in silico Algorithm Advancements for Cancer Chemotherapy Treatment and Coagulopathy Diagnosis

by

Joseph T. Liparulo

B.S. Chemical and Petroleum Engineering, University of Pittsburgh, 2019

Submitted to the Graduate Faculty of

the Swanson School of Engineering in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2022

UNIVERSITY OF PITTSBURGH

SWANSON SCHOOL OF ENGINEERING

This thesis was presented

by

Joseph T. Liparulo

It was defended on

November 10, 2022

and approved by

Jason Shoemaker, Ph.D., Professor, Department of Chemical and Petroleum Engineering

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

Thesis Advisor:

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

***in silico* Algorithm Advancements for Cancer Chemotherapy Treatment and Coagulopathy Diagnosis**

Joseph T. Liparulo, M.S.

University of Pittsburgh, 2022

The interdisciplinary field of systems medicine fills a gap in clinical translation with computational models to help inform decisions made by physicians. This approach takes advantage of modern computational methods to examine interactions between components within the models to predict the cause-and-effect relationship between systems and potential treatment strategies. These mathematical models typically incorporate various pharmacokinetic and pharmacodynamic models to help personalize medicine to individual patients, rather than a 'one-size-fits-all' mentality. The research herein focuses on improving chemotherapy treatment as well as the prediction of coagulopathies in trauma patients.

The general use of mathematical models for cancer chemotherapy treatment has focused primarily on tumor kill while using constraints on dose magnitude to explicitly mitigate toxicity. By incorporating pharmacodynamic toxicity models in this work into chemotherapy treatment schedule design, the physician is able to specify toxicity explicitly. The pharmacokinetic model of drug distribution throughout the body and pharmacodynamic models of both the antitumor efficacy and drug toxicity are incorporated into optimization of dose scheduling. The performance results of an optimal schedule under clinical constraints are clinically indiscernible from the mathematically-optimal solutionUsing nonlinear least squares and clinical measurements, actual patient toxicity/tumor sensitivities can be calculated and the optimal schedule updated. In a clinical setting this algorithm would enable the clinician to prioritize patient quality-of-life through the minimization of individual toxicity while maximizing tumor eradication.

Predicting trauma patient coagulopathies presents a challenge to clinicians; the patient may experience excessive or insufficient clotting and it is imperative that the correct intervention be given as soon as possible. By implementing models of the coagulation cascade with clinical assays such as the thromboelastogram (TEG), it is possible to determine sub-

populations of patients with different coagulopathies, therefore allowing the clinician to make more informed decisions much faster than using the TEG tracing alone. By decreasing the time required to identify abnormal lysis by approximately one-half, the risk of mortality and permanent damage for the patients could be decreased significantly.

Table of Contents

List of Tables

List of Figures

Preface

The journey to completing my education has been exciting to say the least. There was once a time where I did not think I would complete my undergraduate education, but thanks to determination, support, and the team at UPMC, I was able to. I never thought graduate school would be for me, but I am glad to have filled in some of the gaps that were left after undergrad. This work is not just the result of my work, but also the unconditional support from my family: my parents always pushed me to be the best person I could, no matter how difficult the situation and I feel this gave me the perseverance to complete my education. I am so grateful that my decision to attend graduate school introduced me to my fiancée Tatum, who has always been in my corner, making the most difficult times more bearable. Without your support and sunny demeanor, it would have been infinitely more difficult to get through this. My advisor Bob has always pushed me, expecting more than I thought I was capable of, which pushed me harder to excel. I am grateful for all of this.

1.0 Model Based Control in a Clinical Setting

Model based control has been implemented and theorized for many applications including disease transmission, weight control, diabetes management, and treatment improvement. Through the combination of mathematical models of complex biological processes and model-based frameworks, it is possible to optimize and predict treatments based on actual patient inputs [20, 33, 66, 26, 46]. In the more than 40 years since this method has been proposed for use in medicine, the models developed are significantly more accurate and the computational efficiency required to implement them has drastically improved, thanks to the increase in available data and ever-evolving computational systems [64]. This led to the implementation of more complex models that not only more accurately represent the actual process, but also increase the precision and usability of the models. In order to be implemented in a clinical setting, the model-based control must be computationally efficient as to expedite results in time-sensitive scenarios and provide clinicians with inputs and outputs that are easily understood, to avoid false results and interpretations. The work herein presents the use of model-based control for use in two different situations: cancer chemotherapy treatment optimization and coagulopathy prediction (specifically abnormal lysis rate) in trauma patients.

1.1 Cancer Chemotherapy Treatment

Cancer is a disease that is characterized by the uncontrolled growth and spread of malignant cells; more than 1.9 million new cases and 609,000 deaths are expected in 2022 as a direct result of this disease [58]. The cause of death in cancer is either growth of the primary tumor or metastasis, which is the translocation of cancer cells to a different area of the body away from the primary tumor; unfortunately, the exact mechanism behind this process is still unidentified [22]. Depending on the type and location of the primary tumor, the treatment strategies vary. If possible, the tumor will be surgically removed, but if the location is

inoperable (*e.g.* the brain) or the loss of organ function unsustainable (*e.g.* the pancreas), surgical excision is not possible. For non-localized cancers such as leukemia, surgery is impossible. Treatment options in these cases can include radiotherapy, where the patient is given localized high doses of radiation to kill tumor cells and shrink the tumor, and biologic treatments such as immunotherapy which helps the patient's immune system fight the cancer. However, metastases undetectable through imagine are likely to be present once the primary tumor is detected, which then requires systemic treatment such as chemotherapy.

One of the primary chemotherapeutic agents that is used in both single- and multi-agent cancer treatment is docetaxel [61, 47]. Chemotherapy agents work by targeting rapidly dividing cells, such as cancer cells; the selectivity of chemotherapeutics toward cancer cells is always desired, but not always achieved, which can lead to patient toxicities. A common chemotherapy treatment toxicity is depletion of white blood cells. This creates a dichotomy for clinicians in balancing maximal elimination of diseased cells, while at the same time minimizing the toxic side-effects of chemotherapy when selecting a dose schedule (defined by dose magnitude and frequency) that considers these two competing sides. Currently, these clinical schedules are guided by empirical evidence from preclinical trials and then refined through clinical testing. Model-based optimization methods provide insight to this treatment balancing-act though the use of mathematical models.

The work herein is composed of three primary components: pharmacokinetic (PK) models, pharmacodynamic (PD) models, and treatment algorithms. These three components work in tandem to help determine the optimal dosing schedule. PK models are used to describe the drug distribution and concentrations throughout the body, while PD models use these informed concentrations to represent the subsequent effects of the drug concentrations. The treatment algorithms are used to standardize the selection of patient care.

PK models are composed of theoretical or physiological compartments, in which the drug concentration cascades through the different compartments of the body [13, 35]. The primary types of these models are noncompartmental, where there is basically a 'black-box' in which the drug enters and is cleared; compartmental, with the number of compartments ranging from one to many; and physiological, in which the compartments are representative of actual parts of the body. A one-compartment model implies that the drug enters from

outside the body to the central compartment, where it then leaves this central compartment (*i.e.* the body) and drug recirculation does not occur once it leaves the body. A multi-compartment model also assumes the drug enters the central compartment from outside the body, where it then leaves this compartment by either leaving the body, or traveling to the next compartment; this can continue for the number of compartments that are implemented. Once the drug enters these peripheral compartments it can travel back to the primary compartment where the drug originated, but once it leaves the body it does not recirculate. Physiologically-based (PB) multicompartment PK models are becoming more preferred over one-compartment modes for their reliability and usefulness following the advent of increasing computing power [67, 4].

After the concentration of drug is determined through the use of PK models, PD models are used to quantify the drug's therapeutic and/or toxic effect on different parts of the body of interest. Like PK models, these are typically compartmental models as well and are usually one or more of the following types: direct effect, biophase distribution, and indirect response [69]. The direct effect model first observed by Levy describes the direct effect of the drug when there is no time delay between plasma concentration and response [38]. However, there is typically a delay between plasma concentration and response, which requires either a biophase or indirect response model. The biophase models attribute the time delay between drug concentration and therapeutic effect to the time it takes for the plasma concentration of the drug to distribute to the target [54, 12]. Alternatively, indirect response models are used to describe the delayed response caused by an indirect mechanism, *i.e.* the drug could inhibit/stimulate the dissipation/production of drug response which causes the delay [53, 36].

By combining these types of models with treatment algorithms, a clinician is able to make a more informed decision regarding the best course of care for the patient. These algorithms come in many forms including computational tools, but one of the most commonly used is flowcharts. Flowcharts are an easy tool that provide a clinician with guidance on what steps should be taken based on the patient's current condition including the timing of events and can be implemented for a myriad of conditions ranging from hypertension to COVID-19 treatment [17, 18]. With the increasing computational power that is becoming readily available more computational treatment algorithms are being used that implement machine

learning and artificial intelligence, coupled with mathematical models [32].

By targeting tumor growth (and elimination of cancerous cells) and fitting model parameters to actual patient data, several model-based methods have been implemented to determine optimal treatment schedules [16, 43]. These optimal control methods focus on minimizing tumor volume only at a final, fixed time point for treatment by imposing constraints on dose magnitude [37, 41, 60]. However, actual patient treatment does not have a final, fixed time point defined *a priori*. In practice, near real-time or cycle-wise decisions are made based on patient response to drug toxicity and efficacy which often change over time, in a specific patient and across other patients, which leads to the treatment endpoint or remission not being predictable. For this reason, treatment is given in cycles to let clinicians evaluate patient response and use feedback to adjust the treatment strategy as necessary. Here, a model-based framework is developed that would allow clinicians to balance toxic side-effects with drug efficacy *i.e.* tumor shrinkage.

In a previous treatment algorithm based on preclinical animal models, the problem was solved using mixed integer linear programming (MILP) [28, 25]. The work herein employs physiologically-based nonlinear pharmacokinetic (PK), pharmacodynamic (PD), and toxicity models within a receding-horizon based treatment design framework similar to model predictive control (MPC). The toxicity limits and clinical logistics are implemented explicitly as constraints. The tumor volume is minimized over the treatment horizon of individual or multiple treatment cycles in accordance with the aforementioned constraints. Since patient response in toxicity and efficacy can change over time, and the tumor volume and toxicity measurements are taken after every other cycle and each cycle, respectively, these are used to update the patient-specific model as measurements become available. The algorithm provides a rigorous model-based approach that has the potential to minimize the dichotomy experienced by clinicians in balancing toxicity and efficacy, ultimately providing improved patient quality-of-life.

1.2 Coagulopathy Prediction

For middle aged and younger Americans, trauma is the second leading cause of death. This creates challenges in a clinical setting, particularly in the case of coagulopathies (a condition where the blood's ability to clot is affected, leading to either a hypo- or hypercoagulable state) [50]. Since about one-third of all trauma patients experience some type of clotting disorder, this creates an area of potential improvement in personalized and expedited intervention strategies [40]. These strategies include targeted fluid resuscitation which would provide coagulopathic patients with the necessary blood components to restore normal clotting function [19, 21]. Viscoelastic assays, such as the thromboelastogram (TEG), are being used in clinical settings where the patient may be at risk for coagulopathy. This device works by adding a small sample of blood (and possibly a reagent such as kaolin during the use of a rapid TEG to speed up initiation of the clot) to an oscillating stand that simulates sluggish blood flow and driving thrombus formation, which is measured by the displacement of a torsion pin [24]. There are several extractable parameters from the TEG: R-time (time until clot activation), activated clotting time (ACT, transformed from R-time), K-time (time until the TEG reaches 20mm worth of clot strength), α-angle (angle between horizontal and tangent line to the TEG, which indicates the speed of clot formation), maximum amplitude (MA, which indicates maximum clot strength), and lysis rate (LY, which indicates the percentage decrease in clot strength at various time points beyond MA), which is typically measured 30 minutes following MA (LY30). It can be seen in Figure 1 that while parameters such as R-time, ACT, K-time, and α-angle can be determined shortly after the TEG is initialized, parameters such as MA and LY30 take significantly longer to determine before the clinician is able to make an informed decision regarding the clot strength and lysis rate [24]. By using this assay to inform intervention protocols, not only can coagulopathies be corrected, but the 28-day survival rate is also improved compared to more conventional coagulation assays such as prothrombin-time, which typically results in the use of more blood products [19]. Faster point-of-care diagnostics result in faster patient-specific treatment protocols; therefore expediting the results and interpretation of the TEG would help improve the outcome of patients and reduce wasted resources.

The time for a TEG to reach MA can exceed 20-30 minutes, and LY30 is not measured until 30 minutes after MA is achieved; this means that from the time the TEG assay is started the clinician must wait over an hour before the results come back and a decision regarding treatment can be made. Although assays such as the rapid TEG, which rely on the addition of kaolin, decrease the time of clot initiation and determination of parameters such as MA, α-angle, and R-time (ACT), this does not speed up the determination of lysis rate. Clinically, the most concerning results are low MA (MA $<$ 50mm) and high LY30 ($>$ 3%). High LY30 is characteristic of high fibrinolysis and typically an antifibrinolytic, such as tranexamic acid (TXA), is given to patients in this case. Since determining LY30 typically takes at least 40 minutes from TEG initialization, predicting the lysis rate prior to assay completion would improve the time required before a clinician gives the intervention, in turn improving patient outcome.

The following chapter will discuss the methods used in determining a clinically-optimal chemotherapy dosing schedule, the models used in doing so and with the results in Section 2.3. Chapter 3 focuses on determining if a patient is going to experience an abnormal lysis rate faster than the current standard of care, with the results in Section 3.2. Finally, 4 contains a summary of each of these topics followed by future work that should be explored in each area.

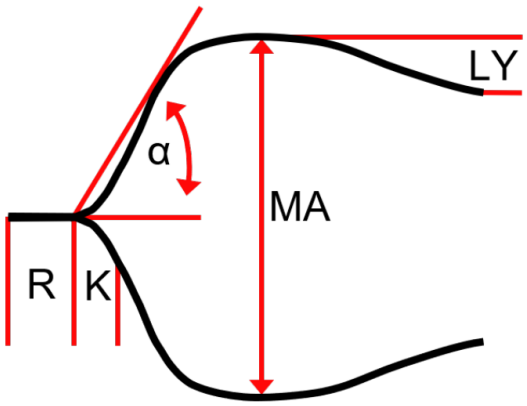

Figure 1: TEG tracing (black) with extractable parameters that are used to guide clinical decisions (red) ,adapted from [48]

2.0 Cancer Treatment via Clinically-Optimal Dosing Schedule

With the cure for cancer still out of reach, it is important to prioritize improvements in the quality-of-life for the patients that endure treatments such as chemotherapy. Since there is a dichotomy between maximizing the tumor cell eradication while minimizing the cytotoxic effect of the drug, it is important that the patient receive the dose that will produce the best outcome. By incorporating a decision support system (DSS) into the treatment program, the disparity of this dichotomy can be decreased[1].

2.1 Decision Support System

A decision support system (DSS) is a program that incorporates engineering tools to support determinations, judgements, and courses of action with the goal of improving clinician treatment decisions by providing guidance and the underlying reasoning in a user-friendly design. As seen in Figure 2, patient data and clinical inputs are implemented in the DSS, where the mathematical model representation of the patient and treatment design algorithm are used to provide guidance on treatment decisions. Through the patient-specific and patient-tailorable mathematical models, the DSS is able to predict the trajectory of disease as well as the response to interventions over the user-specified time period.

2.2 Methods

By incorporating logistical and toxicity constraints with PK models of drug distribution and PD models of drug sensitivity and effectiveness, it is possible to pose the problem in a model predictive control (MPC) framework to determine a clinically-optimal chemotherapy

[1]This work has been previously published in [39]

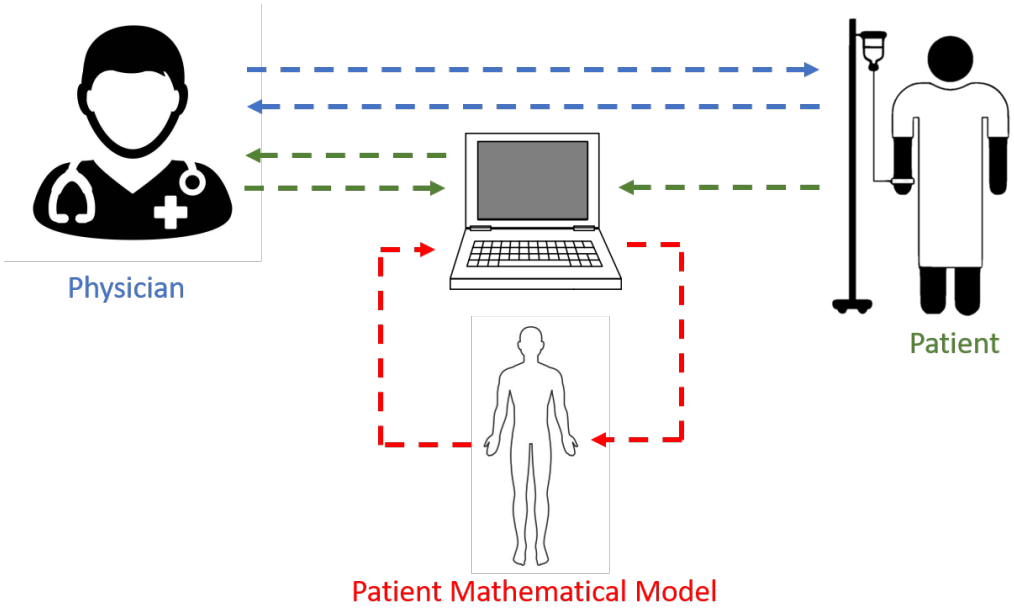

Figure 2: Decision support system role in systems medicine. Adapted from [28]

dosing schedule. The goal of this algorithm is to minimize the tumor volume while keeping the patient within the toxicity limits as to maintain a high quality-of-life.

2.2.1 Treatment Design Algorithm

Since the endpoint of treatment is not readily known *a priori*, using an algorithm such as simplex or other optimization tools is not clinically-relevant. To overcome this MPC must be implemented to predict the optimal treatment schedule over the entire horizon. MPC uses the past behavior of a system over a finite window to predict the future behavior. These predictions and the current measurements are then used to determine an optimal control input with respect to the defined objective and constraints. After a predetermined time interval, the measurement, estimation, and control calculation is repeated with a shifted horizon [7]. This strategy requires the use of predictive models (composed of the PK and PD models) and an objective function (minimizing the tumor volume given a set of constraints). By using an MPC framework, the chemotherapy treatment scheduling problem can be posed

in the following way:

$$\min_{D_d(q)} \sum_{d=1}^{N_d} (N(d)) + \Gamma_u \sum_{q=1}^{m_q} D_d(q)^2 + \Gamma_{cyc} N(N_d) \tag{2.1}$$

Subject to:

$$\text{Drug Pharmacokinetics} \tag{2.2}$$

$$\text{PD: Tumor Kill} \tag{2.3}$$

$$\text{PD: Toxicity} \tag{2.4}$$

$$\sum_{k=1}^{5} b_d(k) \leq 1 \forall d \in \{1, 2, 3\} \tag{2.5}$$

$$D_d^{min} \leq D_d(q) \leq D_d^{max} \tag{2.6}$$

$$\sum_{q=1}^{m_q} \leq D_d^{total} \tag{2.7}$$

The tumor volume ($N(d)$) is minimized each week over the number of weeks of the cycle (N_d); with the problem being solved for one or more multi-week cycles. The dose of drug ($D_d(q)$, which has a maximum number m_q) can be penalized (by the weight of Γ_u) which leads to small magnitude doses not being used since they contribute to the toxicity the patient experiences but have minimal effect on cancer cell eradication. Large doses however, are able to effectively kill the tumor cells and therefore overcome the penalty. The end of cycle tumor volume is typically of most concern so a terminal penalty (Γ_{cyc}) can be used to give preferential weighting to this measurement. Equations (2.2)-(2.4) limit the maximum amount of drug toxicity the patient experiences though the PK drug distribution equations while also maximizing the number of tumor cells that are eradicated. Details of the physiological models and constraints incorporated within the algorithm can be found in sections 2.2.2.1, 2.2.2.2, 2.2.2.3. Equation (2.5) represents the logistical constraint that the patient must only be dosed with chemotherapy once a week, Monday through Friday, and for

no more than three weeks consecutively in a given treatment cycle. This limits the number of times per week a patient will have to visit the hospital to only once, which will benefit the patient quality-of-life, as well as decreasing clinic and insurer costs and staffing concerns by only dosing during the work week and workday. The final two constraints, Equations (2.6) and (2.7) limit the amount of drug than can be given in a single dose and per cycle, respectively.

2.2.2 Case Study and Models

The PK and PD models focus on docetaxel administration for the treatment of solid-tumor cancers such as non-small cell lung, head and neck, and androgen-independent (castrate-resistant) prostate cancers [11]. The most common clinical schedules for this drug are 100 $\frac{mg}{m^2}$ every 21 days, and 35 $\frac{mg}{m^2}$ per week consecutively for three weeks out of four. The key associated toxicity is neutropenia, or a low absolute neutrophil count (ANC), with toxicity grades given in Table 1.

Table 1: Toxicity grades and neutrophil counts

Grade	Cell Count
0 (Normal)	ANC $\geq 2.0 \times 10^6$/ mL
1	1.5×10^6/ mL $\leq$ ANC $< 2.0 \times 10^6$/ mL
2	1.0×10^6/ mL $\leq$ ANC $< 1.5 \times 10^6$/ mL
3	0.5×10^6/ mL $\leq$ ANC $< 1.0 \times 10^6$/ mL
4	ANC $< 0.5 \times 10^9$ / L

2.2.2.1 Physiologically-Based Pharmacokinetic Model

The data used to develop the physiologically-based pharmacokinetic (PBPK) model of drug distribution throughout the body was obtained from a PK case study that administered docetaxel intravenously (IV) to severe combined immunodeficient (SCID) mice bearing SOV-3 human ovarian cancer xenografts [59, 68, 16]. A compartmental model of docetaxel PBPK

was developed by taking mass balances around all tissues and can be seen in Figure 3. This model was then scaled to humans by altering the tissue flowrates and volumes [55] and allowing the "Other" compartment-intercompartment transfer rates to increase, since docetaxel is lipophilic and humans have a higher percentage of body fat than mice. All other rate parameters were held constant. Each tissue compartment, aside from the gut and bone marrow, are composed of vascular and extravascular compartments, where the extravascular compartments include a protein-bound docetaxel concentration state. This was needed because of the lipophilicity and protein binding characteristics of docetaxel, which is also accounted for in the bloodstream. The metabolism of the drug is modeled by first-order kinetics within the liver based on the extravascular concentration. The resulting model is a 35-state linear dynamical model of docetaxel. This model creates challenges when using optimization, where the model is included in a set of dynamic constraints because of the number of states and independent parameters. Therefore, this ODE model was reduced using balanced truncation via the *balmr* command in MATLAB (©2022, The Mathworks, Natick, MA). The reduced model has four states, and the outputs are: plasma concentration to be compared to patient data, tumor concentration to model the drug efficacy on tumor cell eradication, and bone marrow concentration to determine the drug-induced toxicity to the neutrophils.

