovered close

to the true fibrinogen concentration, as changes in its concentration would conceivably alter all of the terms in our objective function. Since these curves were generated at physiological levels of tPA, which are much lower than the 2 nM level used in the parameter estimation case, it also follows that the initial level of plasminogen was robust, since very little of it is activated by the trace amount of tPA present in blood in the absence of injury or pathology. In the future, it may be interesting to investigate what information is contained within each ROTEM metric. It is also possible that a different optimization technique would be able to find better solutions than particle swarm.

## 6.5 Conclusions

In this work, we have presented a model of coagulation, fibrin generation, and fibrinolysis which was validated against thrombin generation data and ROTEM curves. We demonstrated that the model describes clot formation in the presence and absence of tPA, a key protein in the clot degradation process. We found that control parameters for FXIII, the precursor of the protein that assists in fibrin polymerization plays a key role in a number of ROTEM curve metrics. By solving the inverse problem, we found that we could (on average) recover a ROTEM curve and discovered which initial conditions (specifically, the initial concentration of fibrinogen present) must be tightly constrained to recover a curve with the desired metrics.

The model codes described in this paper are available at [https://github.com/varnerlab/Fibrinolysis\\_model\\_w\\_platelets](https://github.com/varnerlab/Fibrinolysis_model_w_platelets)

## CHAPTER 7

### PHYSIOLOGICALLY BASED PHARMACOKINETIC MODELING OF TRAUMA

#### 7.1 A Brief History of Physiologically Based Pharmacokinetic Modeling

<sup>1</sup> PBPK (physiologically based pharmacokinetic modeling) arose from anaesthesiologists seeking to understand the role of ventilation and blood flow rates on the effects of anaesthetics on their patients in the 1920s [258]. In the 1930s, Torirell wrote a set of equations to describe all the key processes in drug metabolism: uptake, distribution, and elimination [331]. However, the computational tools necessary to solve the equations used to describe the concentrations over time had not yet been developed (resulting in the interim use of simplified models), and it wasn't until 1963 that Mapleson used an analog computer to solve for the distribution of various gases within human tissues that the equations describing a PBPK model became tractable [202]. With the increasing availability of digital computers, data based models were developed to model the metabolism of toxic species in rats, and these models were successfully scaled up to predict time course response curves with humans dosed with the same substance [83]. PBPK models are widely used in chemical risk assessments and can be used to create falsifiable predictions of drug metabolism so that models can be revised when data does not fit predictions, resulting in an improved model.

---

<sup>1</sup>Under preparation as "A Validated Physiologically Based Pharmacokinetic Model of Trauma" R. LeCover, M. Bravo, T. Orfeo, K.E. Brummel-Ziedins, J. Varner

## 7.2 Mathematical Formulation

PBPK models arise from the basic principle of conservation of mass.

$$\frac{d(c_{ij} * V_j)}{dt} = \sum_k Q_k c_{ik} - \sum_m Q_m c_{ij} + R_{ij} * V_j \quad (7.1)$$

where  $c_{ij}$  is the concentration of species  $i$  in compartment  $j$ ,  $V_j$  is the volume of compartment  $j$ ,  $k$  represents the set of all compartments connected to compartment  $j$  with inflows,  $m$  is the set of all outflows from the compartment, and  $R_{ij}$  is the rate at which  $i$  is produced or consumed in compartment  $j$  based on the kinetics of the model describing the species of interest. In many PBPK applications, the  $R_{ij}$  terms are assumed to be zero in every compartment except for the liver, in which species are metabolized. Additionally, in many cases, the volume of all compartments is assumed to be constant, allowing for the mass balances to be rewritten as:

$$\frac{d(c_{ij})}{dt} = \frac{1}{V_j} \left( \sum_k Q_k c_{ik} - \sum_m Q_m c_{ij} + R_{ij} * V_j \right) \quad (7.2)$$

However, in the case of trauma, where blood is being lost, this simplification is not suitable to the change in compartment volume as blood is lost and as resuscitation fluids are given.

Within our model, the  $R_{ij}$  terms are described using the reduced order model of coagulation and fibrinolysis described in the previous chapter.

## 7.3 Physiologically Based Pharmacokinetic Modeling of Trauma

PBPK models have been previously used to simulate trauma to a limited extent. Ho et al used a one compartment PBPK model to simulate hemoatocrit



concentrations following blood loss and resuscitation, but only considered the loss of factors through blood loss, not through consumption through coagulation [141]. A slightly more complex model by Hirshberg et al. modeled blood as having three compartments: red cells, plasma, and water, and allowed flow between these compartments based on systolic blood pressure [140]. They used prothrombin time to quantify if a patient was at risk for dilutional coagulopathy, and did not consider the biological mechanisms behind coagulation. A model without compartments, but containing more physiological functions was developed by Simpson et al. [309]. Simpson's model allows for both blood pressure and bleed out rate to change over time, with the bleed out rate decreasing as blood pressure decreases. This model predicts hematocrit levels over time, but not the levels of specific proteins. Reisner et al. re-purposed their model of the cardiovascular system (which was originally created to predict how the cardiovascular system responds to orthostatic stress) to study hemodynamic responses to haemorrhage [260]. Their model includes the heart and pulmonary circulation as well as four peripheral tissue compartments representing the upper body, legs, viscera and kidneys, each of which received the same fraction of the cardiac output. They modeled blood by separating it into two components: red blood cells and plasma. This model includes transcapillary fluid exchange and lymphatic flow but groups all proteins together. The model developed by Peng and Sweeny is very complete, using a 15 organ simulated body and was validated against supplemental fibrinogen given to trauma patients [252]. Brown and co-workers used a three compartment PBPK model to model portions of the immune response to trauma, but the data they used to fit their model was sampled infrequently, and their model fits at some of the later time points were poor [39]. However, all of these models lack a detailed description

of coagulation and fibrinolysis, a shortcoming we aim to remedy.

## **7.4 Modeling different treatments using Seheult bleeding model**

We sought to quantify the effects of different fluid resuscitation treatments on a patient's ability to clot. The PROMMTT study showed that high ratios of plasma to RBCs and platelets to RBCs were independently associated with a decreased 6-hour mortality rate [144], and another study showed that during damage control resuscitation, higher volumes of crystalloids were associated with decreased survival and low volumes of colloids were associated with increased survival [127]. Although the effects of different resuscitation schedules on mortality had been measured, we wished to see how they effected clotting ability, which plays a role in mortality outcomes. To perform these simulations, we used a timeline and model developed by Seheult et al [299]. In this scheme, the treatment a patient receives is divided into five phases, as shown in Figure 7.1.

At minute 0, the patient begins bleeding out at 135 mL/min, and 10 minutes later, EMTs arrive and begin resuscitating the patient. At 30 minutes, the patient arrives at the hospital, and 30 minutes later (at minute 60), the patient receives surgery to repair the damages, and which concludes at minute 160, and then bleeding finally stops at minute 240, four hours after the initial injury. Within this scheme, the patient's bleeding rate is a function of his blood pressure, so when fluids are given, the blood pressure increases, also increasing the bleeding

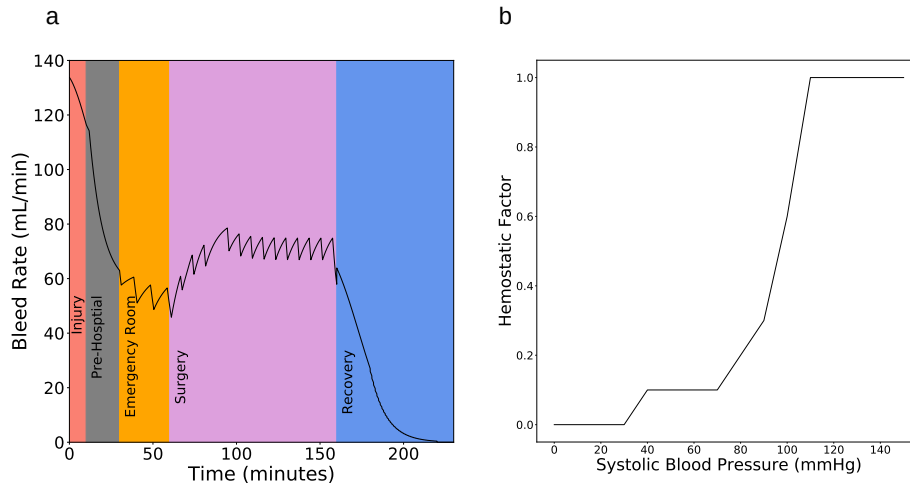


Figure 7.1: (a) Bleeding rate as a function of time under the Sheult et al schedule. (b) Hemostatic factor as a function of blood pressure. As blood pressure decreases, so does the hemostatic factor.

rate, as seen in Figure 7.1a. Mathematically,

$$SBP = (130 * (1 - ((\frac{70 * \text{weight} - \text{total blood volume}}{70 * \text{weight} * 0.67})^{1.6}))^{0.625}) \quad (7.3)$$

which empirically relates blood lost to blood pressure. The bleeding rate was calculated as a function of time and blood pressure.

```

1 if  $t < \text{tend operating room}$  then
2   |   bleed rate =  $\frac{\text{initial bleed rate} * SBP}{130 * \text{hemostatic factor}}$  ;
3 else if  $t > \text{tend operating room}$  AND  $t < \text{tend operating room} + 20$  then
4   |   bleed rate =  $\frac{\text{bleed rate at minute 159}}{(1 + .1 * e^{1 + (t - 160)/10})}$  ;
5 else if  $t \geq \text{tend operating room} + 60$  then
6   |   bleed rate = 0.0 ;
7 else
8   |   return .9 * calculateBleedRate(t-1, x, datadictionary) ;
9 end

```

The hemostatic factor used in the above algorithm attempts to account for how much blood has been lost as a function of blood pressure, as shown in Figure 7.1b.

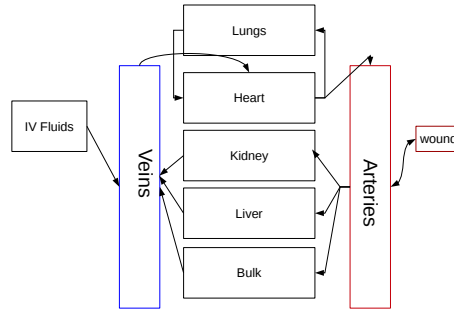


Figure 7.2: The simplified body used in the simulations.

We then combined this model of bleeding with a simplified body, as shown in Figure 7.2. Within this body, blood flows from the arteries, through the the bulk, liver, and kidney compartments, to the veins, where we can inject IV fluids. Blood flows from the veins, to the heart, through the lungs, and back to the heart before returning to the arteries, from where the patient is bleeding. Since these simulations were only run for four hours, net synthesis of proteins was not included.

We simulated three different courses of treatment. The most realistic case is the one in which the patient receives saline when the EMTs arrive, in the pre-hospital phase, and then begins to receive conventional component therapy (CCT) (units of plasma, packed red blood cells, and platelets). In the second case, the patient receives saline when the EMTs arrive, but once he arrives at the hospital, he begins to receive whole blood. In the third case, when the EMTs arrive, they are carrying whole blood, and immediately begin resuscitation with it, and the patient continues to receive whole blood throughout the remainder

of the simulation. To assess the effects of the different treatments, we ran a simulated ROTEM every ten minutes (using the previously described kinetic model), taking the blood composition from the patient's veins.

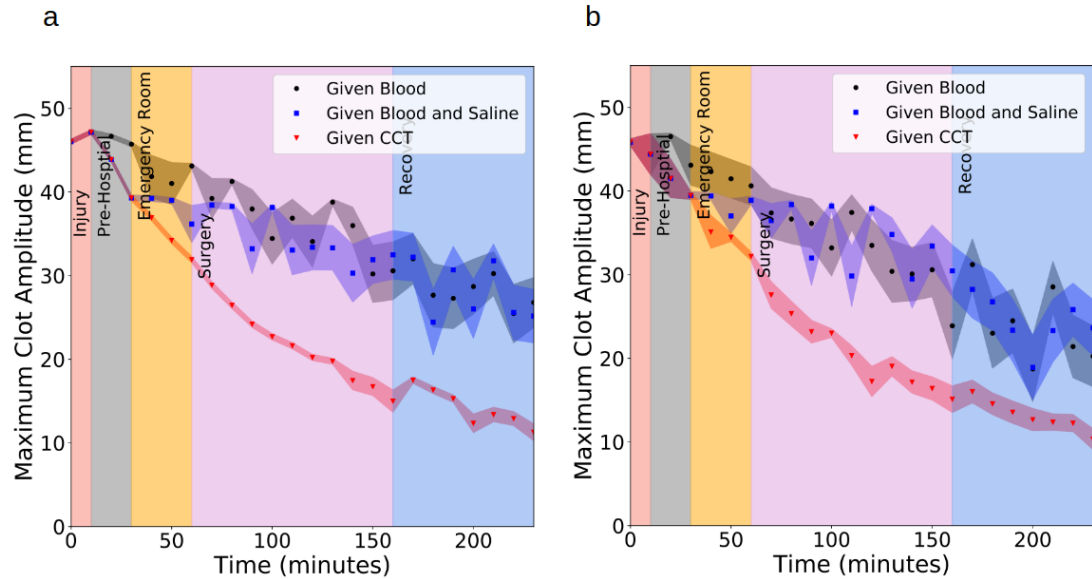


Figure 7.3: (a) The resulting maximum clot amplitude measurement from drawing blood from the simulated patient's veins throughout injury, pre-hospital, emergency room, surgical and recovery phases. The dots represent the mean maximum clot amplitude, and the shaded area represents the 95% confidence interval over fifteen selected parameter families. (b) The resulting maximum clot amplitude measurement from drawing blood from the simulated patient's veins throughout injury, pre-hospital, emergency room, surgical and recovery phases while the patient suffers from acidosis and/or hypothermia. The dots represent the mean maximum clot amplitude, and the shaded area represents the 95% confidence interval over fifteen selected parameter families.

As shown in Figure 7.3a, the differences in treatment are most apparent during the emergency room and surgical phases of the simulation. During these periods, the benefits of whole blood appear, as the MCF of the patient has not declined nearly as much as it has in the case where the patient was initially resuscitated with saline. However, regardless of the fluids used to resuscitate the patient, by the end of the simulation, their maximum clot amplitude has declined, most dramatically in the case when CCT was given, representing a

reduction in clotting capability. This reduction in clotting capability is due to the fact that the initiator of coagulation, the "trigger" complex remained at the same concentration throughout the simulation, causing the coagulation system to keep running as the clot degradation system in turn degraded the formed clot. In short, the system was running in place, with both clot production and clot degradation systems competing with each other and eating up the proteins used to form a clot, even those infused when blood was given. From this set of simulations we learned that if the signal to form a blood clot is not damped, regardless of the resuscitation fluid given, the patient will lose a non-trivial portion of their coagulation capacity. This decline is most dramatic when CCT is given, as by volume, it contains the most additives, which preserve the lifespan of the products, but dilute their coagulation boosting capabilities. Additionally, it appears that the dilutional effects of the initial liter of saline given are still significant even when compared to the coagulation system churning in place over the duration of the simulation, as the patients who received blood and those who received saline and then blood both have declines in their estimated MCF by the time they conclude with surgery, but this drop is far more dramatic in the patients who initially received saline. This finding is agreement with a study that found there was a significant increase in 30 day mortality among patients who received pre-hospital crystalloid and blood products compared to patients who only received blood products [128]. In a study with swine, the swine who received resuscitation fluids (either lactated ringer's solution or Hextend) had higher levels of pro-inflammatory mRNA levels (for IL-6, G-CSF and TNF $\alpha$ ) in their lungs compared to swine who did not receive any resuscitation fluid [354], so it is possible that this increase in inflammation from fluids also occurs in humans, and reduced inflammation may be a benefit of allowing a patient to

remain somewhat hypotensive.

While it would be impossible to measure the rate at which tissue factor exposure declines in humans after injury, this parameter could possibly be estimated in a porcine or primate model. While I do not believe "trigger" being stuck in the on state is the cause of all reported cases of ATC, it could certainly account for some of the cases, where the patients possess a mutation where tissue factor pathway inhibitor (TFPI) is produced at low levels, which makes damping the signal to clot difficult. Alternatively, the TFPI produced could not bind securely to the exposed tissue factor, allowing the signal to clot to persist for longer than it should, resulting in the development of ATC.

To assess the effect of acidosis or hypothermia on our patient, we ran the same set of simulations while modifying the rate constants of the reactions to simulate slower reactions due to either hypothermia or acidosis, the results of which are shown in Figure 7.3b. In these simulations, we assumed that the effects of acidosis and hypothermia reduced the reaction rates by 10%, that is, kinetic rate values were set to 90% of the values used in the previous simulation. We also assumed that this hypothermia/acidosis was not corrected when the viscoelastic assay was run, and that it persisted at this level throughout the simulation. With this reduction in kinetic rate constants, we can see a decline in MCF if we compare similar cases from Figures 7.3b and 7.3a with the cases with the patient receiving whole blood while suffering from acidosis and hypothermia coming closer to approaching the reduction in MCF observed when a patient is treated with CCT. It appears that untreated hypothermia and acidosis negate some of the benefits of blood as a resuscitation fluid, as it does the patient less good to have coagulation factors available to form a clot if their

reaction rates are slowed by cold and an acidic environment.

## 7.5 Simulation of Fibrinogen Supplementation

Fibrinogen is the protein, once activated, that will polymerize and form a clot, and as such, plays an important role in coagulation. Fibrinogen supplementation may boost survival rates in trauma patients: in a small study performed on 31 pairs of propensity matched trauma patients, those who received fibrinogen concentrate within one hour of arrival at the emergency department had a significantly higher in hospital survival rate and a significantly lower 28 day in hospital mortality rate [156]. In a study conducted at the AUVA Trauma Centre, Salzburg, Schlimp and coworkers retroactively assessed the effects of giving fibrinogen concentrate to trauma patients [294]. Within this study, patients were first divided into two groups: those who received fibrinogen concentrate (242) and those who did not (193), and then the group that received fibrinogen concentrate was further subdivided by the amount of fibrinogen concentrate they received. We used the resuscitation fluids and protein concentrations reported in this study to validate the performance of our model over a longer time period. We used the average amount of each type of fluid given within a group to run our simulations. Additionally, since the patients who received no fibrinogen concentrate were less severely injured than those who did (average reported injury severity score of 22 versus 34 in the group that did receive fibrinogen concentrate), we used a lower initial wound volume and trigger concentration in those simulations.

Within these simulations, we used vascular resistances to determine how



the blood volume distributed during and post trauma. The resistances used are shown in Table 7.1.

Organ	Nominal Vascular Resistance	Units of nominal resistance	Source	Trauma Vascular Resistance	Units	Source	Notes
veins	0.00187	mmHg/(ml·min)	[35]	0.00349	mmHg/(ml·min)	[35]	
heart	69.8	mmHg gm min/mL	[27]	57.1	mmHg gm min/mL	[27]	
lungs	577	Dynes sec cm <sup>-5</sup>	[33]	889	Dynes sec cm <sup>-5</sup>	[33]	
kidney	4332	dyne-sec/cm <sup>2</sup>	[132]	9103	dyne-sec/cm <sup>2</sup>	[132]	
liver	288	mmHg gm min/mL	[27]	121	mmHg gm min/mL	[27]	
bulk	16	mmHg min/L	[293]	24	mmHg min/L	[293]	Taken from figure 9, low tolerance
arteries	16	mmHg min/L	[293]	24	mmHg min/L	[293]	using same as bulk

Table 7.1: Vascular resistances used in the body model.

With these resistances, we could calculate the flows using  $Q = \frac{\Delta P}{R}$ . We switched between using nominal vascular resistances and the trauma vascular resistances when the blood volume of the simulated patient dropped below 80% of its original value, and used our patient's mean arterial pressure as the pressure drop in the flow calculations.

Since these simulations were run over a time period of 48 hours, we did include coagulation factor synthesis and degradation in these simulations. Synthesis and degradation base rates were taken from literature, when available, and when not, we calculated from the literature half-lives, where  $k_d = \frac{\log(2)}{\tau_{\frac{1}{2}}}$ . The majority of protein synthesis occurred in the liver, however, tPA was synthesized in the veins, as according to [297]. We also permitted acute phase synthesis of coagulation factors, where after a lag phase, protein synthesis would ramp up to a higher level for a period of time before returning to basal levels. Additionally, the trigger levels were set to decay following an initial phase where they remained at their initial levels (Equation 7.4). The rate of trigger delay as well as the rate of protein synthesis during the acute phase, and the duration of the acute phase were tuned so that our data matched experimental measurements.

$$\text{trigger}(t) = \text{trigger}_0 e^{\frac{t - (\text{trigger begin decline time})}{\tau}} \quad (7.4)$$

As shown in Figure 7.4a, regardless of the amount of supplemental fibrinogen given, the patients that received some fibrinogen (3, 6, or 12 g), are all in about the same place 48 hours later, where the supplemental fibrinogen has all been chewed up and degraded. We found that we were able to fit the fibrinogen trends with a lag phase of two hours, after which synthesis was increased to 5.8 times the normal synthesis rate and remained at this increased synthe-

sis rate throughout the duration of the simulation. It would be interesting to have D-dimer levels from this or a similar study to confirm that the supplemental fibrinogen is being degraded, and not sequestered in another portion of the body.

Figure 7.4b shows how platelets are consumed following injury, with the largest drop in platelet concentration in the most severely injured patients. As platelet concentration does not appear to rebound within the first 48 hours after injury, we assumed that the platelet production rate remained at its nominal level after injury. To fit the platelet concentrations observed, we added a term that accounted for platelets being incorporated into a clot in addition to the platelets lost to bleed out. The consumption term was written as

$$\text{platelets consumed by clot} = \text{overall platelet concentration} * \text{Wound Volume} * A_0 * e^{\frac{t}{\tau}} \quad (7.5)$$

where  $A_0 = .5$  and  $\tau$  describes how quickly platelets stop being incorporated into the clot.

Figure 7.4c, shows that that at the first time point (assumed to be one hour after injury) and a day later, ATIII levels appear to be a function of injury severity, as the most injured patients received the most supplemental fibrinogen. To fit the ATIII data, we used a one hour lag phase, after which ATIII synthesis increased to five times its nominal rate, before returning to the normal synthesis rate 36 hours after injury.

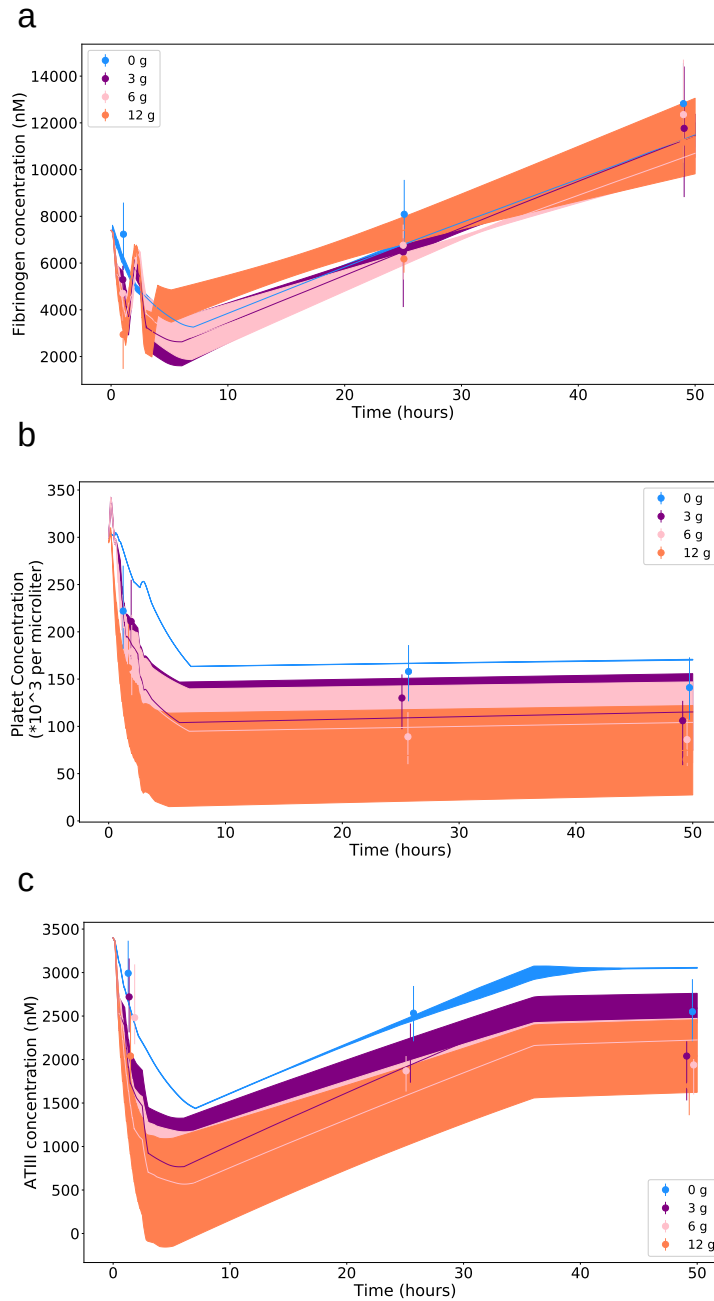


Figure 7.4: Experimental measurements are shown as dots, with the IQR represented as error bars, and the simulated mean levels are shown as lines, with the 95% confidence interval shaded over a family of 24 kinetic parameter sets. (a) Fibrinogen concentrations during the first 48 hours following injury. (b) Platelet count during the first 48 hours following injury. (c) ATIII concentration during the first 48 hours following injury.

## 7.6 Simulation of Cryoprecipitate Dosing

Cryoprecipitate, a product that can be used as a resuscitation fluid following injury can either be derived from whole blood or fresh-frozen plasma (FFP), is a rich source of fibrinogen, factors VIII and XIII as well as von Willebrand factor, and fibronectin, which is known to bind to fibrin [247]. We used data from a feasibility study in rapid cryoprecipitate administration to further validate our model [74]. Within this study, adult patients either received standard haemorrhage therapy or the standard therapy supplemented with two pools cryoprecipitate. The standard treatment consisted of 6 units of red blood cells and 4 units of fresh frozen plasma. Patients in both arms received 1 gram of tranexamic acid as a bolus followed by another gram over 8 hours as infusion. However, patients in the no-cryoprecipitate group recieved on average 50 mL of crystalloid prior to randomization, compared to an average of 250 mL of crystalloid in the cryoprecipitate group prior to randomization.