2.2.2.2 Neutrophil Toxicity Pharmacodynamic Model

A mechanism-based neutrophil trafficking model originally developed in [29] is employed that contains a set of 16 ODEs and 55 parameters. The entire model is able to capture responses to both chemotherapy and inflammation challenges, but for the purpose of this work, the inflammation components were removed. Model reduction was performed to reduce the number of states and to focus the sensitivity of the model on a set of more identifiable parameters, as the model is not *a priori* identifiable [28]. The reduced 9-state model shown in Figure 4 is represented by equations (2.8)-(2.16).

$$\frac{dPr(t)}{dt} = \frac{B_{min}k_G + B_{max}G_{CSF}(t)}{k_G + G_{CSF}(t)} - k_{tox}\frac{E_{max}C_{BM}(t)}{EC_{50} + C_{BM}(t)}Pr(t) - k_{tr}Pr(t) \tag{2.8}$$

$$\frac{dT_1(t)}{dt} = k_{tr}Pr(t) - k_{tr}T_1(t) \tag{2.9}$$

$$\frac{dT_2(t)}{dt} = k_{tr}T_1(t) - k_{tr}T_2(t) \tag{2.10}$$

$$\frac{dT_3(t)}{dt} = k_{tr}T_2(t) - k_{tr}T_3(t) \tag{2.11}$$

$$\frac{dN_c(t)}{dt} = k_{tr}T_3(t) - k_dN_c(t) \tag{2.12}$$

$$\frac{dIL_{17}(t)}{dt} = \frac{B_{G_{max}} * k_N}{k_N + N_c(t)} - k_{d_{IL_{17}}}IL_{17}(t) \tag{2.13}$$

$$\frac{dC_{BM}(t)}{dt} = -k_{bmv}C_{BM}(t) + k_{vbm}C_{plasma}(t) \tag{2.14}$$

$$\frac{dG_{CSF_T}(t)}{dt} = -k_{sc_{bl}}G_{CSF_T}(t) - k_{d_{GCSF_T}}G_{CSF_t}(t) \tag{2.15}$$

$$\frac{dG_{CSF}(t)}{dt} = k_{IL_{17}}IL_{17}(t) - k_{d_{GCSF}}G_{CSF}(t) + k_{sc_{bl}}G_{CSF}(t) \tag{2.16}$$

The progenitor cell population ($Pr(t)$) generates cells that mature through a maturation train ($T_1(t)$,$T_2(t)$) to become mature neutrophils ($T_3(t)$) within the bone marrow. The mature neutrophils then migrate to the vascular space and become circulating neutrophils ($N_c(t)$), which are measured as ANC. $IL-17$ is produced if the ANC begins to decrease through a series of cell types and signaling cascade that is not explicitly represented in this reduced model. In the presence of elevated $IL-17$, granulocyte colony-stimulating factor (G-CSF), which is a powerful simulator of progenitor cell production that also accelerates the migration of mature neutrophils to circulating neutrophils, is produced. Exogenous G-CSF administration, often used to rescue neutrophil toxicity, is released into the bloodstream from a subcutaneous injection site and rapidly equilibrates with the bone marrow. The systemic nature of chemotherapy kills progenitor cells and also slows the release of mature neutrophils from the bone marrow, which leads to the toxic side-effects.

2.2.2.3 Docetaxel Pharmacodynamic Efficacy Model

Solid tumors are typically characterized by exponential growth when they are small, but growth rate decreases as their size increases. Norton [43] has developed models of varying complexity that describe this growth of solid tumors such as breast cancer. This growth pattern is caused by the irregular blood vessel network which not only causes resource limitations, but also challenges in chemotherapy delivery as well. More recent work from Norton *et al.* [44] implements a self-seeding hypothesis that follows the same growth trajectory. A comparison between the different growth models can be seen in Figure 5. Both the Gomperztian and power-law models exhibit exponential growth as the tumor is small, but the growth rate decreases as the size decreases; the growth rate then approaches zero as it reaches the carrying-capacity of tumor volume (1L or 10^{12} cells), generally taken as the patient's death [43]. In order to effectively model the chemotherapy effectiveness, it is important that the model incorporates not only the natural growth and apoptosis, but also the chemotherapy concentration within the tumor. The power-law model can be seen in Equation (2.17).

$$\frac{dN}{dt} = k_{gr}N(t)^{\frac{a}{c}} - k_{die}N(t)^{\frac{b}{c}} - k_{eff}N(t)D(t) \tag{2.17}$$

In this equation, $N(t)$ is the total number of cancer cells, k_{gr} and k_{die} are the growth and death rates of the cells, respectively, and a, b, and c are constants. By setting the values of these constants such that $3 \geq b \geq c \geq a \geq 2$, higher growth rate occurs in areas with lower fractal dimensions than the region where apoptosis occurs [44]. The tumor kill efficacy is size-dependent and implemented using a bilinear kill term where $D(t)$ is the drug concentration within the tumor and k_{eff} is the drug effectiveness constant.

2.2.3 Generating the Feasible Set of Solutions

To generate the clinically-optimal chemotherapy dose schedule for implementation in a GPU for parallel simulation, the feasible set of all possible dosing schedules must first be populated. This is done so that rather than the algorithm selecting the values to implement based on previous iterations, the entire set of possible values is predefined for computational efficiency. These schedules must conform to logistical constraints such as the limit of allow-

able doses in a cycle, allowable treatment times, and the maximum total dose over the cycle. The logistical constraints are set such that a patient can only receive a maximum of three doses during a single cycle and they must be dosed between 8AM and 5PM Monday through Friday; doses will either be given over one hour or 30 minutes depending on the magnitude of the dose. The set of feasible solutions are generated by choosing the resolution for the discretization of time and dose so that all possible combinations of dose magnitude and frequency are populated. The resolutions set the minimum difference in administration time and dose magnitude between the different dosing schedules. These resolutions result from clinical experience, such as the typical length of treatment (*e.g.*, one or 30-minute dosing durations) or the minimum difference between dose magnitudes (*e.g.*, 5 mg). If a coarser discretization is used, the number of feasible sets will decrease and lead to faster simulation times. However, the discretization should be clinically-relevant as there is no reason to be overly precise (*e.g.*, using the nearest second for treatment time and nearest μg for dose magnitude) since this will not produce clinically-significant variations in results and would significantly decrease computational efficiency.

The feasible set was generated using a recursive Python function based on timing constraints and a second recursive function was used to find all the possible dose regimens within the maximum number of doses. A dose of zero is allowed so that the optimal solution can contain fewer doses than the maximum allowable number of doses. The resulting combinatorial set of feasible administration times and dose magnitudes is then generated and saved as a matrix in a binary format for use in the GPU simulation step.

2.2.4 Implementing the Feasible Set in GPU Simulation

By implementing C++ in conjuction with VEXCL (`https://github.com/ddemidov/vexcl`), the GPU kernel can be compiled via a function. The kernel uses the matrix of all feasible schedules and, in parallel at each time step, solves the PKPD ODEs that make up the PK model from Section 2.2.2.1 and the PD Equations (2.8)-(2.17). The ODE solver used in the GPU kernel is from the Boost odeint library (`https://www.boost.org/doc/libs/1_66_0/libs/numeric/odeint/doc/html/index.html`) and employs a Runge-Kutta DOPRI5

method. By using a larger time step, the speed on the solution will increase, but accuracy will be lost, and care must be taken to choose a δt that ensures all feasible chemotherapy administrations fall exactly on a time step. During runtime, after each time step, the GPU kernel updates the objective function, which for the Docetaxel case study presented herein, is given by Equation (2.1).

2.2.5 Dose Schedule Design

By using the optimization problem governed by equations (2.1)-(2.6), physicians are able to add any additional constraints to mitigate toxicity. For this study, which uses docetaxel, the following constraints are added for each day i:

$$N_c(i) \geq 1.0 \times 10^6/mL \tag{2.18}$$

$$\max\left([N_c(i), N_c(i+1) \dots N_c(i+6)]\right) \geq 1.5 \times 10^6/mL \tag{2.19}$$

By using these two additional toxicity constraints, no grade 3 ANC toxicity is guaranteed and any grade 2 toxicity will last no more than 7 consecutive days. Violation of these constraints are tested for at the end of each GPU simulation. If a particular schedule violates one of these toxicity constraints, all further simulations using the same schedule and higher doses are not simulated. Then the schedule is advanced to the next solution in the feasible set and the simulation continues again. Once the full set of acceptably toxic schedules is evaluated, the optimal schedule is determined to be the one that ends the cycle with the lowest objective function value from Equation (2.1).

2.2.6 Updating the Model from Clinical Measurements During Treatment

Since the drugs are administered intravenously, the patient must return to the clinic for each dose. At the end of each cycle, ANC is measured from a serum sample. The efficacy of the drug on the tumor is typically determined at the end of every other cycle via magnetic resonance imaging (MRI), or other imaging methods. Since this model is not specific to any one patient, the actual patient response may vary from the model predictions. By using

nonlinear least squares (NLS), two model parameters are updated after patient measurements become available to reflect actual patient response. The two parameters of interest are drug sensitivity (k_{tox} from Equation (2.8)) and drug efficacy on the tumor (k_{eff} from Equation (2.17)), which are updated after every cycle and every other cycle of treatment, respectively. The updated parameters are tabulated using NLS via the following objective function:

$$\min_{k_{tox},k_{eff}} (ANC_{meas} - N_c(N_d))^2 + w_e * (V_{meas} - N(N_d))^2 \tag{2.20}$$

End-of-cycle measurements (N_d) of ANC and estimated tumor volume are compared to the model prediction; $N_c(t)$ is driven by k_{tox}, while $N(t)$ is affected by k_{eff}; w_e is a binary variable that is equal to one on the cycle where tumor volume is measured and zero otherwise.

2.3 Results

By implementing the logistical and clinical constraints in conjunction with the PK and PD models for docetaxel distribution, efficacy, and toxicity a clinically-optimal dosing schedule for chemotherapy patients is created. After the PK model of docetaxel distribution was compared to samples previously taken during phase 1 clinical trials,, the optimal treatment schedule is compared to current clinical practice and interpatient variability is explored.

2.3.1 Model Fitting

In order to ensure the model accurately fits the docetaxel concentration within the blood stream (and subsequently tumor and bone marrow), time series data is captured after docetaxel administration during phase 1 trials at the University of Pittsburgh Cancer Institute (UPCI) and the Memorial Sloan-Kettering Cancer Center (MSKCC) [8, 56, 31]. During the phase 1 clinical trials at UPCI, 25 patients were dosed with docetaxel at three different dosing magnitudes (50, 60, and 75 $\frac{mg}{m^2}$) via intravenous infusion over one hour and plasma concentrations were measured at fifteen predetermined intervals ranging from 0.5 to 25 hours following infusion [8]. The phase 1 trial performed as MSKCC involved 50 patients being

dosed either weekly with 35 $\frac{mg}{m^2}$, or every three weeks with 55, 70, or 75 $\frac{mg}{m^2}$ and 6-9 plasma concentration measurements per patient were taken at times out to 50.5 hours following infusion [56]. Additionally, a PK study involving elderly patients performed at MSKCC involved 20 patients being dosed weekly with 35 $\frac{mg}{m^2}$ over a 30-minute infusion and 10 plasma samples were taken ranging from 0.25 to 168.5 hours post-infusion [31]. A sample patient fit of the PBPK model, using data from UPCI study 01-150, is shown in Figure 6

It is clear that this model can capture the single-dose concentration profile of docetaxel within the bloodstream, which demonstrates that the SCID mice-derived model can fit human patient data by adjusting only the flowrates, volumes, and the "Other" compartment-intercompartment transfer rates.

2.3.2 Optimal Treatment Compared to Clinical Practice

The most commonly used docetaxel dosing schedules are 100 $\frac{mg}{m^2}$ every 21 days (100 q21d), or 35 $\frac{mg}{m^2}$ every week for three weeks followed by one week off (35 qw 3 of 4)[61]. However, these dosing schedules do not provide every patient with the best quality-of-life; the 100 q21d schedule is often overly-toxic and the 35 qw 3 of 4 requires the patient to go to the clinic three times a month and while failing to have as much tumor cell eradication as the 100 q21d schedule [27, 51]. For this reason, an optimal treatment schedule was determined that can be augmented over time after patient response is measured.

After simulating the set of all feasible chemotherapy schedules and using the change in number of tumor cells as the objective function to maximize, the schedule simulation that resulted in the largest difference of tumor cells that was tolerably toxic is determined to be the mathematically optimal solution. The mathematically-optimal solution was determined to be 55, 45, and 5 $\frac{mg}{m^2}$ on days 0, 11, and 18, respectively of a four-week dosing cycle. This is vastly different than the current clinical dosing schedules mentioned previously. This solution, although mathematically optimal, is not clinically-optimal because of the small dose given on day 18 will not provide significant tumor eradication, will likely increase the patient's resistance to the drug, and requires one more clinic visit [1]. By combining the last two doses of the mathematically-optimal dosing schedule, a more clinically-optimal

(though mathematically sub-optimal) solution would be 55 $\frac{mg}{m^2}$ and 50 $\frac{mg}{m^2}$ on days 0 and 11, respectively. A comparison of the two standard clinical dosing schedules and the clinically-optimal solution can be seen in Figure 7.

Over the course of 84 days, the 100 $\frac{mg}{m^2}$ every 21 days clearly shows more tumor cell eradication than both the 35 $\frac{mg}{m^2}$ qw 3 of 4 and the clinically-optimal DSS solution. Although there may be more tumor eliminated on this dosing schedule, the patient has received four cycles of treatment, compared to the other two schedules that only received three. The patient has also received 25% more drug than the other two schedules. Another drawback to the 100 $\frac{mg}{m^2}$ q 21 day schedule is the noticeable grade 3 neutropenia, which is undesired by the clinician and patient alike. Since the 35 $\frac{mg}{m^2}$ qw 3 of 4 and the DSS solution demonstrate similar toxicity by ANC grade and tumor-killing effect and the DSS (sub)optimal solution has one fewer visit per cycle, the DSS solution is preferred. This treatment strategy would not only improve patient quality-of-life but also decrease costs to the hospital, insurer, and patient with no increase in toxicity. Although the DSS clinically-optimal solution (green) is slightly more toxic than the 35 qw 3 of 4 (red) schedule, the difference in toxicity is insignificant and could be attributed to interpatient sensitivity differences. However, if the patient's toxicity response is slightly more sensitive, the DSS solution would have to be altered as it would likely cause grade 3 neutropenia..

2.3.3 Managing Interpatient Variability

Since each patient is unique, and sensitivities in both in drug efficacy and toxicity often change during the course of treatment, the DSS model must also be able to update to provide individualized optimal treatment. Figure 8 demonstrates how the DSS is able alter the optimal treatment schedule sensitivities of k_{tox} and k_{eff} from Equations (2.8) and (2.17), respectively.

As a simulated case study, a patient begins treatment with $k_{tox} = 1.6$, which is moderately sensitive compared to the DSS nominal value of $k_{tox} = 1.0$. This causes the first cycle (days 0-28) of treatment to be highly toxic to the patient, which can be seen by the patient experiencing grade 3 neutropenia (top Figure 8), which would have been far worse had the

patient received the clinical standard of 100 $\frac{mg}{m^2}$ every 21 days compared to the DSS clinically-optimal solution. In a clinical setting, if the patient had experienced this toxicity, the second dose would most likely be withheld until their ANC returned near baseline, as well as having the following dose magnitude reduced, which is the clinical standard for patient-tailored chemotherapy treatment. By taking an algorithmic approach to determining the dosing schedule it is possible to use clinical measurements to inform the DSS of actual patient response. If measurements were taken after the first dose, rather than at the end of the cycle the actual patient response could inform the next dose; if the patient showed severe toxicity after the first dose the second dose on day 11 could either be reduced or withheld at the clinician's discretion. The DSS is currently able to update k_{tox} following each cycle by comparing the ANC measurement and model prediction according to Equation (2.20); the new value returned is $k_{tox} = 1.6$. The next two cycles (days 29-56 and 57-84) show the patient is held within the toxicity limits using the revised DSS solution of three smaller doses rather than two larger doses; the patient received 40, 30, and 20 $\frac{mg}{m^2}$ on days 4, 16, and 24 (33, 45, and 53 for cycle 2 and 61, 73, and 81 for cycle 3), respectively. Since the patient is now receiving three doses rather than two, the clinical costs will increase because of clinician time and number of patient visits, but this is a direct trade off considering the cost of toxicity rescue via G-CSF ($\sim$\$10,000/cycle, not including patient visit needs). Following the third cycle (days 57-84) the patient develops further increased toxicity ($k_{tox} = 1.9$), while the DSS remains unaware with a $k_{tox} = 1.6$. After the patient experiences another significant toxicity event, the algorithm detects the mismatch and updates accordingly to match the actual patient response. This update causes the optimal dosing schedule to change for cycles five and six to 30, 25, and 25 $\frac{mg}{m^2}$ on days 7, 17, and 25 (days 119, 129, and 137 for cycle five and days 148, 158, and 164 for cycle 6), respectively. By taking measurements during chemotherapy treatment and comparing the model predicted value to actual measurements, it is possible to create an algorithm that will evolve to continually produce the optimal dosing schedule even as patient response changes over time.

2.3.4 Discussion

Chemotherapy has always been a difficult treatment to prescribe, given the dichotomy of minimizing both tumor volume and cytotoxicity. Clinicians have been prescribing standard doses of chemotherapy based on the patients' surface area, weight, and renal function since passing clinical trials for FDA approval [23]. By creating and implementing PK/PD models of the concentration distribution and the effect of said concentration, it is possible to predict patient response to these drugs. It has been shown that a reduced, computationally efficient, PK model can accurately predict the plasma concentration of drug [8]. Implementing these PK models in conjunction with PD models of both tumor eradication and myelosuppression, or suppression in neutrophil production within the bone marrow, allows for the creation of constraints that can be used to create a treatment algorithm.

By describing the treatment algorithm within a GPU architecture, it is possible to minimize a simulated tumor volume while staying within toxicity constraints during treatment. Although this solution may be mathematically-optimal, there are other factors that would aid in making it more clinically-optimal, such as considering the number of times the patient has to visit the clinic during a given treatment cycle. This clinically-optimal solution was shown to increase a simulated patient's quality of life by minimizing the tumor volume while maintaining an acceptable toxicity, and decreasing the number of times they must visit the clinic. This clinically-optimal solution can be tailored to each patient though a nonlinear least squares comparison of actual measurements taken at the end of treatment (*i.e.* ANC, tumor volume), and modeling predicted responses following each measurement to ensure the dose remains optimal.

Since the efficacy of chemotherapy drugs often decrease over time in patients, drug efficacy on the tumor is another important consideration to patient sensitivity [1]. For example, if the k_{eff} of the drug on tumor is set to 0, the optimization problem returns dosing magnitudes of 0 since the drug will have no effect on tumor cells, but rather only decrease drug efficacy; if the dosing penalty is increased, this will also drive the optimization to return optimal dose magnitudes of 0. This allows small dose magnitudes that have minimal effect on tumor volume to be eliminated as to not increase resistance.

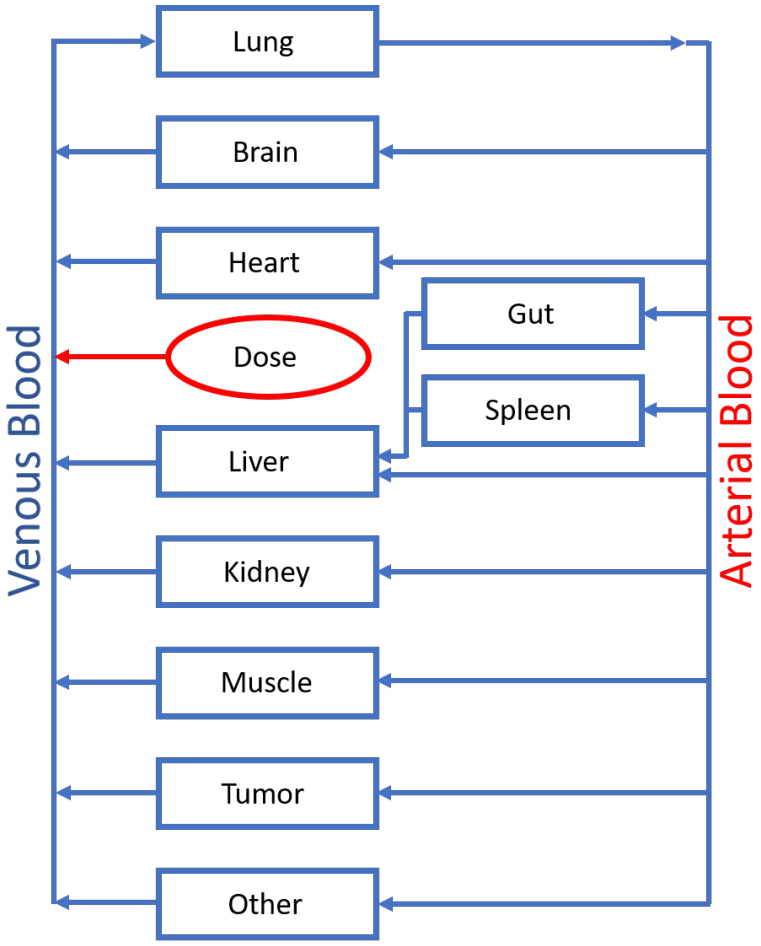

Figure 3: PBPK model of docetaxel distribution throughout the body [16].

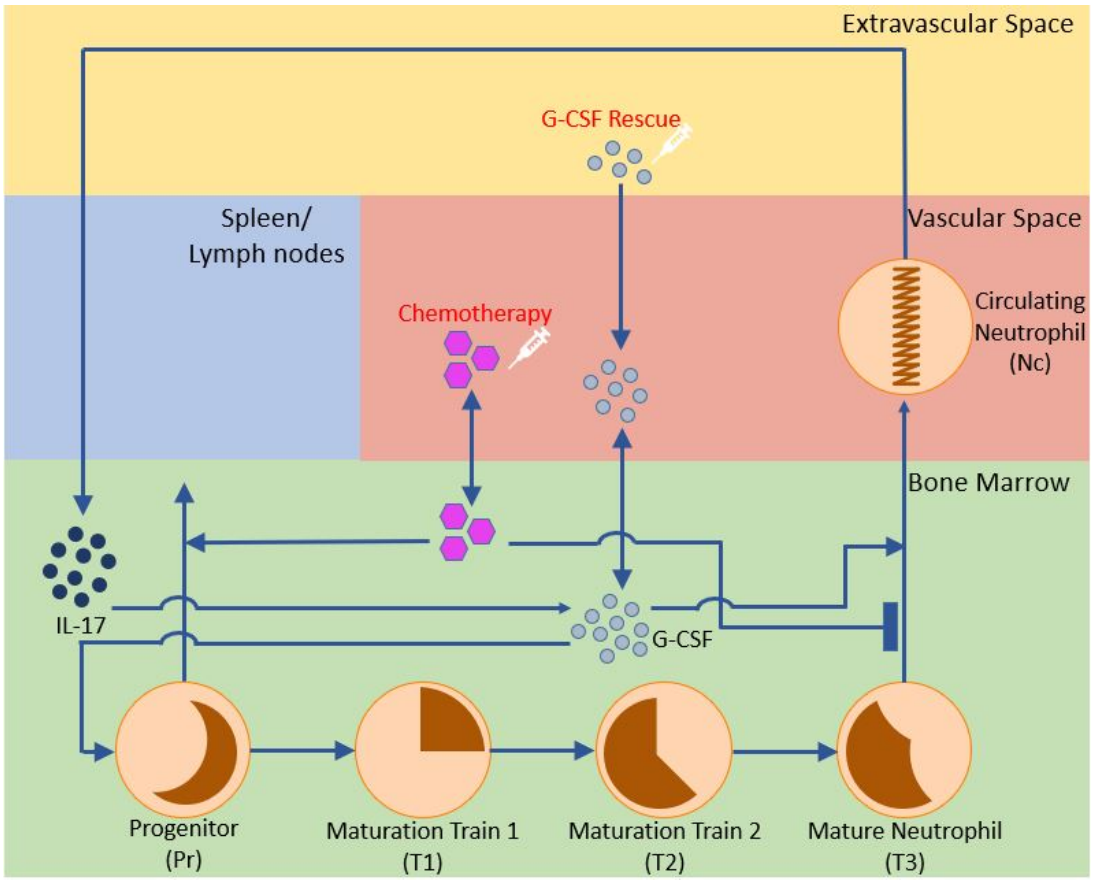

Figure 4: Reduced neutrophil model with G-CSF [28].

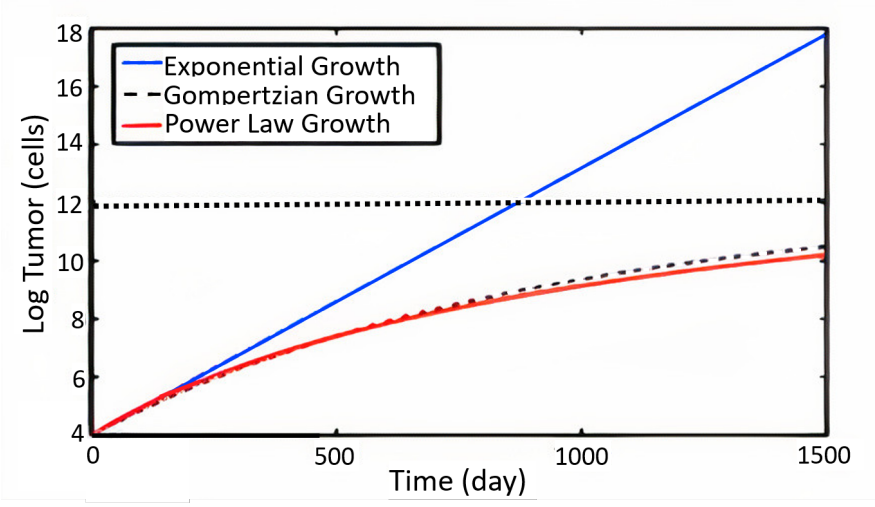

Figure 5: Tumor growth trajectory comparison for exponential, Gompertzian, and power-law models

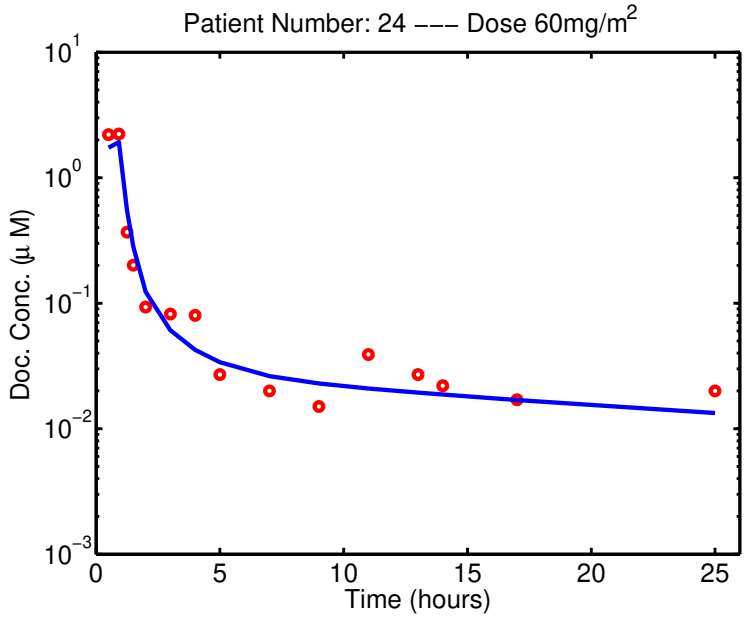

Figure 6: Example of model fit to human patient data. A single docetaxel dose of 60 $\frac{mg}{m^2}$ was administered, and serum concentrations were measured at intervals following administration

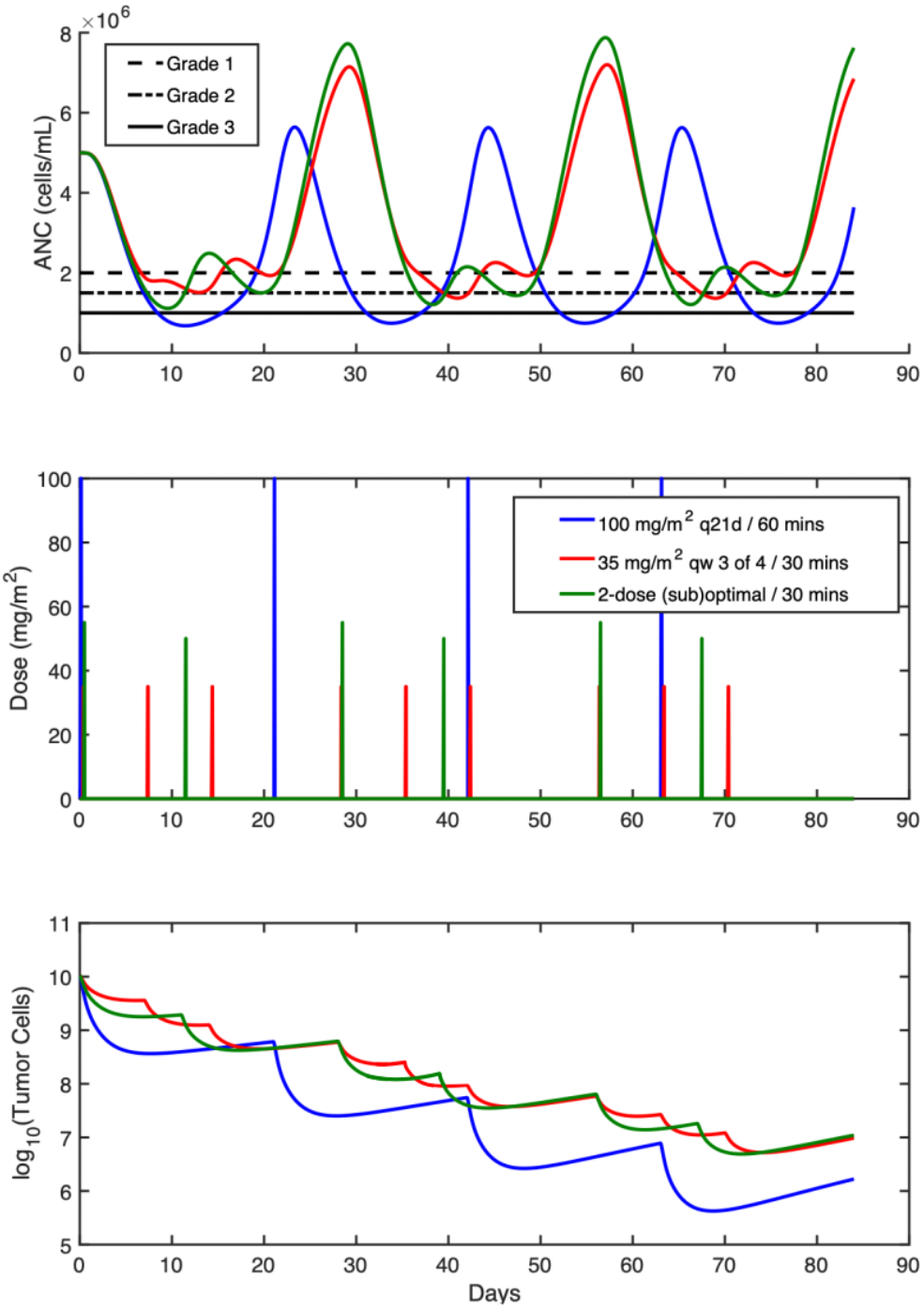

Figure 7: Comparison of standard clinical dosing schedules (100 $\frac{mg}{m^2}$ every 3 weeks, blue; 35 $\frac{mg}{m^2}$ 3 weeks of 4, red) and DSS solution for docetaxel administration (green). **Top:**. ANC over treatment time with toxicity grades shown as horizontal lines. **Middle:**. Dosing day and magnitude ($\frac{mg}{m^2}$). **Bottom:** Number of tumor cells over time.

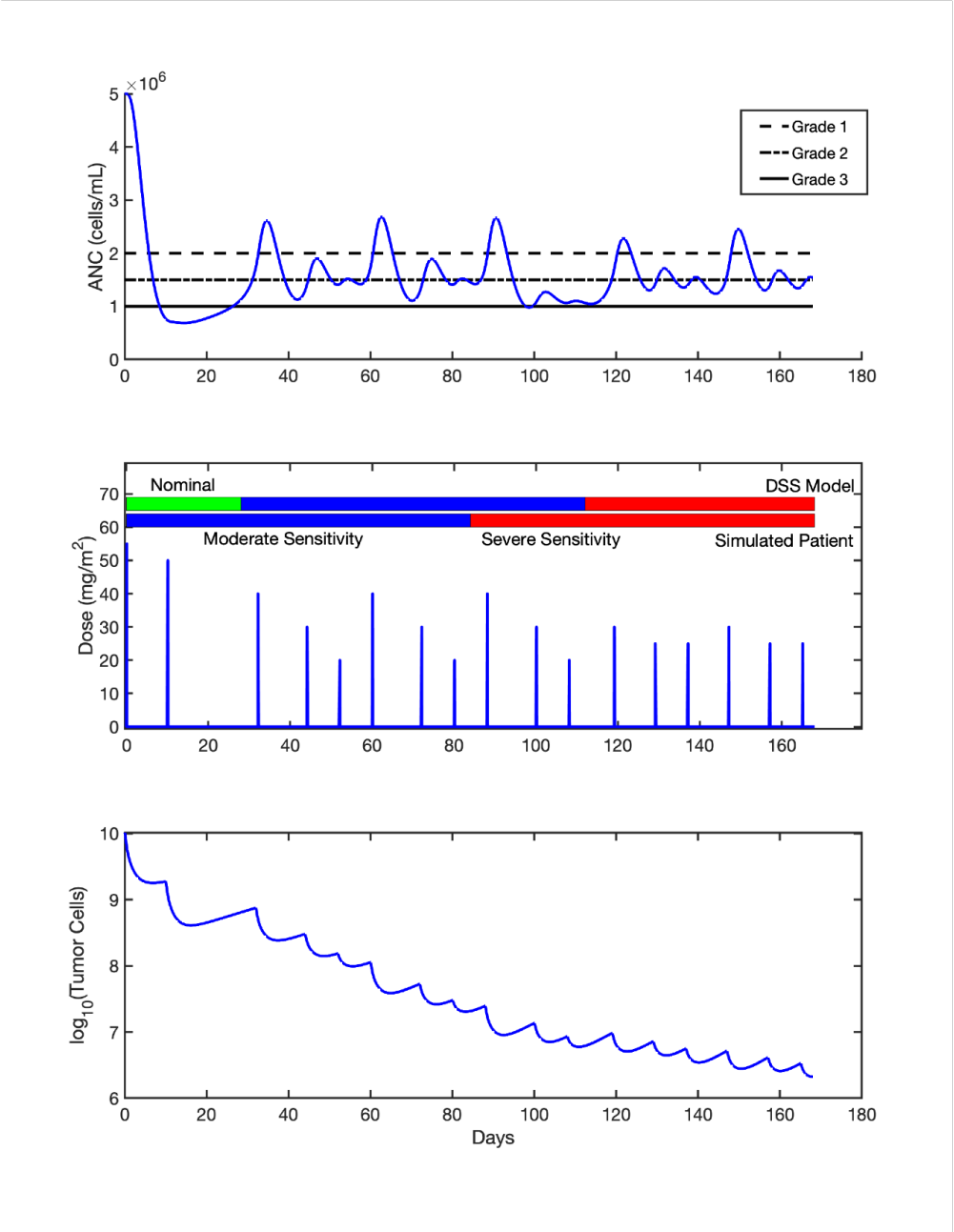

Figure 8: Simulated patient response to treatment in the presence of model-patient mismatch. **Top:** Patient ANC over time, with toxicity grades shown by horizontal lines. **Middle:** Dosing day and magnitude, with the toxicity sensitivity for the algorithm-expected patient (top bar) and actual patient (bottom bar). Sensitivities (k_{tox}) are nominal (1.0), moderate (1.6) and severe (1.9). Doses are calculated each cycle (28 days), with algorithm parameter updates after each cycle. **Bottom:**. number of tumor cells over time

3.0 Early Estimation of Lysis from Patient TEG

The TEG is used to assess different aspects of the patient's clotting ability. This assay, which can range in duration anywhere from 30 to 80 minutes, works by a stand oscillating to simulate slow moving blood which then initiates the clot formation which is measured by the displacement of a torsion pin [24]. This measurement relies on the changing viscoelasticity of the blood as the clot forms and then lyses (breaks down). The processes of most interest in the clotting cascade are the platelet activation, clot growth, and lysis of the clot. Additionally, activators that promote or inhibit certain parts of the coagulation cascade can be added to the assay to either increase the speed of clot formation or to isolate certain parts of the coagulation cascade to help make a more informed decision regarding treatment. Although most of the parameters of the TEG can be reported within minutes of initialization, the key parameters of maximum amplitude (MA) and lysis percentage 30 minutes after maximum amplitude is achieved (LY30) require more time to be report clot strength and lysis rate. These help to inform the clinician of the state of the patient's hemostasis, both primary and secondary. Primary hemostasis can be described as a conglomeration of platelets that form a plug at the endothelial cells of an injury site to prevent further bleeding. Secondary hemostasis is composed of the intrinsic and extrinsic clotting pathways, which combine to form the primary clotting pathway. This combined pathway activates fibrinogen into fibrin units (which has an affinity for itself), which combine to form strands that bind the platelets together and stabilize the plug at the injury site. The intrinsic pathway activation occurs when there is exposed endothelial collagen, while the extrinsic pathway activation occurs through tissue factor released by the endothelial cells following injury; the intrinsic pathway is the longer pathway of the two in terms of time [45, 10]. The rapidTEG is currently one of the fastest assays that activates both the intrinsic and extrinsic clotting pathways and can be completed in as little as 30 minutes [52]. It is important that both pathways are activated during the assay so the clinician can get a comprehensive interpretation of the patient's clotting ability; if only one pathway were activated the clinician's interpretation of the results would be an inaccurate representation of the patient's actual clotting ability.

The work herein focuses on expediting the prediction of whether the patient's LY30 values are above or below a threshold to help the clinician determine if the clot will be lysing abnormally, and therefore intervene before the situation worsens. Previous work has focused on expediting the prediction of high or low MA, but LY30 can only be determined after MA occurs which is the key motivation for this work [49, 48].

3.1 Methods

By using patient TEG tracings in conjunction with machine learning techniques such as K-nearest neighbors and logistic regression, it is possible to predict whether a patient is going to experience different severities of abnormal lysis rates. This is done by including the four key ODE model parameters (β, k_4, k_5, and Ψ from Equations (3.1)-(3.9)) as inputs into the logistic regression and the levels of lysis (above or below a threshold value) as the classes of each patient.

3.1.1 Data Implementation

TEG time-series data was collected from subjects enrolled in the Study of Tranexamic Acid during Air Medical Prehospital transport (STAAMP) and Prehospital Air Medical Plasma (PAMPer) clinical trials [6, 5, 57]. This included a total of 1,045 TEG tracings from 293 patients. The data was then deidentified to remove any personal information that could be used to correlate a particular patient with a TEG tracing.

3.1.2 Challenges in Predicting LY30 of Patients with Extended ACT

Since there is no delay function explicitly incorporated into ODE model, it is nearly impossible for the function to predict accurate nearest neighbors when the ACT of the patient TEG exceeds the standard range of 80-140 seconds [42]. This may be possible to overcome by introducing various temporal delays into each of the simulated TEG tracings, but this would increase the simulation and optimization time significantly. Since there is

significant delay before the initialization of the TEG, the neighbors that the function predicts are not accurate representations of the patient TEG tracing. Examples of the TEG tracings that have longer than standard ACT can be seen in Figure 9. There are a total of 26 patients out of the 874 filtered patients that have an ACT longer than 150 seconds, or roughly 3% of the filtered data set. The median ACT of the filtered data set, excluding the 26 patients previously mentioned, is 80 seconds. The normal distribution of ACT can be seen in Figure 10. Since the model is unable to capture these extended ACT times, these patients were removed from the machine learning analysis because using the neighbors determined by the algorithm would negatively impact accuracy of the predictor.

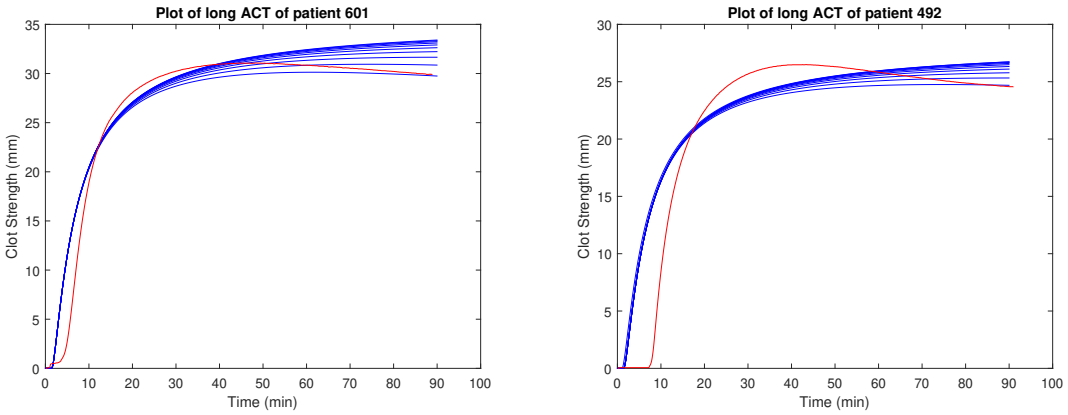

Figure 9: Examples of TEG tracings with ACT longer than 150 seconds. The nearest neighbors are shown in blue, while the patient TEG is shown in red.

3.1.3 Determining the True MA Time

Since LY30 is measured by comparing the maximum amplitude, it is important to accurately determine the time at which MA occurs. Being able to predict time of MA will also aid in selecting the range of time in which data will be used for prediction of LY30. TEG data is inherently noisy, and has very minute fluctuations that are insignificant clinically, but can create challenges when doing numerical analyses. The TEG tracing data measures the clot strength down to the ten-thousandth of a millimeter, which is more significant digits than is necessary to make clinical judgements regarding the clotting characteristics of the

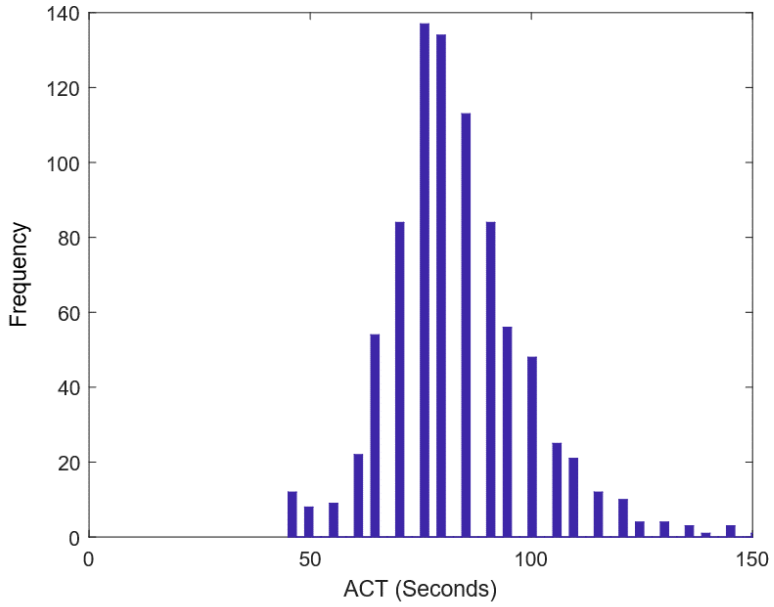

Figure 10: Histogram of ACT of filtered data set, excluding any ACT longer than 150 seconds.

patient [63]. For example, The TEG operator may thing the assay has reached MA since the next point decreases, albeit by a very small interval which would cause inaccurate interpretation of MA time and could lead to incorrect LY30 calculation leading to improper patient treatment. The Figure below, Figure 11 highlights how the TEG may still be increasing, even after previous points have shown a negative change in clot strength. It is imperative therefore for this algorithm to correctly predict the global maximum of the TEG tracing.

Being able to predict the time of MA accurately is crucial in determining where to start the analysis for determination of nearest neighbors, and noise in the tracing make it difficult. Additionally, this enables real-time implementation of the algorithm simultaneously with the TEG running in the clinical setting such that MA can be identified as soon as it happens. In order to account for the noise in the TEG while determining MA, a four-point concavity

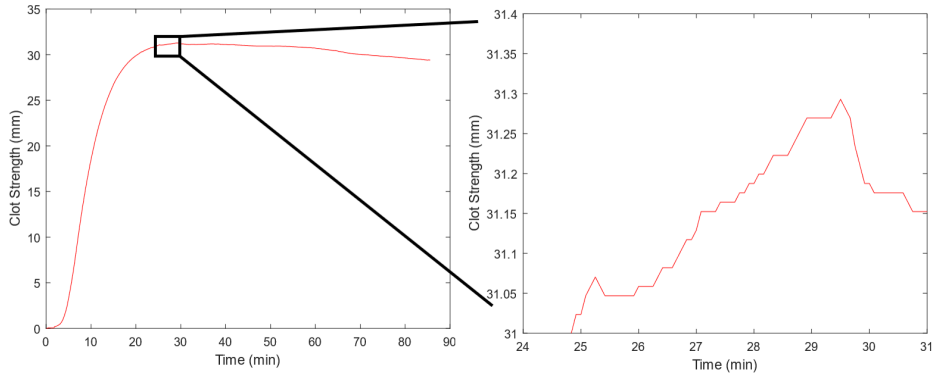

Figure 11: Example of TEG tracing that exhibits oscillation before reaching MA.

analysis was done that looked at the values of the TEG tracing from 30 minutes pre-MA to MA. This test involved looking at the change between successive points and determining the slope. The change in these successive slopes were examined, and if the change was decreasing, the inflection would be concave, whereas increasing would indicate a convex inflection. Two separate values of difference in magnitude for the noise threshold were used between the first and fourth successive points: 0.1mm and 0.025mm. To verify MA has been reached the test needed to satisfy the fact that the inflection was concave, and the magnitude difference between the first and fourth successive points had to be greater than the noise threshold values.