Our model succeeds in capturing the fibrinogen trajectory with and without cryoprecipitate being given. In general, our model appears to predict a less steep decline in fibrinogen concentration than was found to be the mean concentration found in the twenty patients in the study. During the first hour or so of the simulation, the control group has a more steep decline in initial fibrinogen concentration due to the fact that they received more crystalloid than the cryoprecipitate group. It is interesting to note there appears to be a wide variation in fibrinogen concentrations throughout the first day following injury, especially at the 25 hour timepoint. Some of this variation may be due to the fact that there is variation in how the signal to produce fibrinogen gets transduced to the liver, variations in the proportion of each of the subchains of fibrinogen that are pro-

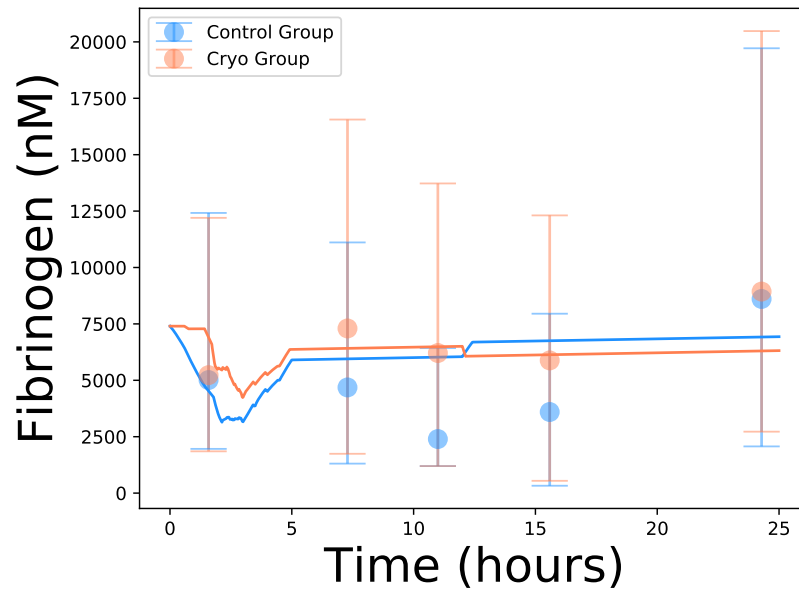


Figure 7.5: Fibrinogen concentrations over the first day following injury with and without cryoprecipitate supplementation. The reported concentrations from [74] are shown in dots with error bars, the simulated concentrations are shown as lines.

duced, or due to variations of the reserve of fibrinogen available intracellularly. It may also be due to variations in clot structure, with some people producing more compact clots which are more resistant to fibrinolysis [4]. Through genome wide association studies, 41 significant loci affection fibrinogen have been identified, and it would be interesting to see if any of these 41 loci could explain some of the variation in fibrinogen levels following injury [81].

## 7.7 Comparison of Whole Blood and Hextend Resuscitation in Monkeys

To examine what is occurring within trauma patients at a more granular scale and to validate our model further, we used data published by Sheppard et al in a

monkey model of trauma [303]. In this study, male rhesus macaques, weighing between 6 and 13 kg were anaesthetized, and then given a 15 cm laparotomy as well as a femur fracture, and then blood was removed until the mean arterial pressure reached 20 mmHg, to simulate haemorrhage [304]. Ninety minutes after injury, the monkeys were resuscitated with either citrated whole blood or Hextend, equivalent to half of the shed blood volume. Following whole blood or Hextend resuscitation, the monkeys were given additional lactated Ringer's solution, equivalent to double the shed blood volume.

For the purposes of our simulation, we simulated a monkey with a weight of 7 kg, with trigger being exposed both in the wound compartment and within the liver, however, the trigger factor in the liver was initially expressed at a concentration of 3 pM as compared to the 5 pM trigger in the wound compartment. We assumed that the whole blood used for resuscitation contained 90% of the coagulation factors measured at baseline, as the study stated that shed blood was collected in anticoagulant citrate-phosphate-dextrose-adenine solution at a 10:1 ratio [303].

We found if we used the same acute synthesis rates and lag phase durations as found in humans based on our estimation from the fibrinogen supplementation study, we recovered the correct trends for monkeys, but underestimated the amount of most of the factors present by the end of the study. However, if we adjusted the lag phases and synthesis rates, we could achieve much better fits. In general, as shown in Table 7.2, the estimated synthesis rates in monkeys are much higher than those in humans. This raises the question as to if studies performed in animals to simulate traumatic injury can easily be generalized to humans, if the acute phase protein responses are indeed markedly



Protein	Study	Lag Time (minutes)	Acute Phase Multiplier
FII	Fibrinogen Supplementation in Humans	30	2
Fibrinogen	Fibrinogen Supplementation in Humans	120	5.8
ATIII	Fibrinogen Supplementation in Humans	60	5
FII	Hexend vs Blood in Monkeys	90	9
Fibrinogen	Hexend vs Blood in Monkeys	120	11.5
ATIII	Hexend vs Blood in Monkeys	120	30

Table 7.2: A comparison of estimated human and rhesus macaque acute phase multipliers and lag times.

different. Humans and the rhesus macaque diverged from a common ancestor about 25 million years ago, and aligned sequences between humans and rhesus macaques share 90.6% of identity, allowing for a large number of differences between the two species to develop [118]. Although swine are commonly used as models for haemorrhagic shock and trauma, the levels of cytokines  $\text{TNF-}\alpha$ , IL-10, and IL-6 in swine in a polytrauma model do not correspond to the time course levels in humans [337]. The human transcriptome changes that occur after trauma are just starting to be characterized [366], but as of yet, the gene regulatory network which controls these changes is poorly understood, and it is unclear how well this regulatory network is preserved between species. Additionally, animals are usually anaesthetized before they are injured, and it is difficult to quantify how anaesthesia effects the response to injury. Based on our simulation results, we would take any results generated in an animal model with a grain of salt, until adjustments are made to account for differing simulation rates and lag phases between different species. Given these unmeasured differences between species, it might prove worthwhile to develop an understanding of the gene regulatory network of trauma in animals that are commonly used to simulate trauma so that we can be more certain as to how well the animal results generalize to humans.

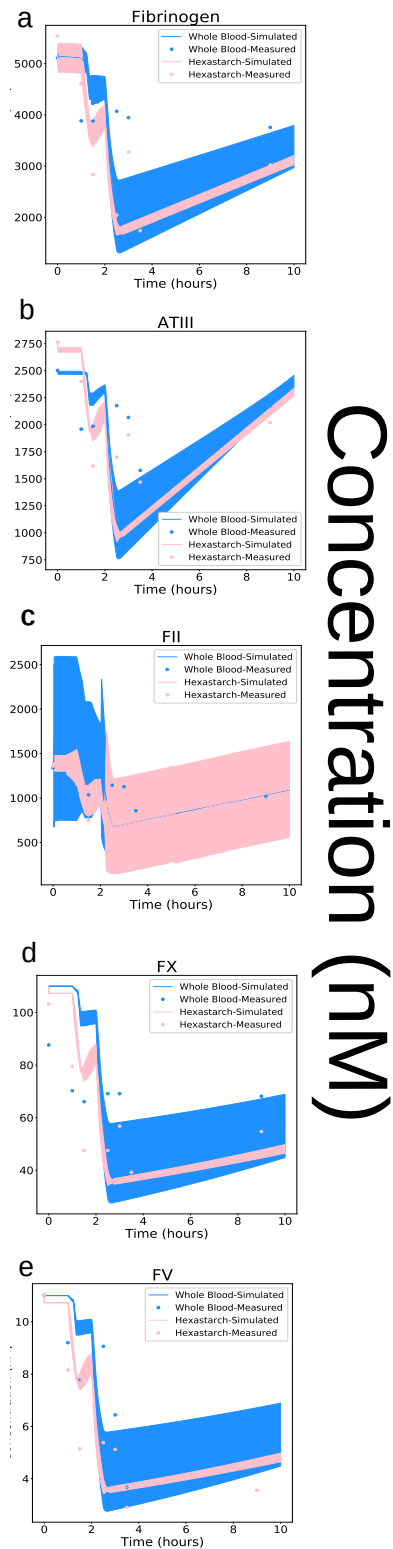


Figure 7.6: The pink corresponds to initial resuscitation with blood, and the blue corresponds to initial resuscitation with Hextend. Experimental data shown as dots, simulations shown as lines with one standard deviation shaded. (a) Fibrinogen (b) ATIII (c) FII (d) FX (e) FV

## 7.8 Dosing with Tranexamic Acid (TXA)

Following the CRASH-2 trial, which showed that patients which received tranexamic acid less than one hour after injury had a significantly lower risk of death due to bleeding [71], there was significant interest in treating all trauma patients with TXA soon after injury. Within this study, patients received 2g of TXA in total-1 gram as a loading dose, administered in 10 minutes, and then the second gram was administered over 8 hours [362]. The reason why TXA was so effective is an open question, and many aspects of its mechanism of action are still an active area of investigation. Interestingly, while TXA reduced the risk of bleeding if it was given within three hours of injury, if it was given more than three hours after injury, it increased the risk of death due to bleeding [274]. TXA binds to plasminogen at several lysine binding sites, one with high affinity and others with lower affinity [151], and when it binds, it prevents plasminogen from activating and degrading fibrin. TXA inhibits plasmin, the active form of plasminogen, but much higher doses are needed for this inhibition to occur [363]. TXA also inhibits uPA (which also activates plasminogen), with half of the activity of uPA being blocked at a TXA concentration of  $3.63 \pm .16$  mM, compared with  $86.79 \pm 2.3$  nM necessary to inhibit half of the activity of plasminogen [365]. TXA is also thought to interact with IL-6 and  $\text{TNF}\alpha$ , both components of the immune response, and may result in higher levels of these proteins, along with others associated with the pro-inflammatory immune response [17].

So that we could investigate the effects of a patient receiving TXA, we used a previously developed two compartment PBPK model to simulate the concentration of TXA over time. Within this model, the two compartments are the well perfused compartment and the poorly perfused compartment, where the

parameters of the model are scaled by bodyweight [123].

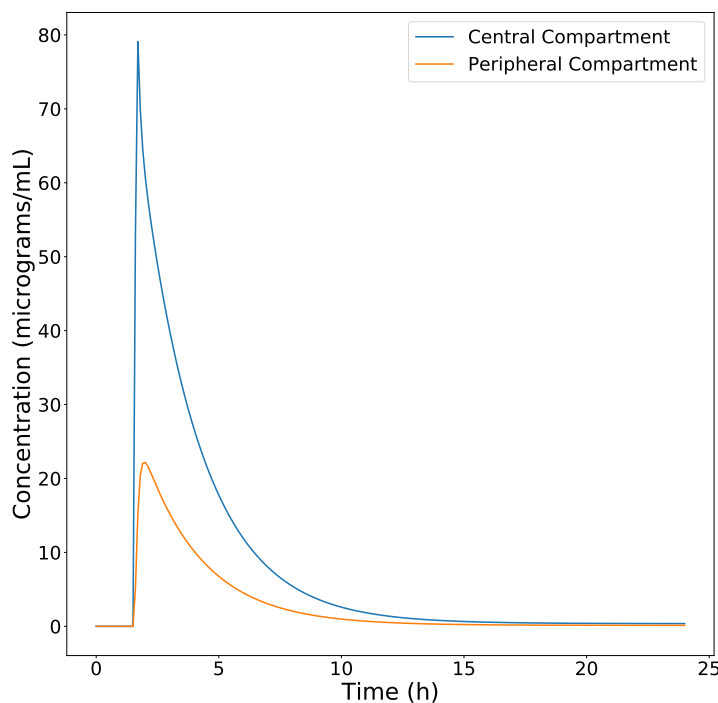


Figure 7.7: TXA concentration in the first day after trauma. The 75 kg patient receives 1 gram of TXA at 90 minutes, and then another 1 gram in the following 8 hours.

We can then use the concentration of TXA in the well perfused compartment as the concentration in the blood, and add an additional reaction to the full body PBPK to account for the deactivation of plasminogen when it complexes with TXA. To accurately model this process, we need the rate constant which describes how fast this plasminogen-TXA complex forms. While I am unable to find an exact value for this rate constant, it was estimated in a computational study, and we used this estimated value of  $1E-6M^{-1}s^{-1}$  value in our simulations [308]. With a rate constant this large, within our model, once TXA is given, the amount of plasminogen available drops to near zero very rapidly. Since

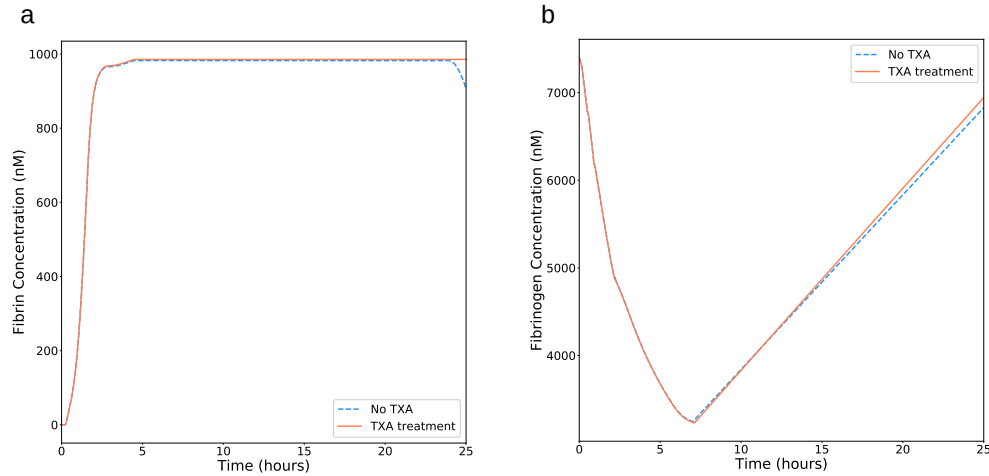


Figure 7.8: Fibrin concentration within the wound compartment.

Figure 7.9: The effects of dosing a trauma patient with TXA over the first 25 hours following injury at time zero. The dashed blue lines indicate a patient that did not receive TXA, and the solid pink lines indicate a patient who received TXA. (a) Fibrin concentration within the wound compartment. (b) Fibrinogen concentration within the vein compartment.

the inhibition of plasmin is thought to occur at much higher concentrations of TXA, I modeled that interaction as occurring one thousand times slower than the interaction between TXA and plasminogen.

We used the same conditions as the trauma patient who received no supplemental fibrinogen to investigate the effects of receiving TXA on a trauma dosing schedule. Within our simulations, we found that there was a negligible effect of TXA on fibrinogen concentration within the first 20 hours, as shown in Figure 7.8b. This is an expected result, as TXA within our model (and as far as is known), does not interact directly with thrombin, the protein responsible for converted fibrinogen into its active form. While TXA reacts quickly to deactivate plasminogen, and the physiological levels of tPA present within

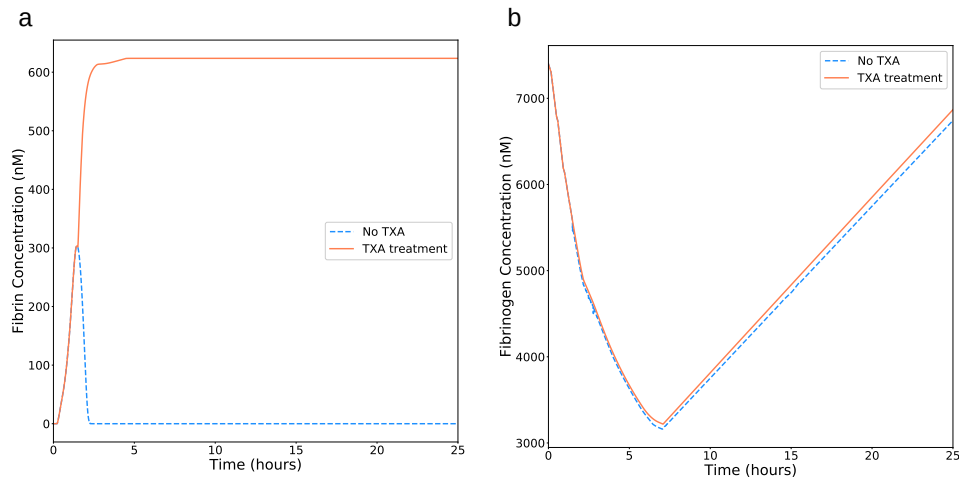


Figure 7.10: The effects of dosing a trauma patient with TXA over the first 25 hours following injury at time zero with the initial tPA concentration set to ten times the nominal concentration. The dashed blue lines indicate a patient that did not receive TXA, and the solid pink lines indicate a patient who received TXA. (a) Fibrin concentration within the wound compartment. (b) Fibrinogen concentration within the vein compartment.

our simulated trauma patient, the formed clot is not broken down, as insufficient plasminogen is activated, which would be necessary to degrade the fibrin clot, as seen in 7.8a, where the fibrin concentration within the wound compartment increases and then remains flat until 20 hours. Since the clot remains intact throughout the course of the simulation owing to the low levels of tPA present, the observed effects of TXA are negligible.

Since we believed that the negligible effects of TXA were due to the low amounts of tPA present, we simulated a trauma patient with ten times the amount of nominal tPA present to see if the effects of TXA would become more pronounced. We found with this increase in available tPA, there were dramatic changes both within fibrin and fibrinogen concentrations, as seen in Figure 7.10.

Giving TXA helped preserve the the fibrin clot for a longer period of time, and increased the maximum fibrin concentration achieved in the wound compartment (Figure 7.10a). However, this dose of TXA did not effect the fibrinogen concentration with in the patient's veins (Figure 7.10b). TXA appears to be helping to preserve the initial clot that forms rather than preserving the pool of available fibrinogen. The fact that that TXA was so effective in the CRASH-2 trial may be because this sort of dysfunction is common following injury, where too much tPA is released, causing the over-activation of plasminogen and the premature degradation of a clot. However, this is only one possible path leading to the production of excess plasmin-plasmin concentrations could also become elevated if plasminogen synthesis is elevated or if anti-plasmin levels are reduced, either due to reduced synthesis of anti-plasmin or increased.

Since late dosing has the apparently paradoxical effect of increasing the risk of death by bleeding if it is given more than three hours post-injury, we compared late doing with on-time dosing of TXA in the case of elevated tPA, where TXA appears to make a significant difference. To simulate the late dosing, we used the same dosing schedule, but shifted it forward in time three hours, so that the late dosed patient received his first dose of TXA four and half hours following injury. As seen in Figure 7.11, the late dosing of TXA very nearly resembles the case when no TXA was given, with a rapid decline of fibrin within the wound compartment (Figure 7.11a) and large dip in venous fibrinogen pool (Figure 7.11b). Although nearly as soon as the patient receives TXA, the amount of plasminogen is reduced to nearly zero, a significant amount of it has already been converted to plasminogen by the high concentrations of tPA present, which TXA does inhibit, but not before this plasminogen has already degraded the clot that formed in the wound compartment when the dosing of TXA is de-

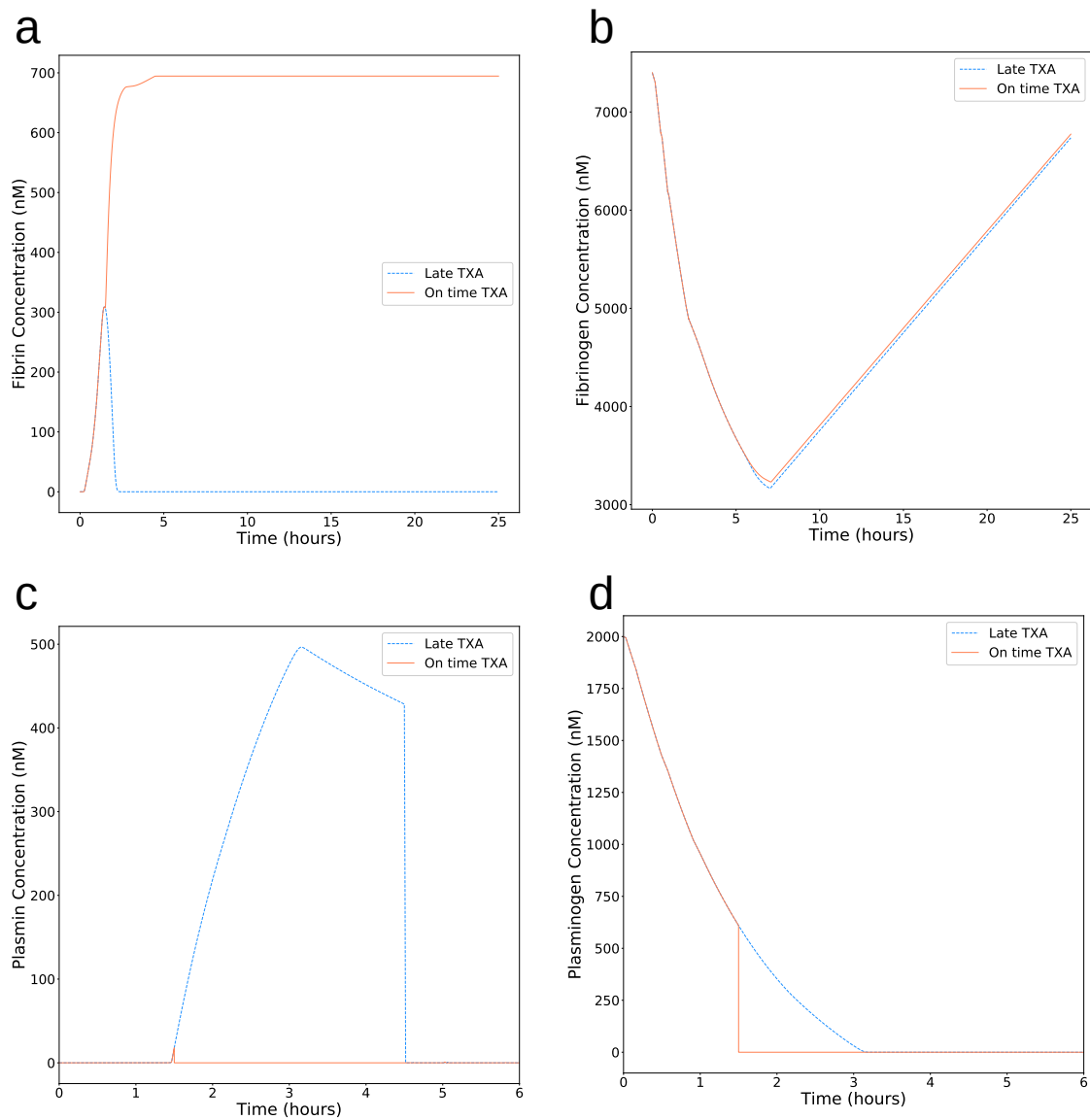


Figure 7.11: The effects of dosing a trauma patient with TXA over the first 25 hours following injury at time zero both with the initial tPA concentration set to ten times the nominal concentration. The dashed blue lines indicate a patient that received his first dose of TXA three hours and half hours after injury, and the solid orange line indicates a patient who received his first dose of TXA 90 minutes following injury. (a) Fibrin concentration within wound compartment (b) Fibrinogen concentration within vein compartment (c) plasmin concentration within vein compartment (d) Plasminogen concentration in the vein compartment



layed. However, based on our simulations, it is unclear why this would put the patient at a higher risk of death, because the late dosing appears to resemble the case when no TXA is given. A possible reason why our late dosed patient resembles the patient who did not receive TXA is the lack of uPA-TXA interactions within our model, and it is believed that uPA is responsible for the lack of TXA efficacy outside of the three hour window [365].

While these simulations highlight the potentially dramatic role TXA can play in altering the state of a patient's coagulation, they should be treated with caution. The rate constant for plasmin activation was estimated with a much higher concentration of tPA (2 nM) versus the approximately .07 nM that is found physiologically [1], so it is possible that there are significant errors in estimating that rate. Additionally, both the small kinetic model and the PBPK model lack rate constants for uPA related interactions, and these interactions may significantly contribute to TXA's efficacy. Both the kinetic model of coagulation and fibrinolysis and the PBPK model would be significantly strengthened by more information about how quickly uPA activates plasminogen (in a way similar to the tPA titrations that were used to estimate the tPA-plasminogen rate constants) and how quickly TXA activates uPA [190]. In a recent study, a correlation was found between peak tPA correlation and injury severity score [69], further highlighting why TXA may be so effective: in severely injured patients enough tPA is present to activate large amounts of plasminogen, so the addition of TXA can noticeably inhibit this activation. Much research remains to be done on how exactly TXA interacts with all portions of the coagulation system, and how quickly these side reactions occur. It should also be confirmed that TXA does not significantly effect synthesis of coagulation and immune related proteins, because if it does, that could partially explain some of its effectiveness.

## 7.9 Comparison of Massive Transfusion Protocols

A severely injured patient who is admitted to an emergency room will almost certainly require massive transfusion, however, what fluids are given will vary depending on which hospital to which they are admitted. We used our simulated person to examine the efficacy of two different massive transfusion protocols: one from Stanford and another from Parkland. Under the Stanford protocol, a patient receives blood upon admission, followed by packed red blood cells, fresh frozen plasma and platelets in a 6:4:1 ratio [271]. Under the Parkland protocol, a patient first receives packed red blood cells and plasma (in a 5:2 ratio), then the next set of fluids includes platelets, plasma, and packed blood cells (in a 1:2:5 ratio), and the third shipment includes packed red blood cells, plasma, and cryoprecipitate (in a 5:2:10 ratio) [248]. We simulated resuscitation with either one of these protocols following an initial in-transit resuscitation with 500 mL of saline, with in hospital resuscitation beginning 45 minutes after injury.

Initially, the Stanford protocol preserves far more of patient's clotting capacity, as these patients are receiving blood, which contains far less filler than segmented blood products (Figure 7.12). However, this initial advantage is lost around the third hour, when the Parkland patients receive a large dose of cryoprecipitate, which is a very potent source of coagulation factors which dramatically boosts the patients clotting capacity. These simulations show that the benefits of blood as a resuscitation fluid will be diluted once fluids that contain more filler are given, but that cryoprecipitate can make up for these dilute fluids being used. While we only compared two massive transfusion protocols, many more are used across the United States and the globe, and if a hospital wishes

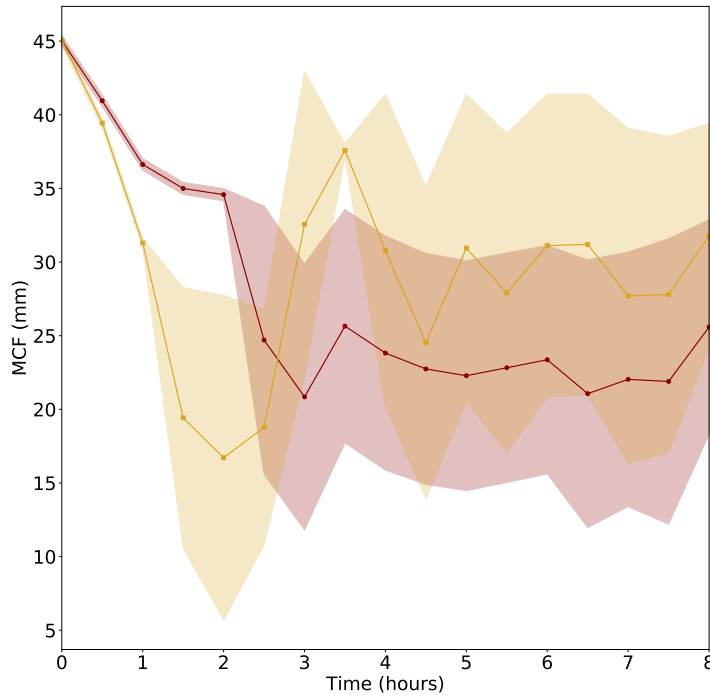


Figure 7.12: Simulated maximum ROTEM amplitude under Stanford (red) or Parkland (yellow) resuscitation schedules. The mean amplitude is shown as a point every 30 minutes, and the shaded area represents one standard deviation over  $N = 6$  parameter families.

to change their protocol, our model could be used as a first step in showing non-inferiority. This approach could be further validated by performing experiments as in [295], where they simulated treatment by diluting healthy people's blood and then added fibrinogen, FFP, FXIII, and Ringer's acetate, however, in some respects, our in-silico approach is more complete, as it includes the depletion of the available coagulation factors that would occur in trauma which is not captured in the use of diluted blood from healthy volunteers.

## 7.10 Conclusions

Our PBPK model makes it possible for treatments to be run in-silico before they are attempted in either an animal in the field or human patients. To our knowledge, this is the first PBPK model to deal combine a detailed description of coagulation with TXA, a drug that has shown promise in reducing deaths due to trauma. It also highlights the cautions that need to be taken when using animal models of trauma, as their acute phase responses may not correspond to human responses. We would encourage those who create animal models of trauma to adapt this model for their model animal so that we can better understand and model the differences between human and their model animal responses to trauma and treatment. Our model highlights that more work remains to be done in understanding how hypothermia and acidosis effect coagulation, however, it is clear that patients would benefit from having both of these conditions controlled as soon as is possible. As computers become more efficient, it is possible that this model could be used to solve a scarce resource distribution problem: how should blood or blood products of a limited amount be distributed in the case of mass casualties to best preserve clotting capacity?

CHAPTER 8

**PREDICTION OF ACUTE TRAUMATIC COAGULOPATHY VIA  
MACHINE LEARNING**

### **8.1 Machine Learning in Health Care**

Machine learning has been applied to numerous problems within health care, from the prediction of Alzheimer's from brain MRIs to the prediction of suicide risk in mental health patients from their electronic health records [215]. The idea of applying machine learning to health care is not new—in fact, machine learning was being applied to the problem of determining diagnostic rules for cancer before the term “machine learning” became widespread [32]. As electronic health care records become more widespread and commonly used, the pool of data potentially available for learning also increases. However, applying machine learning to the health care setting can be challenging due to missing or incomplete data, many of which have a temporal dimension but are still very highly dimensional. Furthermore, in many cases, the data is missing not at random, meaning that interpolation of the missing data must be done carefully. In order to apply many machine learning algorithms, the data must undergo extensive preprocessing to fill in missing data points or remove features that are missing for a certain percentage of the population. Despite these challenges, the use of machine learning in health care will continue to proliferate in the twenty-first century and hopefully guide clinicians to more successful outcomes. However, we must proceed cautiously to make sure that machine learning continues to be a boon, and does not reduce the skills of physicians, get taken out of context (especially when applied to textual data), and capture the inherent uncertainty

in dealing with complex systems, such as humans [44].

The intensive care unit represents a rich data source for machine learning, as the patients within it are closely monitored, producing time series data with a very high temporal resolution. This data can then be used to develop a time series classification framework, with decision trees [46], or a Weibull-COX proportional hazard model [239], or to predict mortality [9]. Vranas and coworkers used consensus clustering to see if machine learning was able to group patients by their similarity, and discovered it was successful in grouping the patients into human recognizable groups based on ICU admission data [347]. Eshelman and colleagues used the MIMIC II database to create a RIPER model, which uses a rules based method, to predict if a patient was at high risk of becoming hemodynamically unstable [100]. Using a Bayesian nonparameteric learning approach, Lehman and coworkers found that they could identify the healthiest and unhealthiest patients in the ICU based on extracted features from blood pressure time series data [183]. As an increasing amount of ICU data becomes of publicly available, machine learning will continue to flourish in this space.

Much work has been done to develop classification algorithms which use health care data to determine if a patient does or does not have a disease, however, the time dimension of these predictions is just starting to be explored [216]. Time series prediction is not a new field, but many of the time series prediction methods which have been traditionally used in economics assume that there exist a seasonality to the data, a pattern which repeats several times within the measured time period [233], which does not hold true for time series prediction over a span of days within an ICU, and therefore renders techniques like ARIMA and SARIMA ill-suited [251]. While there exists a collection of datasets

for benchmarking time series prediction techniques, none of the datasets within this collection, the only medical time series included are measured at very high frequency, and not on the scale of an hour to hours, which is the rate at which the results of laboratory tests would make it back to the treating physician, and this repository focuses on univariate prediction [250]. Gaussian process multivariate regression has been used to develop MedGP, an algorithm to predict several health conditions, including sepsis, however, they severely restricted the variables used for their predictions and only included patients who had measurements for their specified variables [59].

Sequence prediction can be divided into three types: many to many, many to one, and one to many. In many to many, many time series are used to predict many time series many outputs, in many to one, many inputs are used to predict one time series, and in one to many, one time series is used to predict many outputs. In this case, we need a many to many predictor, which can take in the high dimensional input that is an ICU patient and output a prediction of their state as a function of time. Although stage recurrent neural networks have shown promise in sequence prediction, at least when applied to financial data, they are a many to one technique and not suitable for application to populations of patients [257]. The GRU-ODE-Bayes technique, which combined gated recurrent units with the Fokker-Planck differential equation outperformed a number of other GRU-ODE methods as well as a variational autoencoder at predicting several medical time series, however, in this comparison, the data was preprocessed to severely reduce the number of features in the dataset [80].

## 8.2 Machine Learning in Trauma

A number of groups have been building machine learning models which aim to predict survival following traumatic injury and other trauma related metrics. In an early use of machine learning in trauma, Demsar and colleagues analysed which patients benefited from damage control surgery by growing classification trees and through the use of a naive Bayes classifier [84]. Through the use of SuperLearner, which combines several machine learning algorithms, Christie and co-workers were able to predict death as well as transfusion with area under the curve approaching .9, which indicates good accuracy and precision [67]. Using a different hybrid algorithm which combined basic detection rules with a multi-layer perceptron, Liu and coworkers were able to predict which patients would require a life-saving intervention of 89.8% accuracy [188]. Li and coworkers compared logistic regression and random forests for the prediction of ATC and found that both techniques were fairly successful in predicting ATC ( $F_1$  score > 90% for both techniques), however, it is unclear how they dealt with the time aspect of the data [182]. They also only included patients with complete data in their study, which could bias their results, as it is expected that there will be significant missingness in medical data, especially in an emergency setting, as not all data may be collected for patients who die soon after arrival. Talbert and coworkers have proposed to include machine learning in triage via CATT, however, the results that they report indicate that traditional triage may be more effective than their predictive model [327]. Machine learning has also been applied to predict outcomes and severity of traumatic brain injury, using multivariable logistic regression in children [66], a support vector machine to predict intracranial pressure [58], with a decision tree with features extracted



from CT images of the brain [122], just to list a few. A recent review of machine learning in trauma highlighted that artificial neural networks show promise in predicting mortality and other outcomes in trauma patients [189].

## **8.3 Data**

### **8.3.1 MIMIC-III**

To predict which patients were at high risk of ATC, we needed a large set of trauma patients and their health care records. The MIMIC-III database [163] contains deidentified health care data from approximately 60,000 intensive care unit admissions from the Beth Israel Deaconess Medical Center in Boston, Massachusetts. MIMIC contains a wealth of information, including treatments given, physiological information, lab results, and notes from attending health care professionals. We selected 3025 unique hospital admissions, based on the criteria in Figure 8.1.

#### **Pre-Processing**

We built a PostgreSQL database to store the MIMIC-III database and made all queries using this system [163]. We first limited our population to patients with traumatic injury. To construct our  $x$  vector, we concatenated all laboratory test values, chart event values, input events, and drugs prescribed. We next binned the data trying multiple window sizes (30 minutes, 1 hour, and 2 hours), where the first recorded event was defined as time  $t = 0$ , as shown in Figure 8.2. For

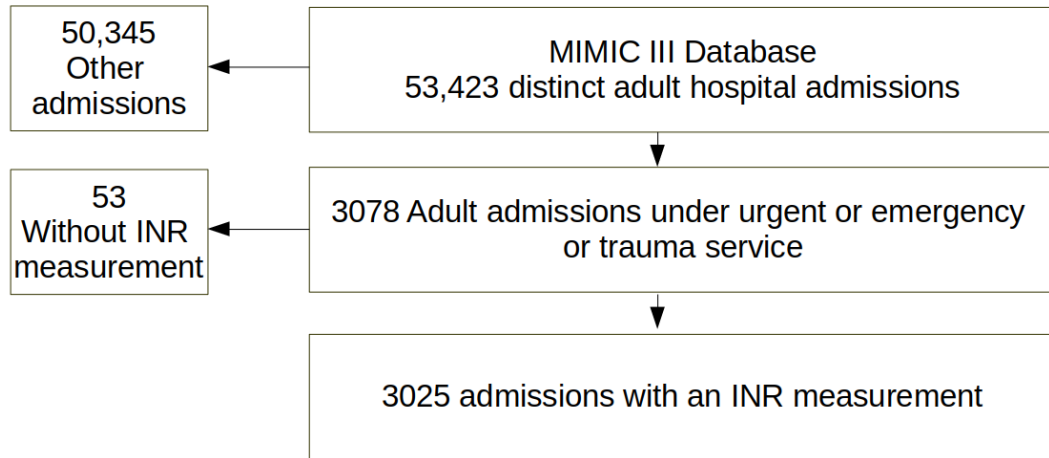


Figure 8.1: The details of the patient selection process. Beginning from all the adult patient admissions in MIMIC-III, we selected only patients who were in the ICU for unplanned reasons and had at least one INR measurement.

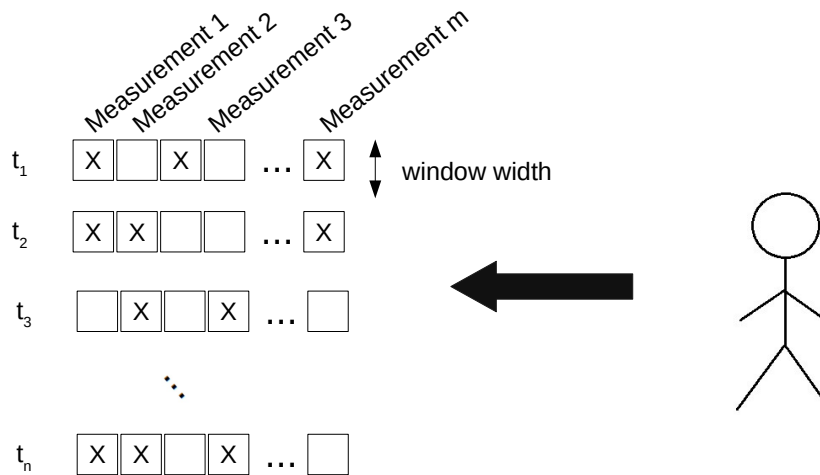


Figure 8.2: A graphical representation of how patient feature vectors were constructed. The Xs denote measurements that were available at time point  $t$ , and the empty boxes represent that that measurement was not available at a time point for a patient. This missingness pattern was used to construct masks to describe the missingness of data.

laboratory and chart events, if more than one event was recorded within a window, we computed the average value. For drugs, if more than one administration was recorded within the window, we took the sum of all of the doses within a window. During data preparation, we also constructed additional masks, as described by [57], which recorded if a measurement was present within a window, and if so, at what time it was recorded. With these masks, we aimed to capture the missingness of the data, which may contain additional information, since which tests have been performed or not have been extremely predictive [185]. We constructed our  $y$  vector based on internationalized normal ratio (INR) tests, which is a measure of the body's ability to form blood clots. The normal value for a healthy person is 1.0. Specifically, we encoded patients with an  $\text{INR} \geq 1.5$  as +1 (having ATC) and the patients with an  $\text{INR} < 1.5$  as -1 (not having ATC). While pre-processing the data, we also removed tests that are sometimes used interchangeably with INR to diagnose ATC, including PTT, aPTT, and PT. To deal with the missingness of the data, we used linear interpolation to fill values that were missing in the time series.

### **8.3.2 Activation of Coagulation and Inflammation in Trauma Study Dataset**

The Activation of Coagulation and Inflammation (ACIT) dataset was collected between February 2005 and April 2015 originally collected for a single-center prospective cohort study of severely injured trauma patients [67]. Patients were included in the dataset if they were over the age of 15, not pregnant, not incarcerated, and not transferred from another hospital. The dataset contains clinical

and laboratory measurements collected at admission as well as 2, 3, 4, 6, 12, 24, 48, 72, 96, and 120 hours after injury. However, the number of INR measurements in this dataset declines dramatically after admission, as at admission, 1380 patients have an INR measurement, whereas the next highest number of INR measurements are recorded at 6 hours post admission, but only 228 patients have an INR measurement at a timepoint. Thus, this dataset is much less rich than MIMIC in terms measurements over time.

### **Pre-Processing**

We characterized the data within the ACIT dataset into 4 categories: demographic (items that would not evolve with time), hourly measurements, daily measurements, and summary measurements (measurements that summarized, for example, the amount of fluids given over a 24 hours period). Since only the upon admission timepoint had a large number of INR measurements, we decided to focus our analysis upon the 1380 patients who had an admission INR measurement. Since there were very few patients with complete data, we filled missing numerical data with the mean of the patients who at a measurement, and categorical data with the most common case within the category. We then removed all columns containing long blocks of text, and rescaled the data. For our admission classification of ATC, we included the demographic data, the 0 hour measurements, the admission day measurements, and the 0 to 24 hour summary measurements.

## 8.4 Classification via Logistic Regression

Logistic regression allows for the separation of data into two groups, which allows us to divide patients into groups based on those who never have ATC during the entirety of their ICU visit and those that do have an INR that classifies them as having ATC during their visit. Formally, logistic regression corresponds to the model

$$\log\left(\frac{p(x)}{1 - p(x)}\right) = \beta_0 + x \cdot \beta \quad (8.1)$$

which can be solved for  $p(x)$ , giving

$$p(x, b, w) = \frac{e^{\beta_0 + x \cdot \beta}}{1 + e^{\beta_0 + x \cdot \beta}} = \frac{1}{1 + e^{-(\beta_0 + x \cdot \beta)}} \quad (8.2)$$

We can then classify patients based on Equation 8.2, where they are in class 1 (having ATC) if  $p \geq .5$ , and in class 0 (not having ATC) otherwise. While this technique is simple and implemented in `sklearn.linear_model.LogisticRegression`, it can only separate the data if a separating hyperplane exists, meaning that the data must be linearly separable. We can also see which variables are important to the classification process by looking at which are more heavily weighted, if logistic regression does a good job of classifying patients. We found that logistic regression was not very successful at separating patients who had ATC from those that did not, with logistic regression performing on par with just classifying by the base rate, as seen in Table 8.1. This poor performance is obvious when we examine the  $F_1$  scores, none of which exceed .2. The  $F_1$  score is the harmonic mean of precision and recall, such that  $F_1 = \frac{2(\text{precision} \cdot \text{recall})}{\text{precision} + \text{recall}}$ , and as such incorporates the true positive, false negative, false positive, and false negative rates. Another downside of this technique is that it is incapable of dealing with time series data

Window Width	Base Rate	Accuracy Train	Accuracy Test	$F_1$ Score
30	16.9%	84.2%	79.4%	.111
60	20.8%	83.6%	72.6%	.109
120	21.9%	81.3%	76.0%	.131

Table 8.1: Accuracy of logistic regression for predicting ATC in the MIMIC-III dataset as a function of bin width. While the accuracy of the training and test sets are far above the base rate, one can observe they are very close to 1-base rate, which means if we simply predicted a patient’s likelihood of having ATC at the first time point at the rate 1-base rate, we would perform nearly as well as logistic regression.

of different lengths or in dealing with the time dimension of data, so we chose to apply it to only the first time point.

Although logistic regression did not perform well on the MIMIC-III dataset, we wanted to see if this technique could be successful on a dataset more focused on trauma, such as ACIT. The ACIT dataset had another potential advantage—it was already much smaller in terms of features, with fewer than 500 features compared to several thousand in MIMIC. However, logistic regression also performed poorly on this dataset, since the base rate was 91.0% (where base rate is the percentage of patients without ATC), and the test accuracy was 91.6%, a negligible improvement over the base rate. The  $F_1$  score was also poor, coming in at an average of .566 using five fold cross validation. We also found it difficult to reduce the dimensionality of the data while preserving its richness, with the first two components principal component analysis only explaining 24% of the variance.

## 8.5 Classification via Support Vector Machine

Since logistic regression performed poorly, we turned to support vector machines. In their most basic forms, support vector machines seek to find a separating hyperplane by solving

$$\begin{aligned} \min_{\|\mathbf{w}\|} \\ \text{s.t. } y_i(\mathbf{w} \cdot (x_i + b)) \geq 1 \end{aligned} \quad (8.3)$$

for all  $i = 1, 2, \dots, n$  over  $n$  different datapoints. Solving this optimization problem means that points are correctly classified: if  $y_i = 1$ , things will be in the class where  $\mathbf{w} \cdot x_i + b \geq 0$  and if  $y_i = -1$ , things will be in the class where  $\mathbf{w} \cdot x_i + b < 0$ , where 1 and -1 represent the two classes. However, this form is only successful if we can create a separating hyperplane. To deal with the possibility that there is not a separating hyperplane, we can pose a slightly different problem:

$$\begin{aligned} \min_{\mathbf{w}, \xi, b} \frac{1}{2} \mathbf{w} \cdot \mathbf{w} + C \sum_{i=1}^n \xi_i \\ \text{s.t. } y_i(\mathbf{w} \cdot (x_i + b)) \geq 1 - \xi_i \\ \xi_i \geq 0 \end{aligned} \quad (8.4)$$

for all  $i = 1, 2, \dots, n$ . This formulation introduces the slack variables  $\xi_i$ , and  $C$ , which controls the trade off between margin (how far away the points are from the separating hyperplane) and the training error. The larger  $C$  is, the more the model is penalized for having points on the wrong side of the hyperplane. To move even further away from requiring linear separability, we can use a

kernel<sup>1</sup> (a mathematical function with some nice properties) and take advantage of strong duality<sup>2</sup> to reformulate the problem yet again. Starting with the dual form, we solve

$$\begin{aligned} \max D(\alpha) &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j (\mathbf{x}_i \cdot \mathbf{x}_j) \\ \text{s.t.} \quad &\sum_{i=1}^n y_i \alpha_i = 0 \\ &0 \leq \alpha_i \leq C \end{aligned} \tag{8.5}$$

where  $\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i$

Now, solving this problem only requires the evaluation of the dot product between  $\mathbf{x}_i$  and  $\mathbf{x}_j$ , and we replace this dot product with the calculation of the kernel between these vectors,  $K(\mathbf{x}_i, \mathbf{x}_j)$ .

$$\begin{aligned} \max D(\alpha) &= \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j) \\ \text{s.t.} \quad &\sum_{i=1}^n y_i \alpha_i = 0 \\ &0 \leq \alpha_i \leq C \end{aligned} \tag{8.6}$$

The use of a kernel allows us to find other separating hypersurfaces other than just a hyperplane. This is especially important for ATC, since it appears that this data is not linearly separable. We investigated classification of a subsample of 500 of the MIMIC-III patients at the first timepoint using a support vector machine with either a linear, polynomial, radial basis function, or sig-

---

<sup>1</sup>One definition of a kernel: Let  $X$  be a non-empty set. A function is a valid kernel in  $X$  if for all  $n$  and all  $x_1, \dots, x_n$  in  $X$ , it produces a Gram matrix that is symmetric and positive semi-definite

<sup>2</sup>This is a convex problem, so strong duality holds, so the duality gap will be zero.



Dataset	Linear	Polynomial (degree 3)	Radial Basis Function	Sigmoid
MIMIC-III Bin Width 30	0.273	0.198	*	0.119
MIMIC-III Bin Width 60	0.387	0.267	0.153	0.037
MIMIC-III Bin Width 120	0.374	0.35	0.108	0.162
ACIT	0.526	0.571	*	*

Table 8.2: The calculated  $F_1$  scores using a SVM on a sample of 500 MIMIC-III patients or ACIT patients at the first timepoint with various kernel types. A \* denotes an ill-defined  $F_1$  score as no patients were predicted to be in one of the classes. The  $F_1$  scores reported are an average over five fold cross validation.

moid kernel.

Overall, support vector machines, regardless of kernel or bin width used performed poorly on classifying patients as ATC or non-ATC at the first timepoint (Table 8.2). A SVM using a 3<sup>rd</sup> degree polynomial performed the best overall, however, this performance comes with a  $F_1$  score of only .571, corresponding to a recall of .62 and a precision of .6, neither of which are high enough to make this type of prediction clinically useful. Overall, SVMs provided a slight boost in performance over logistic regression, but not enough of a boost to serve as a useful predictor or classifier.

## 8.6 Classification via Recurrent Neural Network

Recurrent neural networks (RNNs) offer promise when applied to health care, as they have previously been successful in learning representations of sequential data, such as speech [124]. RNNs have demonstrated success when applied to electronic health care records, such in DoctorAI, for the prediction of diagnosis and medication order, however, in this application, the time windows for data binning are much wider than what would be necessary in an ICU setting, on the order of weeks or months versus hours [64]. A RNN outper-

formed a number of machine learning techniques, including logistic regression and a support vector machine at predicting heart failure at an eighteen month time window [65]. On the time scale of days, long short-term memory (LSTM) RNNs have shown promise in predicting sepsis as well as myocardial infarction [165].

RNNs consist of nodes, arranged in layers, which pass information from layer to layer through connections. Each node applies a non-linear function to the inputs it receives and the weights that is aiming to learn before passing this value to the next layer. A unit with a LSTM RNN traditionally has three gates: the input gate, the output gate, and the forgetting gate [125]. Mathematically, the forgetting gate calculates

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (8.7)$$

where  $x_t$  is our feature vector at time  $t$ ,  $h_{t-1}$  is the output at time  $t - 1$ ,  $W_f$  are the weights associated with the forgetting unit, and  $b_f$  is the bias associated with the forgetting unit, and  $\sigma$  is the activation function of the forgetting gate. The input gate calculates

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (8.8)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (8.9)$$

These are combined to calculate  $C_t$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (8.10)$$

Finally, all of this information is passed to the output gate which calculates

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (8.11)$$

$$h_t = o_t * \tanh(C_t) \quad (8.12)$$

which is fed into the next layer of the RNN.

### 8.6.1 Prediction of Acute Traumatic Coagulopathy with Recurrent Neural Network

We applied a single-layer GRU for binary classification of ATC, as developed in 2014 by [63]. Mathematically, this unit calculates

$$z_t = \sigma(W_Z \cdot [h_{t-1}, x_t]) \quad (8.13)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (8.14)$$

$$\tilde{h} = \tanh(W \cdot [r_t * h_{t-1}, x_t]) \quad (8.15)$$

and then outputs

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (8.16)$$

While this is similar to the “vanilla” LSTM, the main difference is that the forget and input gates have been combined into a single update gate.

Our grid search consisted of different methods both to prepare the data and train the model. We tested different learning rates, including 1e-3, 1e-4, and 1e-5 and used the Adam optimizer. We tested hidden layer initialization using either samples drawn from a normal distribution or zeros. Hidden layer sizes tested included 32, 64, and 128 units. We used a linear classifier atop the final hidden representation, with the binary prediction defined as  $\text{sign}(x)$ . We used soft margin loss as implemented in `PyTorch`. Batch normalization was used. We supplemented our data with additional features so that the neural network could become aware of patterns within the missingness of the data: we created a vector of masks and time-since-last measurement as described by Che et al. and concatenated this to the normalized features to define a feature vector.

We found that throughout the parameters tested in the grid search, the increase in accuracy over the baseline accuracy was poor, however, we found that accuracy improved dramatically if we fed the weights of the last layer of our neural network into a support vector machine and then used those weights to predict into which class the patients fell. Since there was variation in the base rate (what percentage of the patients had ATC during their stay) depending on the method of data preparation used (depending on the minimum sequence length required for training, some patients were excluded), we used accuracy over base (AOB) as our metric to compare different metaparameters, where

$$\text{AOB} = (\text{Accuracy of chained LSTM and SVM}) - \text{fraction of patients with ATC in the current dataset} \quad (8.17)$$

as well as the  $F_1$  score.

As of yet, the features being fed into our machine learning techniques came from the test values, chart event values, input events, and drugs tables within MIMIC-III, and we had done nothing with the text that was available in MIMIC in the noteevents table. We needed a way to convert these notes into features that could also be fed into a LSTM, so we turned to Latent Dirichlet Allocation (LDA), a technique based on a generative model of a corpus (a group of documents). LDA assumes that each document within the corpus is a random mixture of latent topics, where each topic is characterized by a distribution of words [30]. With LDA, we can take a corpus, and learn the topics which compose it, and then use the probabilities assigned to these topics as feature to describe a document within the corpus.

To perform LDA on the free text in MIMIC-III, we once again binned the notes based on when they were entered, and if there were multiple notes within the same time window, they were concatenated together. We then lemmatized

the words with `WordNetLemmatizer`, and added both bigrams and trigrams that appeared at least 20 times within the corpus to our vocabulary, so that LDA would capture terms like "100 bpm" or "70 mL saline". We then filtered out words that were either too rare (they appeared in fewer than 20 documents) or overly common (appearing in more than 50% of documents). After creating tri- and bi-grams, stand alone numbers were removed from the documents. Using default parameters to perform LDA with 50 topics, and adding these 50 features as our input to the chained LSTM-SVM boosted our accuracy, in the best case to 90.3%, or an accuracy over base of .364. This result, in which adding text data boosted the accuracy of a machine learning method is agreement with the results of Weissman and colleagues, who found that adding text data boosted their model's discrimination [356].