3.1.4 ODE Model of Coagulation Cascade

The complete mechanistic model of the clotting cascade contains 76 species and 105 kinetic constants [3]. Since there are a number of factors that can influence the coagulation cascade, including medications, genetics, hypothermia, and inflammation, it is important to create a model that captures the complete resulting function of the key components that can better capture a range of clotting dysfunction. Additionally, the lack of identifiability and the computational time required to fit this complete model would make fitting each patient

TEG impractical and inaccurate in a clinical setting [9]. A diagram of the mechanistic model can be seen in Figure 12 [3]. To make modeling this process and fitting to patient data computationally feasible, the mechanistic clotting cascade model was simplified to estimate key parameters of the TEG using mass conservation principles [34, 30, 2]. Although this sacrifices a significant number of the states in the complete mechanistic model, it retains the key biologically-motivated parameters related to coagulopathies [49, 15].

As seen in Figure 13, the cascade initiates by the conversion of prothrombin to thrombin. Thrombin then interacts with resting platelets to transform them into activated platelets. The activated platelets provide reaction sites that further convert prothrombin to thrombin. The thrombin is continually reacting with fibrinogen, which form a cross-linked network that together with the activated platelets form a clot. This clot forms and is then broken down through fibrinolysis [48]. This simplified low-order ODE model gives a balance between clinical relevance and physiological representation, enabling efficient computations with a high degree of accuracy compared to the complete mechanistic model.

3.1.5 Creating the Library of Simulated Patient TEG's

In order to determine the parameters within the simplified ODE model specific to each patient in an expedited manner, identification of the nearest neighbors between patient and simulated TEG tracings will be implemented. This method is preferred for clinical use over slow optimization algorithms that can take up to an hour to determine the parameters within the ODE model following the TEG assay. Before determining the optimal number of nearest neighbors and finding them, a database of simulated TEG tracings had to be created for comparison to the patient tracings to identify the parameter space of the ODE model. This database included a library of 160,000 different simulated TEG tracings by varying four of the ODE model parameters where each parameter had a discretization of 20. This discretization was chosen because it provided an adequate number of tracings that the encompassed the entire patient database; the number of discretizations could be increased but this would come at the cost of computation time. For the basis of the machine learning model, the key parameters from the ODE model will be used as the input variables in conjunction with

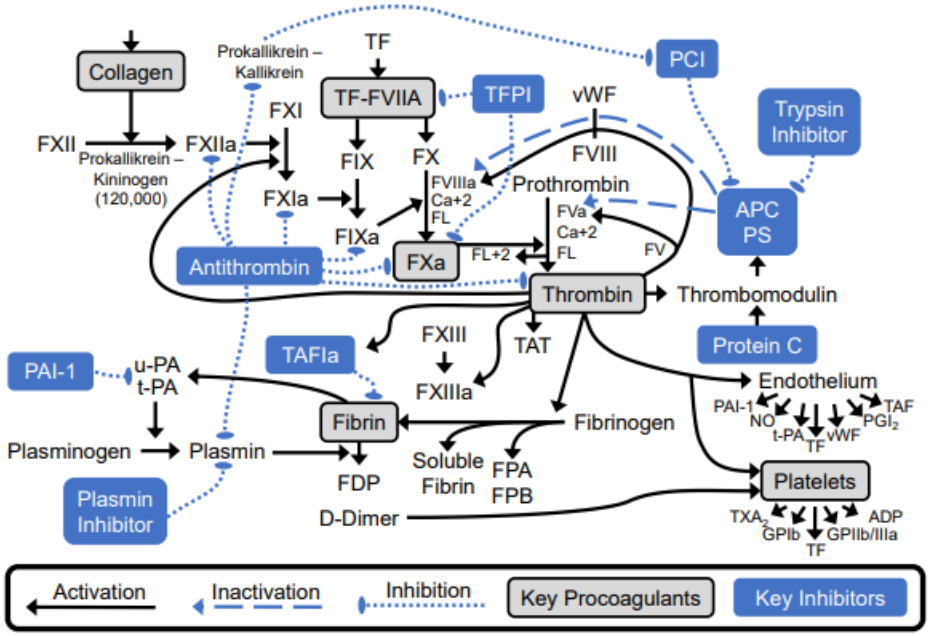

Figure 12: The coagulation pathways leading to clot formation including inactivation, activation, and inhibition, taken from [3]. Clotting factors are denoted by roman numerals in the active and inactive forms (i.e. FXIII is inactive factor 8 while FXIIIa is active). Abbreviations: ADP (adenosine diphosphate), Ca+2 (calcium with +2 charge), FDP(fibrin degradation products), FL (phospholipids), FPA (fibrinopeptide A), FPB (fibrinopeptide B), GP (glycoprotein), PAI-1 (plasminogen activator inhibitor), PCI (protein C inhibitor), PS (protein S), TAFI (thrombin activation fibrinolysis inhibitor), TAT (thrombin-antithrombin complex), TF (tissue factor), TFPI (tissue factor pathway inhibitor, t-PA (tissue-plasminogen activator), TXA2 (thromboxane A2), u-PA (urokinase-plasminogen activator), and vWF (von Willebrand factor).

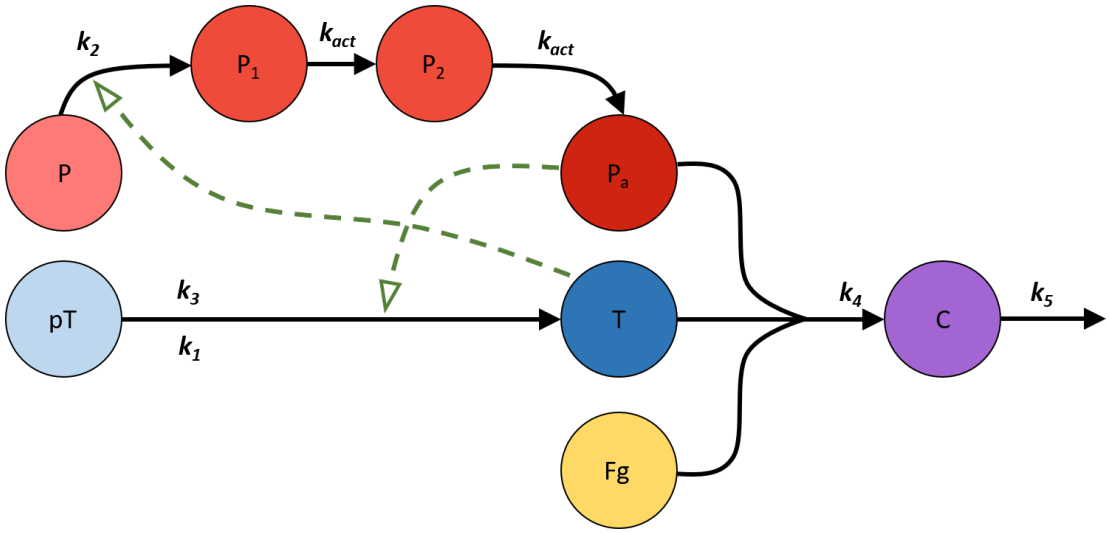

Figure 13: Model schematic of simplified clotting pathway including prothrombin (pT), thrombin (T), resting platelets (P), activated platelet cascade (P_1, P_2, P_a), fibrinogen (Fg), and clot (C) [48].

a binary output variable of high or low lysis rate. This method was chosen because each of the key model parameters represents a particular rate of the clotting cascade so that it will be possible to interpret which of the model parameters most influences the high or low lysis rates. The binary variable of high or low lysis is used so that the machine learning algorithm is able to predict the probability of being in one of the two classes. The model parameters implemented in the ODE model below in Equations (3.1)-(3.9) that will be used in the machine learning model will be β, k_4, k_5, and Ψ; their ranges can be seen in Table 2.

The four estimated parameters represent the platelet and thrombin activation (β), rate of clot formation (k_4), lysis rate (k_5), and a scaling factor (Ψ). Since the rate of clot formation is directly impacted by the activation of thrombin and platelets, it is crucial to include in the analysis. Furthermore, the rate of clot formation and lysis rate are indicative of different types of coagulopathies and must also be included in the model. The scaling factor must also be included in the analysis because a low scaling factor would indicate that the patient will not be able to form a substantial clot. Since the model captures the dynamics

Table 2: Ranges for the four varied ODE model parameters

Parameter	Minimum Value	Maximum Value
β	0.001	10
k_4	1E-4	0.1
k_5	1E-6	0.001
Ψ	0	70

of elimination of chemical compounds from the body, Michaelis-Menten kinetics were widely implemented in this model. The varied parameters can be seen in the ODE model below (Equations (3.1)-(3.9)). This system contains 6 rate coefficients (k_1, k_2, k_3, k_4, k_5, k_{act}) with units of s^{-1}, a saturation constant, μ, with arbitrary units, and eight initial concentrations including the three initiation states ($P(0) = pT(0) = F_g(0) = 1$) and five initially inactive states ($P_1(0) = P_2(0) = P_a(0) = T(0) = C(0) = 0$) also with arbitrary units.The saturation constant μ and activation rate constant k_{act} remained constant for this analysis due to identifiability concerns within the ODE model [48, 15]. The clot state, $C(t)$ was scaled using factor Ψ, with units of mm, in order to relate the model to experimental data.

$$\frac{dpT(t)}{dt} = -k_1 pT(t) - \beta P_a(t) \cdot \frac{pT(t)}{\mu_+ pT(t)}. \tag{3.1}$$

$$\frac{dT(t)}{dt} = k_1 pT(t) + \beta P_a(t) \cdot \frac{pT(t)}{\mu + pT(t)} - k_4 T(t) \cdot \frac{P_a(t)}{\mu + P_a(t)} \cdot \frac{F_g(t)}{\mu + F_g(t)}. \tag{3.2}$$

$$\frac{dP(t)}{dt} = -\beta P(t) \cdot \frac{T(t)}{\mu + T(t)}. \tag{3.3}$$

$$\frac{dP_1(t)}{dt} = \beta P(t) \cdot \frac{T(t)}{\mu + T(t)} - k_{act} P_1(t) \tag{3.4}$$

$$\frac{dP_2(t)}{dt} = k_{act} P_1(t) - k_{act} P_2(t) \tag{3.5}$$

$$\frac{dP_a(t)}{dt} = k_{act} P_2(t) - k_4 T(t) \cdot \frac{P_a(t)}{\mu + P_a(t)} \cdot \frac{F_g(t)}{\mu + F_g(t)} \tag{3.6}$$

$$\frac{dF_g(t)}{dt} = -k_4 T(t) \cdot \frac{P_a(t)}{\mu + P_a(t)} \cdot \frac{F_g(t)}{\mu + F_g(t)} \tag{3.7}$$

$$\frac{dC(t)}{dt} = k_4 T(t) \cdot \frac{P_a(t)}{\mu + P_a(t)} \cdot \frac{F_g(t)}{\mu + F_g(t)} - k_5 C(t) \tag{3.8}$$

$$Y(t) = \Psi \cdot C(t) \tag{3.9}$$

3.1.5.1 Using K-Nearest Neighbors for Patient TEG's in Parameter Space

After creating the library of simulated TEG tracings, the optimal number of nearest neighbors needs to be determined. It is important to pick a number of nearest neighbors (k-value) that is large enough so that machine learning algorithm has enough training sets as an input, yet small enough as to not be too computationally expensive and slow down the prediction. Additionally, having too large of a k-value could cause neighbors to be included in the set that are further than desired from the actual value. To do this, the normalized sum of squared error (NSSE) for each patient compared to each entry in the simulated TEG library was calculated; an example of the formula used to calculate error can be seen below.

$$\text{Error}_{j,i} = \frac{1}{\text{length of assay}_j} \cdot \sum_{i=1}^{160000} \sum_{j=1}^{length} (\text{actual}_j - \text{simulated}_{j,i})^2 \tag{3.10}$$

The error of each simulated TEG tracing for each patient was then sorted by increasing values to determine the optimal number of nearest neighbors. By plotting the sorted error values versus number of nearest neighbors it should be possible to find a "knee" in the curve. This would indicate that the rate of performance loss is increasing per incremental inclusion of another neighbor and would therefore decrease the accuracy of the machine learning algorithm, which is grounds to stop including subsequent nearest neighbors. However, the results of this procedure, highlighted in Figure 14, show that there is not always a defined knee in the curve. This makes determining the optimal number of nearest neighbors more challenging since each patient has a varying "knee", or optimal number of neighbors, when then nearest neighbors are plotted in order of increasing error. The range of the optimal number of nearest neighbors based on the "knee" method was between 1 and 40 across all patients. The bottom plot in Figure 14 shows a different patient's sorted error plot; it is clear that this curve has a knee around 16 neighbors. Ultimately, 10 nearest neighbors were chosen as a good balance between the number of neighbors and minimizing the error for the last

nearest neighbor. By using the NSSE of each of the neighbors, the worse allowable neighbor had a NSSE value of 0.6. An NSSE value above this affected the training of the machine learning algorithm. The average NSSE value of each neighbor throughout the entire dataset (n=8,740) was 0.22 (with a range of 0.0001 to 0.58). Since there was a total of 874 TEG tracings and 10 nearest neighbors were implemented, there are a total of 8,740 parameter groupings used for this analysis.

3.1.6 Predicting LY30

After determining the optimal number of nearest neighbors to be 10, logistic regression was implemented for its ease of use and computational effeciency to predict if a patient was going to experience high or low lysis rates. This algorithm is able to predict which side of the threshold value (*i.e.* above or below) the patient will fall Logistic regression (LR) is a generalization of the general linear regression that defines an outcome as either positive (LY30 $\geq$ threshold) or negative in relation to a set of inputs. Using the four previously discussed model inputs (β, k_4, k_5, Ψ) as the basis for classification, Matlab (©2022 The Mathworks, Natick, MA) was implemented to train, test and validate the model using a $\frac{1}{3}$ split for each set. This was done because there was a sufficient number of data points to have training, testing, and validation cohorts of data.

3.1.7 Balancing Classes of Data with SMOTE

After determining the actual LY30 values for each patient in the data set by comparing the TEG MA to amplitude 30 minutes after MA, over 75% of the classifications were for LY30 $\geq$ 3%. This necessitated balancing the population of each class for all three data sets to ensure accurate and robust model prediction. To accomplish this the synthetic minority oversampling technique (SMOTE) was implemented as to not create duplicate entries that could either decrease the accuracy of the regression or oversample a particular set of parameters [14].

SMOTE works by creating synthetic points (in variable space), rather than duplicating the data. It does this by first selecting random samples from the minority class, specifically

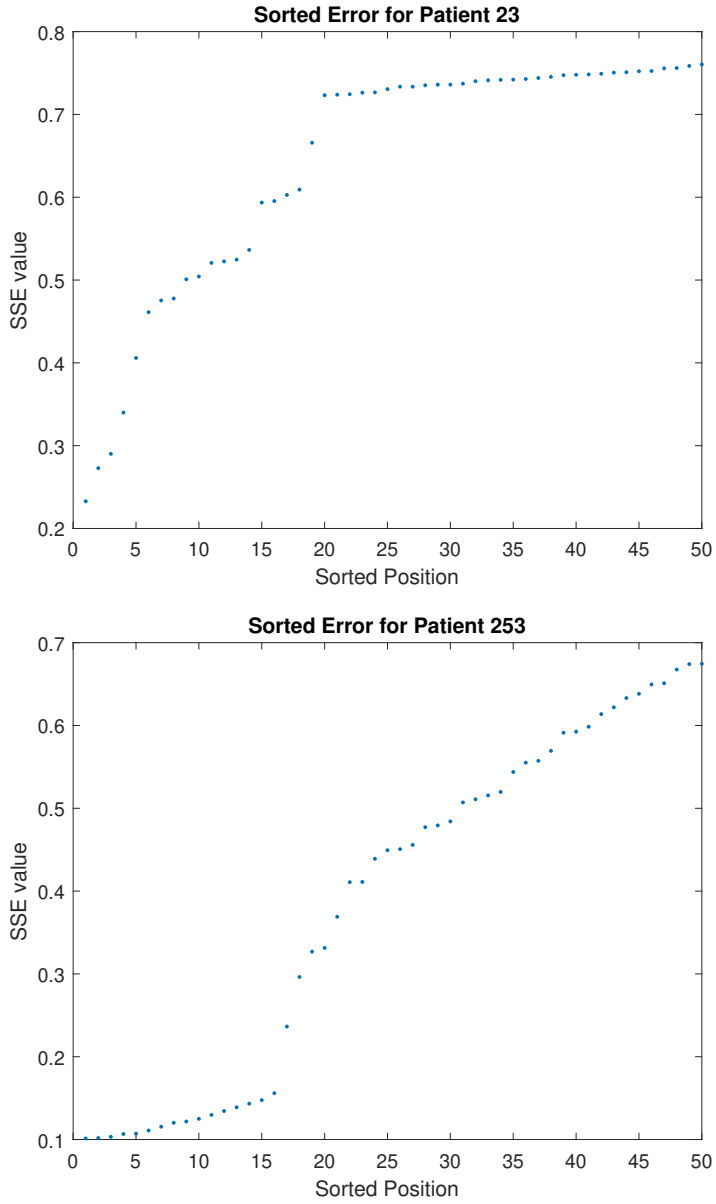

Figure 14: Sorted error values for two different patients; **Top:** patient 23 that shows no discernible knee in the sorted error values; **Bottom:** patient 253 that shows a clear knee at 16 neighbors.

the class that was negative for high LY30. Then, the nearest neighbor(s) to the minority class sample are determined. These neighbor(s) are then used to create a vector between the randomly selected minority sample and the selected neighbor(s); this vector is then multiplied by a random number between 0 and 1 and added to the randomly selected minority sample. This process is repeated for the required number of synthetic samples so that the class populations are approximately equal [14]. This step is needed because the training set requires an approximate even split between the two classes (above and below the threshold values) of data, which would cause the testing and validation sets to be comprised of only about 12% of the lesser class. After running the logistic regression prior to using SMOTE, there was not enough samples in the class below the threshold values to produce any true negative results, thus giving lower values for the area under the curve (AUC) on the receiver operating characteristic (ROC) curve. In Figure 15, you can see the original parameter distribution of the model when using the nearest neighbors is nicely discretized prior to using SMOTE. The red points are negative for high LY30 and the blue is positive. After using SMOTE you and see that there are a lot of redpoints between the nicely discretized blue points in Figure 16. This was able to give me sufficient balance of data that I could then split randomly into training, testing, and validation data.

3.2 Results

The following procedure was used to predict whether a patient is going to experience lysis rates above or below a threshold value:

1. Filter the TEG tracings that would be impossible to model from the original set (n=1,045).
    a. Filter out tracings with less than 1,000 seconds of data (994 TEG tracings remain)
    b. Filter out tracings with an average amplitude of less than or equal to 2mm (932 TEG tracings remain)
    c. Filter out tracings with an average amplitude less than or equal to 5mm and MA less than or equal to 7mm (913 TEG tracings remain)

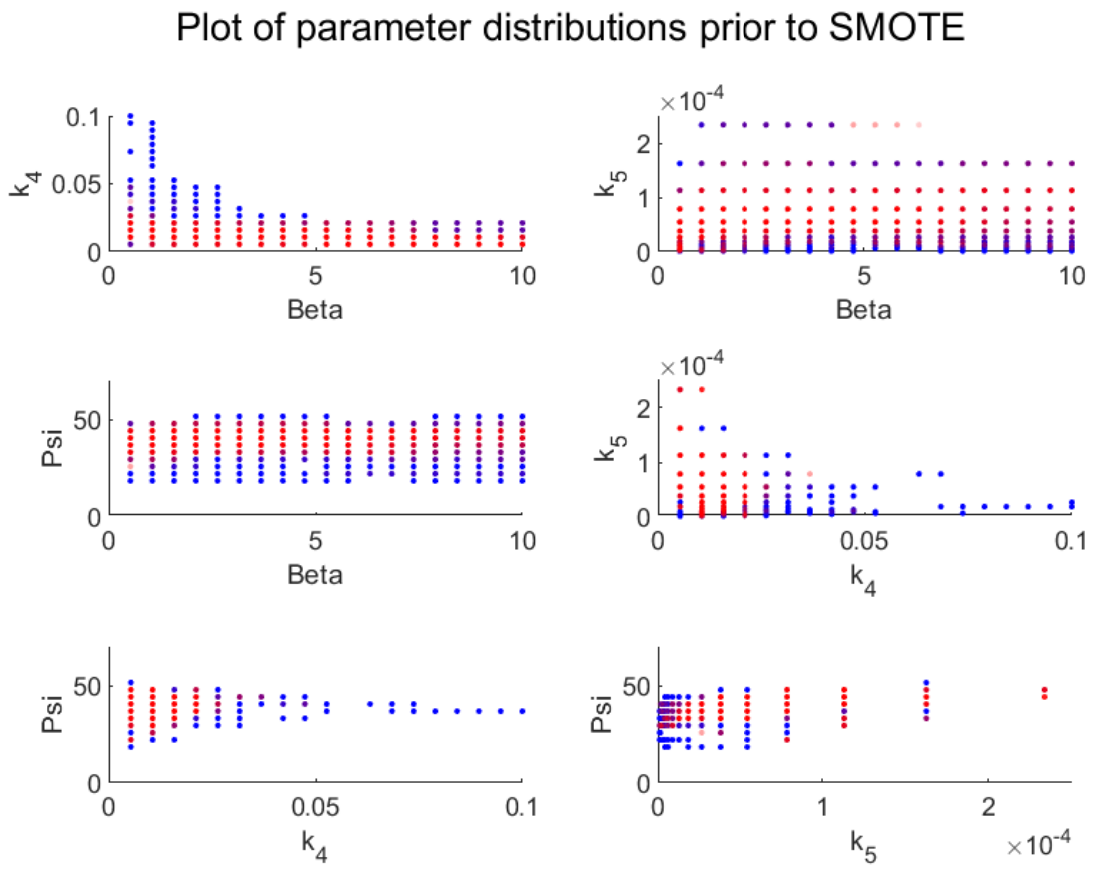

Figure 15: Parameter space distribution of nearest neighbors prior to using SMOTE.

d. Filter out any tracings with a negative jump greater in magnitude than the greatest positive jump (886 TEG tracings remain)

e. Filter out any tracings that have a jump greater than or equal to 2.5mm in magnitude in either direction (874 TEG tracings remain)

f. Filter out any tracings that exhibit an activated clotting time (ACT) longer than 150 seconds (848 TEG tracings remain)

2. Calculate the actual lysis rates of each of the remaining patients and assign them a binary variable based on being above the chosen threshold

3. Create the simulated TEG tracing library of 160,000 different TEG tracings

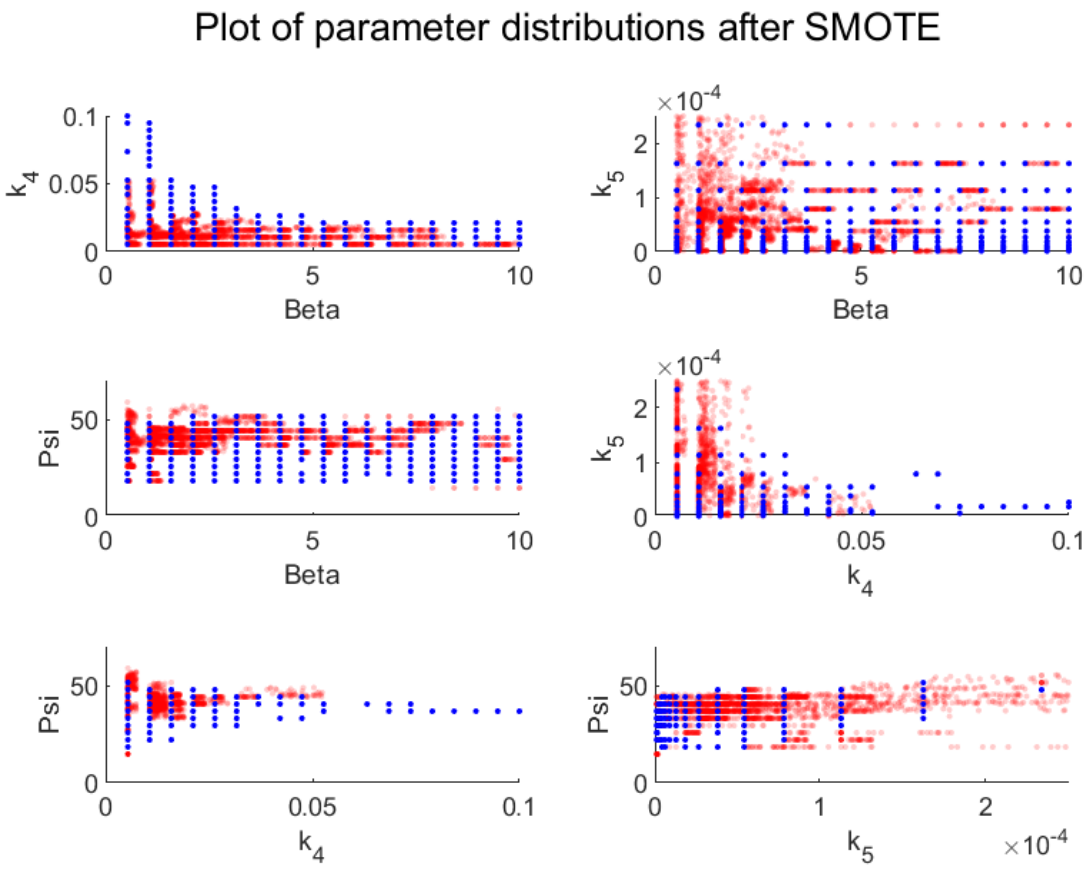

Figure 16: Parameter space distribution of nearest neighbors after using SMOTE.