In order to determine the best meta-parameters for our binary classification system, we used SMAC, a tool which can optimize the parameters or hyper-parameters of any algorithm [154]. SMAC, which stands for sequential model-based algorithm configuration, which constructs a random forest model to predict algorithm performance as it searches through parameter space, and then uses that constructed model to select promising parameters and then compares each parameter configuration to the incumbent parameter set which has had the best performance so far. We can then use CAVE to evaluate the importance each of the meta-parameters to the error of the model [29, 184] using functional ANOVA (fANOVA). fANOVA decomposes a function  $\hat{y}(\theta) = \sum_{U \subseteq N} \hat{f}_U(\theta_U)$  so that we have decomposed  $\hat{y}: \theta_1 \times \dots \times \theta_n \implies \mathbb{R}$  into additive components that only depend on a subset of the  $N$  inputs [147], where  $\hat{y}$  can be the error of the model. Through the calculation of the variance of  $\hat{y}$  across its domain  $\theta$ , the importance of each of the parameters and interactions between the parameters can be quan-

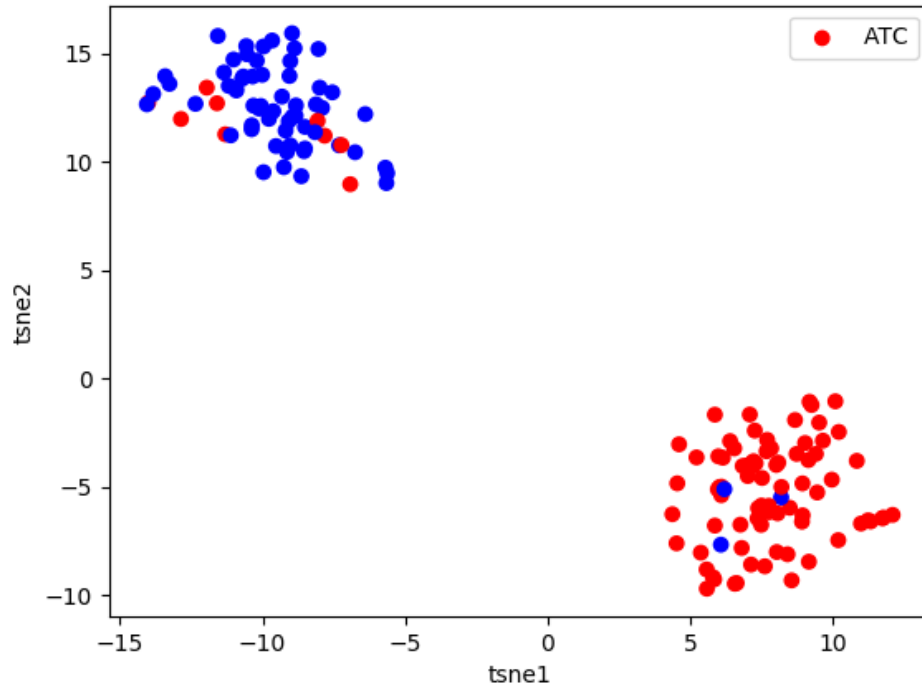


Figure 8.3: A low dimensional representation of the weights of the LSTM via tSNEs which was fairly successful at separating the patients into two groups. The red dots represent patients who developed ATC at some point within their stay, the blue ones did not.

tified. So that we did not bias our fANOVA results by discarding configurations which took a lot of memory to run, we limited the number of patients included in each sample to 500 if more than 500 fit the selected metaparameters.

The metaparameter which contributes the most to the variance of the chained LSTM-SVM model is the padding method used, which controls if the sequences that met the minimum sequence length were either padded to the length of the median of all the sequences or to the length of the longest sequence selected, (Table 8.3). Of all of the options of additional information to add to the model, the text features appear to contribute the most to the variation of the model performance, followed by the use of the masks as in [57], and finally, the

Metaparameter	Importance
pad_method	20.9561 +/- 3.356
use_text_features	4.5818 +/- 3.2456
min_seq_len	2.5183 +/- 1.508
lr	1.2108 +/- 1.0397
use_masks	1.107 +/- 0.7551
hidden_size	0.7842 +/- 0.8512
kernel	0.6223 +/- 1.3248
window_width	0.6046 +/- 0.5643
rare_feat_thresh	0.4409 +/- 0.474
hidden_init_type	0.2637 +/- 0.3021
use_drugs	0.2357 +/- 0.4659

Table 8.3: The estimated importance of the metaparameters in our chained LSTM-SVM model,  $\pm$  one standard deviation, over 1000 metaparameter combinations.

Metaparameter	Best Value
hidden_init_type	ZEROS
hidden_size	47
kernel	sigmoid
lr	0.00719031
pad_method	LONGEST
rare_feat_thresh	0.2875
use_masks	False
use_text_features	True
min_seq_len	7
use_drugs	False
window_width	120

Table 8.4: The best metaparameters found via SMAC for binary classification of MIMIC-III patients using a chained LSTM-SVM via minimizing  $1-F_1$  score, resulting in an average  $F_1$  score of .894 over five fold cross validation.

use of information from the drugs portion of MIMIC-III. The fact that the drugs information did not contribute very much to variance in the model performance suggests that at the initial time point, the drugs that a patient has received do not contribute very much to ATC, perhaps because they have not had time to become distributed throughout the body or have not had sufficient time to alter the state of the coagulation system.

Although the use of drug information only contributed a small portion to the variance of the performance of the chained LSTM-SVM model, they were not included in the overall best performing model identified by SMAC (Table 8.4). The widest bin width was also selected, which suggests that when just trying to classify patients, how frequently measurements are collected does not play a key role in producing accurate predictions. The rare feature threshold selected is rather high-around .3, indicating that it is helpful to remove a large number of features, and that if features are sparsely populated, they do not help boost model performance very much. The separation between ATC and non-ATC patients can be seen in Figure 8.3, where after performing PCA, most of the patients in the ATC and non-ATC groups cluster together. In future analysis, it may be interesting to more closely examine the patients which are misclassified to determine in which dimensions they are outliers.

## 8.7 Sequence Prediction

While being able to predict which patients develop ATC throughout their hospital stay is a step in the right direction, it would be far more useful for clinicians to be able to look at the patient's chart up to that time point, and then be able to predict if the patient was at high risk of developing ATC in the near future, the next day, or in the next time step. This type of problem can be posed as a sequence prediction problem-given a time series of features, can we predict not only the next time step, but several time steps into the future?

Sequence prediction is a class of techniques that aim to do more than just perform static classification, rather, they aim to use the time-series nature of



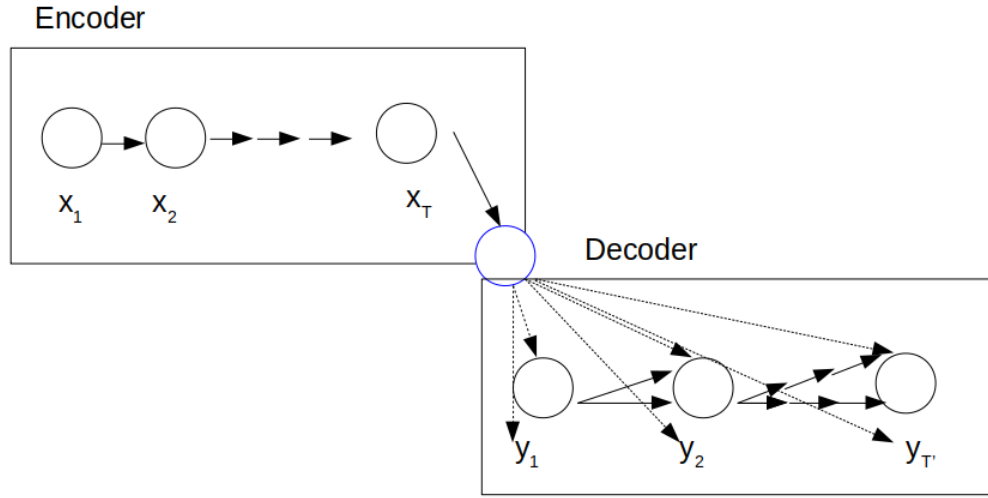


Figure 8.4: A generic schematic of an encoder-decoder model, with the low dimensional representation shown in blue. The inputs are  $x_i$ , and the outputs are  $y_i$  for timepoints  $1, 2, \dots, T$  where  $T$  is the length of the sequence.

the data and the patterns that occur as a function of time to predict the future of a variable, given its past and other related time-series measurements. Deep Neural Networks (DNNs) can be joined together in an encoder-decoder fashion (where one DNN performs the encoding step, and the other performs the decoding step) to perform machine translation, and this set-up allows for the production of a low dimensional representation of the input data, which is potentially human understandable [324]. We modelled our sequence prediction encoder-decoder after one of the most famous architectures-seq2seq, in which LSTMs are used as both the encoder and decoder models [324]. In our case, we fed in the features extracted from MIMIC-III that describe patients and output a sequence of INR values, which we could use to determine if a patient was likely to have ATC at a given timepoint.

Looking at the best metaparmeter results found via SMAC (Table 8.5), it is not surprising to note that our model had the best performance when it was

Parameter	Best Value
hidden_size	318
lambda_r	9.66081e-07
lr	0.0056
min_seq_len	6
num_steps_to_predict	2
pad_method	MEDIAN
rare_feat_thresh	0.167
use_drugs	false
use_masks	false
use_text_features	false
window_width	30
x_fill_method	nominal

Table 8.5: Best metaparameters found after running SMAC for 908 configurations.

Parameter	Estimated Importance via fANOVA
num_steps_to_predict	23.85 +/- 7.40
min_seq_len	24.10 +/- 7.69
use_drugs	00.02 +/- 0.07
use_masks	00.01 +/- 0.05
lr	00.90 +/- 1.80
window_width	00.17 +/- 0.40
x_fill_method	00.13 +/- 0.26
rare_feat_thresh	00.36 +/- 0.63
lambda_r	00.13 +/- 0.22

Table 8.6: Single metaparameter importance as estimated by fANOVA via CAVE for sequence prediction after running SMAC on 908 model configurations with a mean squared error of .26.

only asked to predict only two time steps forward in the future. It is interesting to note that the model performed best with the narrowest bin size, 30 minutes. This is possibly because most of these bins would have had a unique measurement in them, unlike the bin width of 120 minutes, many of which would have in many cases the average of more than one measurement in it. It may be worth exploring to see if model performance would be even better with narrower bins. The rare feature threshold found to be the best was around 18%, a fairly low bar,

Pairwise Parameters	Estimated Importance via fANOVA
min_seq_len & num_steps_to_predict	29.3379 +/- 7.837
num_steps_to_predict & lr	1.0465 +/- 1.5091
min_seq_len & lr	1.0182 +/- 1.0597
hidden_size & lr	0.8743 +/- 3.3682
min_seq_len & hidden_size	0.7967 +/- 0.9077
num_steps_to_predict & hidden_size	0.6171 +/- 0.9414
min_seq_len & rare_feat_thresh	0.447 +/- 0.8691
num_steps_to_predict & rare_feat_thresh	0.3524 +/- 0.764
hidden_size & rare_feat_thresh	0.1576 +/- 0.3914
lr & rare_feat_thresh	0.0866 +/- 0.188

Table 8.7: Pairwise metaparameter importance as estimated by fANOVA for sequence prediction via CAVE after running SMAC.

meaning that features were included if they were missing in about 82% of patients. It is also interesting to note in this case that adding the text features didn't improve the accuracy of the sequence prediction, possibly because they are not entered on the same timescale as ATC. We believe this is one of the first cases to investigate if adding text based features boosted sequence prediction performance in an emergency medical setting, in which the time scale is in hours, not weeks or months.

From our fANOVA results (Table 8.6), it appears that the level that the rare feature threshold is set at plays a large role in the variance of our predictive model. A higher rare feature threshold removes more features from our  $X$  array, as it requires that this feature be present in at least this percentage of cases, so by changing the rare feature threshold, features that only have measurements for a few patients are discarded. This tuning appears to be important in our sequence to sequence model, as does the learning rate, the rate at which the parameters of the model are updated during training. This parameter can be thought of as analogous to the temperature in simulated annealing or the step size in Euler integration: if it is too small, you may not get anywhere within a

reasonable amount of time, but if it is too large, you may skip over possibly good areas of parameter space, and therefore, it requires careful tuning. Interestingly, it appears that the window width that we used to bin the patient data had a minimal effect on the variance of the model, however, the window width behind the scenes determines the minimum sequence length that can be produced for a patient, so it is difficult to untangle this effect.

If we examine the pairwise metaparameter contributions to the variance of the model (Table 8.7), we can see that together, the minimum sequence length and the number of steps into the future to predict contribute a good deal. Both of these parameters are indirectly tied to how long a patient's sequence is, which is a function of how wide the bins are, which further highlights that perhaps the effect of bin width should be investigated in a synthetic dataset, where sequence length is held constant and bin widths are varied with different frequency data to tease out if there is an ideal bin width.

In the ideal case, clinicians would be able to predict who is at high risk of developing ATC from measurements that are already taken commonly, frequently, and non-invasively from patients in an ICU setting, such as blood pressure and oxygen saturation. Right now, we pulled the majority of the features used in our model directly from the events that were charted for a patient in the ICU, but if we were to focus on readily available measurements which are taken at high frequency (every minute or more frequently), we may want to create artificial features, such as the rate of change of oxygen saturation over an interval, or the absolute difference between the maximum and minimum mean arterial pressure measured within an interval, as these constructed features may contain additional information relevant to the progression of ATC. Unfortunately,

of our approximately 3,000 trauma patients with an INR measurement, fewer than 40 of them have matched waveform data available in the MIMIC-III Waveform Database Matched Subset [120], giving us too small of a sample to peruse this question, however, if more trauma patients made it into the matched cohort, this would be an avenue worth perusing. In future work in predicting acute traumatic coagulopathy, practitioners must take care to include a representative sample of patients while training their classifiers or predictors, as by removing patients with missing measurements, they are likely removing the more severely injured patients who died soon after admission, perhaps of ATC or related complications, and not including these patients would result in a biased classifier that would be unlikely to perform well if applied in real time in an emergency care setting.

## CHAPTER 9

### CONCLUSIONS

This work presents the development of a multiscale model of the human body which can be applied to the investigation of ATC as well as data-science prospective on the topic. The multiscale model consists of a validated model of coagulation and fibrinolysis, which is then embedded into a PBPK model, which was validated against both human and rhesus macaque data. This multiscale model allowed us to investigate a number of topics that are currently debated in the field of trauma, among which, the ideal resuscitation fluid, the impact of hypothermia and acidosis, and the impact of TXA on a trauma patient. While we only investigated one drug, our model could easily be extended to include any coagulation altering drug with known kinetics and distribution patterns.

While much work has been done in describing the molecular interactions between proteins and molecules involved in coagulation and fibrinolysis and animal models have been created to simulate ATC, very few models have been created to link the two levels, a shortcoming that this work aims to address. The human body is a complex system, and in formulating a mathematical model of it, we have considerably simplified it and potentially removed details that may be later revealed to play key roles in the development and resolution of ATC. Within this model, there is no definition of fatality, and so long as the volume of none of the compartments goes to zero, the simulation will continue, regardless if in reality, the patient would have died due to shock or other causes. Transcapillary refill is assumed to occur at a constant rate, when in reality, this rate will be effected by the osmality of the given resuscitation fluids. Trauma is known

to raise plasma levels of epinephrine and norepinephrine [368], which lead to vasoconstriction [235], however, our model does not include the vasoconstrictive effects of these chemicals nor any others that are known to alter vascular tone. While platelets are included within the coagulation model, we have not considered that platelets may decline in effectiveness due to platelet exhaustion [18]. Our model does not take into account any of the metabolic changes that may be occurring at either the cellular or organism level following trauma, which may alter organ perfusion or heart rate. We have not modelled any of the complex signalling that may occur after trauma which leads to the change in production of coagulation related proteins, rather, they have just been assigned a lag time and an acute synthesis rate. I believe that an understanding of how these feedback systems work will greatly augment our understanding of the development of ATC and perhaps allow us to identify which patients are at high risk of developing it before injury occurs.

At the present, the term ATC refers to any clinically observed impairment of the coagulation system following injury, but since this system has many points of feedback, some redundant mechanisms, and many points of cross-talk with other biological cascades, I believe that there are many ways in which the coagulation system can break, many of which then fall under the diagnosis of ATC. If, for instance, a patient has both reduced levels of a procoagulant protein and an anticoagulant protein, due to genetic variation, they may live their entire life unaware of this due to the competing effects of these two alterations. However, in the case of injury, if a point of the feedback system which instructs the liver to produce more of a certain factor is broken or signaling goes awry within the endothelium, resulting in the dumping of large amounts of tPA hours after injury, both of these signal transduction errors may result in ATC, although for differ-

ent reasons. We are just beginning to suss out some of the different dysfunctional states of the coagulation system that may exist through the progression of ATC, from hypofibrinolysis to fibrinolysis shutdown to hyperfibrinolysis, but our understanding of what leads to these states in some patients is very shallow [224]. While these signalling pathways may be difficult to untangle in humans, chips with several organoids on them could allow investigation into the key players of these pathways in a non-invasive and easily tunable manner. Using organoids grown from different animals would potentially allow us to determine if these signalling pathways are common to all mammals, or if they vary from species to species, which would inform future work using animal models of traumatic injury. In-vitro models of trauma exist [88, 89], so as a first step to building the gene regulatory network of trauma, we can use these existing model systems along with RNA-seq to construct a preliminary map of what is occurring in one tissue type [353].

Our work highlighted that efficacy of whole blood over component therapy and showed that a reduction in reaction rates, whether due to hypothermia, acidosis, or other factors reduces the clot formation enhancing effects of whole blood, however, we did not determine what the optimal dosing of components would be, if whole blood was unavailable. These simulations could be strengthened with data showing exactly how each coagulation factor's activity is effected by both a reduction in temperature, a drop in pH, and both of these perturbations, both in in plasma (no platelets) and in whole blood, where platelets may also be effected by these changes. However, it appears that minimizing the amount of non-coagulation active fluid (fillers) may be a good rule of thumb to preserve clotting capability. It may prove informative to explore with this model if there are any cases where it may be better to give no resuscitation



fluid rather than component therapy owing to its dilutional effects. These simulations could guide the practice of permissive hypotension. While most hospitals have a massive transfusion protocol in place, these protocols vary from hospital to hospital [334], so our model could be use to compare the outcomes of differing massive transfusion protocols and could be used to asses the cost-benefit ratio of changes to an already existing massive transfusion protocol or as a first step in demonstrating non-inferiority.

We simulated the effects of one drug that is known to interact with fibrinolysis-tranexamic acid, however, there are a number of coagulation interfacing drugs available on the market, as well as recombinant coagulation factors which are traditionally proscribed for treating hemophilia, but which could perhaps be used to alter the progression of ATC [326]. Our PBPK model could be expanded to simulate the use of these recombinant coagulation factors or to simulate drug interactions before clinical trials were undertaken to estimate effect sizes and recruit an appropriately sized population in order to get statistically sound outcomes.

In summary, trauma is a significant perturbation to a number of the body's systems, including the coagulation system. If a person, either due to genetic or environmental factors, is robust to this perturbation, they will not go on to develop ATC, and will recover from their injuries. If a person's coagulation/immune systems are susceptible to this perturbation, they will go on to develop ATC. The shape of this perturbation may determine which point of control fails leading to the development of ATC, as their are many possibilities for uncorrected dysregulation which could lead to ATC, in various forms: disseminated intravascular coagulation, hyperfibrinolysis, or fibrinolysis shutdown, or

other modalities which have yet to be characterized.

The data science portion of this work revealed that ATC can be discerned in trauma patients without looking at traditional coagulation assays through the application of machine learning. We found that we were better able to separate patients with ATC from patients who never developed ATC during their hospital stay with an accuracy of greater than ninety percent using a LSTM as an input to a SVM, and that the addition of features constructed from the free text data boosted our accuracy. Our text features were constructed using latent dirichlet allocation over 50 topics, however, it is possible that accuracy could be boosted even further through the use of another technique to construct features from the text, such as ELMo [253], ULMFiT [149], or BERT [86]. It might prove interesting to try to predict ATC just using text based data and make a comparison between all of these natural language processing techniques, as a surprising amount of numerical information about the patient makes it into the doctor's notes.

Although logistic regression is a powerful technique for separating data into groups, it failed at separating patients with ATC from those without ATC, with two different datasets, which indicates that this data is not linearly separable nor separable by a hyperplane, and that non-linear techniques will probably be more successful in the future, such as using a support vector machine with a kernel. This lack of linear separability partially explains why the prediction of ATC is a difficult problem.

We found that using a neural network in combination with a support vector machine is effective in separating patients who never develop ATC from those do. Its possible that this combination of techniques is successful because

the neural network highlights which features are useful, and then the support vector machine can make use of these features to separate the patients into two groups. We also had some success with the sequence prediction problem, but much validation could be needed before this type of method could be deployed in an emergency room setting.

Although we did not explore the direct application of machine learning to discovering the mechanisms behind ATC, it might prove illuminating to use a human understandable technique to see what features this technique finds important. One possible way of approaching this problem would be to try to back calculate which factor concentrations lead to a ROTEM curve. If this approach is successful, it would allow clinicians to quickly identify how a patient's coagulation system has misfired and treat that specific mode of dysfunction to treat ATC.

To conclude, it is my belief that ATC is a manifestation of one or more ways that the coagulation system can misfire following a large perturbation, such as injury. Treatments that are broad (giving blood or cryoprecipitate) are likely to be successful, as they contain a large amount of several coagulation related proteins, allowing for treatment of many possible dysfunctions. Machine learning, if applied carefully, shows potential in separating patients with ATC from those without, and could potentially allow for faster diagnosis and treatment.

## CHAPTER 10

### APPENDIX

#### Additional file 10.1-(Data fits for CHO metabolism problem)

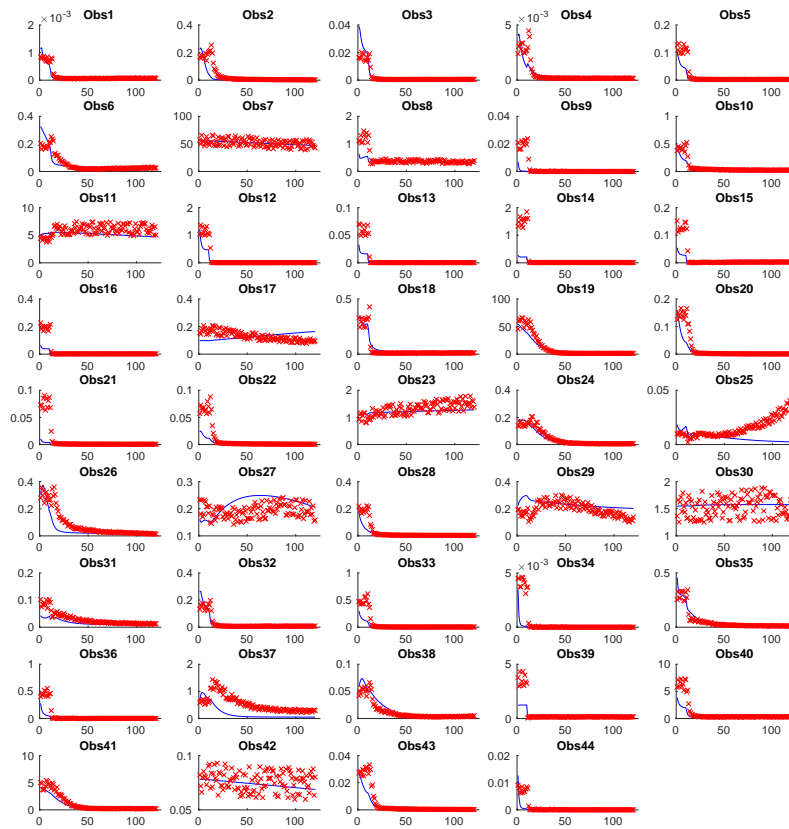


Figure 10.1: **(Data fits for CHO metabolism problem)** Pseudo-experimental data (red x) vs. optimal solution obtained using DOPS (solid blue lines) for the 44 observed states. X axis: time [s]; Y axis: metabolite concentrations [mM].

## Additional file 10.2-(Data fits for *S.cerevisiae* metabolism problem)

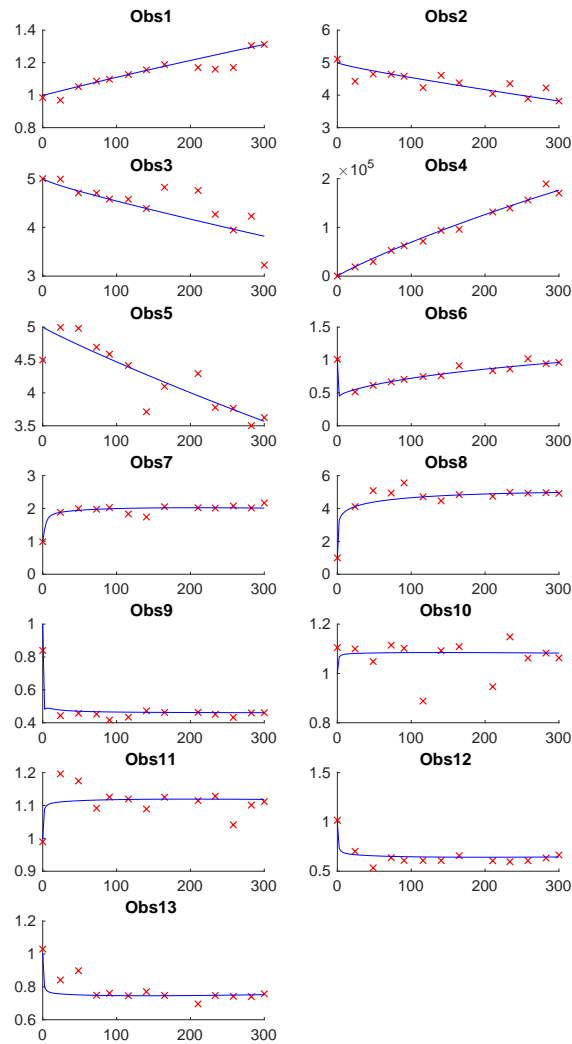


Figure 10.2: **(Data fits for *S.cerevisiae* metabolism problem)** Pseudo-experimental data (red x) vs. optimal solution obtained using DOPS (solid blue lines) for the 13 observed states. X axis: time [s]; Y axis: metabolite concentrations [mM].

### Additional file 10.3-(Comparison of states and parameters)

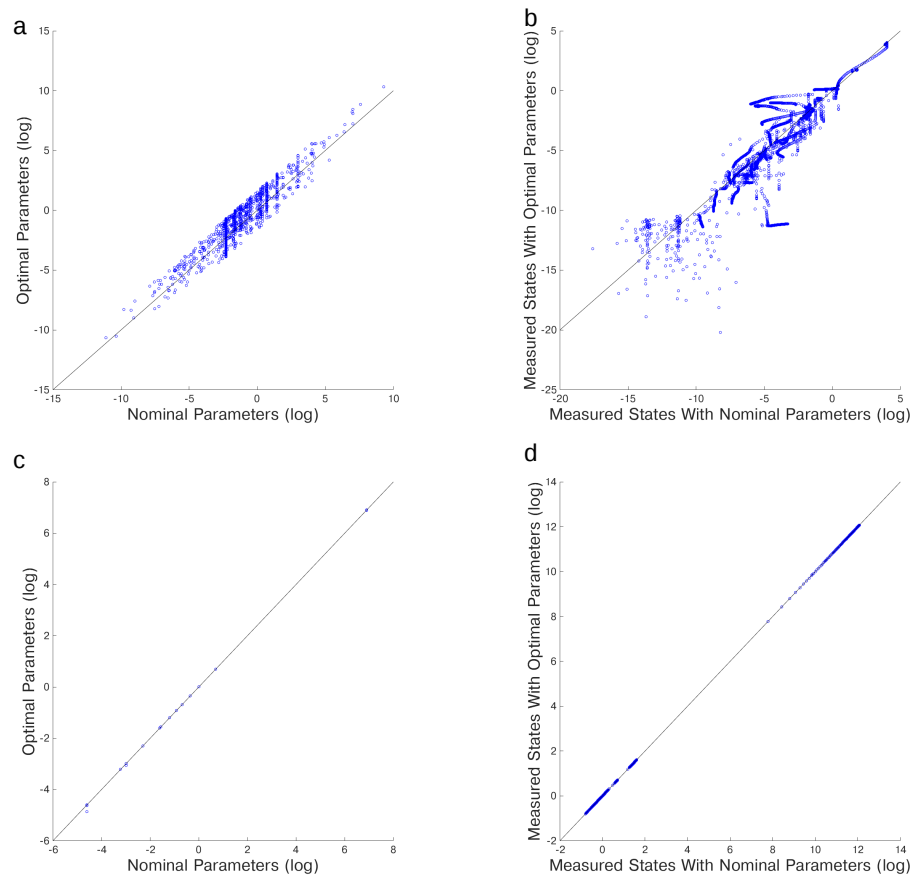


Figure 10.3: **(A)** Difference between nominal and optimal parameters for genome wide kinetic model of *S.cerevisiae* with 1759 unknown parameters using DOPS. **(B)** Difference between experimental (measured) data and data simulated with optimal parameters for genome wide kinetic model of *S.cerevisiae* with 1759 unknown parameters using DOPS. **(C)** Difference between nominal and optimal parameters for metabolic model of Chinese Hamster Ovary Cells (CHO) cells with 117 parameters using DOPS. **(D)** Difference between experimental (measured) data and data simulated with optimal parameters for metabolic model of Chinese Hamster Ovary Cells (CHO) cells with 117 parameters using DOPS.

### Additional file 10.4-(Time Comparison)

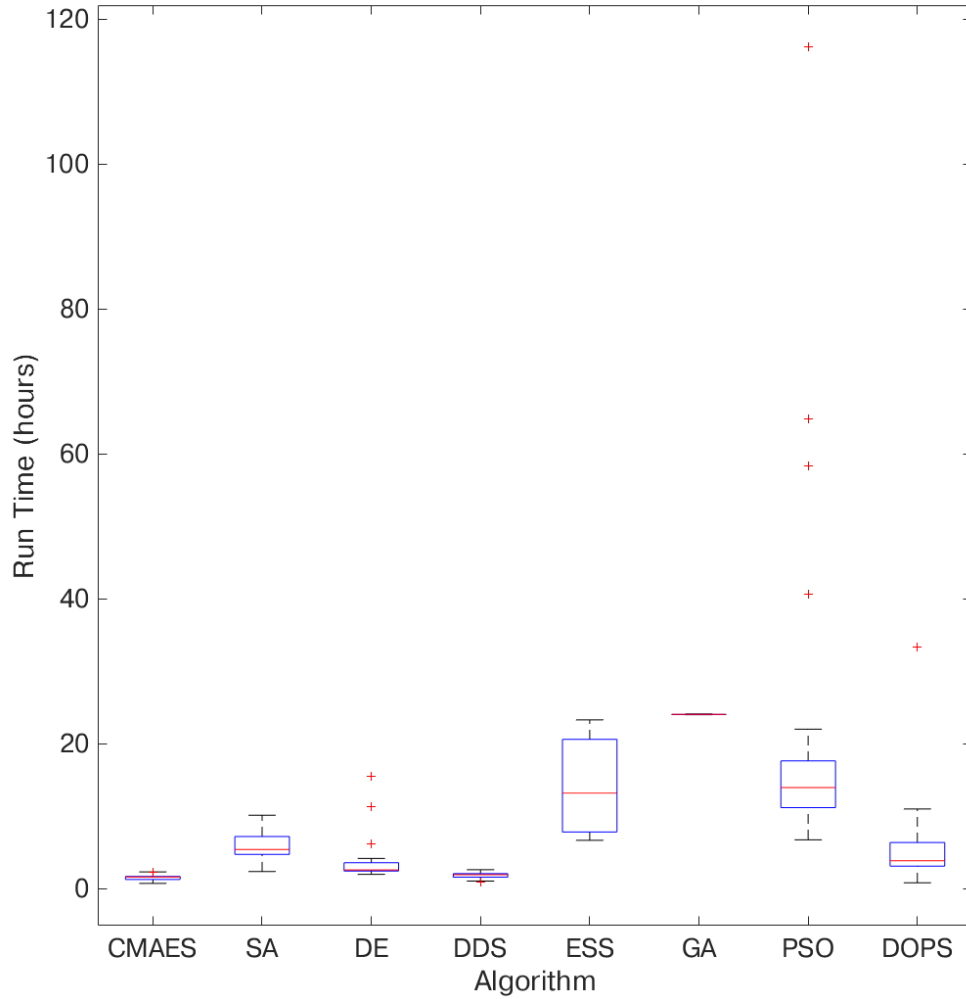


Figure 10.4: Comparison of the runtime of the different optimization methods used for comparison with  $\mathcal{T} = 25$  trials per method. All methods used take about the same amount of time to perform 4,000 function evaluations on the coagulation problem, as this problem is very stiff, so the majority of the time is spent solving the system of differential equations.

## **Additional file 10.5-(Convergence Curves)**



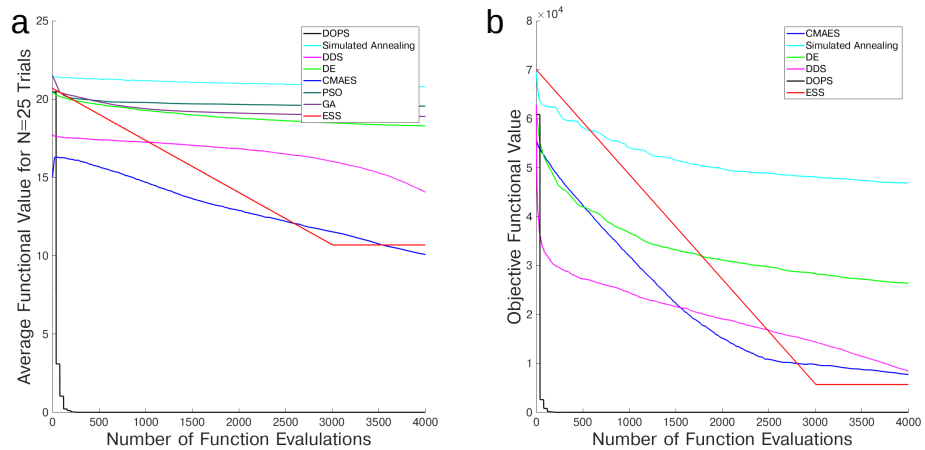


Figure 10.5: Mean convergence curves for different metaheuristics for (a) Ackley 300 dimensional and (b) Rastrigin 300 dimensional with  $\mathcal{T} = 25$  trials per method. DOPS not only finds a better solution than any other technique, it finds it with fewer function evaluations

### Additional file 10.6-(Comparison of DOPS to ESS)

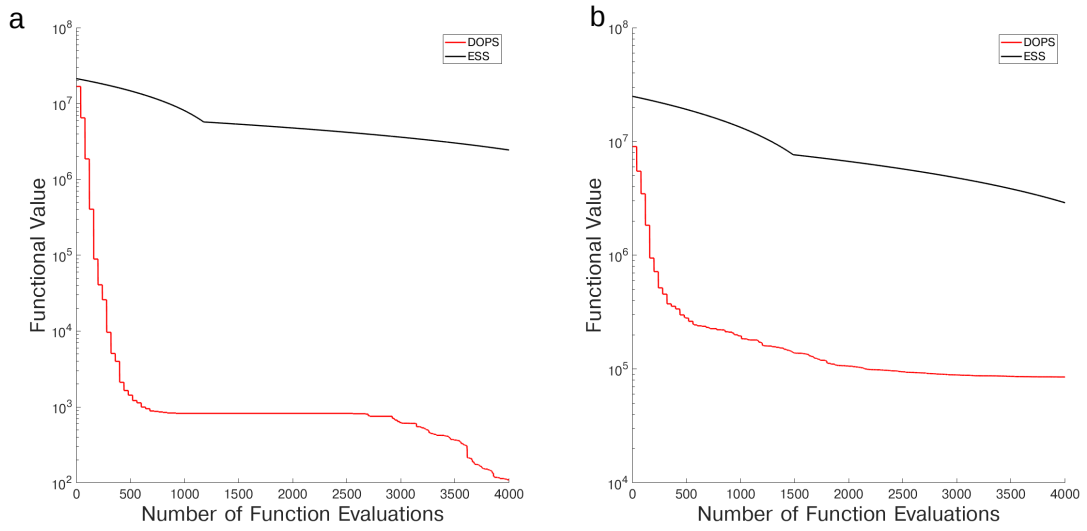


Figure 10.6: Mean convergence curves for DOPS and ESS for the (a) CHO model and (b) the coagulation model with  $\mathcal{T} = 25$  trials per method.

### Additional file 10.7-(Comparison of functional values)

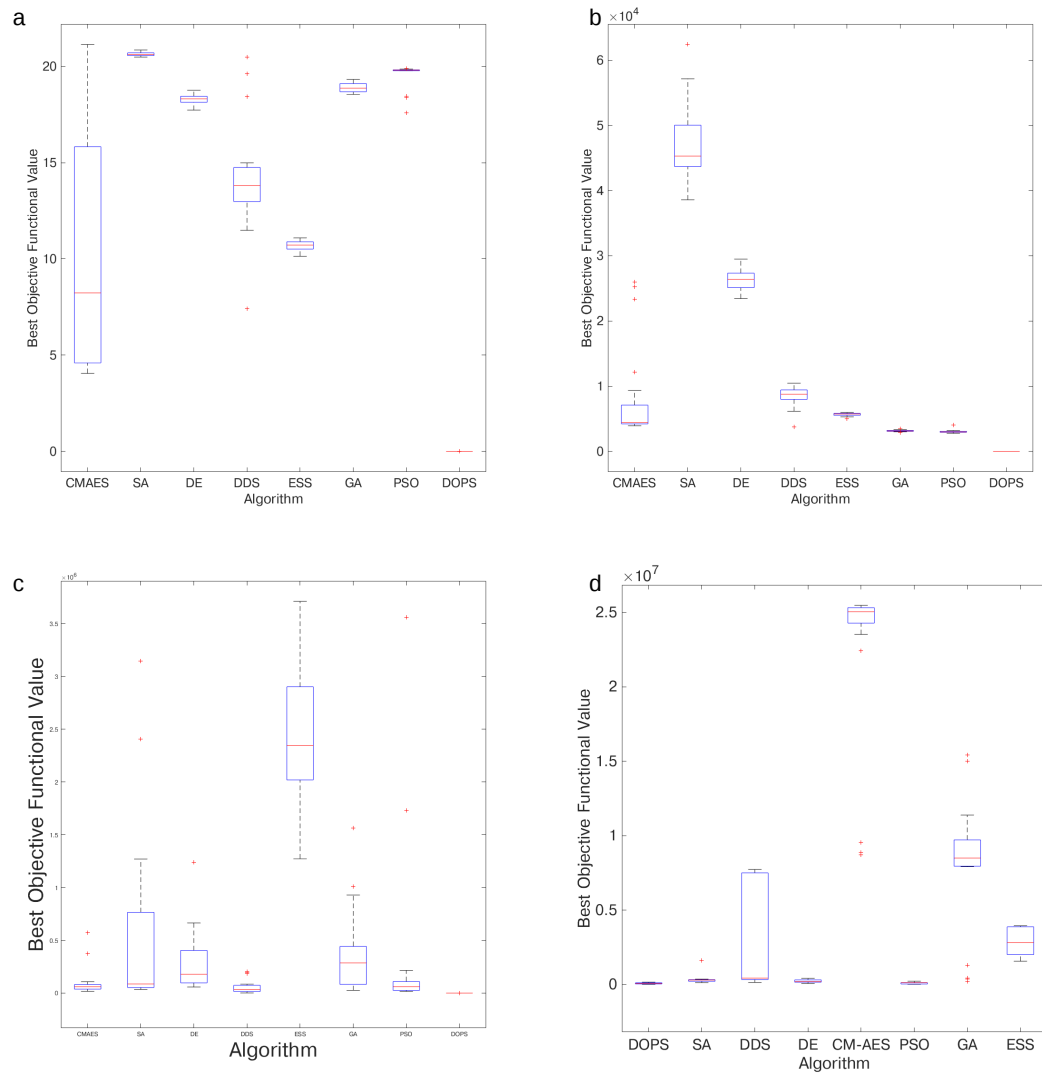


Figure 10.7: Variability analysis in best objective value for  $\mathcal{T} = 25$  trials. (a) Ackley 300 dimensional (b) Rastrigin 300 dimensional (c) CHO model (d) coagulation.

### Additional file 10.8-(Dispersion Curves)

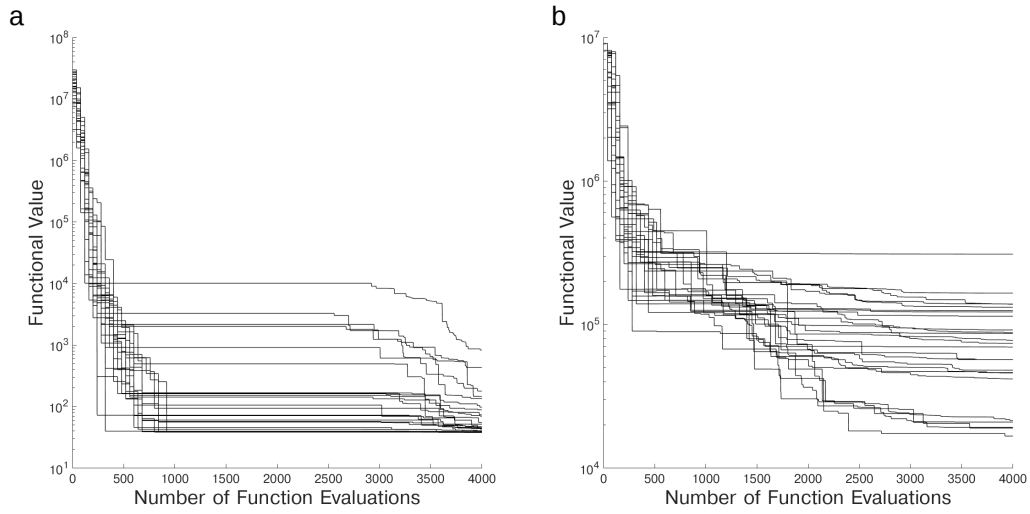


Figure 10.8: Dispersion curves for DOPS on (a) CHO model (b) coagulation with  $\mathcal{T} = 25$  trials per problem .

## BIBLIOGRAPHY

- [1] Tissue plasminogen activator (tpa).
- [2] Ajith Abraham, He Guo, and Hongbo Liu. Swarm intelligence: foundations, perspectives and applications. *Swarm Intelligent Systems*, pages 3–25, 2006.
- [3] Arash Afshari, Anne Wikkelsø, Jesper Brok, Ann Merete Møller, and Jørn Wetterslev. Thrombelastography (teg) or thromboelastometry (rotem) to monitor haemotherapy versus usual care in patients with massive transfusion. *Cochrane database of systematic reviews*, (3), 2011.
- [4] Ramzi Ajjan, Bernard CB Lim, Kristina F Standeven, Robert Harrand, Sarah Dolling, Fladia Phoenix, Richard Greaves, Radwa H Abou-Saleh, Simon Connell, D Alastair M Smith, et al. Common variation in the c-terminal region of the fibrinogen  $\beta$ -chain: effects on fibrin structure, fibrinolysis and clot rigidity. *Blood, The Journal of the American Society of Hematology*, 111(2):643–650, 2008.
- [5] H. Akaike. A new look at the statistical model identification. *IEEE Trans. Auto. Cont.*, 19:716 – 723, 1974.
- [6] Bree B Aldridge, John M Burke, Douglas A Lauffenburger, and Peter K Sorger. Physicochemical modelling of cell signalling pathways. *Nat Cell Biol*, 8(11):1195–203, Nov 2006.
- [7] Umme Amara, Michael A Flierl, Daniel Rittirsch, Andreas Klos, Hui Chen, Barbara Acker, Uwe B Brückner, Bo Nilsson, Florian Gebhard, John D Lambris, et al. Molecular intercommunication between the complement and coagulation systems. *The Journal of Immunology*, 185(9):5628–5636, 2010.
- [8] M Anand, K Rajagopal, and KR Rajagopal. A model incorporating some of the mechanical and biochemical factors underlying clot formation and dissolution in flowing blood: review article. *Journal of Theoretical Medicine*, 5(3-4):183–218, 2003.
- [9] Rajsavi S Anand, Paul Stey, Sukrit Jain, Dustin R Biron, Harikrishna Bhatt, Kristina Monteiro, Edward Feller, Megan L Ranney, Indra Neil Sarkar, and Elizabeth S Chen. Predicting mortality in diabetic icu patients using

- machine learning and severity indices. *AMIA Summits on Translational Science Proceedings*, 2018:310, 2018.
- [10] Maksat Ashyraliyev, Yves Fomekong-Nanfack, Jaap A Kaandorp, and Joke G Blom. Systems biology: parameter estimation for biochemical models. *Febs Journal*, 276(4):886–902, 2009.
  - [11] Heike E Assmus, Ralf Herwig, Kwang-Hyun Cho, and Olaf Wolkenhauer. Dynamics of biological systems: role of systems biology in medical research. *Expert Review of Molecular Diagnostics*, 2006.
  - [12] B stedt, I Lecander, and Tor Ny. The placental type plasminogen activator inhibitor, pai 2. *Fibrinolysis*, 1(4):203–208, 1987.
  - [13] MG Baines, KG Millar, and P Mills. Studies of complement levels in normal human pregnancy. *Obstetrics & Gynecology*, 43(6):806–810, 1974.
  - [14] Julio R Banga. Optimization in computational systems biology. *BMC systems biology*, 2(1):47, 2008.
  - [15] Brittany E Bannish, James P Keener, and Aaron L Fogelson. Modelling fibrinolysis: a 3d stochastic multiscale model. *Mathematical Medicine and Biology*, page dqs029, 2012.
  - [16] Rekha Bansal. Humanized and chimeric anti-properdin antibodies. US Patent 8,664,362, March 2014.
  - [17] Christopher D Barrett, Yi Wen Kong, and Michael B Yaffe. Influence of tranexamic acid on inflammatory signaling in trauma. In *Seminars in Thrombosis and Hemostasis*, volume 46, pages 183–188. Thieme Medical Publishers, 2020.
  - [18] Ashley N Bartels, Cory Johnson, Julie Lewis, James W Clevenger, Stephen L Barnes, Richard D Hammer, and Salman Ahmad. Platelet adenosine diphosphate inhibition in trauma patients by thromboelastography correlates with paradoxical increase in platelet dense granule content by flow cytometry. *Surgery*, 160(4):954–959, 2016.
  - [19] Diana Barthel, Susann Schindler, and Peter F Zipfel. Plasminogen is a complement inhibitor. *Journal of Biological Chemistry*, 287(22):18831–18842, 2012.

- [20] David Bassen, Michael Vilkhovoy, Mason Minot, Jonathan T Butcher, and Jeffrey D Varner. JuPOETs: A Constrained Multiobjective Optimization Approach to Estimate Biochemical Model Ensembles in the Julia Programming Language. *bioRxiv*, 10.1101/056044, 2016.
- [21] David M Bassen, Michael Vilkhovoy, Mason Minot, Jonathan T Butcher, and Jeffrey D Varner. Jupoets: a constrained multiobjective optimization approach to estimate biochemical model ensembles in the julia programming language. *BMC Syst Biol*, 11(1):10, 01 2017.
- [22] David M Bassen, Michael Vilkhovoy, Mason Minot, Jonathan T Butcher, and Jeffrey D Varner. Jupoets: a constrained multiobjective optimization approach to estimate biochemical model ensembles in the julia programming language. *BMC systems biology*, 11(1):10, 2017.
- [23] Charles Bayly-Jones, Doryen Bubeck, and Michelle A Dunstone. The mystery behind membrane insertion: a review of the complement membrane attack complex. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1726):20160221, 2017.
- [24] James V Beck and Keith A Woodbury. Inverse problems and parameter estimation: integration of measurements and analysis. *Measurement Science and Technology*, 9(6):839, 1998.
- [25] James Vere Beck and Kenneth J Arnold. *Parameter estimation in engineering and science*. James Beck, 1977.
- [26] Mandy J Bell. A historical overview of preeclampsia-eclampsia. *Journal of Obstetric, Gynecologic & Neonatal Nursing*, 39(5):510–518, 2010.
- [27] RF Bellamy, DC Pedersen, and LR DeGuzman. Organ blood flow and the cause of death following massive hemorrhage. *Circulatory shock*, 14(2):113–127, 1984.
- [28] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B. Shah. Julia: A fresh approach to numerical computing. *arXiv*, arXiv:1411.1607v4, July 2015.
- [29] A. Biedenkapp, J. Marben, M. Lindauer, and F. Hutter. CAVE: Configuration assessment, visualization and evaluation. In *Proceedings of the International Conference on Learning and Intelligent Optimization (LION’18)*, 2018.

- [30] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.
- [31] Anna M Blom, Lena Kask, and Björn Dahlbäck. Structural requirements for the complement regulatory activities of c4bp. *J Biol Chem.*, 276(29):27136–27144, 2001.
- [32] Ivan Bratko and P Mulec. *An experiment in automatic learning of diagnostic rules*. 1980.
- [33] CARL E Bredenberg, SHINNOSUKE Nomoto, and WATTS R Webb. Pulmonary and systemic hemodynamics during hemorrhagic shock in baboons. *Annals of surgery*, 192(1):86, 1980.
- [34] Benjamin Brenner. Haemostatic changes in pregnancy. *Thrombosis research*, 114(5-6):409–414, 2004.
- [35] GUENTHER F Brobmann, HARVEY B Ulano, George Moore, Paul Meyer, GERHARD A Brecher, and EUGENE D Jacobson. Effect of acute hemorrhage on arterial and venous resistance. *American Journal of Physiology-Legacy Content*, 220(1):12–15, 1971.
- [36] Karim Brohi, Mitchell J Cohen, Michael T Ganter, Michael A Matthay, Robert C Mackersie, and Jean-François Pittet. Acute traumatic coagulopathy: initiated by hypoperfusion: modulated through the protein c pathway? *Annals of surgery*, 245(5):812, 2007.
- [37] Karim Brohi, Mitchell J Cohen, Michael T Ganter, Marcus J Schultz, Marcel Levi, Robert C Mackersie, and Jean-François Pittet. Acute coagulopathy of trauma: hypoperfusion induces systemic anticoagulation and hyperfibrinolysis. *Journal of Trauma and Acute Care Surgery*, 64(5):1211–1217, 2008.
- [38] Karim Brohi, Jasmin Singh, Mischa Heron, and Timothy Coats. Acute traumatic coagulopathy. *Journal of Trauma and Acute Care Surgery*, 54(6):1127–1130, 2003.
- [39] David Brown, Rami A Namas, Khalid Almahmoud, Akram Zaaqoq, Joydeep Sarkar, Derek A Barclay, Jinling Yin, Ali Ghuma, Andrew Abboud, Gregory Constantine, et al. Trauma in silico: individual-specific mathematical models and virtual clinical populations. *Science translational medicine*, 7(285):285ra61–285ra61, 2015.



- [40] Kathleen E Brummel, Sara G Paradis, Saulius Butenas, and Kenneth G Mann. Thrombin functions during tissue factor-induced blood coagulation. *Blood*, 100(1):148–152, 2002.
- [41] Roland Brun, Peter Reichert, and Hans R Künsch. Practical identifiability analysis of large environmental simulation models. *Water Resources Research*, 37(4):1015–1030, 2001.
- [42] Saulius Butenas, Cornelis van’t Veer, and Kenneth G Mann. “normal” thrombin generation: Presented in part at the xvth congress of the international society on thrombosis and haemostasis, june 6-12, 1997, florence, italy (abstr ps-1653), at the 15th international congress on thrombosis, october 16-21, 1998, antalya, turkey (abstr 234), and at the 40th annual meeting of the american society of hematology, december 4-8, 1998, miami beach, fl (abstr 151). *Blood*, 94(7):2169–2178, 1999.
- [43] James R Byrnes and Alisa S Wolberg. Newly-recognized roles of factor xiii in thrombosis. In *Seminars in thrombosis and hemostasis*, volume 42, pages 445–454. Thieme Medical Publishers, 2016.
- [44] Federico Cabitza, Raffaele Rasoini, and Gian Franco Gensini. Unintended consequences of machine learning in medicine. *Jama*, 318(6):517–518, 2017.
- [45] Yves Cadroy, Hélène Grandjean, Josiane Pichon, Roland Desprats, Alain Berrebi, Alain Fournié, and Bernard Boneu. Evaluation of six markers of haemostatic system in normal pregnancy and pregnancy complicated by hypertension or pre-eclampsia. *BJOG: An International Journal of Obstetrics & Gynaecology*, 100(5):416–420, 1993.
- [46] Daniel Calvelo, Marie-C Chambrin, Denis Pomorski, and Pierre Ravaux. Icu patient state characterization using machine learning in a time series framework. In *Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making*, pages 356–360. Springer, 1999.
- [47] James E Campbell, Michael Adam Meledeo, and Andrew P Cap. Comparative response of platelet fv and plasma fv to activated protein c and relevance to a model of acute traumatic coagulopathy. *PloS one*, 9(6):e99181, 2014.
- [48] Francesca Campolongo, Jessica Cariboni, and Andrea Saltelli. An effective screening design for sensitivity analysis of large models. *Environmental modelling & software*, 22(10):1509–1518, 2007.