4. Determine the normalized sum of squared error between each of the 848 patients to each of the 160,000 simulated TEG tracings
5. Take the 10 parameter sets with the lowest error values to be the nearest neighbors and assign each of these neighbors to have the same binary variable as the patient they are the neighbor to
6. Implement SMOTE to balance the classes of data so there is a near 50/50 balance between positive and negative classes
7. Randomly split the data in training, testing, and validating sets
8. Initialize the logistic regression to predict whether a patient is going to expeiernce high

or low lysis rates

3.2.0.1 Data Preprocessing

Before the data can be analyzed it must be filtered to exclude any TEG tracings that will be impossible for the ODE model to replicate. Since the TEG measurement device is highly sensitive to outside vibrations, there are a number of exclusionary criteria that must be considered when removing TEGs from the analysis.

The first exclusion criterion is to remove all TEG tracings that have less than 1,000 seconds worth of data to allow for enough data for the algorithm to analyze; operators can stop the TEG early if there is an issue, such as someone bumping into the table while the TEG is running. There were 51 tracings that had to be removed for this reason, and examples of them can be seen in Figure 17.

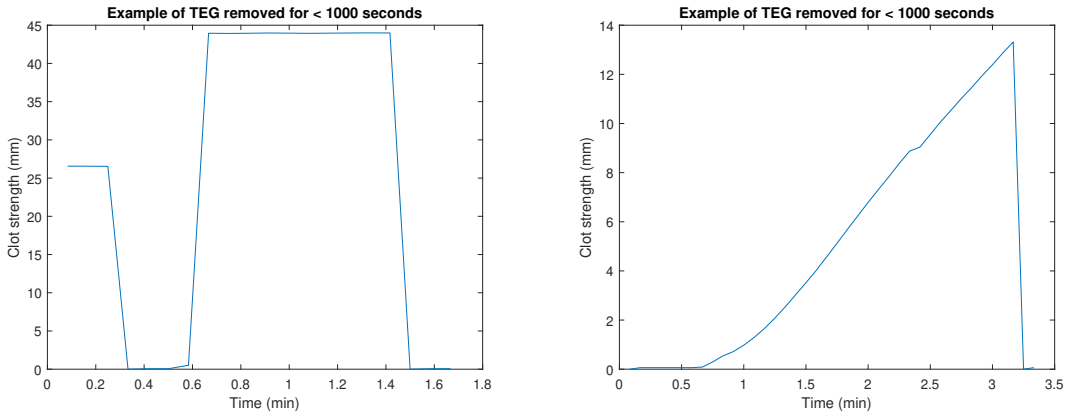

Figure 17: Examples of TEG tracings removed for having less than 1000 seconds of data

The second exclusion criterion is to remove TEG tracings based on average amplitude. Any tracing that had an average amplitude of 2 mm or less was removed because the TEG never initialized (*i.e.* the clot never began to form within the oscillating specimen cup) and would not contain any useful information regarding patient status. 62 TEGs were removed for this reason, and examples can be seen in Figure 18.

The third exclusion criterion was any TEG that had an average amplitude of less than 5 mm and a max amplitude of 7 mm or less. These tracings never reached a reasonable

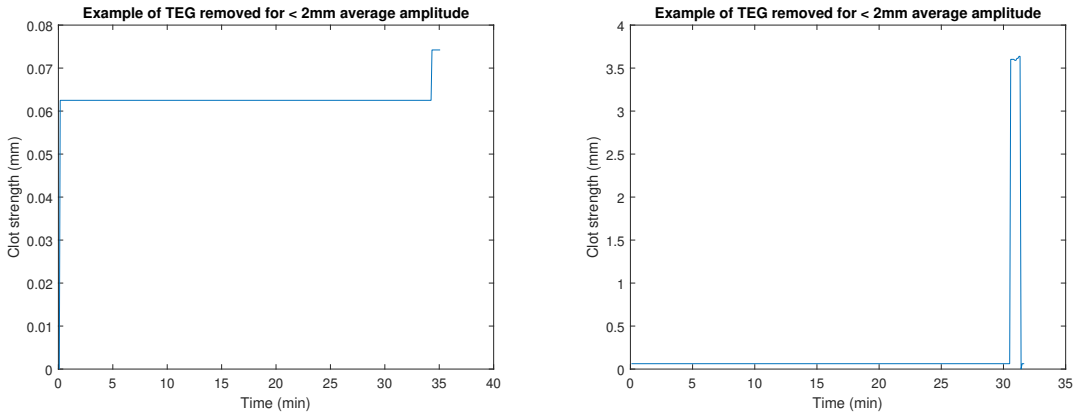

Figure 18: Examples of TEG tracings removed for having less than 2 mm average amplitude.

maximum amplitude. A total of 19 tracings had to be removed for this reason, and examples can be seen in Figure 19.

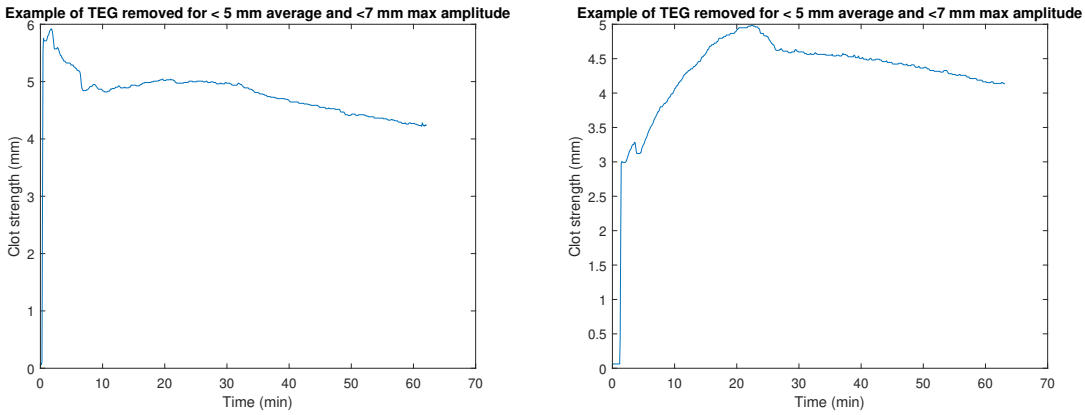

Figure 19: Examples of TEG tracings removed for having less than 5 mm average amplitude or less than 7 mm maximum amplitude.

Additionally, any tracing that had a large negative jump greater in magnitude than the maximum positive jump between data points was removed because it is unrealistic for the TEG tracing to exhibit these discontinuities. 27 TEG tracings were removed for this reason, and examples can be seen in Figure 20.

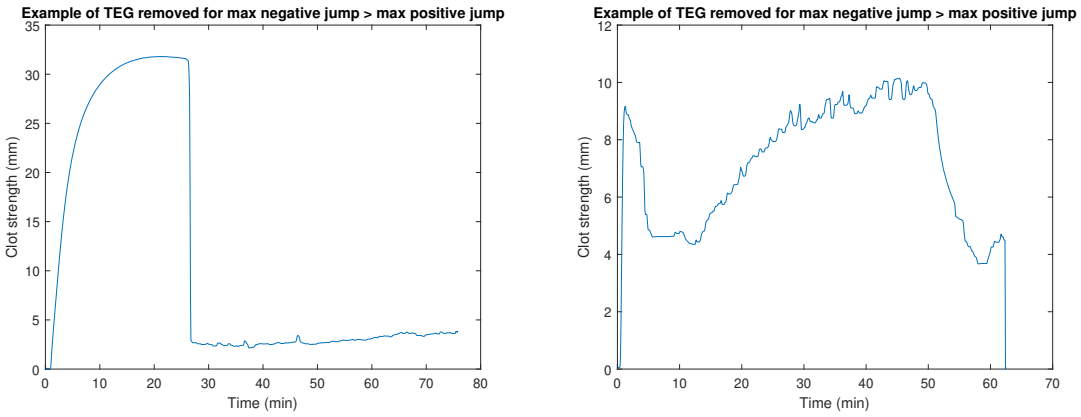

Figure 20: Examples of TEG tracings removed for having a negative jump greater in magnitude than the maximum positive jump.

Finally, any TEG that had a jump greater than or equal to 2.5 mm in either direction was removed because the model cannot reasonably fit discontinuities this large. A total of 12 TEG tracings were removed for this reason, and examples can be seen in Figure 21.

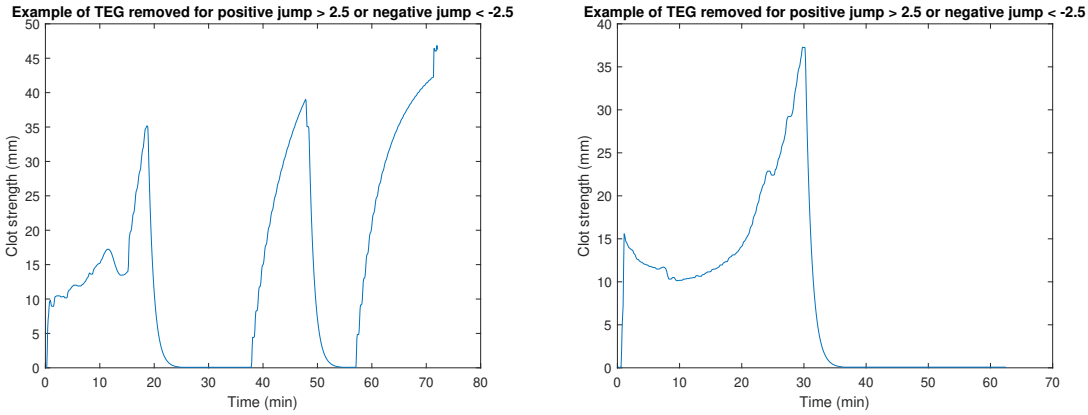

Figure 21: Examples of TEG tracings removed for having a positive jump greater than 2.5 mm or a anegative jump less than 2.5 mm.

After filtering the data by removing TEG tracings without enough sampling time, low amplitudes, large discontinuities, and longer than standard ACT times, of the original 1,045

TEG tracings there were a total of 848 tracings that are able to be analyzed.

3.2.1 Results of Determining the Real MA Time

The results of this can be seen in Figure 22. It is clear that using the threshold value for noise to be 0.1mm resulted in very few tracings that are still exhibiting significant noise; regardless of the starting time for MA identification, there are at most two patient TEGs with noise above the threshold value. However, using the smaller threshold of 0.025mm resulted in a maximum of seven tracings that show significant noise when starting MA identification at seven minutes pre-MA. Since the data set for this analysis contains 848 tracings, the number of patients that exhibit noise from MA to 30 minutes pre-MA is 36 and 11 for the lower and upper noise thresholds, respectively. Therefore, it is reasonable to expect MA has been reached if the tracing is no longer continually increasing by more than 0.025mm across the four successive points. At roughly five minutes before MA occurs there is an insignificant portion of patients with noise higher than the threshold value, and at two minutes pre-MA no patients exhibit noise above either threshold. Determining real MA time is important because the logistic regression will be done at various time intervals both from the beginning of the TEG tracing and also at specific intervals following MA.

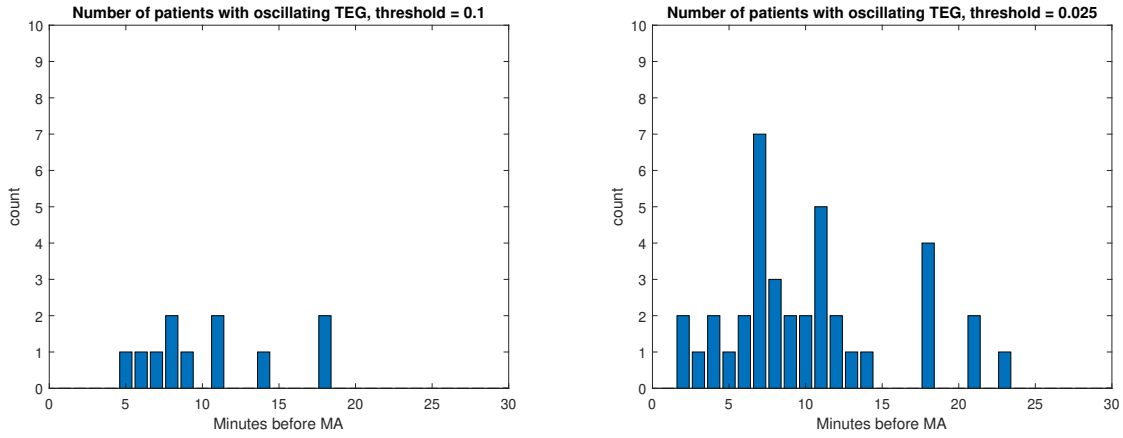

Figure 22: Histogram of patient TEG tracings that oscillate prior to MA.

3.2.2 Predicting LY30 at Multiple Time Intervals from TEG

The goal of this research is to determine how soon an accurate prediction of high or low lysis rates can be determined from the TEG tracing. Currently, the clinician waits at least 30 minutes post-MA before making an inference regarding the LY30 value of the patient and intervening, typically with tranexamic acid (TXA) [65]. The current standard threshold for high LY30 varies, but 3% is regarded as the *lower* high threshold for intervention and 8% as the *upper* high threshold. [62]. Univariate logistic regressions were performed on both the lower and upper threshold values for LY30 at various time intervals before and after MA to determine how soon clinicians can predict abnormal lysis rates.

3.2.2.1 Univariate Logistic Regressions

Using both the lower and upper high threshold LY30 values, nine different time intervals were used to see how early the algorithm could predict abnormal LY30 values. One set intervals are from the start of the TEG tracing to five minutes, 30 minutes, 35 minutes, time of MA plus five minutes, time of MA plus 30 minutes. A second set of intervals start at from time of MA and look ahead three minutes, five minutes, and 30 minutes beyond MA. Since the median time for MA is 32 minutes, as shown in the normal distribution in Figure 23, the accuracy for 35 minutes and start to MA plus five minutes should be similar, as should start to MA plus 30 minutes and start to 60 minutes since they are taken at approximately the same time on the TEG tracings. The differences in the nearest neighbors that are populated based on using different time intervals is highlighted in Figures 24-25.

3.2.2.2 Predicting LY30 $\geq$ 3%

By using the *lower* high LY30 threshold value of 3% in conjunction with the various time intervals of the TEG tracing, logistic regression can be used to predict high and low LY30 values as well as the relative importance of each parameter of the ODE model in determining the prediction. A receiver operating characteristic (ROC) curve can be seen in Figure 26. It is clear that starting at MA, without the beginning rise of the TEG tracing,

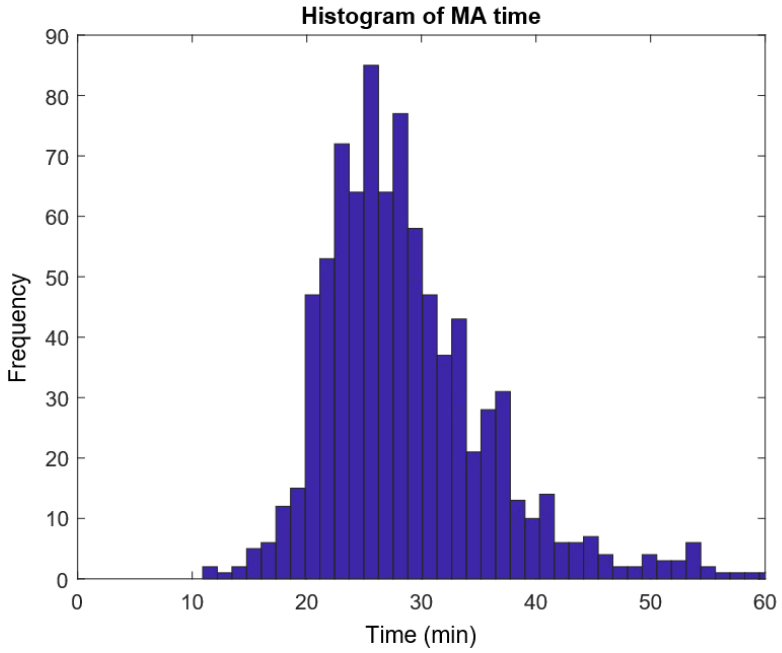

Figure 23: Histogram of patient TEG time to MA.

provides the most accurate results. This is because the algorithm is not concerned about the rise, but only the fall after reaching MA. It is clear from Table 3, that starting at MA provides more accurate predictions compared to starting from the beginning of the tracing; it is also evident that k_5, the parameter in the ODE model that governs the rate of lysis, plays the most important role in the prediction of high LY30. In accordance with the results of median MA time, the AUC values for start to MA + 30 and 60 minutes, as well as start to MA + 5 and 35 minutes, have very similar AUC values because they occur at roughly the same point on the TEG tracing. Although using the time interval of MA + 30 provides the highest AUC, this analysis would take approximately 60 minutes. The AUC of the ROC at MA + 5 provides a reasonably accurate prediction and saves 25 minutes over the MA + 30 time interval. The entire set of ROC curves and confusion matrices for this *lower* high LY30 threshold can be seen in 6.1-6.2; the "0" classifier represents patients who fall below

the threshold (do not have elevated lysis rates), while "1" represents those that are above the threshold.

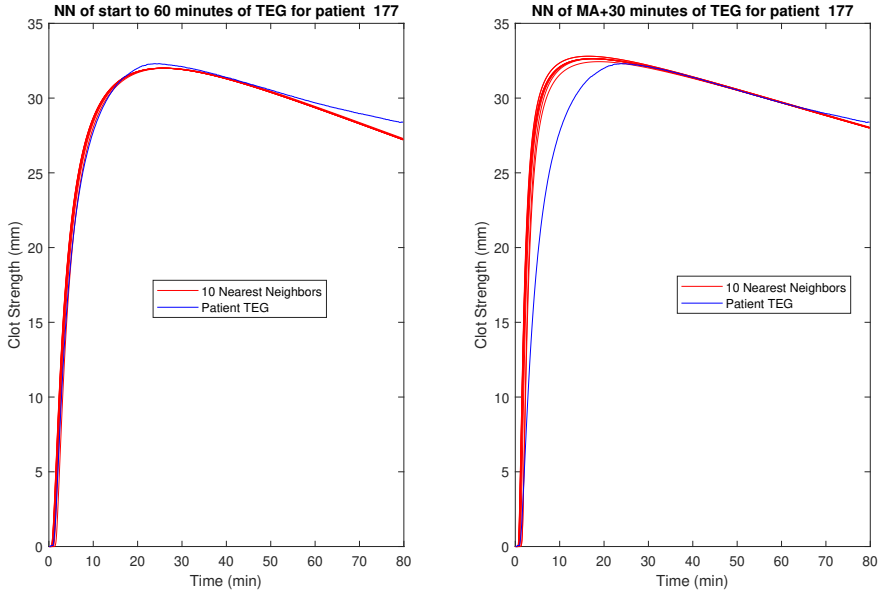

Figure 24: Nearest neighbors to patient 177

3.2.2.3 Predicting LY30 $\geq$ 8%

By using the *upper* threshold for LY30 ($>$ 8%), with the various time intervals of the TEG tracing and logistic regression LY30 values above the upper threshold can be predicted as well as the relative importance of each parameters' impact on the prediction. The ROC curve can be seen in Figure 27. Agreeing with the *lower* LY30 threshold value of 3%, it is clear that starting at MA, rather than using the whole tracing from the beginning provides more accurate predictions. However, this prediction is slightly less accurate, as evidence by the lower AUC values in Table 4; thiscould be due to the low number of patients within the data set (n=95) and the significant overlap between the k_5 values of the different classes of datathat experience LY30 $\geq$ 8This is highlighted in Figure 28. Also, in agreement with the *lower* threshold value for LY30, k_5 is by far the most important parameter when predicting LY30 values through logistic regression. A patient that has an LY30 value above this *upper*

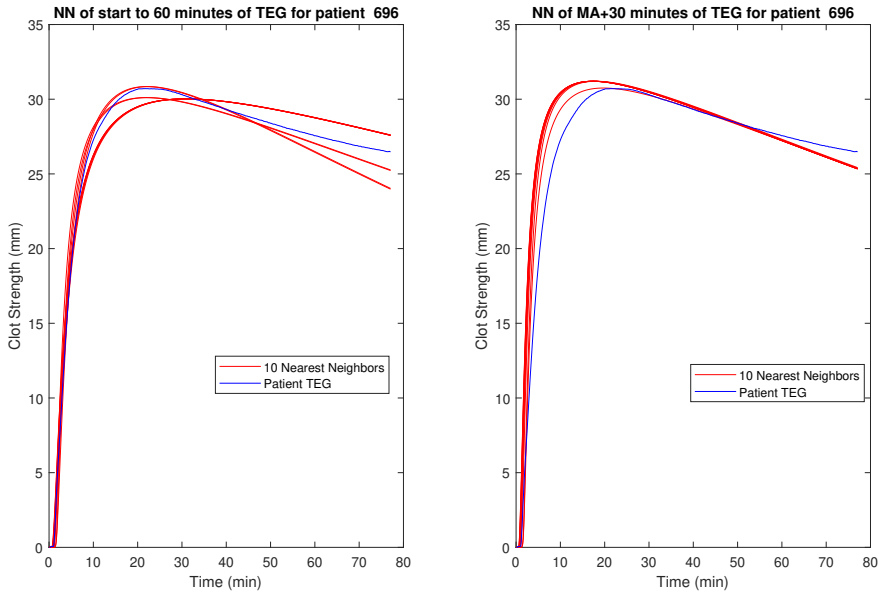

Figure 25: Nearest neighbors to patient 696.

threshold will require more expeditious intervention than the *lower* threshold value to prevent mortality and permanent injury. The entire set of ROC curves and confusion matrices for this *upper* high LY30 threshold can be seen in 6.3-6.4; the "0" classifier represents patients who fall below the threshold, while "1" represents those that are above the threshold.