- [49] A Cap and BJ Hunt. The pathogenesis of traumatic coagulopathy. *Anaesthesia*, 70(s1):96, 2015.
- [50] Andrew P Cap, Heather F Pidcock, Philip Spinella, Geir Strandenes, Matthew A Borgman, Martin Schreiber, John Holcomb, Homer Chin-Nan Tien, Andrew N Beckett, Heidi Doughty, Tom Woolley, Joseph Rappold, Kevin Ward, Michael Reade, Nicolas Prat, Sylvain Ausset, Bijan Kheirabadi, Avi Benov, Edward P Griffin, and Jason B Corley. Damage control resuscitation. *Military Medicine*, 183:36 – 43, 2018.
- [51] K M Cawthorn, C van ’t Veer, J B Lock, M E DiLorenzo, R F Branda, and K G Mann. Blood coagulation in hemophilia a and hemophilia c. *Blood*, 91(12):4581–92, Jun 1998.
- [52] Kevin M Cawthorn, Cornelis van’t Veer, Jennifer B Lock, Maria E DiLorenzo, Richard F Branda, and Kenneth G Mann. Blood coagulation in hemophilia a and hemophilia c. *Blood*, 91(12):4581–4592, 1998.
- [53] Gabriela Cesarman-Maus and Katherine A Hajjar. Molecular mechanisms of fibrinolysis. *British journal of haematology*, 129(3):307–321, 2005.
- [54] John C Chapin and Katherine A Hajjar. Fibrinolysis and the control of blood coagulation. *Blood reviews*, 29(1):17–24, 2015.
- [55] Manash S Chatterjee, William S Denney, Huiyan Jing, and Scott L Diamond. Systems biology of coagulation initiation: kinetics of thrombin generation in resting and activated human blood. *PLoS computational biology*, 2010.
- [56] AK Chauhan and TL Moore. Presence of plasma complement regulatory proteins clusterin (apo j) and vitronectin (s40) on circulating immune complexes (cic). *Clin Exp Immunol.*, 145(3):398–406, 2006.
- [57] Zhengping Che, Sanjay Purushotham, Kyunghyun Cho, David Sontag, and Yan Liu. Recurrent neural networks for multivariate time series with missing values. *Scientific reports*, 8(1):6085, 2018.
- [58] Wenan Chen, Charles Cockrell, Kevin R Ward, and Kayvan Najarian. Intracranial pressure level prediction in traumatic brain injury by extracting features from multiple sources and using machine learning methods. In *2010 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 510–515. IEEE, 2010.

- [59] Li-Fang Cheng, Gregory Darnell, Bianca Dumitrascu, Corey Chivers, Michael E Draugelis, Kai Li, and Barbara E Engelhardt. Sparse multi-output gaussian processes for medical time series prediction. *arXiv preprint arXiv:1703.09112*, 2017.
- [60] Leon C Chesley. Hypertensive disorders in pregnancy. *Journal of nurse-midwifery*, 30(2):99–104, 1985.
- [61] Ngaam J Cheung, Xue-Ming Ding, and Hong-Bin Shen. Optifel: A convergent heterogeneous particle swarm optimization algorithm for takagi-sugeno fuzzy modeling. *IEEE Transactions on Fuzzy Systems*, 22(4):919–933, 2014.
- [62] Andreia Z Chignalia, Feliz Yetimakman, Sarah C Christiaans, Sule Unal, Benan Bayrakci, Brant M Wagener, Robert T Russell, Jeffrey D Kerby, Jean-Francois Pittet, and Randal O Dull. The glycocalyx and trauma: a review. *Shock (Augusta, Ga.)*, 45(4):338, 2016.
- [63] Kyunghyun Cho, Bart van Merriënboer, Çağlar Gülçehre, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *CoRR*, abs/1406.1078, 2014.
- [64] Edward Choi, Mohammad Taha Bahadori, Andy Schuetz, Walter F Stewart, and Jimeng Sun. Doctor ai: Predicting clinical events via recurrent neural networks. In *Machine Learning for Healthcare Conference*, pages 301–318, 2016.
- [65] Edward Choi, Andy Schuetz, Walter F Stewart, and Jimeng Sun. Using recurrent neural network models for early detection of heart failure onset. *Journal of the American Medical Informatics Association*, 24(2):361–370, 2016.
- [66] Shu-Ling Chong, Nan Liu, Sylvaine Barbier, and Marcus Eng Hock Ong. Predictive modeling in pediatric traumatic brain injury using machine learning. *BMC medical research methodology*, 15(1):22, 2015.
- [67] S Ariane Christie, Amanda S Conroy, Rachael A Callcut, Alan E Hubbard, and Mitchell J Cohen. Dynamic multi-outcome prediction after injury: Applying adaptive machine learning for precision medicine in trauma. *PloS one*, 14(4):e0213836, 2019.
- [68] Maurice Clerc. *Particle swarm optimization*. ISTE, London, 2006.

- [69] Timothy J Coats and Mohamed Morsy. Biological mechanisms and individual variation in fibrinolysis after major trauma. *Emergency Medicine Journal*, 2020.
- [70] Jared Cohen, Thomas Scorer, Zachary Wright, Ian J Stewart, Jonathan Sosnov, Heather Pidcoke, Chriselda Fedyk, Hana Kwan, Kevin K Chung, Kelly Heegard, et al. A prospective evaluation of thromboelastometry (rotem) to identify acute traumatic coagulopathy and predict massive transfusion in military trauma patients in afghanistan. *Transfusion*, 59(S2):1601–1607, 2019.
- [71] Crash-2 Collaborators et al. The importance of early treatment with tranexamic acid in bleeding trauma patients: an exploratory analysis of the crash-2 randomised controlled trial. *The Lancet*, 377(9771):1096–1101, 2011.
- [72] JP Collet, D Park, C Lesty, J Soria, C Soria, G Montalescot, and JW Weisel. Influence of fibrin network conformation and fibrin fiber diameter on fibrinolysis speed dynamic and structural approaches by confocal microscopy. *Arteriosclerosis, thrombosis, and vascular biology*, 20(5):1354–1361, 2000.
- [73] Sean J Cone, Andrew T Fuquay, Justin M Litofsky, Taylor C Dement, Christopher A Carolan, and Nathan E Hudson. Inherent fibrin fiber tension propels mechanisms of network clearance during fibrinolysis. *Acta Biomaterialia*, 2020.
- [74] N Curry, C Rourke, R Davenport, S Beer, L Pankhurst, A Deary, H Thomas, C Llewelyn, L Green, H Doughty, et al. Early cryoprecipitate for major haemorrhage in trauma: a randomised controlled feasibility trial. *British journal of anaesthesia*, 115(1):76–83, 2015.
- [75] John Danziger. Vitamin k-dependent proteins, warfarin, and vascular calcification. *Clinical journal of the American society of nephrology*, 3(5):1504–1510, 2008.
- [76] Ross Davenport. Pathogenesis of acute traumatic coagulopathy. *Transfusion*, 53:23S–27S, 2013.
- [77] Ross Davenport, Joanna Manson, Henry De’Ath, Sean Platton, Amy Coates, Shubha Allard, Daniel Hart, Rupert Pearse, K John Pasi, Peter MacCallum, et al. Functional definition and characterisation of acute traumatic coagulopathy. *Critical care medicine*, 39(12):2652, 2011.

- [78] Ross A Davenport and Karim Brohi. Cause of trauma-induced coagulopathy. *Current Opinion in Anesthesiology*, 29(2):212–219, 2016.
- [79] Carrie A Davis, Benjamin C Hitz, Cricket A Sloan, Esther T Chan, Jean M Davidson, Idan Gabdank, Jason A Hilton, Kriti Jain, Ulugbek K Baymurov, Aditi K Narayanan, et al. The encyclopedia of dna elements (encode): data portal update. *Nucleic acids research*, 46(D1):D794–D801, 2018.
- [80] Edward De Brouwer, Jaak Simm, Adam Arany, and Yves Moreau. Grude-bayes: Continuous modeling of sporadically-observed time series. In *Advances in Neural Information Processing Systems*, pages 7377–7388, 2019.
- [81] Paul S De Vries, Daniel I Chasman, Maria Sabater-Lleal, Ming-Huei Chen, Jennifer E Huffman, Maristella Steri, Weihong Tang, Alexander Teumer, Riccardo E Marioni, Vera Grossmann, et al. A meta-analysis of 120 246 individuals identifies 18 new loci for fibrinogen concentration. *Human molecular genetics*, 25(2):358–370, 2016.
- [82] Shirley Uitte de Willige, Kristina F Standeven, Helen Philippou, and Robert AS Ariëns. The pleiotropic role of the fibrinogen  $\gamma'$  chain in hemostasis. *Blood*, 114(19):3994–4001, 2009.
- [83] Robert L Dedrick. Animal scale-up. *Journal of pharmacokinetics and biopharmaceutics*, 1(5):435–461, 1973.
- [84] Janez Demšar, Blaž Zupan, Noriaki Aoki, Matthew J Wall, Thomas H Granchi, and J Robert Beck. Feature mining and predictive model construction from severe trauma patient’s data. *International journal of medical informatics*, 63(1-2):41–50, 2001.
- [85] Zoltán Derzsy, Zoltán Prohászka, János Rigó Jr, George Füst, and Attila Molvarec. Activation of the complement system in normal pregnancy and preeclampsia. *Molecular immunology*, 47(7-8):1500–1506, 2010.
- [86] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.
- [87] Scott L Diamond. Systems biology of coagulation. *Journal of Thrombosis and Haemostasis*, 11(s1):224–232, 2013.
- [88] Lawrence N Diebel, Mark E Diebel, Jonathan V Martin, and David M

- Liberati. Acute hyperglycemia exacerbates trauma-induced endothelial and glycocalyx injury: an in vitro model. *Journal of Trauma and Acute Care Surgery*, 85(5):960–967, 2018.
- [89] Lawrence N Diebel, Jonathan V Martin, and David M Liberati. Early tranexamic acid administration ameliorates the endotheliopathy of trauma and shock in an in vitro model. *Journal of Trauma and Acute Care Surgery*, 82(6):1080–1086, 2017.
- [90] Geoffrey P Dobson, Hayley L Letson, Rajiv Sharma, Forest R Sheppard, and Andrew P Cap. Mechanisms of early trauma-induced coagulopathy: The clot thickens or not? *Journal of Trauma and Acute Care Surgery*, 79(2):301–309, 2015.
- [91] Geoffrey P Dobson, Jodie L Morris, Lisa M Davenport, and Hayley L Letson. Traumatic-induced coagulopathy as a systems failure: A new window into hemostasis. In *Seminars in Thrombosis and Hemostasis*. Thieme Medical Publishers, 2020.
- [92] Nathan Dow, Julia R Coleman, Hunter Moore, Zachary Osborn, Adrian M Sackheim, Grant Hennig, Saulius Butenas, Mark T Nelson, Ernest E Moore, and Kalev Freeman. Dense and dangerous: The tissue plasminogen activator (t-pa)-resistant fibrinolysis shutdown phenotype is due to abnormal fibrin polymerization. *Journal of Trauma and Acute Care Surgery*, 2019.
- [93] Luci M Dusse, Maria G Carvalho, Kathryn Getliffe, David Voegeli, Alan J Cooper, and Bashir A Lwaleed. Increased circulating thrombomodulin levels in pre-eclampsia. *Clinica chimica acta; international journal of clinical chemistry*, 387(1-2):168, 2008.
- [94] Luci M Dusse, Danyelle RA Rios, Melina B Pinheiro, Alan J Cooper, and Bashir A Lwaleed. Pre-eclampsia: relationship between coagulation, fibrinolysis and inflammation. *Clinica Chimica Acta*, 412(1-2):17–21, 2011.
- [95] Richard P Dutton, Colin F Mackenzie, and Thomas M Scalea. Hypotensive resuscitation during active hemorrhage: impact on in-hospital mortality. *Journal of Trauma and Acute Care Surgery*, 52(6):1141–1146, 2002.
- [96] Jose A Egea, María Rodríguez-Fernández, Julio R Banga, and Rafael Martí. Scatter search for chemical and bio-process optimization. *Journal of Global Optimization*, 37(3):481–503, 2007.

- [97] David Epstein and Jeffrey C Kurz. Complement binding aptamers and anti-c5 agents useful in the treatment of ocular disorders. US Patent App. 12/224,708, March 2007.
- [98] Offer Erez, Roberto Romero, Edi Vaisbuch, Juan Pedro Kusanovic, Shali Mazaki-Tovi, Tinnakorn Chaiworapongsa, Francesca Gotsch, Pooja Mittal, Samuel S Edwin, Chia-Ling Nhan-Chang, et al. The pattern and magnitude of “in vivo thrombin generation” differ in women with preeclampsia and in those with sga fetuses without preeclampsia. *The Journal of Maternal-Fetal & Neonatal Medicine*, 31(13):1671–1680, 2018.
- [99] Bengt I Eriksson, Elsa Eriksson, Erika Gyzander, Ann-Catrine Teger-nilsson, and Bo Risberg. Thrombosis after hip replacement: relationship to the fibrinolytic system. *Acta orthopaedica Scandinavica*, 60(2):159–163, 1989.
- [100] Larry J Eshelman, KP Lee, Joseph J Frassica, Wei Zong, Larry Nielsen, and Mohammed Saeed. Development and evaluation of predictive alerts for hemodynamic instability in icu patients. In *AMIA Annual Symposium Proceedings*, volume 2008, page 379. American Medical Informatics Association, 2008.
- [101] Charles T Esmon. The roles of protein c and thrombomodulin in the regulation of blood coagulation. *J Biol Chem*, 264(9):4743–4746, 1989.
- [102] Charles T Esmon. The interactions between inflammation and coagulation. *Br J Haematol*, 131(4):417–30, Nov 2005.
- [103] William R Esposito and Christodoulos A Floudas. Deterministic global optimization in nonlinear optimal control problems. *Journal of Global Optimization*, 17(1-4):97–126, 2000.
- [104] Amparo Estellés, Juan Gilabert, Salvador Grancha, Koji Yamamoto, Terri Thinnes, Francisco España, Justo Aznar, and David Loskutoff. Abnormal expression of type 1 plasminogen activator inhibitor and tissue factor in severe preeclampsia. *Thrombosis and haemostasis*, 79(03):500–508, 1998.
- [105] David H Farrell.  $\gamma$  fibrinogen as a novel marker of thrombotic disease. *Clinical chemistry and laboratory medicine*, 50(11):1903–1909, 2012.
- [106] Rebecca A Fleck, L Vijaya Mohan Rao, Samuel I Rapaport, and Nissi Varki. Localization of human tissue factor antigen by immunostaining

- with monospecific, polyclonal anti-human tissue factor antibody. *Thrombosis research*, 59(2):421–437, 1990.
- [107] D Fogel. Artificial intelligence through simulated evolution. *Wiley-IEEE Press*, 2009.
  - [108] Aaron L Fogelson and Nessay Tania. Coagulation under flow: the influence of flow-mediated transport on the initiation and inhibition of coagulation. *Pathophysiology of haemostasis and thrombosis*, 34(2-3):91–108, 2005.
  - [109] Centers for Disease Control, Prevention, et al. Web-based injury statistics query and reporting system (wisqars). atlanta, ga: National center for injury prevention and control. retrieved sept 20,2016, 2012.
  - [110] The American Association for the Surgery of Trauma. Cost of injury, 2005.
  - [111] Alexandre Fort, Christelle Borel, Eugenia Migliavacca, Stylianos E Antonarakis, Richard J Fish, and Marguerite Neerman-Arbez. Regulation of fibrinogen production by micrnas. *Blood*, 116(14):2608–2615, 2010.
  - [112] D Frith, JC Goslings, C Gaarder, M Maegele, MJ Cohen, S Allard, PI Johansson, S Stanworth, C Thiemermann, and K Brohi. Definition and drivers of acute traumatic coagulopathy: clinical and experimental investigations. *Journal of Thrombosis and Haemostasis*, 8(9):1919–1925, 2010.
  - [113] Gerald M Fuller and Zhixin Zhang. Transcriptional control mechanism of fibrinogen gene expression. *Annals of the New York Academy of Sciences*, 936(1):469–479, 2001.
  - [114] Bruce Furie and Barbara C Furie. The molecular basis of blood coagulation. *Cell*, 53(4):505–518, 1988.
  - [115] Attila Gábor, Alejandro F Villaverde, and Julio R Banga. Parameter identifiability analysis and visualization in large-scale kinetic models of biosystems. *BMC systems biology*, 11(1):54, 2017.
  - [116] Don A Gabriel, Kathleen Muga, and Emily M Boothroyd. The effect of fibrin structure on fibrinolysis. *Journal of Biological Chemistry*, 267(34):24259–24263, 1992.
  - [117] SATOSHI Gando, ICHIRO Tede, and MUNEHITO Kubota. Posttrauma coagulation and fibrinolysis. *Critical care medicine*, 20(5):594–600, 1992.



- [118] Richard A Gibbs, Jeffrey Rogers, Michael G Katze, Roger Bumgarner, George M Weinstock, Elaine R Mardis, Karin A Remington, Robert L Strausberg, J Craig Venter, Richard K Wilson, et al. Evolutionary and biomedical insights from the rhesus macaque genome. *Science*, 316(5822):222–234, 2007.
- [119] David E Goldberg. Genetic algorithms. *Pearson Education India*, 2006.
- [120] Ary L Goldberger, Luis AN Amaral, Leon Glass, Jeffrey M Hausdorff, Plamen Ch Ivanov, Roger G Mark, Joseph E Mietus, George B Moody, Chung-Kang Peng, and H Eugene Stanley. Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *circulation*, 101(23):e215–e220, 2000.
- [121] Marsha G Goldfarb, Gloria J Bazzoli, and Rosanna M Coffey. Trauma systems and the costs of trauma care. *Health services research*, 31(1):71, 1996.
- [122] Tianxia Gong, Ruizhe Liu, Chew Lim Tan, Neda Farzad, Cheng Kiang Lee, Boon Chuan Pang, Qi Tian, Suisheng Tang, and Zhuo Zhang. Classification of ct brain images of head trauma. In *IAPR International Workshop on Pattern Recognition in Bioinformatics*, pages 401–408. Springer, 2007.
- [123] S Grassin-Delyle, OM Theusinger, R Albrecht, S Mueller, DR Spahn, S Urien, and P Stein. Optimisation of the dosage of tranexamic acid in trauma patients with population pharmacokinetic analysis. *Anaesthesia*, 73(6):719–729, 2018.
- [124] Alex Graves and Navdeep Jaitly. Towards end-to-end speech recognition with recurrent neural networks. In *International conference on machine learning*, pages 1764–1772, 2014.
- [125] Alex Graves and Jürgen Schmidhuber. Framewise phoneme classification with bidirectional lstm and other neural network architectures. *Neural networks*, 18(5-6):602–610, 2005.
- [126] Cary P Gross, Gerard F Anderson, and Neil R Powe. The relation between funding by the national institutes of health and the burden of disease. *New England Journal of Medicine*, 340(24):1881–1887, 1999.
- [127] Chrissy Guidry, Elizabeth Gleeson, Eric R Simms, Lance Stuke, Peter Meade, Norman E McSwain Jr, and Juan C Duchesne. Initial assessment

on the impact of crystalloids versus colloids during damage control resuscitation. *journal of surgical research*, 185(1):294–299, 2013.

- [128] Francis X Guyette, Jason L Sperry, Andrew B Peitzman, Timothy R Biliyar, Brian J Daley, Richard S Miller, Brian G Harbrecht, Jeffrey A Claridge, Tyler Putnam, Therese M Duane, et al. Prehospital blood product and crystalloid resuscitation in the severely injured patient: a secondary analysis of the prehospital air medical plasma trial. *Annals of surgery*, 2019.
- [129] Tsonwin Hai, Curt D Wolfgang, Derek K Marsee, Amy E Allen, and Uma-sundari Sivaprasad. Atf3 and stress responses. *Gene Expression, The Journal of Liver Research*, 7(4-5):321–335, 1999.
- [130] Jan Hartmann, Matthew Murphy, and Joao D Dias. Viscoelastic hemostatic assays: Moving from the laboratory to the site of care—a review of established and emerging technologies. *Diagnostics*, 10(2):118, 2020.
- [131] Jan Hartmann, Mark Walsh, Anne Grisoli, Anthony V Thomas, Faisal Shariff, Ross McCauley, Stefani Vande Lune, Nuha Zackariya, Shivani Patel, Michael S Farrell, et al. Diagnosis and treatment of trauma-induced coagulopathy by viscoelastography. In *Seminars in Thrombosis and Hemostasis*, volume 46, pages 134–146. Thieme Medical Publishers, 2020.
- [132] Dennis F Hayes, Michael H Werner, Irwin K Rosenberg, Charles E Lucas, Melvyn Westreich, and Volker Bradley. Effects of traumatic hypovolemic shock on renal function. *Journal of Surgical Research*, 16(5):490–497, 1974.
- [133] M Hellgren and M Blombäck. Studies on blood coagulation and fibrinolysis in pregnancy, during delivery and in the puerperium. *Gynecologic and obstetric investigation*, 12(3):141–154, 1981.
- [134] Margareta Hellgren. Hemostasis during normal pregnancy and puerperium. In *Seminars in thrombosis and hemostasis*, volume 29, pages 125–130. Copyright© 2003 by Thieme Medical Publishers, Inc., 333 Seventh Avenue, New ... , 2003.
- [135] JD Herman. Salib: Sensitivity analysis library in python (numpy). contains sobol, morris, fractional factorial and fast methods. available online: <https://github.com/jdherman/salib>.
- [136] Jon Herman and Will Usher. Salib: an open-source python library for sensitivity analysis. *The Journal of Open Source Software*, 2(9), 2017.

- [137] D L Higgins and K G Mann. The interaction of bovine factor v and factor v-derived peptides with phospholipid vesicles. *J Biol Chem*, 258(10):6503–8, May 1983.
- [138] Alan C. Hindmarsh, Peter N. Brown, Keith E. Grant, Steven L. Lee, Radu Serban, Dan E. Shumaker, and Carol S. Woodward. Sundials: Suite of nonlinear and differential/algebraic equation solvers. *ACM Trans. Math. Softw.*, 31(3):363–396, September 2005.
- [139] Hirohumi Hirayama, Kiyono Yoshii, Hidetomo Ojima, Norikazu Kawai, Shintaro Gotoh, and Yuzo Fukuyama. Linear systems analysis of activating processes of complement system as a defense mechanism. *Biosystems*, 39(3):173–185, 1996.
- [140] Asher Hirshberg, Mark Dugas, Eugenio I Banez, Bradford G Scott, Matthew J Wall Jr, and Kenneth L Mattox. Minimizing dilutional coagulopathy in exsanguinating hemorrhage: a computer simulation. *Journal of Trauma and Acute Care Surgery*, 54(3):454–463, 2003.
- [141] Anthony MH Ho, Peter W Dion, Claudia AY Cheng, Manoj K Karmakar, Gregory Cheng, Zhiyong Peng, and Yu Wai Ng. A mathematical model for fresh frozen plasma transfusion strategies during major trauma resuscitation with ongoing hemorrhage. *Canadian journal of surgery*, 48(6):470, 2005.
- [142] Gerald Hochleitner, Ken Sutor, Caroline Levett, Harald Leyser, Christoph J Schlimp, and Cristina Solomon. Revisiting hartert’s 1962 calculation of the physical constants of thrombelastography. *Clinical and Applied Thrombosis/Hemostasis*, 23(3):201–210, 2017.
- [143] Matthew F Hockin, Kenneth C Jones, Stephen J Everse, and Kenneth G Mann. A model for the stoichiometric regulation of blood coagulation. *Journal of Biological Chemistry*, 277(21):18322–18333, 2002.
- [144] John B Holcomb, Deborah J Del Junco, Erin E Fox, Charles E Wade, Mitchell J Cohen, Martin A Schreiber, Louis H Alarcon, Yu Bai, Karen J Brasel, Eileen M Bulger, et al. The prospective, observational, multicenter, major trauma transfusion (promtt) study: comparative effectiveness of a time-varying treatment with competing risks. *JAMA surgery*, 148(2):127–136, 2013.
- [145] Leroy Hood, James R Heath, Michael E Phelps, and Biaoyang Lin. Sys-

tems biology and new technologies enable predictive and preventative medicine. *Science*, 306(5696):640–643, 2004.