3.2.2.4 Predicting 3% $\leq$ LY30 $<$ 8% vs. LY30 $\geq$ 8%

Determining the magnitude of a patient's LY30 value is of interest to clinicians so they can expedite treatment and/or have other interventions on-hand if the LY30 value is $\geq 8\%$. This regression was done on patients only with LY30 values $\geq 3\%$, as to eliminate those patients who do not experience abnormal LY30 values. The ROC curve for the various time intervals of this analysis can be seen in Figure 29, and the AUC and relative importance of k_5 can be seen in Table 5. The AUC values of this analysis are comparable to the values of predicting LY30 values $\geq 8\%$, as are the relative importance of k_5. As evident in Table 5, although the time interval from MA+30 provides the highest AUC of the ROC curve,

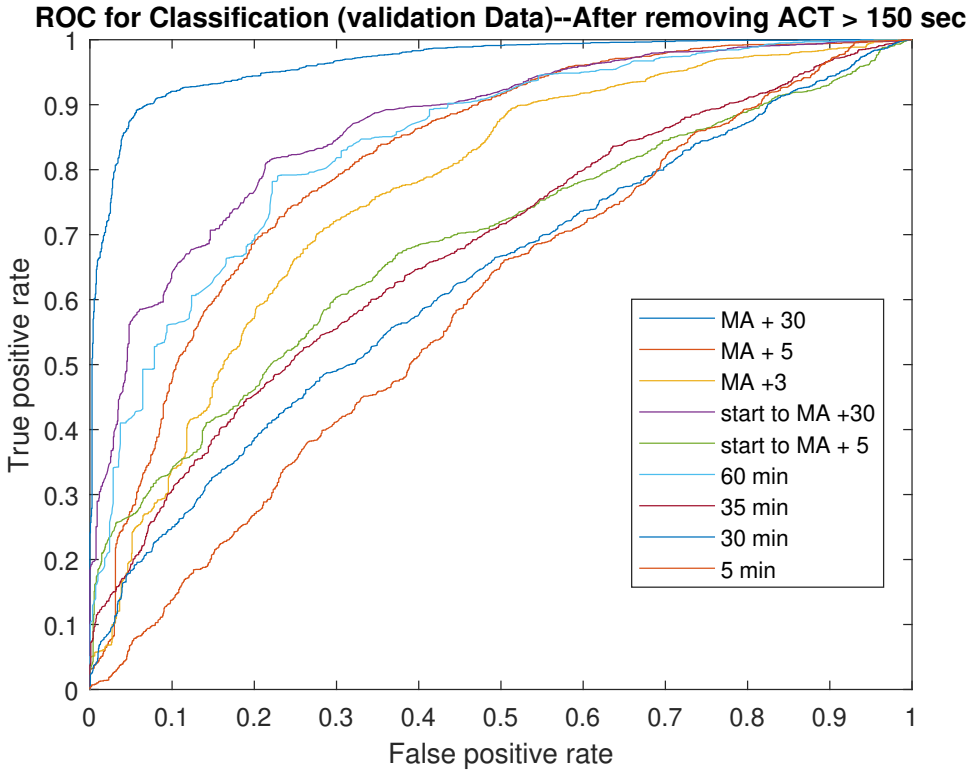

Figure 26: ROC curve for validation data at various time intervals in TEG tracing ($\geq 3\%$).

with start to 60 minutes being the second best, both of these analyses would take roughly 60 minutes to accomplish. Starting from MA and going five minutes beyond provides nearly the same AUC compared to start to 60 minutes and the analysis would require 25 less minutes to compete. This indicates that starting from MA would still be preferrable in accelerating the prediction of abnormal lysis. This is the primary reason why it is imperative to know when MA occurs, so the regression can be started at that particular point. Alternatively, higher accuracy may be achieved by first predicting if the patient will experience LY30 $\geq 3\%$, then using the positive results from that analysis to predict the severity of the abnormal LY30. The entire set of ROC curves and confusion matrices for this *upper* high LY30 threshold can be seen in 6.5-6.6; the "0" classifier represents patients who fall below the threshold ($3\% \leq$ LY30 $< 8\%$), while "1" represents those that are above the threshold ($\geq 8\%$).

Table 3: AUC of ROC curve and Relative Importance of LY30 $\geq$ 3%

Time Interval	AUC	Relative Importance of k_5	Approximate Time (min)
MA + 30	0.9653	0.9998	60
MA + 5	0.8393	0.9996	35
MA + 3	0.7652	0.9995	33
Start to MA + 30	0.8680	0.9985	60
Start to MA + 5	0.6850	0.9971	35
Start to 60	0.8562	0.9983	60
Start to 35	0.6773	0.9964	35
Start to 30	0.6428	0.9948	30
Start to 5	0.6186	0.9893	5

3.2.3 Multivariate Logistic Regression

Following the univariate regressions, a multivariate regression was performed to determine if the model can accurately predict multiple severity levels of LY30 at the same time; this analysis focused on discerning LY30 values $< 3\%$, $3\% \leq$ LY30 $< 8\%$, and LY30 $\geq 8\%$. The results of using the time interval from MA + 30 minutes can be seen in Figure 30. It is clear that the model has a more difficult time predicting different levels of LY30 values, compared to the binary prediction. The entire set of confusion matrices for this *upper* high LY30 threshold can be seen in 6.7. Here, the "0" classifier represents those patients with LY30 $\geq 8\%$; "1" represents those patients with LY30 values such that $3\% \leq$ LY30 $< 8\%$; and "2" represents those patients with LY30 $< 3\%$. It is clear that the predictor is much more successful at predicting patients with LY30 $\geq 8\%$ and LY30 $< 3\%$, rather than $3\% \leq$ LY30 $< 8\%$. Rather than using a multivariate regression to determine the severity of LY30, it would be more accurate to first run a regression to determine if the patient's LY30 value $\geq 3\%$, then running another regression to determine if the patient's LY30 is $\geq 8\%$, provided the result is positive on the first regression as evidence by Tables 3-

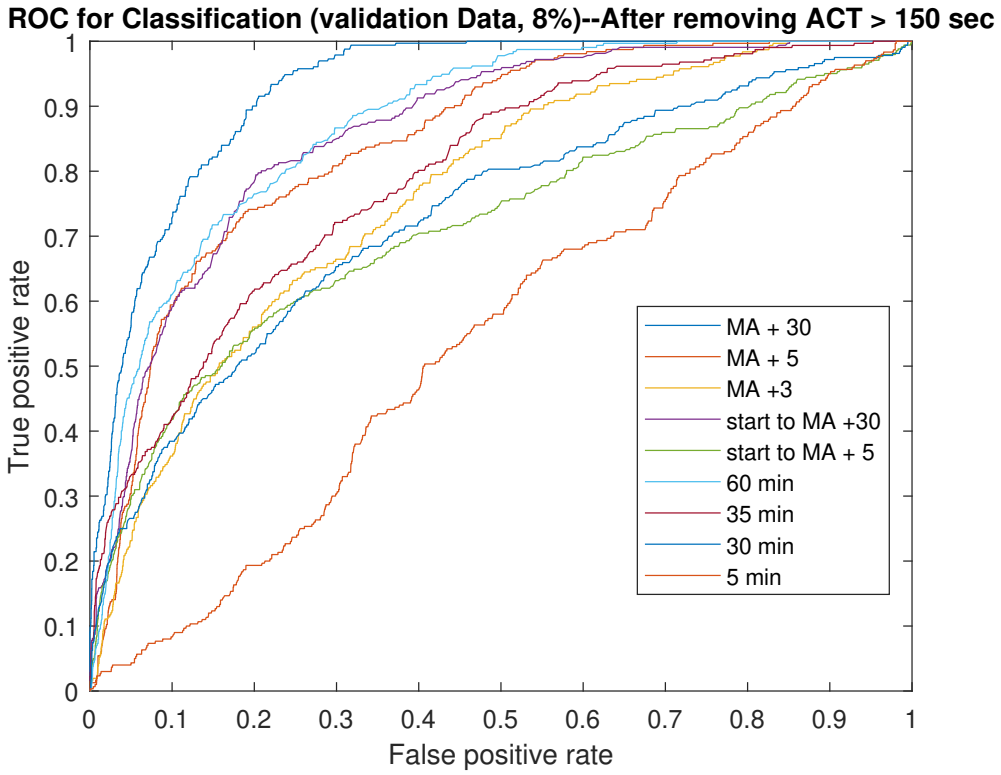

Figure 27: ROC curve for validation data at various time intervals in TEG tracing ($\geq 8\%$).

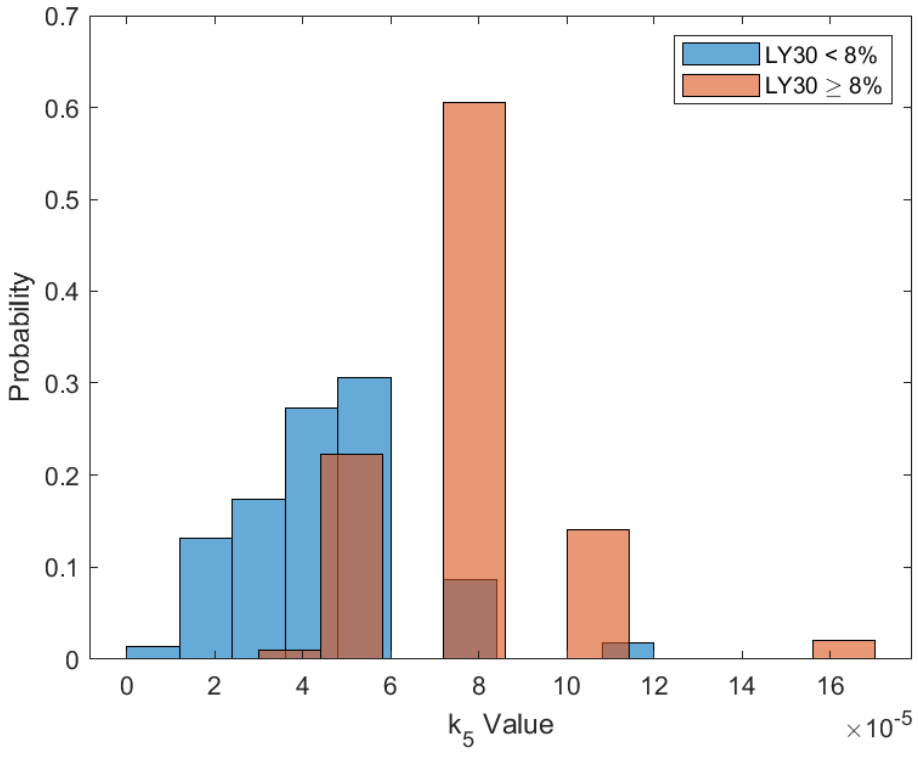

Figure 28: Probability density histogram of the k_5 values for LY30 $< 8\%$ and LY30 $\geq 8\%$.

Table 4: AUC of ROC curve and Relative Importance of LY30 $\geq 8\%$

Time Interval	AUC	Relative Importance of k_5	Approximate Time (min)
MA + 30	0.9416	0.9997	60
MA + 5	0.8472	0.9996	35
MA + 3	0.8135	0.9996	33
Start to MA + 30	0.8649	0.9979	60
Start to MA + 5	0.7299	0.9951	35
Start to 60	0.8741	0.9984	60
Start to 35	0.8206	0.9973	35
Start to 30	0.7470	0.9959	30
Start to 5	0.5407	0.9748	5

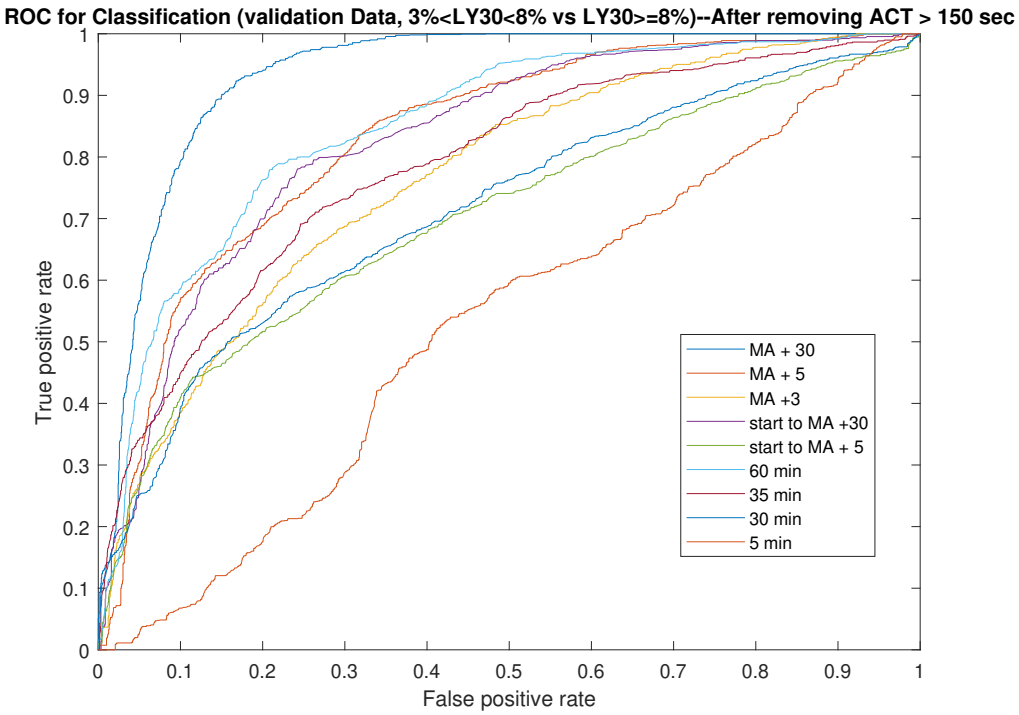

Figure 29: ROC curve for validation data at various time intervals in TEG tracing ($3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$).

Table 5: AUC of ROC curve and Relative Importance of $3\% \leq \text{LY30} < 8\%$ vs. $\text{LY30} \geq 8\%$.

Time Interval	AUC	Relative Importance of k_5	Approximate Time (min)
MA + 30	0.9485	0.9997	60
MA + 5	0.8431	0.9996	35
MA + 3	0.7787	0.9996	33
Start to MA + 30	0.8374	0.9970	60
Start to MA + 5	0.7076	0.9943	35
Start to 60	0.8510	0.9976	60
Start to 35	0.7878	0.9961	35
Start to 30	0.7292	0.9952	30
Start to 5	0.5475	0.9336	5

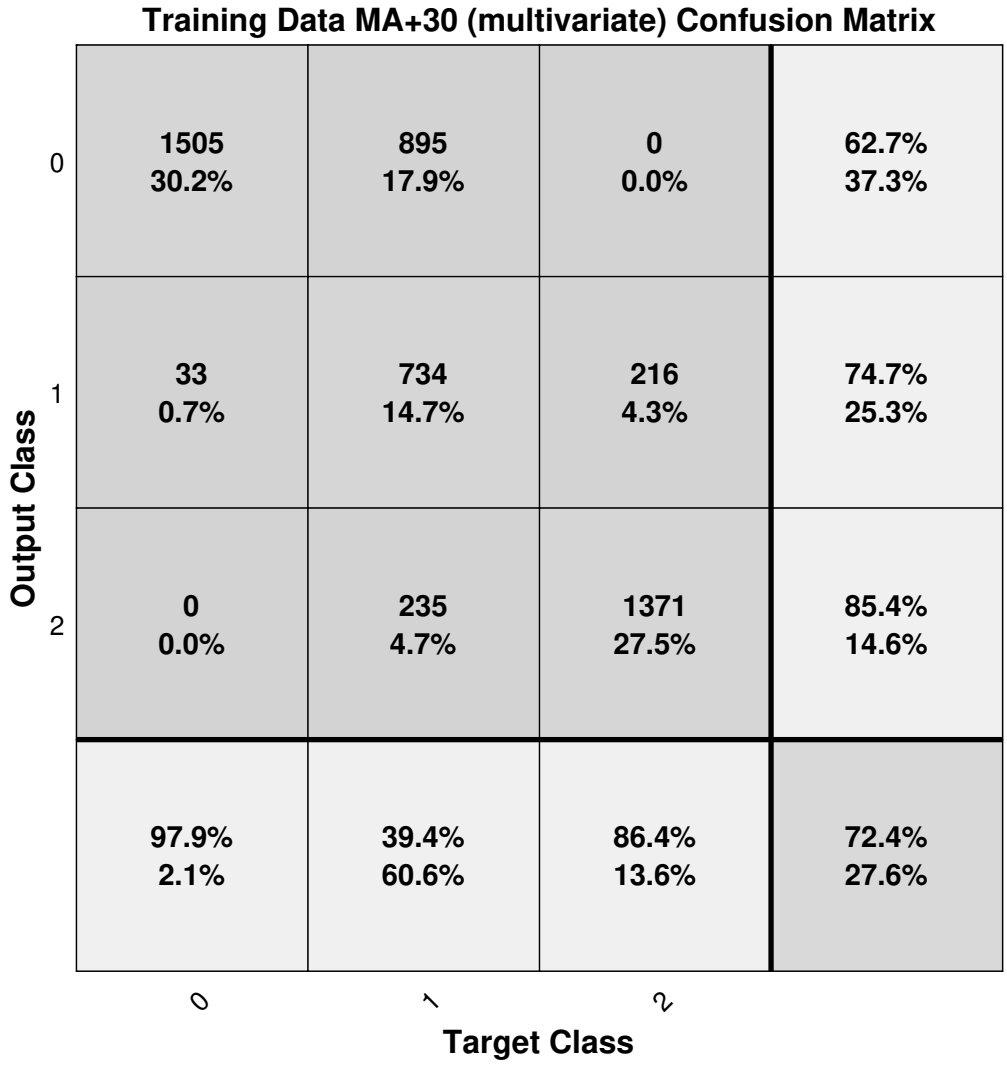

Figure 30: Confusion matrix for multivariate regression using MA + 30 time interval of TEG tracing.

4.0 Summary

By taking a systems medicine approach to complex problems such as optimal dose scheduling of chemotherapy treatment and predicting abnormal lysis rates in trauma patients, it is possible to distill these problems down to their core to create realizable solutions. The use of physiologically-based, simplified ODE models, provides computationally solutions that allow implementation in a clinical setting where they can be used to support the treatment decisions clinicians make regarding their patients, while also improving the quality-of-life and long-term prognosis of said patients.

4.1 Cancer Chemotherapy Treatment Algorithm

By combining PK models of drug distribution via a compartmental model and PD models of neutrophil toxicity and tumor cell growth/eradication, a cyclic chemotherapy treatment scheduling algorithm was synthesized. By implementing a model predictive control framework, the receding-horizon control is mathematically suboptimal, but does allow for more flexible constraint handling and the use of nonlinear models that would be difficult to implement in optimal control scenarios. This mathematically suboptimal solution is guided by the use of clinically-relevant constraints that would be easy for a clinician to understand and implement by mapping them onto the PD toxicity model in conjunction with logistic constraints such as time of dosing and number of doses that impact both the patient and the clinic/insurer. By using a parallel simulation architecture in a GPU, the calculations are fast and computationally efficient which allows for real-time use in a clinical setting, rather than waiting hours or even days for an optimization algorithm to return a (sub)optimal solution. Additionally, the algorithm can be updated at key points during treatment (*e.g.* at the end of each cycle or as patient measurements are collected) to reflect the actual patient responses of drug toxicity to neutrophils and drug efficacy on the tumor. This allows for subsequent clinically-optimal schedules to be tailored to what the patient is experiencing, rather than

solely based on what the model predicts. This treatment design algorithm could serve to improve patient quality-of-life, as well as serve as a framework for creating additional single-, and even multi-agent treatment schedule designs as other PK and, and more importantlyPD models are available.

4.2 Predicting Coagulopathies from TEG

Through the use of TEG data, a simplified clotting cascade ODE model, and multiple machine learning techniques (including K-nearest neighbors and logistic regression), it is possible to predict whether a trauma patient is going to experience various threshold values of LY30. The complete TEG tracing data set required pre-processing as the model is unable to capture discontinuities and TEG tracings that are unrealistic (*e.g.* TEG too short to provide patient insight, insufficient amplitude, and physically-impossible jumps in either direction). A simulated TEG tracing library was created by discretizing key variables within the simplified ODE model and iterating through to create 160,000 different TEG tracings for use in K-nearest neighbors identification. The nearest neighbors in parameter space are determined via normalized error function and the simulated TEG tracing library. Although this model has its challenges in capturing patients that experience longer than normal ACT times and overly noisy TEG tracing profiles, this the overall number of excluded TEGs is a small fraction of the total patient data. While MA+30 minutes provides the highest accuracy in prediction, this analysis would take roughly 60 minutes to complete. However, by implementing the time segment from MA+5, which has a similar predictive capability, the clinician could save 25 minutes before making an inference regarding treatment. Clearly not incorporating the initial rise of the TEG, but rather only from MA forward, the prediction accuracy increases significantly. This algorithm can be implemented in a clinical setting where it has the potential to not only improve time to intervention, but also patient quality-of-life. K-nearest neighbors and logistic regression, which takes a couple minutes to run, are significantly less computationally-intensive than other parameter estimation tools such as affine parallel tempering Markov chain Monte Carlo methods, which takes hours to run.

4.3 Future Work

The cancer chemotherapy work presented herein is modular and can be expanded to include multiple drug treatments, in both single- and multi-agent treatments. This is because the objective function for optimal treatment can be augmented to include multiple toxicity classes (*e.g.* nephrotoxicity due to elevated blood urea nitrogen and creatinine levels during treatment with Cisplatin) through models akin to the PD neutrophil toxicity model and compartmental PK model of drug distribution throughout the body. By including multiple agents in treatment, the tumor will become less resistant to any one drug, and the patient's optimal dose should not have to increase as fast due to resistance to any one drug, which also alleviates excess toxicity concerns. However, incorporating multiple drugs could present challenges algorithmically because combination drug interactions would need to be incorporated into the model. This could also lead to toxicities in other parts of the body, that the current PK model has been reduced to exclude.