- [146] Robert Hooke and To A Jeeves. “direct search” solution of numerical and statistical problems. *Journal of the ACM (JACM)*, 8(2):212–229, 1961.
- [147] Holger Hoos and Kevin Leyton-Brown. An efficient approach for assessing hyperparameter importance. In *International conference on machine learning*, pages 754–762, 2014.
- [148] Reiner Horst and Hoang Tuy. *Global optimization: Deterministic approaches*. Springer Science & Business Media, 2013.
- [149] Jeremy Howard and Sebastian Ruder. Universal language model fine-tuning for text classification. *arXiv preprint arXiv:1801.06146*, 2018.
- [150] PW Howie, CRM Prentice, and GP McNicol. Coagulation, fibrinolysis and platelet function in pre-eclampsia, essential hypertension and placental insufficiency. *BJOG: An International Journal of Obstetrics & Gynaecology*, 78(11):992–1003, 1971.
- [151] Marc Hoylaerts, Henri Roger Lijnen, and Désiré Collen. Studies on the mechanism of the antifibrinolytic action of tranexamic acid. *Biochimica et Biophysica Acta (BBA)-General Subjects*, 673:75–85, 1981.
- [152] Xianzhen Hu, V Michael Holers, Joshua M Thurman, Trent R Schoeb, Theresa N Ramos, and Scott R Barnum. Therapeutic inhibition of the alternative complement pathway attenuates chronic EAE. *Mol Immunol.*, 54:302–308, 2013.
- [153] Harriet Hunt, Simon Stanworth, Nicola Curry, Tom Woolley, Chris Cooper, Obioha Ukoumunne, Zhivko Zhelev, and Chris Hyde. Thromboelastography (teg) and rotational thromboelastometry (rotem) for trauma-induced coagulopathy in adult trauma patients with bleeding. *Cochrane Database of Systematic Reviews*, (2), 2015.
- [154] Frank Hutter, Holger H Hoos, and Kevin Leyton-Brown. Sequential model-based optimization for general algorithm configuration (extended version). *Technical Report TR-2010–10, University of British Columbia, Computer Science, Tech. Rep.*, 2010.
- [155] Kei Ikeda, Kohei Nagasawa, Takahiko Horiuchi, Tomomi Tsuru, Hiroaki

- Nishizaka, and Yoshiyuki Niho. C5a induces tissue factor activity on endothelial cells. *Thrombosis and haemostasis*, 77(02):394–398, 1997.
- [156] Yuki Itagaki, Mineji Hayakawa, Kunihiro Maekawa, Tomoyo Saito, Akira Kodate, Yoshinori Honma, Asumi Mizugaki, Tomonao Yoshida, Takayoshi Ohyasu, Kenichi Katabami, et al. Early administration of fibrinogen concentrate is associated with improved survival among severe trauma patients: a single-centre propensity score-matched analysis. *World Journal of Emergency Surgery*, 15(1):1–10, 2020.
- [157] Andra H James. Pregnancy-associated thrombosis. *ASH Education Program Book*, 2009(1):277–285, 2009.
- [158] Momin Jamil and Xin-She Yang. A literature survey of benchmark functions for global optimization problems. *Int. Journal of Mathematical Modelling and Numerical Optimisation*, 4(2):150 – 194, 2013.
- [159] Momin Jamil and Xin-She Yang. A literature survey of benchmark functions for global optimisation problems. *International Journal of Mathematical Modelling and Numerical Optimisation*, 4(2):150–194, 2013.
- [160] Khuloud Jaqaman and Gaudenz Danuser. Linking data to models: data regression. *Nat Rev Mol Cell Biol*, 7(11):813–9, Nov 2006.
- [161] Pär I Johansson, Trine Stissing, Louise Bochsén, and Sisse R Ostrowski. Thrombelastography and tromboelastometry in assessing coagulopathy in trauma. *Scandinavian journal of trauma, resuscitation and emergency medicine*, 17(1):45, 2009.
- [162] Alexander E John and Nathan J White. Platelets and fibrinogen: emerging complexity in trauma-induced coagulopathy. In *Seminars in Thrombosis and Hemostasis*, volume 46, pages 125–133. Thieme Medical Publishers, 2020.
- [163] Alistair EW Johnson, Tom J Pollard, Lu Shen, H Lehman Li-wei, Mengling Feng, Mohammad Ghassemi, Benjamin Moody, Peter Szolovits, Leo Anthony Celi, and Roger G Mark. MIMIC-III, a freely accessible critical care database. *Scientific data*, 3:160035, 2016.
- [164] Gregory J Jurkovich, William B Greiser, Arnold Luterman, and P WILLIAM Curreri. Hypothermia in trauma victims: an ominous predictor of survival. *The Journal of trauma*, 27(9):1019–1024, 1987.

- [165] Deepak A Kaji, John R Zech, Jun S Kim, Samuel K Cho, Neha S Dangayach, Anthony B Costa, and Eric K Oermann. An attention based deep learning model of clinical events in the intensive care unit. *PloS one*, 14(2):e0211057, 2019.
- [166] Minoru Kanehisa, Yoko Sato, Miho Furumichi, Kanae Morishima, and Mao Tanabe. New approach for understanding genome variations in kegg. *Nucleic acids research*, 47(D1):D590–D595, 2019.
- [167] Ioannis Karalis, Sunil K Nadar, Eman Al Yemeni, Andrew D Blann, and Gregory YH Lip. Platelet activation in pregnancy-induced hypertension. *Thrombosis research*, 116(5):377–383, 2005.
- [168] Kenneth J Katschke, Ping Wu, Rajkumar Ganesan, Robert F Kelley, Mary A Mathieu, Philip E Hass, Jeremy Murray, Daniel Kirchhofer, Christian Wiesmann, and Menno van Lookeren Campagne. Inhibiting alternative pathway complement activation by targeting the factor d exosite. *J Biol Chem.*, 287:12886–12892, 2012.
- [169] Christoph R Kaufmann, Kevin M Dwyer, John D Crews, Sheila J Dols, and Arthur L Trask. Usefulness of thrombelastography in assessment of trauma patient coagulation. *Journal of Trauma and Acute Care Surgery*, 42(4):716–722, 1997.
- [170] M.A. Khanin and V.V. Semenov. A mathematical model of the kinetics of blood coagulation. *Journal of Theoretical Biology*, 136(2):127 – 134, 1989.
- [171] Scott Kirkpatrick, C Daniel Gelatt, Mario P Vecchi, et al. Optimization by simulated annealing. *science*, 220(4598):671–680, 1983.
- [172] Hiroaki Kitano. Systems biology: a brief overview. *Science*, 295(5560):1662–1664, 2002.
- [173] Lucy Z Kornblith, Matthew E Kutcher, Brittney J Redick, Carolyn S Calfee, Ryan F Vilardi, and Mitchell Jay Cohen. Fibrinogen and platelet contributions to clot formation: implications for trauma resuscitation and thromboprophylaxis. *The journal of trauma and acute care surgery*, 76(2):255, 2014.
- [174] Andrey A Korotaevskiy, Leonid G Hanin, and Mikhail A Khanin. Non-linear dynamics of the complement system activation. *Mathematical biosciences*, 222(2):127–143, 2009.

- [175] E G Krug, G K Sharma, and R Lozano. The global burden of injuries. *Am J Public Health*, 90(4):523–6, Apr 2000.
- [176] Thomas Lang, Anne Bauters, Siegmund L Braun, Bernd Pötzsch, Klaus-Werner von Pape, Hans-Jürgen Kolde, and Meret Lakner. Multi-centre investigation on reference ranges for rotem thromboelastometry. *Blood coagulation & fibrinolysis*, 16(4):301–310, 2005.
- [177] Robert D Langdell, Robert H Wagner, and Kenneth M Brinkhous. Effect of antihemophilic factor on one-stage clotting tests: a presumptive test for hemophilia and a simple one-stage antihemophilic factor assay procedure. *The Journal of laboratory and clinical medicine*, 41(4):637–647, 1953.
- [178] Rachel LeCover, Thomas Orfeo, Kathleen Brummel-Ziedins, Maria Bravo, Anthony Pusateri, and Jeffrey Varner. Kinetic modeling of coagulation and fibrinolysis. *IFAC-PapersOnLine*, 52(26):94–100, 2019.
- [179] Marcel Levi, Tymen T Keller, Eric van Gorp, and Hugo ten Cate. Infection and inflammation and the coagulation system. *Cardiovascular research*, 60(1):26–39, 2003.
- [180] Sumner N Levine. Enzyme amplifier kinetics. *Science*, 152(3722):651–653, 1966.
- [181] Dawei Li, Fang Fang, and Gang Zhao. Multiple trauma severity associated with levels of plasma thrombomodulin, activated protein c, and d-dimer. *Int J Clin Exp Med*, 13(3):1973–1978, 2020.
- [182] Kaiyuan Li, Huitao Wu, Fei Pan, Li Chen, Cong Feng, Yihao Liu, Hui Hui, Xiaoyu Cai, Hebin Che, Yulong Ma, et al. A machine learning-based model to predict acute traumatic coagulopathy in trauma patients upon emergency hospitalization. *Clinical and Applied Thrombosis/Hemostasis*, 26:1076029619897827, 2020.
- [183] H Lehman Li-wei, Shamim Nemati, Ryan P Adams, and Roger G Mark. Discovering shared dynamics in physiological signals: Application to patient monitoring in icu. In *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5939–5942. IEEE, 2012.
- [184] M. Lindauer, K. Eggenberger, M. Feurer, A. Biedenkapp, J. Marben, P. Müller, and F. Hutter. Boah: A tool suite for multi-fidelity bayesian optimization & analysis of hyperparameters. *arXiv:1908.06756 [cs.LG]*.

- [185] Zachary C Lipton, David C Kale, and Randall Wetzel. Modeling missing data in clinical time series with rnns. *arXiv preprint arXiv:1606.04130*, 2016.
- [186] M Kathryn Liszewski, Timothy C Farries, Douglas M Lublin, Isabelle A Rooney, and John P Atkinson. Control of the complement system. *Adv Immunol.*, 61:201–283, 1995.
- [187] Bing Liu, Jing Zhang, Pei Yi Tan, David Hsu, Anna M Blom, Benjamin Leong, Sunil Sethi, Bow Ho, Jeak Ling Ding, and PS Thiagarajan. A computational and experimental study of the regulatory mechanisms of the complement system. *PLoS Comput Biol*, 7(1):e1001059, 2011.
- [188] Nehemiah T Liu, John B Holcomb, Charles E Wade, Andriy I Batchinsky, Leopoldo C Cancio, Mark I Darrah, and José Salinas. Development and validation of a machine learning algorithm and hybrid system to predict the need for life-saving interventions in trauma patients. *Medical & biological engineering & computing*, 52(2):193–203, 2014.
- [189] Nehemiah T Liu and Jose Salinas. Machine learning for predicting outcomes in trauma. *Shock*, 48(5):504–510, 2017.
- [190] C Longstaff and M Locke. Increased urokinase and consumption of  $\alpha$ 2-antiplasmin as an explanation for the loss of benefit of tranexamic acid after treatment delay. *Journal of Thrombosis and Haemostasis*, 17(1):195–205, 2019.
- [191] Colin Longstaff and Craig Thelwell. Understanding the enzymology of fibrinolysis and improving thrombolytic therapy. *FEBS letters*, 579(15):3303–3309, 2005.
- [192] Deyan Luan, Fania Szlam, Kenichi A Tanaka, Philip S Barie, and Jeffrey D Varner. Ensembles of uncertain mathematical models can identify network response to therapeutic interventions. *Mol Biosyst*, 6(11):2272–86, Nov 2010.
- [193] Deyan Luan, Michael Zai, and Jeffrey D Varner. Computationally derived points of fragility of a human cascade are consistent with current therapeutic strategies. *PLoS computational biology*, 3(7):e142, 2007.
- [194] Deyan Luan, Michael Zai, and Jeffrey D Varner. Computationally derived points of fragility of a human cascade are consistent with current therapeutic strategies. *PLoS Comput Biol*, 3(7):e142, Jul 2007.



- [195] P Lukacik, P Roversi, J White, D Esser, GP Smith, J Billington, PA Williams, PM Rudd, MR Wormald, DJ Harvey, et al. Complement regulation at the molecular level: the structure of decay-accelerating factor. *Proc Natl Acad Sci USA.*, 101(5):1279–1284, 2004.
- [196] Jana BA MacLeod, Mauricio Lynn, Mark G McKenney, Stephen M Cohn, and Mary Murtha. Early coagulopathy predicts mortality in trauma. *Journal of Trauma and Acute Care Surgery*, 55(1):39–44, 2003.
- [197] Kenneth G Mann. Thrombin formation. *Chest*, 124(3):4S–10S, 2003.
- [198] Kenneth G Mann. Thrombin generation in hemorrhage control and vascular occlusion. *Circulation*, 124(2):225–235, 2011.
- [199] Kenneth G Mann, Kathleen Brummel-Ziedins, Thomas Orfeo, and Saulius Butenas. Models of blood coagulation. *Blood Cells, Molecules, and Diseases*, 36(2):108–117, 2006.
- [200] Kenneth G Mann, Saulius Butenas, and Kathleen Brummel. The dynamics of thrombin formation. *Arteriosclerosis, thrombosis, and vascular biology*, 23(1):17–25, 2003.
- [201] KG Mann, K Brummel, and S Butenas. What is all that thrombin for? *Journal of Thrombosis and Haemostasis*, 1(7):1504–1514, 2003.
- [202] WW Mapleson. An electric analogue for uptake and exchange of inert gases and other agents. *Journal of applied physiology*, 18(1):197–204, 1963.
- [203] F Martorana and A Moro. On the kinetics of enzyme amplifier systems with negative feedback. *Mathematical Biosciences*, 21(1-2):77–84, 1974.
- [204] Dimitrios C Mastellos, Despina Yancopoulou, Petros Kokkinos, Markus Huber-Lang, George Hajishengallis, Ali R Biglarnia, Florea Lupu, Bo Nilsson, Antonio M Risitano, Daniel Ricklin, and John D Lambris. Compstatin: a c3-targeted complement inhibitor reaching its prime for bedside intervention. *Eur J Clin Invest*, 45(4):423–40, Apr 2015.
- [205] Colin Mathers. 1updated who projectionsof mortality and causes of death 2016-2060.
- [206] Nena Matijevic, Yao-Wei W Wang, Charles E Wade, John B Holcomb, Bryan A Cotton, Martin A Schreiber, Peter Muskat, Erin E Fox, Debo-

- rah J Del Junco, Jessica C Cardenas, et al. Cellular microparticle and thrombogram phenotypes in the prospective observational multicenter major trauma transfusion (prommtt) study: correlation with coagulopathy. *Thrombosis research*, 134(3):652–658, 2014.
- [207] Paul L McCormack. Tranexamic acid. *Drugs*, 72(5):585–617, 2012.
- [208] Jonathan P Meizoso, Charles A Karcutskie, Juliet J Ray, Nicholas Namias, Carl I Schulman, and Kenneth G Proctor. Persistent fibrinolysis shutdown is associated with increased mortality in severely injured trauma patients. *Journal of the American College of Surgeons*, 224(4):575–582, 2017.
- [209] Michael A Meledeo, Maryanne C Herzig, James A Bynum, Xiaowu Wu, Anand K Ramasubramanian, Daniel N Darlington, Kristin M Reddoch, and Andrew P Cap. Acute traumatic coagulopathy: the elephant in a room of blind scientists. *Journal of Trauma and Acute Care Surgery*, 82(6S):S33–S40, 2017.
- [210] Joost PM Melis, Kristin Strumane, Sigrid R Ruuls, Frank J Beurskens, Janine Schuurman, and Paul WHI Parren. Complement in therapy and disease: Regulating the complement system with antibody-based therapeutics. *Mol Immunol.*, 67:117–130, 2015.
- [211] Pedro Mendes and D Kell. Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation. *Bioinformatics*, 14(10):869–883, 1998.
- [212] Zhi Hong Meng, Alisa S Wolberg, Dougald M Monroe III, and Maureane Hoffman. The effect of temperature and ph on the activity of factor viia: implications for the efficacy of high-dose factor viia in hypothermic and acidotic patients. *Journal of Trauma and Acute Care Surgery*, 55(5):886–891, 2003.
- [213] Gena Middleton and Bette Ruzevick. Alteplase (cathflo activase). *Clinical journal of oncology nursing*, 8(4):417–418, 2004.
- [214] Fiona Milne, Chris Redman, James Walker, Philip Baker, Julian Bradley, Carol Cooper, Michael de Swiet, Gillian Fletcher, Mervi Jokinen, Deirdre Murphy, et al. The pre-eclampsia community guideline (precog): how to screen for and detect onset of pre-eclampsia in the community. *Bmj*, 330(7491):576–580, 2005.

- [215] Riccardo Miotto, Fei Wang, Shuang Wang, Xiaoqian Jiang, and Joel T Dudley. Deep learning for healthcare: review, opportunities and challenges. *Briefings in bioinformatics*, 19(6):1236–1246, 2017.
- [216] Riccardo Miotto, Fei Wang, Shuang Wang, Xiaoqian Jiang, and Joel T Dudley. Deep learning for healthcare: review, opportunities and challenges. *Briefings in bioinformatics*, 19(6):1236–1246, 2018.
- [217] Alexander Y Mitrophanov, Fania Szlam, Roman M Sniecinski, Jerrold H Levy, and Jaques Reifman. Controlled multifactorial coagulopathy: Effects of dilution, hypothermia, and acidosis on thrombin generation in vitro. *Anesthesia & Analgesia*, 2019.
- [218] Alexander Y Mitrophanov, Alisa S Wolberg, and Jaques Reifman. Kinetic model facilitates analysis of fibrin generation and its modulation by clotting factors: implications for hemostasis-enhancing therapies. *Molecular BioSystems*, 10(9):2347–2357, 2014.
- [219] Charin Modchang, Wannapong Triampo, and Yongwimon Lenbury. Mathematical modeling and application of genetic algorithm to parameter estimation in signal transduction: Trafficking and promiscuous coupling of g-protein coupled receptors. *Computers in Biology and Medicine*, 38(5):574–582, 2008.
- [220] Patrick Kofod Mogensen and other contributors. Optim.jl. <https://julianlsolvers.github.io/Optim.jl/stable/#>, 2017.
- [221] Patrick Kofod Mogensen and Asbjørn Nilsen Riseth. Optim: A mathematical optimization package for Julia. *Journal of Open Source Software*, 3(24):615, 2018.
- [222] Carmen G Moles, Pedro Mendes, and Julio R Banga. Parameter estimation in biochemical pathways: a comparison of global optimization methods. *Genome research*, 13(11):2467–2474, 2003.
- [223] Francesco Montefusco, Ozgur E Akman, Orkun S Soyer, and Declan G Bates. Ultrasensitive negative feedback control: A natural approach for the design of synthetic controllers. *PLoS One*, 11(8):e0161605, 2016.
- [224] Hunter B Moore and Ernest E Moore. Temporal changes in fibrinolysis following injury. In *Seminars in Thrombosis and Hemostasis*, volume 46, pages 189–198. Thieme Medical Publishers, 2020.

- [225] Hunter B Moore, Ernest E Moore, Benjamin R Huebner, Monika Dzieciatkowska, Gregory R Stettler, Geoffrey R Nunns, Peter J Lawson, Arsen Ghasabyan, James Chandler, Anirban Banerjee, et al. Fibrinolysis shutdown is associated with a fivefold increase in mortality in trauma patients lacking hypersensitivity to tissue plasminogen activator. *The journal of trauma and acute care surgery*, 83(6):1014–1022, 2017.
- [226] Hunter B Moore, Ernest E Moore, Ioannis N Liras, Eduardo Gonzalez, John A Harvin, John B Holcomb, Angela Sauaia, and Bryan A Cotton. Acute fibrinolysis shutdown after injury occurs frequently and increases mortality: a multicenter evaluation of 2,540 severely injured patients. *Journal of the American College of Surgeons*, 222(4):347–355, 2016.
- [227] Hassan OJ Morad, Samuel C Belete, Thomas Read, and Andrew M Shaw. Time-course analysis of c3a and c5a quantifies the coupling between the upper and terminal complement pathways in vitro. *J Immunol Methods*, 427:13–18, 2015.
- [228] Jorge J Moré. The levenberg-marquardt algorithm: implementation and theory. In *Numerical analysis*, pages 105–116. Springer, 1978.
- [229] B Paul Morgan and Claire L Harris. Complement, a target for therapy in inflammatory and degenerative diseases. *Nat Rev Drug Discov*, 14:857–877, 2015.
- [230] Max D Morris. Factorial sampling plans for preliminary computational experiments. *Technometrics*, 33(2):161–174, 1991.
- [231] Hans J Muller-Eberhard. The membrane attack complex of complement. *Annual review of immunology*, 4(1):503–528, 1986.
- [232] Hans J Müller-Eberhard. Molecular organization and function of the complement system. *Annual review of biochemistry*, 57(1):321–347, 1988.
- [233] R Murray, J Spiegel, R Schiller, et al. Schaum’s outline of probability and statistics, 2013.
- [234] NJ Mutch, L Thomas, NR Moore, KM Lisiak, and NA Booth. Tafia, pai-1 and  $\alpha$ 2-antiplasmin: complementary roles in regulating lysis of thrombi and plasma clots. *Journal of Thrombosis and Haemostasis*, 5(4):812–817, 2007.

- [235] Lenny W Naftalin and John A Yagiela. Vasoconstrictors: indications and precautions. *Dental Clinics*, 46(4):733–746, 2002.
- [236] Satyaprakash Nayak, Saniya Salim, Deyan Luan, Michael Zai, and Jeffrey D Varner. A test of highly optimized tolerance reveals fragile cell-cycle mechanisms are molecular targets in clinical cancer trials. *PLoS One*, 3(4):e2016, 2008.
- [237] Matthew D Neal, Joshua B Brown, Ernest E Moore, Joseph Cuschieri, Ronald V Maier, Joseph P Minei, Timothy R Billiar, Andrew B Peitzman, Mitchell J Cohen, Jason L Sperry, et al. Pre-hospital use of non-steroidal anti-inflammatory drugs (nsaids) is associated with a reduced incidence of trauma-induced coagulopathy. *Annals of surgery*, 260(2):378, 2014.
- [238] John A Nelder and Roger Mead. A simplex method for function minimization. *The computer journal*, 7(4):308–313, 1965.
- [239] Shamim Nemati, Andre Holder, Fereshteh Razmi, Matthew D Stanley, Gari D Clifford, and Timothy G Buchman. An interpretable machine learning model for accurate prediction of sepsis in the icu. *Critical care medicine*, 46(4):547–553, 2018.
- [240] Michael Nesheim, Wei Wang, Michael Boffa, Mariko Nagashima, John Morser, and Laszlo Bajzar. Thrombin, thrombomodulin and tafi in the molecular link between coagulation and fibrinolysis. *Thrombosis and haemostasis*, 78(01):386–391, 1997.
- [241] Vance G Nielsen. A comparison of the thrombelastograph and the rotem. *Blood Coagulation & Fibrinolysis*, 18(3):247–252, 2007.
- [242] RE Nieman, DG Fisher, and DE Seborg. A review of process identification and parameter estimation techniques†. *International Journal of Control*, 13(2):209–264, 1971.
- [243] Sarah E Niles, Daniel F McLaughlin, Jeremy G Perkins, Charles E Wade, Yuanzhang Li, Philip C Spinella, and John B Holcomb. Increased mortality associated with the early coagulopathy of trauma in combat casualties. *Journal of Trauma and Acute Care Surgery*, 64(6):1459–1465, 2008.
- [244] Nasimul Noman and Hitoshi Iba. Inferring gene regulatory networks using differential evolution with local search heuristics. *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, 4(4):634–647, 2007.

- [245] Marina Noris, Miriam Galbusera, Sara Gastoldi, Paolo Macor, Federica Banterla, Elena Bresin, Claudio Tripodo, Serena Bettoni, Roberta Donadelli, Elisabetta Valoti, et al. Dynamics of complement activation in aHUS and how to monitor eculizumab therapy. *Blood*, pages 1715–1726, 2014.
- [246] G Nuttall. Experimente über die bacterienfeindlichen Einflüsse des thierischen Körpers. *Z. Hyg. Infektionskr.*, 4:353–394, 1888.
- [247] Alexander Olausson, Mark C Fitzgerald, Gim A Tan, and Biswadev Mitra. Cryoprecipitate administration after trauma. *European Journal of Emergency Medicine*, 23(4):269–273, 2016.
- [248] Terence O’Keeffe, Majed Refaai, Kathryn Tchorz, John E Forestner, and Ravi Sarode. A massive transfusion protocol to decrease blood component use and costs. *Archives of surgery*, 143(7):686–691, 2008.
- [249] Michael K Pangburn and Hans J Müller-Eberhard. The alternative pathway of complement. *Springer Semin Immunopathol*, 7:163–192, 1984.
- [250] Antonio Rafael Sabino Parmezan, GEAPA Batista, et al. Icmc-usp time series prediction repository. *Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, Brasil*. URL <https://goo.gl/uzxGZJ>, 2014.
- [251] Antonio Rafael Sabino Parmezan, Vinicius MA Souza, and Gustavo EAPA Batista. Evaluation of statistical and machine learning models for time series prediction: Identifying the state-of-the-art and the best conditions for the use of each model. *Information Sciences*, 484:302–337, 2019.
- [252] Henry Peng and Andrew Sweeny. Development of physiologically-based mathematical models for hemostatic resuscitation in trauma. 2016.
- [253] Matthew E Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. Deep contextualized word representations. *arXiv preprint arXiv:1802.05365*, 2018.
- [254] W Plasencia, N Maiz, S Bonino, C Kaihura, and KH Nicolaides. Uterine artery doppler at 11+ 0 to 13+ 6 weeks in the prediction of pre-eclampsia. *Ultrasound in Obstetrics and Gynecology*, 30(5):742–749, 2007.
- [255] L Poller, AMHP Van den Besselaar, J Jespersen, A Tripodi, D Houghton, and European Concerted Action on Anticoagulation. The effect of sample

- size on fresh plasma thromboplastin isi determination. *British journal of haematology*, 105(3):655–663, 1999.
- [256] Robert R Proctor and Samuel I Rapaport. The partial thromboplastin time with kaolin: a simple screening test for first stage plasma clotting factor deficiencies. *American Journal of Clinical Pathology*, 36(3):212–219, 1961.
  - [257] Yao Qin, Dongjin Song, Haifeng Chen, Wei Cheng, Guofei Jiang, and Garrison Cottrell. A dual-stage attention-based recurrent neural network for time series prediction. *arXiv preprint arXiv:1704.02971*, 2017.
  - [258] Micaela Reddy, RS Yang, Melvin E Andersen, and Harvey J Clewell III. *Physiologically based pharmacokinetic modeling: science and applications*. John Wiley & Sons, 2005.
  - [259] Christopher WG Redman, Gavin P Sacks, and Ian L Sargent. Preeclampsia: an excessive maternal inflammatory response to pregnancy. *American journal of obstetrics and gynecology*, 180(2):499–506, 1999.
  - [260] Andrew T Reisner and Thomas Heldt. A computational model of hemorrhage and dehydration suggests a pathophysiological mechanism: Starling-mediated protein trapping. *American Journal of Physiology-Heart and Circulatory Physiology*, 304(4):H620–H631, 2013.
  - [261] Peter Rhee, Bellal Joseph, Viraj Pandit, Hassan Aziz, Gary Vercruysse, Narong Kulvatunyou, and Randall S Friese. Increasing trauma deaths in the united states. *Annals of surgery*, 260(1):13–21, 2014.
  - [262] Nicklaus T Rice, Fania Szlam, Jeffrey D Varner, Peter S Bernstein, Arthur D Szlam, and Kenichi A Tanaka. Differential contributions of intrinsic and extrinsic pathways to thrombin generation in adult, maternal and cord plasma samples. *PLoS One*, 11(5):e0154127, 2016.
  - [263] Nicklaus T Rice, Fania Szlam, Jeffrey D Varner, Peter S Bernstein, Arthur D Szlam, and Kenichi A Tanaka. Differential contributions of intrinsic and extrinsic pathways to thrombin generation in adult, maternal and cord plasma samples. *PloS one*, 11(5), 2016.
  - [264] J. Richards, T. Gauss, and P. Bouzat. *Vasopressors for Post-traumatic Hemorrhagic Shock: Friends or Foe?*, pages 413–426. Springer International Publishing, Cham, 2020.