The coagulation work can also be expanded through the prediction of multiple TEG variables simultaneously. For example, one regression could be performed to predict abnormal LY30 values, while simultaneous regressions can be run to predict abnormal MA and α-angle values. This could be used to make better inferences as to what intervention is required. However, predicting multiple assay parameters simultaneously can present problems if the first prediction is inaccurate; this would lead to exacerbated inaccuracy of the following prediction. Additionally, some type of delay, either explicitly incorporated or generated by the use of a different function (*e.g.* a Hill function) could be incorporated into the ODE model to allow for the capture of patient data that has longer than normal ACT. Alternatively, incorporating a discretized temporal delay into each of the simulated TEG tracings would allow for better results when implementing the ODE model, but these options would increase computation time and could present identifiability concerns. Furthermore, this application can be applied to more novel assays such as rotational thromboelastometry (ROTEM), where more aspects of the clotting cascade can be quantified though assay variables. This would also aid in making clinical decisions regarding what intervention would most benefit the patient.

5.0 Supporting Equations for Cancer Model

5.1 35-state docetaxel PBPK model developed by Florian *et al.*[16]

The complete physiologically-based PK model of docetaxel distribution throughout the body:

Venous Blood:

$$\frac{dC_{ven}}{dt} = \frac{1}{V_{ven}(1-f_{hem})}(F_{in}C_{in} - F_{tot}C_{ven}) + \frac{f_{hem}}{1-f_{hem}}k_{rbcplas}C_{rbcv} - k_{plasrbc}f_{unb}C_{ven} + \frac{u(t)}{V_{ven}(1-f_{hem})} \quad (5.1)$$

$$\frac{dC_{rbcv}}{dt} = -k_{plasrbc}C_{rbcv} + \frac{1-f_{hem}}{f_{hem}}k_{plasrbc}f_{unb}C_{ven} \quad (5.2)$$

Lung:

$$\frac{dC_{lv}}{dt} = \frac{F_{tot}}{V_{lv}}(C_{ven} - C_{lv}) - k_{lve}f_{unb}C_{lv} + \frac{V_{le}}{V_{lv}}k_{lev}C_{le} \quad (5.3)$$

$$\frac{dC_{le}}{dt} = \frac{V_{lv}}{V_{le}}k_{lve}C_{lv} - k_{lev}C_{le} + k_{bind_{out}}C_{lb} - k_{bind_{in}}C_{le} \quad (5.4)$$

$$\frac{dC_{lb}}{dt} = -k_{bind_{out}}C_{lb} + k_{bind_{in}}C_{le} \quad (5.5)$$

Arterial Blood:

$$\frac{dC_{art}}{dt} = \frac{1}{V_{art}(1-f_{hem})}(F_{tot}C_{lv} - F_{tot}C_{art}) + \frac{f_{hem}}{1-f_{hem}}k_{rbcplas}C_{rbca} - k_{plasrbc}f_{unb}C_{art} \quad (5.6)$$

$$\frac{dC_{rbca}}{dt} = -k_{rbcplas}C_{rbca} + \frac{1-f_{hem}}{f_{hem}}k_{plasrbc}f_{unb}C_{art} \quad (5.7)$$

Gut:

$$\frac{dC_{gv}}{dt} = \frac{F_g}{V_{gv}}(C_{art} - C_{gv}) \quad (5.8)$$

Brain:

$$\frac{dC_{bv}}{dt} = \frac{F_b}{V_{bv}}(C_{art} - C_{bv}) - k_{bve}f_{unb}C_{bv} + \frac{V_{be}}{V_{bv}}k_{bev}C_{be} \tag{5.9}$$

$$\frac{dC_{be}}{dt} = \frac{V_{bv}}{V_{be}}k_{bve}f_{unb}C_{bv} - k_{bev}C_{be} + k_{bind_{out}}C_{bb} - k_{bind_{in}}C_{be} \tag{5.10}$$

$$\frac{dC_{bb}}{dt} = -k_{bind_{out}}C_{bb} + k_{bind_{in}}C_{be} \tag{5.11}$$

Spleen:

$$\frac{dC_{sv}}{dt} = \frac{F_s}{V_{sv}}(C_{art} - C_{sv}) - k_{sve}f_{unb}C_{sv} + \frac{V_{se}}{V_{sv}}k_{sev}C_{se} \tag{5.12}$$

$$\frac{dC_{se}}{dt} = \frac{V_{sv}}{V_{se}}k_{sve}f_{unb}C_{sv} - k_{sev}C_{se} + k_{bind_{out}}C_{sb} - k_{bind_{in}}C_{se} \tag{5.13}$$

$$\frac{dC_{sb}}{dt} = -k_{bind_{out}}C_{sb} + k_{bind_{in}}C_{se} \tag{5.14}$$

Liver:

$$\frac{dC_{liv}}{dt} = \frac{1}{V_{liv}}(F_{li}C_{art} + F_gV_{gv} + F_sC_{sv} - (F_g + F_s + F_{li})C_{liv}) \\ - k_{live}f_{unb}C_{liv} + \frac{V_{lie}}{V_{liv}}k_{liev}C_{lie} \tag{5.15}$$

$$\frac{dC_{lie}}{dt} = \frac{V_{liv}}{V_{lie}}k_{live}C_{liv} - k_{liev}C_{lie} + k_{bind_{out}}C_{lib} - k_{bind_{in}}C_{lie} - k_{clli}C_{lie} \tag{5.16}$$

$$\frac{dC_{lib}}{dt} = -k_{bind_{out}}C_{lib} + k_{bind_{in}}C_{lie} \tag{5.17}$$

Kidney:

$$\frac{dC_{kv}}{dt} = \frac{F_k}{V_{kv}}(C_{art} - C_{kv}) - k_{kve}f_{unb}C_{kv} + \frac{V_{ke}}{V_{kv}}k_{kev}C_{ke} \tag{5.18}$$

$$\frac{dC_{ke}}{dt} = \frac{V_{kv}}{V_{ke}}k_{kve}f_{unb}C_{kv} - k_{kev}C_{ke} + k_{bind_{out}}C_{kb} - k_{bind_{in}}C_{ke} \tag{5.19}$$

$$\frac{dC_{kb}}{dt} = -k_{bind_{out}}C_{kb} + k_{bind_{in}}C_{ke} \tag{5.20}$$

Muscle:

$$\frac{dC_{mv}}{dt} = \frac{F_m}{V_{mv}}(C_{art} - C_{mv}) - k_{mve}f_{unb}C_{mv} + \frac{V_{me}}{V_{mv}}k_{mev}C_{me} \tag{5.21}$$

$$\frac{dC_{me}}{dt} = \frac{V_{mv}}{V_{me}}k_{mve}f_{unb}C_{mv} - k_{mev}C_{me} + k_{bind_{out}}C_{mb} - k_{bind_{in}}C_{me} \tag{5.22}$$

$$\frac{dC_{mb}}{dt} = -k_{bind_{out}}C_{mb} + k_{bind_{in}}C_{me} \tag{5.23}$$

Fat:

$$\frac{dC_{fv}}{dt} = \frac{F_f}{V_{fv}}(C_{art} - C_{fv}) - k_{fve}f_{unb}C_{fv} + \frac{V_{fe}}{V_{fv}}k_{fev}C_{fe} \tag{5.24}$$

$$\frac{dC_{fe}}{dt} = \frac{V_{fv}}{V_{fe}}k_{fve}f_{unb}C_{fv} - k_{fev}C_{fe} + k_{bind_{out}}C_{fb} - k_{bind_{in}}C_{fe} \tag{5.25}$$

$$\frac{dC_{fb}}{dt} = -k_{bind_{out}}C_{fb} + k_{bind_{in}}C_{fe} \tag{5.26}$$

Tumor:

$$\frac{dC_{tv}}{dt} = \frac{F_t}{V_{tv}}(C_{art} - C_{tv}) - k_{tve}f_{unb}C_{tv} + \frac{V_{te}}{V_{tv}}k_{tev}C_{te} \tag{5.27}$$

$$\frac{dC_{te}}{dt} = \frac{V_{tv}}{V_{te}}k_{tve}f_{unb}C_{tv} - k_{tev}C_{te} + k_{bind_{out}}C_{tb} - k_{bind_{in}}C_{te} \tag{5.28}$$

$$\frac{dC_{tb}}{dt} = -k_{bind_{out}}C_{tb} + k_{bind_{in}}C_{te} \tag{5.29}$$

Heart:

$$\frac{dC_{hv}}{dt} = \frac{F_h}{V_{hv}}(C_{art} - C_{hv}) - k_{hve}f_{unb}C_{hv} + \frac{V_{he}}{V_{hv}}k_{hev}C_{he} \tag{5.30}$$

$$\frac{dC_{he}}{dt} = \frac{V_{hv}}{V_{he}}k_{hve}f_{unb}C_{hv} - k_{hev}C_{he} + k_{bind_{out}}C_{hb} - k_{bind_{in}}C_{he} \tag{5.31}$$

$$\frac{dC_{hb}}{dt} = -k_{bind_{out}}C_{hb} + k_{bind_{in}}C_{he} \tag{5.32}$$

Other:

$$\frac{dC_{ov}}{dt} = \frac{F_o}{V_{ov}}(C_{art} - C_{ov}) - k_{ove}f_{unb}C_{ov} + \frac{V_{oe}}{V_{ov}}k_{oev}C_{oe} \tag{5.33}$$

$$\frac{dC_{oe}}{dt} = \frac{V_{ov}}{V_{oe}} k_{ove} f_{unb} C_{ov} - k_{oev} C_{oe} + k_{bind_{out}} C_{ob} - k_{bind_{in}} C_{oe} \tag{5.34}$$

$$\frac{dC_{ob}}{dt} = -k_{bind_{out}} C_{ob} + k_{bind_{in}} C_{oe} \tag{5.35}$$

5.2 State-space reduced model for docetaxel distribution

Reduced physiologically-based PK model of docetaxel distribution that preserves blood, tumor, and bone-marrow drug concentrations:

$$\dot{x} = \underline{\underline{A}} x + \underline{B} u \tag{5.36}$$

$$y = \underline{\underline{C}} x \tag{5.37}$$

$$\underline{\underline{\mathrm{A}}} = \begin{bmatrix} -1.258 \times 10^{-4} & -1.258 \times 10^{-4} & 5.422 \times 10^{-4} & -1.825 \times 10^{-5} & -1.147 \times 10^{-5} \\ 3.651 \times 10^{-4} & -1.133 \times 10^{-3} & -3.706 \times 10^{-4} & 2.44 \times 10^{-3} & 1.631 \times 10^{-3} \\ -8.15 \times 10^{-4} & 4.922 \times 10^{-3} & -1.027 \times 10^{-2} & -7.773 \times 10^{-3} & -5.653 \times 10^{-3} \\ -4.65 \times 10^{-4} & 2.67 \times 10^{-3} & 1.128 \times 10^{-2} & -3.372 \times 10^{-2} & -3.837 \times 10^{-2} \\ -2.971 \times 10^{-4} & 1.775 \times 10^{-3} & -7.655 \times 10^{-3} & -3.837 \times 10^{-2} & -0.1122 \end{bmatrix}$$

$$\underline{\mathrm{B}} = \begin{bmatrix} 0.1104 \\ -6.455 \times 10^{-2} \\ 0.135 \\ 0.07957 \\ 0.05083 \end{bmatrix}$$

$$\underline{\underline{\mathrm{C}}} = \begin{bmatrix} 3.1975 \times 10^{-4} & -1.529 \times 10^{-3} & 2.332 \times 10^{-3} & 1.984 \times 10^{-3} & -1.267 \times 10^{-3} \\ 7.672 \times 10^{-5} & -1.529 \times 10^{-3} & -6.8348 \times 10^{-5} & -4.0005 \times 10^{-6} & 2.7321 \times 10^{-6} \end{bmatrix}$$

6.0 Supporting Results for LY30 prediction

6.1 ROC curves for LY30 $\geq$ 3%

Supporting ROC curves for prediction of LY30 $\geq$ 3%:

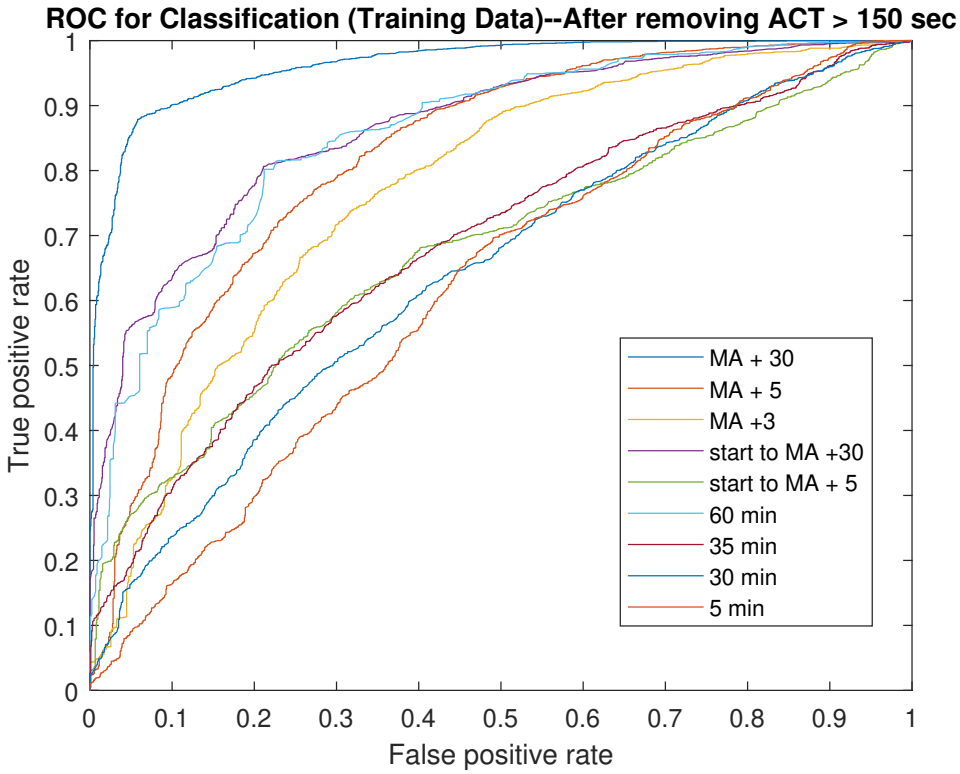

Figure 31: ROC curve for training data at various time intervals in TEG tracing ($\geq$ 3%).

6.2 Confusion matrices for LY30 $\geq$ 3%

Supporting confusion matrices for prediction of LY30 $\geq$ 3%:

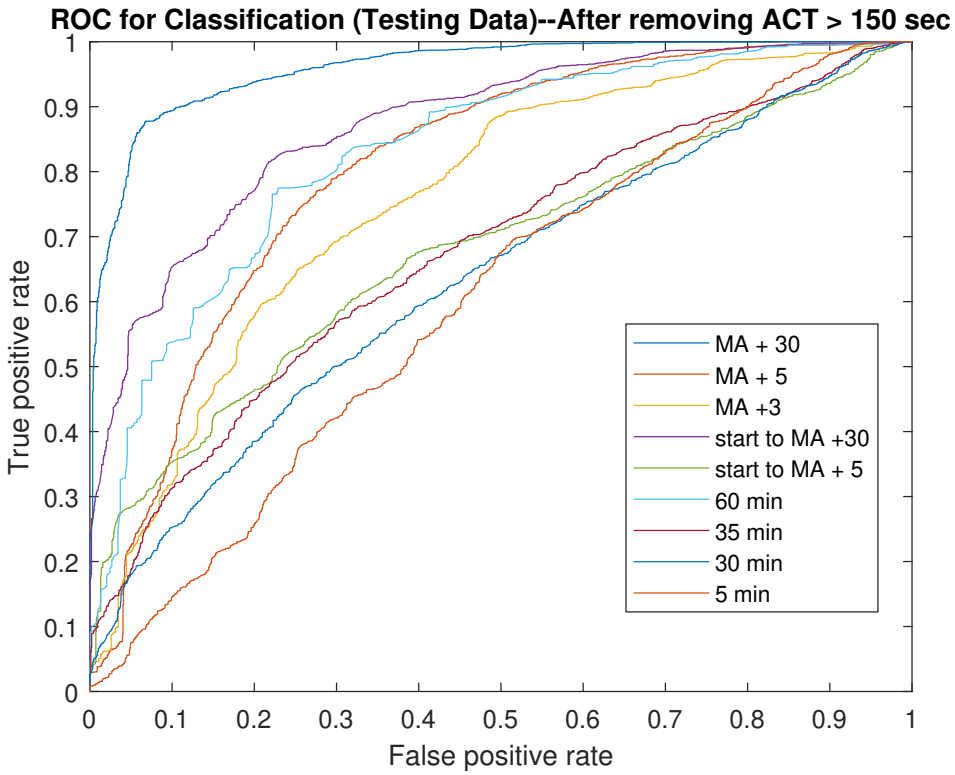

Figure 32: ROC curve for testing data at various time intervals in TEG tracing ($\geq 3\%$).

6.3 ROC curves for LY30 $\geq$ 8%

Supporting ROC curves for the prediction of LY30 $\geq 8\%$:

6.4 Confusion matrices for LY30 $\geq$ 8%

Supporting confusion matrices for the prediction of LY30 $\geq 8\%$

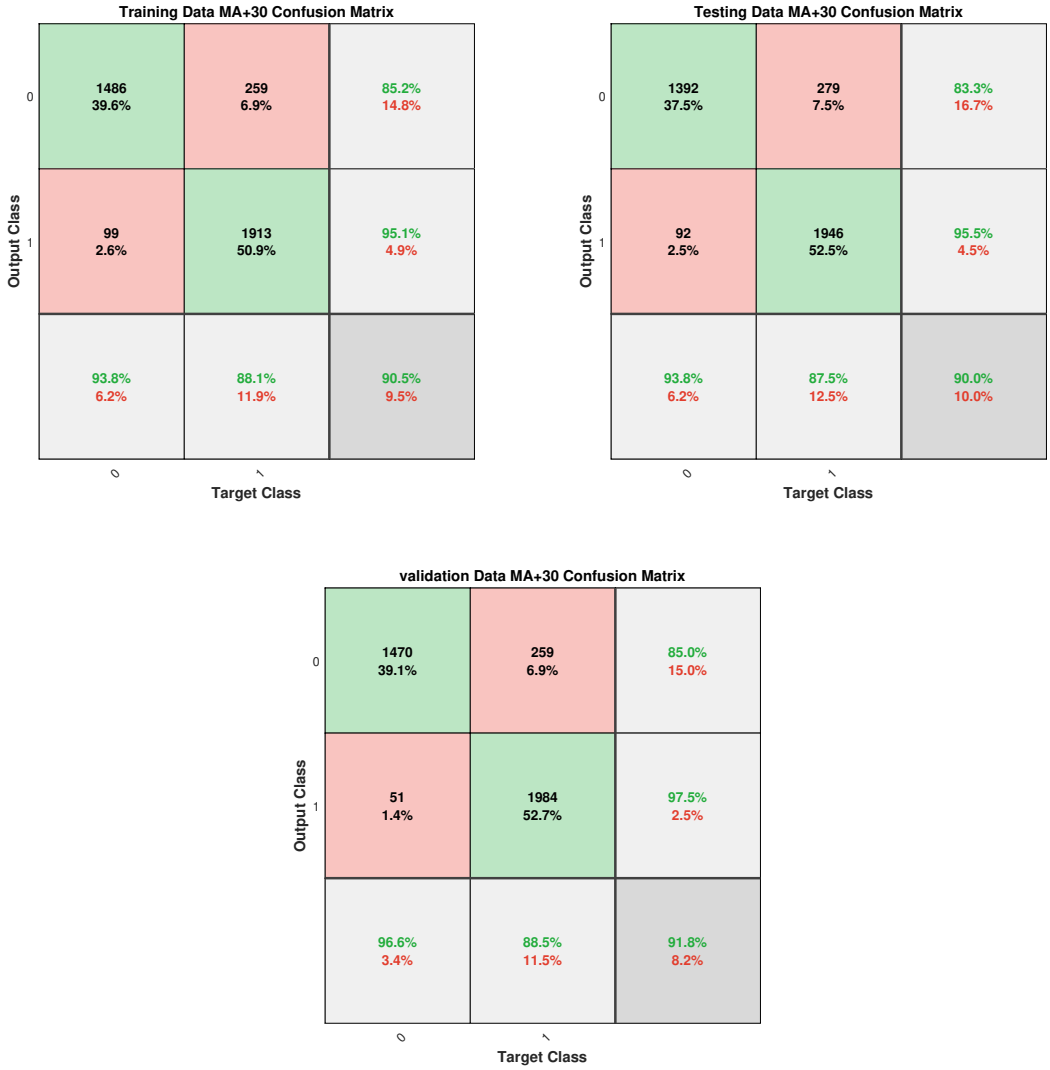

Figure 33: Confusion matrix for univariate logistic regression ($\geq 3\%$) of MA+30 data.

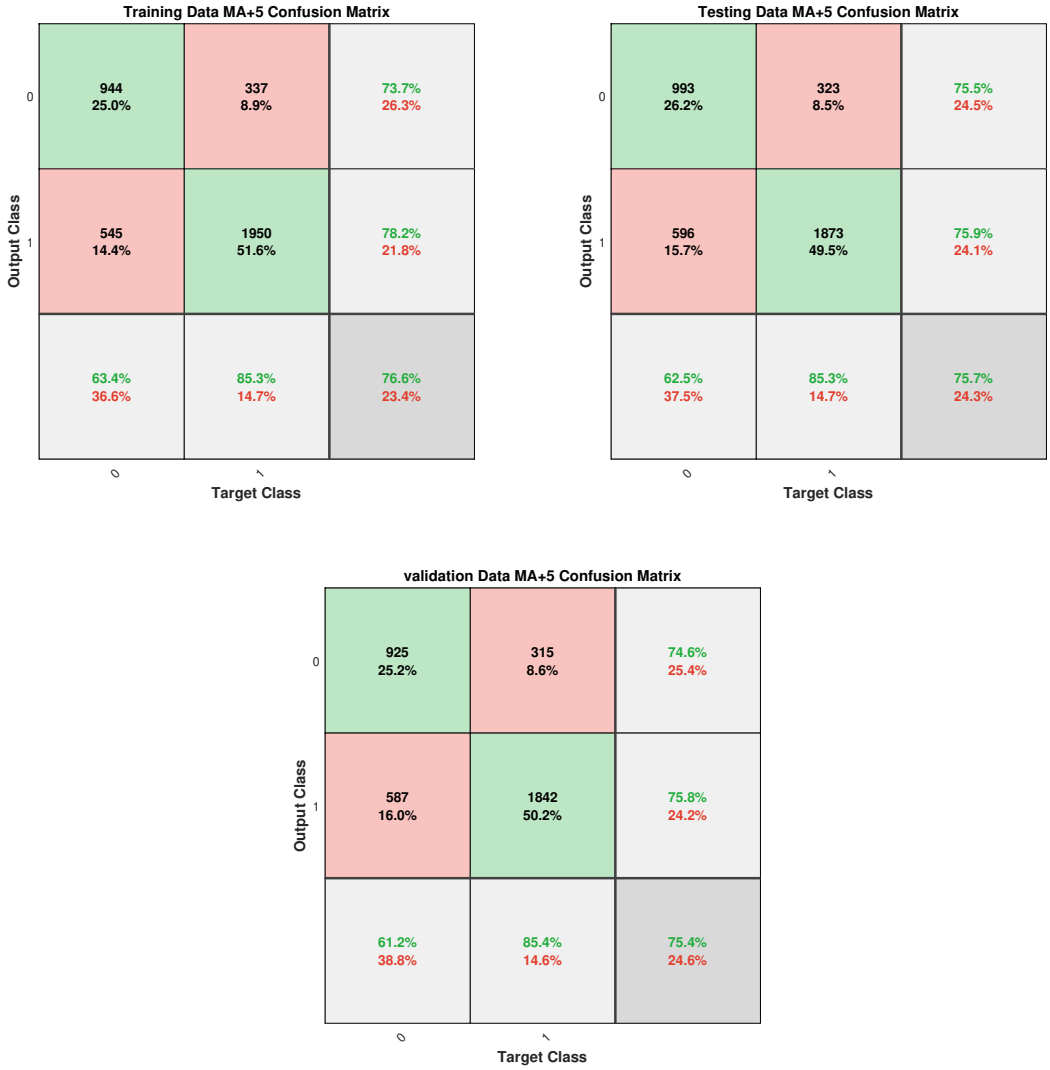

Figure 34: Confusion matrix for univariate logistic regression ($\geq 3\%$) of MA+5 data.