- [265] Daniel Ricklin, George Hajishengallis, Kun Yang, and John D Lambris. Complement: a key system for immune surveillance and homeostasis. *Nature immunology*, 11(9):785, 2010.
- [266] Daniel Ricklin and John D Lambris. Complement-targeted therapeutics. *Nat Biotechnol*, 25(11):1265–75, Nov 2007.
- [267] Daniel Ricklin and John D Lambris. Complement in immune and inflammatory disorders: pathophysiological mechanisms. *J Immunol*, 190(8):3831–3838, 2013.
- [268] Daniel Ricklin and John D Lambris. Progress and trends in complement therapeutics. *Adv Exp Med Biol*, 735:1–22, 2013.
- [269] Roger S Riley, David Rowe, and Lyman M Fisher. Clinical utilization of the international normalized ratio (inr). *Journal of clinical laboratory analysis*, 14(3):101–114, 2000.
- [270] Rebecca C Riley-Vargas, Darcy B Gill, Claudia Kemper, M Kathryn Liszewski, and John P Atkinson. Cd46: expanding beyond complement regulation. *Trends Immunol*, 25(9):496–503, 2004.
- [271] Daniel J Riskin, Thomas C Tsai, Loren Riskin, Tina Hernandez-Boussard, Maryanne Purtill, Paul M Maggio, David A Spain, and Susan I Brundage. Massive transfusion protocols: the role of aggressive resuscitation versus product ratio in mortality reduction. *Journal of the American College of Surgeons*, 209(2):198–205, 2009.
- [272] Daniel Rittirsch, Michael A Flierl, and Peter A Ward. Harmful molecular mechanisms in sepsis. *Nat Rev Immunol*, 8(10):776–787, 2008.
- [273] G E Rivard, K E Brummel-Ziedins, K G Mann, L Fan, A Hofer, and E Cohen. Evaluation of the profile of thrombin generation during the process of whole blood clotting as assessed by thrombelastography. *J Thromb Haemost*, 3(9):2039–43, Sep 2005.
- [274] I Roberts, H Shakur, T Coats, B Hunt, E Balogun, L Barnetson, L Cook, T Kawahara, P Perel, D Prieto-Merino, et al. The crash-2 trial: a randomised controlled trial and economic evaluation of the effects of tranexamic acid on death, vascular occlusive events and transfusion requirement in bleeding trauma patients. *Health technology assessment (Winchester, England)*, 17(10):1, 2013.



- [275] James M Roberts and Carl A Hubel. The two stage model of preeclampsia: variations on the theme. *Placenta*, 30:32–37, 2009.
- [276] Maria Rodriguez-Fernandez, Jose A Egea, and Julio R Banga. Novel meta-heuristic for parameter estimation in nonlinear dynamic biological systems. *BMC bioinformatics*, 7(1):483, 2006.
- [277] Michael Roguska, Igor Splawski, Beate Diefenbach-Streiber, Elizabeth Dolan, Bijan Etemad-Gilbertson, Jean-Michel Rondeau, and Mark Keating. Generation and Characterization of LFG316, A Fully-Human Anti-C5 Antibody for the Treatment of Age-Related Macular Degeneration. *IOVS*, 55:3433–3433, 2014.
- [278] Michael J Rohrer and ANITA M Natale. Effect of hypothermia on the coagulation cascade. *Critical care medicine*, 20(10):1402–1405, 1992.
- [279] Samuel W Ross, Bradley W Thomas, A Britton Christmas, Kyle W Cunningham, and Ronald F Sing. Returning from the acidotic abyss: Mortality in trauma patients with a ph; 7.0. *The American Journal of Surgery*, 214(6):1067–1072, 2017.
- [280] C Rourke, N Curry, S Khan, R Taylor, I Raza, R Davenport, S Stanworth, and K Brohi. Fibrinogen levels during trauma hemorrhage, response to replacement therapy, and association with patient outcomes. *Journal of Thrombosis and Haemostasis*, 10(7):1342–1351, 2012.
- [281] Peter J Rousseeuw. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65, 1987.
- [282] Adithya Sagar, Wei Dai, Mason Minot, Rachel LeCover, and Jeffrey D Varner. Reduced order modeling and analysis of the human complement system. *PloS one*, 12(11):e0187373, 2017.
- [283] Adithya Sagar, Rachel LeCover, Christine Shoemaker, and Jeffrey Varner. Dynamic optimization with particle swarms (dops): A meta-heuristic for parameter estimation in biochemical models. *BMC systems biology*, 12(1):87, 2018.
- [284] Adithya Sagar, Christine A. Shoemaker, and J. Varner. Dynamic Optimization with Particle Swarms (DOPS): A meta- heuristic for parameter estimation in biochemical models. *Biotechnol. J*, submitted., 2016.

- [285] Adithya Sagar and Jeffrey D Varner. Dynamic modeling of the human coagulation cascade using reduced order effective kinetic models. *Processes*, 3(1):178–203, 2015.
- [286] Adithya Sagar and Jeffrey D. Varner. Dynamic modeling of the human coagulation cascade using reduced order effective kinetic models. *Processes*, 3(1):178, 2015.
- [287] Andrea Saltelli, Paola Annoni, Ivano Azzini, Francesca Campolongo, Marco Ratto, and Stefano Tarantola. Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Comput Phys Commun*, 181(2):259–270, 2010.
- [288] J Vidya Sarma and Peter A Ward. The complement system. *Cell and tissue research*, 343(1):227–235, 2011.
- [289] Hajime Satoh, Fumiki Kushihata, Masahide Hatano, Jota Watanabe, and Yasutsugu Takada. Clinical significance of soluble fibrin in coagulopathy caused by highly invasive surgery. *International Surgery*, 2020.
- [290] A Sauaia, F A Moore, E E Moore, J B Haenel, R A Read, and D C Lezotte. Early predictors of postinjury multiple organ failure. *Arch Surg*, 129(1):39–45, Jan 1994.
- [291] A Sauaia, F A Moore, E E Moore, K S Moser, R Brennan, R A Read, and P T Pons. Epidemiology of trauma deaths: a reassessment. *J Trauma*, 38(2):185–93, Feb 1995.
- [292] S Scarpelini, SG Rhind, B Nascimento, H Tien, PN Shek, HT Peng, H Huang, R Pinto, V Speers, M Reis, et al. Normal range values for thromboelastography in healthy adult volunteers. *Brazilian journal of medical and biological research*, 42(12):1210–1217, 2009.
- [293] Alicia M Schiller, Jeffrey T Howard, and Victor A Convertino. The physiology of blood loss and shock: New insights from a human laboratory model of hemorrhage. *Experimental Biology and Medicine*, 242(8):874–883, 2017.
- [294] Christoph J Schlimp, Martin Ponschab, Wolfgang Voelckel, Benjamin Treichl, Marc Maegele, and Herbert Schöchl. Fibrinogen levels in trauma patients during the first seven days after fibrinogen concentrate therapy: a retrospective study. *Scandinavian journal of trauma, resuscitation and emergency medicine*, 24(1):29, 2016.

- [295] David E Schmidt, Märit Halmin, Agneta Wikman, Anders Östlund, and Anna gren. Relative effects of plasma, fibrinogen concentrate, and factor xiii on rotem coagulation profiles in an in vitro model of massive transfusion in trauma. *Scandinavian journal of clinical and laboratory investigation*, 77(6):397–405, 2017.
- [296] Herbert Schöchl, Ulrike Nienaber, Georg Hofer, Wolfgang Voelckel, Csilla Jambor, Gisela Scharbert, Sibylle Kozek-Langenecker, and Cristina Solomon. Goal-directed coagulation management of major trauma patients using thromboelastometry (rotem®)-guided administration of fibrinogen concentrate and prothrombin complex concentrate. *Critical care*, 14(2):R55, 2010.
- [297] Y Schrauwen, REM De Vries, T Kooistra, and JJ Emeis. Acute release of tissue-type plasminogen activator (t-pa) from the endothelium; regulatory mechanisms and therapeutic target. *Fibrinolysis*, 8:8–12, 1994.
- [298] Hans-Wilhelm Schwaeble, Cordula Margaret Stover, Clark E Tedford, James B Parent, and Teizo Fujita. Methods for treating conditions associated with masp-2 dependent complement activation. US Patent 7,919,094, April 2011.
- [299] Jansen N Seheult, Michelle N Stram, Jason Sperry, Philip C Spinella, Darrell J Triulzi, and Mark H Yazer. In silico model of the dilutional effects of conventional component therapy versus whole blood in the management of massively bleeding adult trauma patients. *Transfusion*, 2018.
- [300] Margrethe H Serres, Shuba Gopal, Laila A Nahum, Ping Liang, Terry Gaasterland, and Monica Riley. A functional update of the escherichia coli k-12 genome. *Genome biology*, 2(9):research0035–1, 2001.
- [301] LT Sharief, AS Lawrie, IJ Mackie, C Smith, F Peyvandi, and RA Kadir. Changes in factor xiii level during pregnancy. *Haemophilia*, 20(2):e144–e148, 2014.
- [302] Liang Shen, Sheida Tabaie, and Natalia Ivascu. Viscoelastic testing inside and beyond the operating room. *Journal of thoracic disease*, 9(Suppl 4):S299, 2017.
- [303] Forest R Sheppard, Thomas A Mitchell, Andrew P Cap, Leasha J Schaub, Antoni R Macko, and Jacob J Glaser. Prehospital whole blood resuscitation prevents coagulopathy and improves acid–base status at hospital

arrival in a nonhuman primate hemorrhagic shock model. *Transfusion*, 2019.

- [304] Forest R Sheppard, Thomas A Mitchell, Antoni R Macko, Darren M Fryer, Leasha J Schaub, Kassandra M Ozuna, and Jacob J Glaser. Whole blood and hextend: Bookends of modern tactical combat casualty care field resuscitation and starting point for multifunctional resuscitation fluid development. *Journal of Trauma and Acute Care Surgery*, 85(1S):S33–S38, 2018.
- [305] Yuhui Shi and Russell C Eberhart. Empirical study of particle swarm optimization. In *Proceedings of the 1999 Congress on Evolutionary Computation*, 1999. CEC 99., volume 3, page 1950, 1999.
- [306] ML Shwartz, SV Pizzo, RL Hill, and PA McKee. Human factor xiii from plasma and platelets. *J. Biol. Chem*, 248:1395–1407, 1973.
- [307] Baha Sibai, Gus Dekker, and Michael Kupferminc. Pre-eclampsia. *The Lancet*, 365(9461):785–799, 2005.
- [308] MMCG Silva, C Thelwell, SC Williams, and C Longstaff. Regulation of fibrinolysis by c-terminal lysines operates through plasminogen and plasmin but not tissue-type plasminogen activator. *Journal of Thrombosis and Haemostasis*, 10(11):2354–2360, 2012.
- [309] SH Simpson, G Menezes, SN Mardel, S Kelly, R White, and T Beattie. A computer model of major haemorrhage and resuscitation. *Medical engineering & physics*, 18(4):339–343, 1996.
- [310] JG Sissons, J Liebowitch, N Amos, and DK Peters. Metabolism of the fifth component of complement, and its relation to metabolism of the third component, in patients with complement activation. *J Clin Invest.*, 59(4):704, 1977.
- [311] Kieran Smallbone and Pedro Mendes. Large-scale metabolic models: From reconstruction to differential equations. *Industrial Biotechnology*, 9(4):179–184, 2013.
- [312] I.M Sobol. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math Comput Simulat*, 55:271 – 280, 2001.
- [313] Justin Sobrino and Shahid Shafi. Timing and causes of death after injuries.

In *Baylor University Medical Center Proceedings*, volume 26, pages 120–123. Taylor & Francis, 2013.

- [314] Cristina Solomon, Marco Ranucci, Gerald Hochleitner, Herbert Schöchl, and Christoph J Schlimp. Assessing the methodology for calculating platelet contribution to clot strength (platelet component) in thromboelastometry and thrombelastography. *Anesthesia and analgesia*, 121(4):868, 2015.
- [315] Priya Soma-Pillay, Nelson-Piercy Catherine, Heli Tolppanen, Alexandre Mebazaa, Heli Tolppanen, and Alexandre Mebazaa. Physiological changes in pregnancy. *Cardiovascular journal of Africa*, 27(2):89, 2016.
- [316] Philip C Spinella and John B Holcomb. Resuscitation and transfusion principles for traumatic hemorrhagic shock. *Blood reviews*, 23(6):231–240, 2009.
- [317] AE St John, JC Newton, EJ Martin, BM Mohammed, D Contaifer Jr, JL Saunders, GM Brophy, BD Spiess, KR Ward, DF Brophy, et al. Platelets retain inducible alpha granule secretion by p-selectin expression but exhibit mechanical dysfunction during trauma-induced coagulopathy. *Journal of thrombosis and haemostasis: JTH*, 2019.
- [318] David B Staab, Victor J Sorensen, John J Fath, SB Raman, H Matilda Horst, and Farouck N Obeid. Coagulation defects resulting from ambient temperature-induced hypothermia. *The Journal of trauma*, 36(5):634–638, 1994.
- [319] Eric AP Steegers, Peter Von Dadelszen, Johannes J Duvekot, and Robert Pijnenborg. Pre-eclampsia. *The Lancet*, 376(9741):631–644, 2010.
- [320] H. B. Stoner. *Thermoregulation after Injury*, pages 495–499. Springer US, Boston, MA, 1973.
- [321] Rainer Storn and Kenneth Price. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optimization*, 11(4):341–359, 1997.
- [322] T Sumner, E Shephard, and IDL Bogle. A methodology for global-sensitivity analysis of time-dependent outputs in systems biology modelling. *Journal of The Royal Society Interface*, 9(74):2156–2166, 2012.

- [323] Jianyong Sun, Jonathan M Garibaldi, and Charlie Hodgman. Parameter estimation using metaheuristics in systems biology: a comprehensive review. *Computational Biology and Bioinformatics, IEEE/ACM Transactions on*, 9(1):185–202, 2012.
- [324] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*, pages 3104–3112, 2014.
- [325] AJG Swaak, A Hannema, C Vogelaar, FA Boom, L van Es, R van Aalst, and LW Statius van Eps. Determination of the half-life of c3 in patients and its relation to the presence of c3-breakdown products and/or circulating immune complexes. *Rheumatol Int.*, pages 161–166, 1982.
- [326] Kamilla Swiech, Virgínia Picanço-Castro, and Dimas Tadeu Covas. Production of recombinant coagulation factors: Are humans the best host cells? *Bioengineered*, 8(5):462–470, 2017.
- [327] Douglas A Talbert, Matt Honeycutt, and Steve Talbert. A machine learning and data mining framework to enable evolutionary improvement in trauma triage. In *International Workshop on Machine Learning and Data Mining in Pattern Recognition*, pages 348–361. Springer, 2011.
- [328] Katerina Tashkova, Peter Korošec, Jurij Šilc, Ljupčo Todorovski, and Sašo Džeroski. Parameter estimation with bio-inspired meta-heuristic optimization: modeling the dynamics of endocytosis. *BMC systems biology*, 5(1):159, 2011.
- [329] Ryan Tasseff, Satyaprakash Nayak, Saniya Salim, Poorvi Kaushik, Noreen Rizvi, and Jeffrey D Varner. Analysis of the molecular networks in androgen dependent and independent prostate cancer revealed fragile and robust subsystems. *PLoS One*, 5(1):e8864, 2010.
- [330] Ryan Tasseff, Satyaprakash Nayak, Sang Ok Song, Andrew Yen, and Jeffrey D Varner. Modeling and analysis of retinoic acid induced differentiation of uncommitted precursor cells. *Integr Biol (Camb)*, 3(5):578–91, May 2011.
- [331] Torsten Teorell. Kinetics of distribution of substances administered to the body, i: the extravascular modes of administration. *Archives internationales de pharmacodynamie et de therapie*, 57:205–225, 1937.

- [332] Bryan A Tolson and Christine A Shoemaker. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Resources Research*, 43(1), 2007.
- [333] Bryan A. Tolson and Christine A. Shoemaker. Dynamically dimensioned search algorithm for computationally efficient watershed model calibration. *Water Res Research*, 43(1):W01413, 2007. W01413.
- [334] Angela B Treml, Jed B Gorlin, Richard P Dutton, and Barbara M Scavone. Massive transfusion protocols: a survey of academic medical centers in the united states. *Anesthesia & Analgesia*, 124(1):277–281, 2017.
- [335] Donald D Trunkey and Robert C Lim. Analysis of 425 consecutive trauma fatalities: an autopsy study. *Journal of the American College of Emergency Physicians*, 3(6):368–371, 1974.
- [336] Kuan-Yao Tsai and Feng-Sheng Wang. Evolutionary optimization with data collocation for reverse engineering of biological networks. *Bioinformatics*, 21(7):1180–1188, 2005.
- [337] Apple P Valparaíso, Diego A Vicente, Benjamin A Bograd, Eric A Elster, and Thomas A Davis. Modeling acute traumatic injury. *Journal of surgical research*, 194(1):220–232, 2015.
- [338] Natal A W van Riel. Dynamic modelling and analysis of biochemical networks: mechanism-based models and model-based experiments. *Brief Bioinform*, 7(4):364–74, Dec 2006.
- [339] Varnerlab. <http://www.varnerlab.org>.
- [340] Varnerlab. JuPOETs: A Constrained Multiobjective Optimization Approach to Estimate Biochemical Model Ensembles in the Julia Programming Language.
- [341] Christopher Vernon and Jennifer L LeTourneau. Lactic acidosis: recognition, kinetics, and associated prognosis. *Critical care clinics*, 26(2):255–283, 2010.
- [342] Alejandro F Villaverde, Sophia Bongard, Klaus Mauch, Dirk Müller, Eva Balsa-Canto, Joachim Schmid, and Julio R Banga. High-confidence predictions in systems biology dynamic models. *8th International Conference*

on *Practical Applications of Computational Biology & Bioinformatics (PACBB 2014)*, pages 161–171, 2014.

- [343] Alejandro F Villaverde, Jose A Egea, and Julio R Banga. A cooperative strategy for parameter estimation in large scale systems biology models. *BMC systems biology*, 6(1):75, 2012.
- [344] Alejandro F Villaverde, David Henriques, Kieran Smallbone, Sophia Bongard, Joachim Schmid, Damjan Cicin-Sain, Anton Crombach, Julio Saez-Rodriguez, Klaus Mauch, Eva Balsa-Canto, et al. Biopredyn-bench: a suite of benchmark problems for dynamic modelling in systems biology. *BMC systems biology*, 9(1):8, 2015.
- [345] Carl-Wilhelm Vogel, David C Fritzinger, Brian E Hew, Mike Thorne, and Holger Bammert. Recombinant cobra venom factor. *Molecular immunology*, 41:191–199, 2004.
- [346] Erwin A Vogler and Christopher A Siedlecki. Contact activation of blood-plasma coagulation. *Biomaterials*, 30(10):1857–1869, 2009.
- [347] Kelly C Vranas, Jeffrey K Jopling, Timothy E Sweeney, Meghan C Ramsey, Arnold S Milstein, Christopher G Slatore, Gabriel J Escobar, and Vincent X Liu. Identifying distinct subgroups of intensive care unit patients: a machine learning approach. *Critical care medicine*, 45(10):1607, 2017.
- [348] DG Walker, O Yasuhara, PA Patston, EG McGeer, and PL McGeer. Complement c1 inhibitor is produced by brain tissue and is cleaved in alzheimer disease. *Brain Res.*, 675(1):75–82, 1995.
- [349] M J Walport. Complement. first of two parts. *N Engl J Med*, 344(14):1058–66, Apr 2001.
- [350] M J Walport. Complement. second of two parts. *N Engl J Med*, 344(15):1140–4, Apr 2001.
- [351] Mark J Walport. Complement. *New England Journal of Medicine*, 344(14):1058–1066, 2001.
- [352] Feng-Sheng Wang, Tzu-Liang Su, and Horng-Jhy Jang. Hybrid differential evolution for problems of kinetic parameter estimation and dynamic optimization of an ethanol fermentation process. *Industrial & engineering chemistry research*, 40(13):2876–2885, 2001.



- [353] Zhong Wang, Mark Gerstein, and Michael Snyder. Rna-seq: a revolutionary tool for transcriptomics. *Nature reviews genetics*, 10(1):57–63, 2009.
- [354] Jennifer M Watters, Brandon H Tieu, S Rob Todd, Terisa Jackson, Patrick J Muller, Darren Malinoski, and Martin A Schreiber. Fluid resuscitation increases inflammatory gene transcription after traumatic injury. *Journal of Trauma and Acute Care Surgery*, 61(2):300–309, 2006.
- [355] Carl P Weiner and John Brandt. Plasma antithrombin iii activity: an aid in the diagnosis of preeclampsia-eclampsia. *American journal of obstetrics and gynecology*, 142(3):275–281, 1982.
- [356] Gary E Weissman, Rebecca A Hubbard, Lyle H Ungar, Michael O Harhay, Casey S Greene, Blanca E Himes, and Scott D Halpern. Inclusion of unstructured clinical text improves early prediction of death or prolonged icu stay. *Critical care medicine*, 46(7):1125, 2018.
- [357] Wynne H Weston-Davies, Miles A Nunn, Fernando O Pinto, Ian J Mackie, Stephen John Richards, Samuel J Machin, Raymond Prudo, and Peter Hillmen. Clinical and immunological characterisation of coversin, a novel small protein inhibitor of complement C5 with potential as a therapeutic agent in PNH and other complement mediated disorders. *Blood*, 124:4280–4280, 2014.
- [358] Patrick Wettstein, André Haeberli, Monika Stutz, Miriam Rohner, Cinzia Corbetta, Konrad Gabi, Thomas Schnider, and Wolfgang Korte. Decreased factor xiii availability for thrombin and early loss of clot firmness in patients with unexplained intraoperative bleeding. *Anesthesia & Analgesia*, 99(5):1564–1569, 2004.
- [359] Matthew F Whelihan, Vicentios Zachary, Thomas Orfeo, and Kenneth G Mann. Prothrombin activation in blood coagulation: the erythrocyte contribution to thrombin generation. *Blood*, 120(18):3837–45, Nov 2012.
- [360] David Whiting and James A DiNardo. Teg and rotem: technology and clinical applications. *American journal of hematology*, 89(2):228–232, 2014.
- [361] E B Williams, S Krishnaswamy, and K G Mann. Zymogen/enzyme discrimination using peptide chloromethyl ketones. *J Biol Chem*, 264(13):7536–45, May 1989.
- [362] JA Williams-Johnson, AH McDonald, G Gordon Strachan, and EW Williams. Effects of tranexamic acid on death, vascular occlusive

events, and blood transfusion in trauma patients with significant haemorrhage (crash-2): a randomised, placebo-controlled trial. *West Indian Medical Journal*, 59(6):612–624, 2010.

- [363] David S Wishart, Yannick D Feunang, An C Guo, Elvis J Lo, Ana Marcu, Jason R Grant, Tanvir Sajed, Daniel Johnson, Carin Li, Zinat Sayeeda, et al. Drugbank 5.0: a major update to the drugbank database for 2018. *Nucleic acids research*, 46(D1):D1074–D1082, 2018.
- [364] Alisa S Wolberg, Zhi Hong Meng, Dougald M Monroe III, and Maureane Hoffman. A systematic evaluation of the effect of temperature on coagulation enzyme activity and platelet function. *Journal of Trauma and Acute Care Surgery*, 56(6):1221–1228, 2004.
- [365] Guojie Wu, Blake A Mazzitelli, Adam J Quek, Matthew J Veldman, Paul J Conroy, Tom T Caradoc-Davies, Lisa M Ooms, Kellie L Tuck, Jonathan G Schoenecker, James C Whisstock, et al. Tranexamic acid is an active site inhibitor of urokinase plasminogen activator. *Blood advances*, 3(5):729, 2019.
- [366] Wenzhong Xiao, Michael N Mindrinos, Junhee Seok, Joseph Cuschieri, Alex G Cuenca, Hong Gao, Douglas L Hayden, Laura Hennessy, Ernest E Moore, Joseph P Minei, et al. A genomic storm in critically injured humans. *Journal of Experimental Medicine*, 208(13):2581–2590, 2011.
- [367] Sumith Yesudasan and Rodney D Averett. Recent advances in computational modeling of fibrin clot formation: A review. *Computational Biology and Chemistry*, page 107148, 2019.
- [368] James G Young and Irving Gray. Biochemical response to trauma: Iii. epinephrine and norepinephrine levels in plasma of rats subjected to tumbling trauma. *American Journal of Physiology-Legacy Content*, 186(1):67–70, 1956.
- [369] Peter Young. Parameter estimation for continuous-time models—a survey. *Automatica*, 17(1):23–39, 1981.
- [370] Nehemiah Zewde, Ronald D Gorham Jr, Angel Dorado, and Dimitrios Morikis. Quantitative modeling of the alternative pathway of the complement system. *PloS one*, 11(3):e0152337, 2016.
- [371] Shi-Zheng Zhao, Jing J Liang, Ponnuthurai N Suganthan, and Mehmet Fatih Tasgetiren. Dynamic multi-swarm particle swarm opti-

mizer with local search for large scale global optimization. *Evolutionary Computation*, 2008. CEC 2008, pages 3845–3852, 2008.

- [372] Zilong Zhao, Min Wang, Ye Tian, Tristan Hilton, Breia Salsbery, Eric Z Zhou, Xiaoping Wu, Perumal Thiagarajan, Eric Boilard, Min Li, et al. Cardiolipin-mediated procoagulant activity of mitochondria contributes to traumatic brain injury–associated coagulopathy in mice. *Blood, The Journal of the American Society of Hematology*, 127(22):2763–2772, 2016.