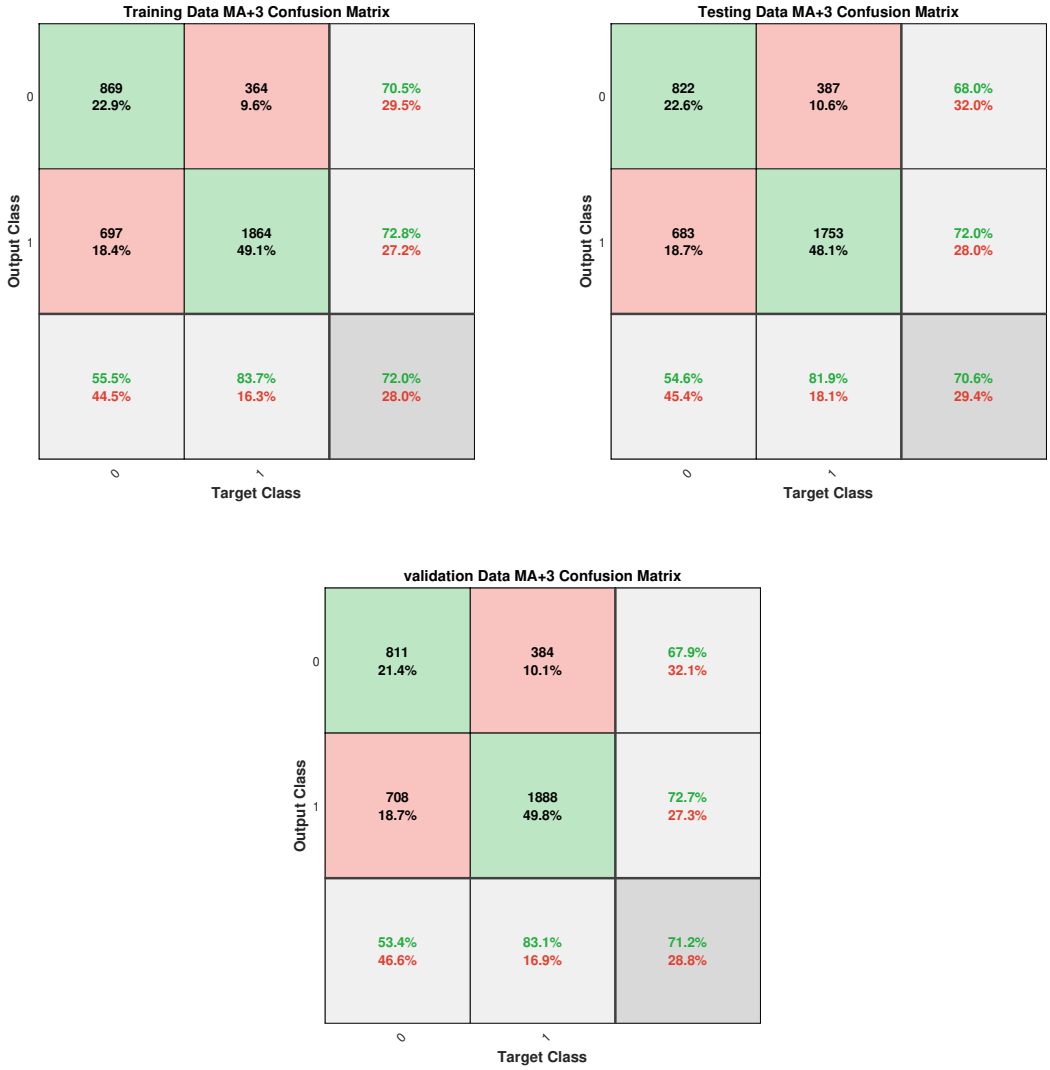

Figure 35: Confusion matrix for univariate logistic regression ($\geq 3\%$) of MA+3 data.

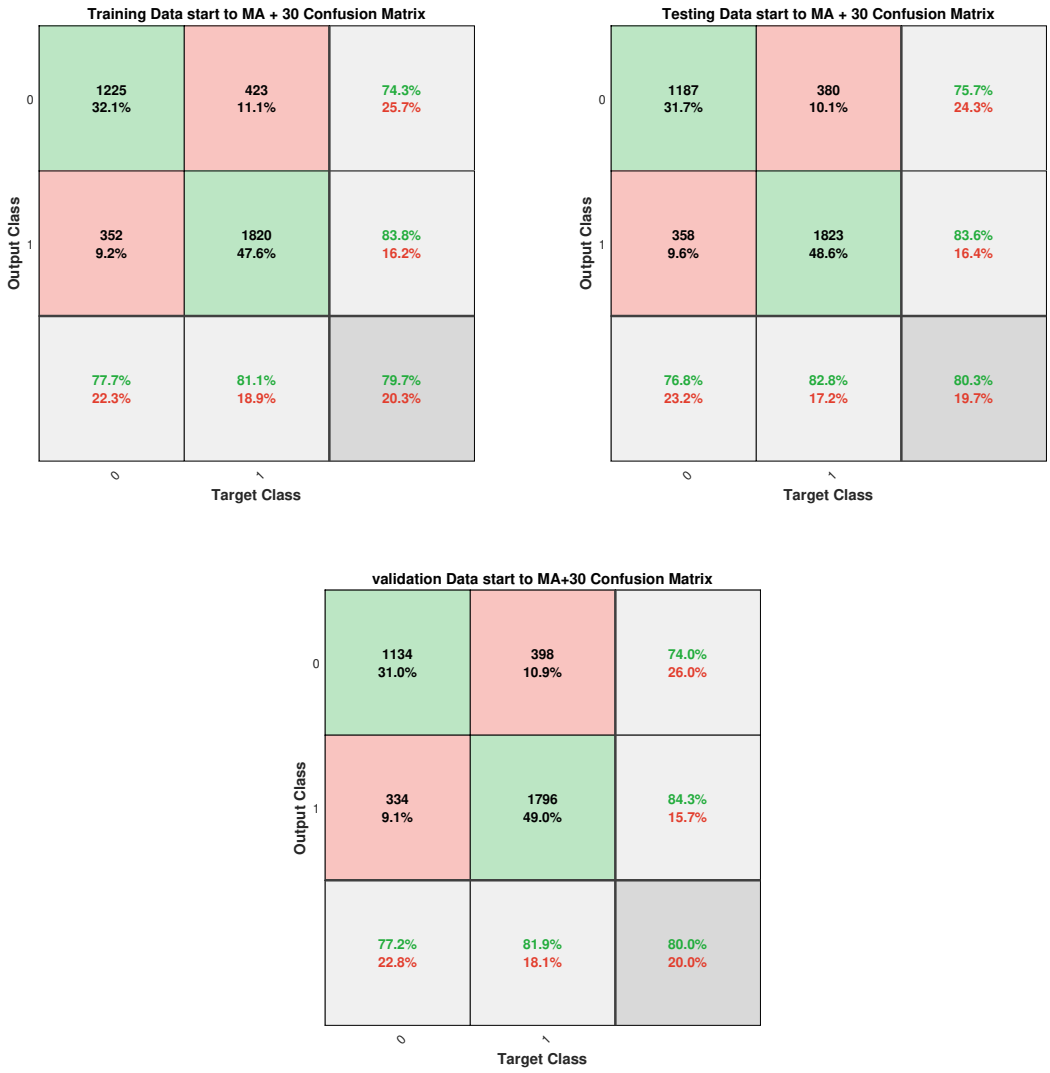

Figure 36: Confusion matrix for univariate logistic regression ($\geq 3\%$) of start to MA+30 data.

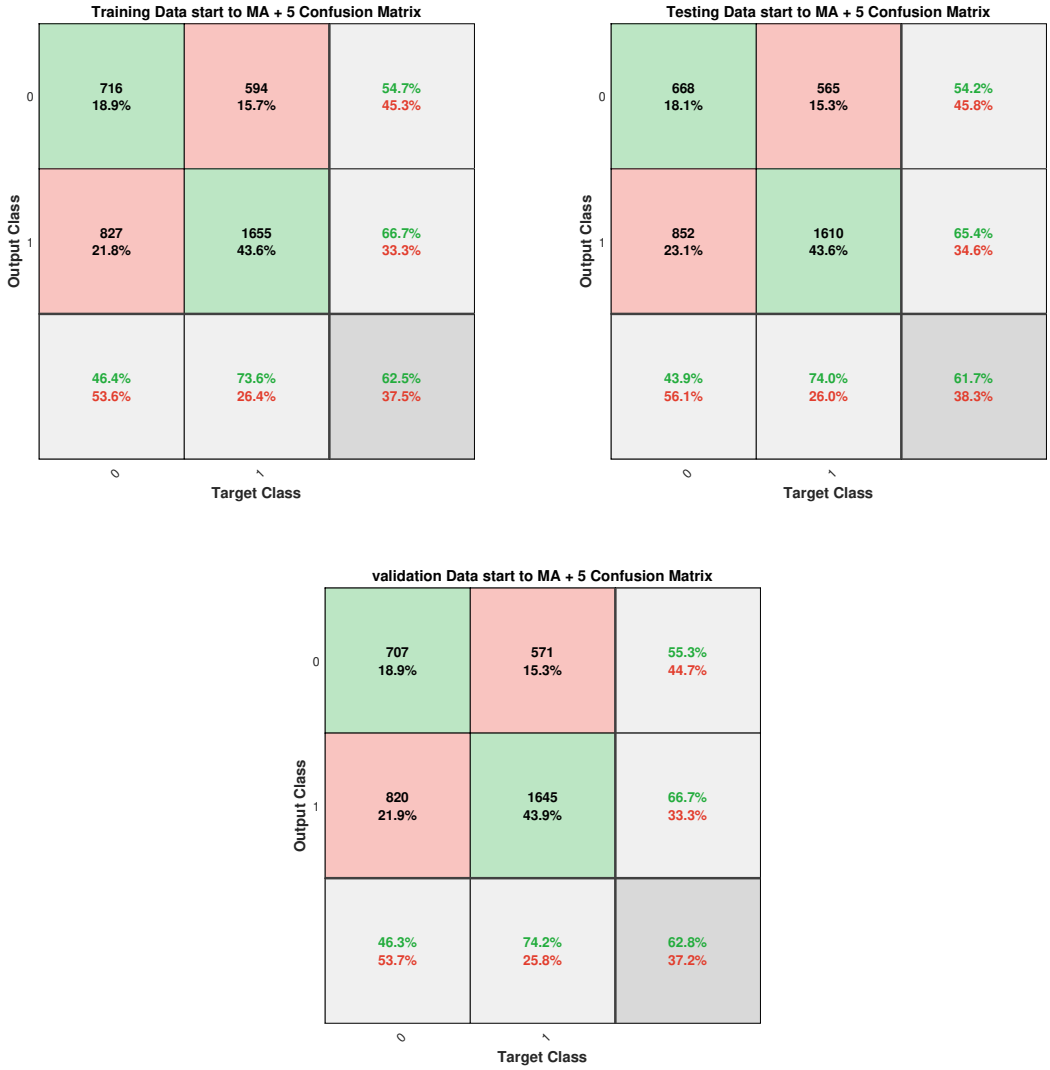

Figure 37: Confusion matrix for univariate logistic regression ($\geq 3\%$) of start to MA+5 data.

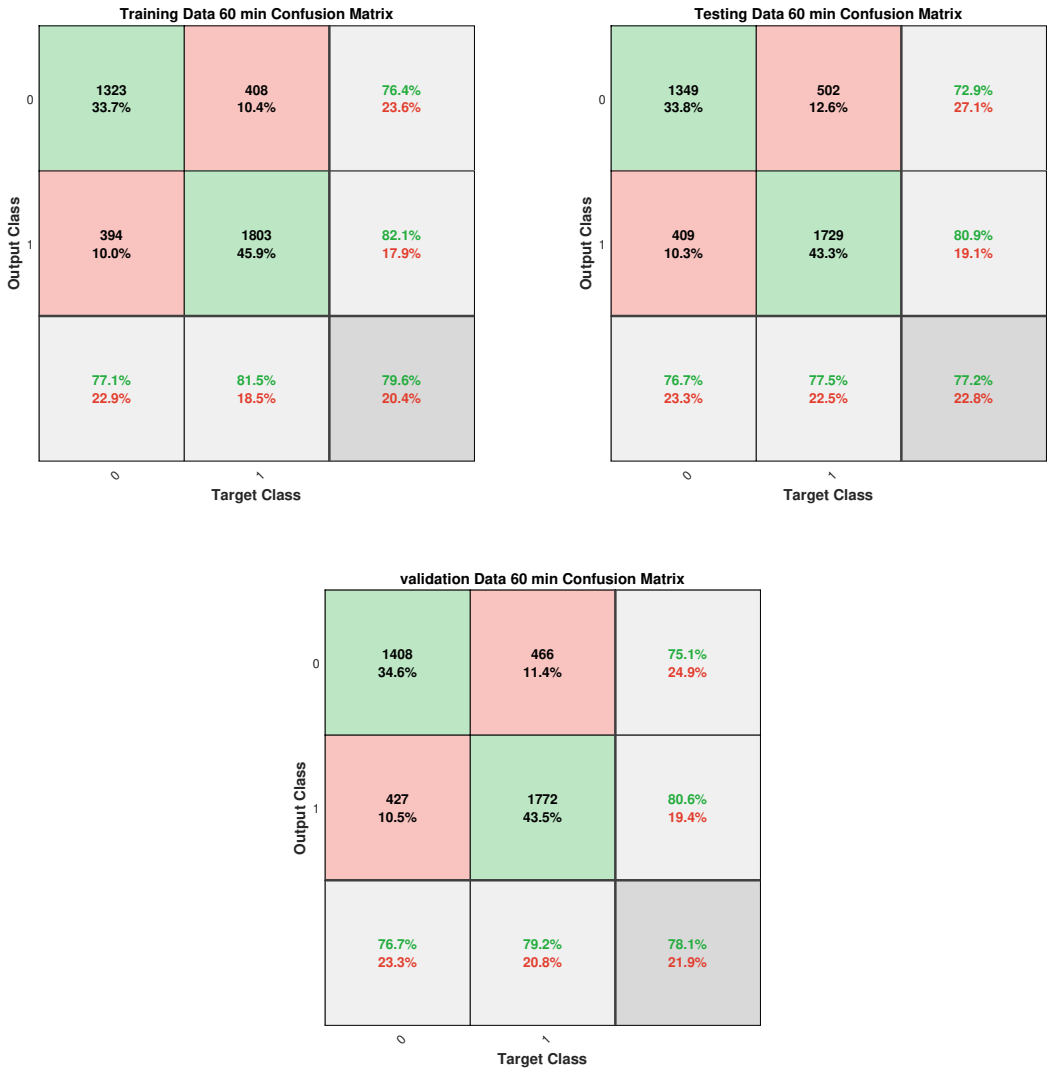

Figure 38: Confusion matrix for univariate logistic regression ($\geq 3\%$) of start to 60 minutes data.

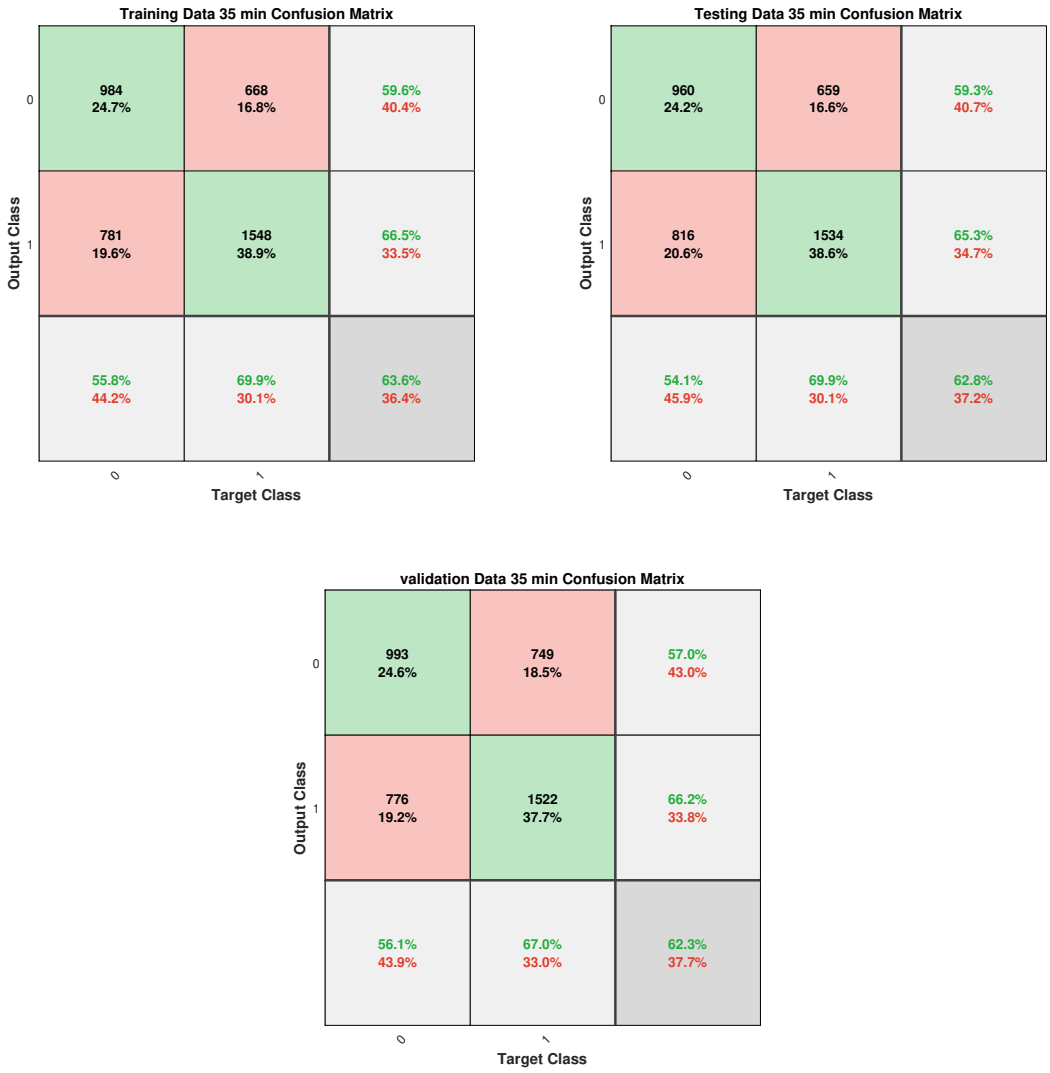

Figure 39: Confusion matrix for univariate logistic regression ($\geq 3\%$) of start to 35 minutes data.

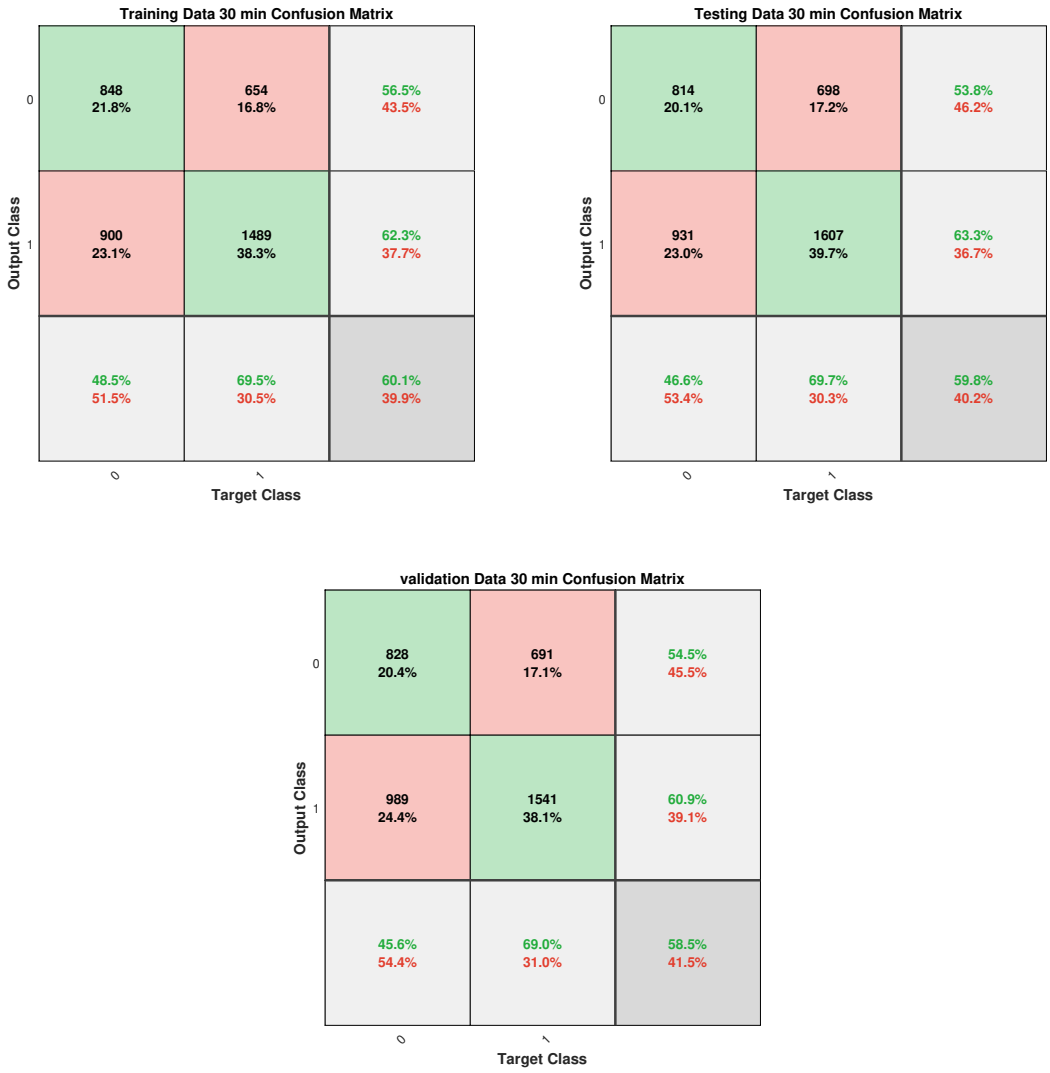

Figure 40: Confusion matrix for univariate logistic regression ($\geq 3\%$) of start to 30 minutes data.

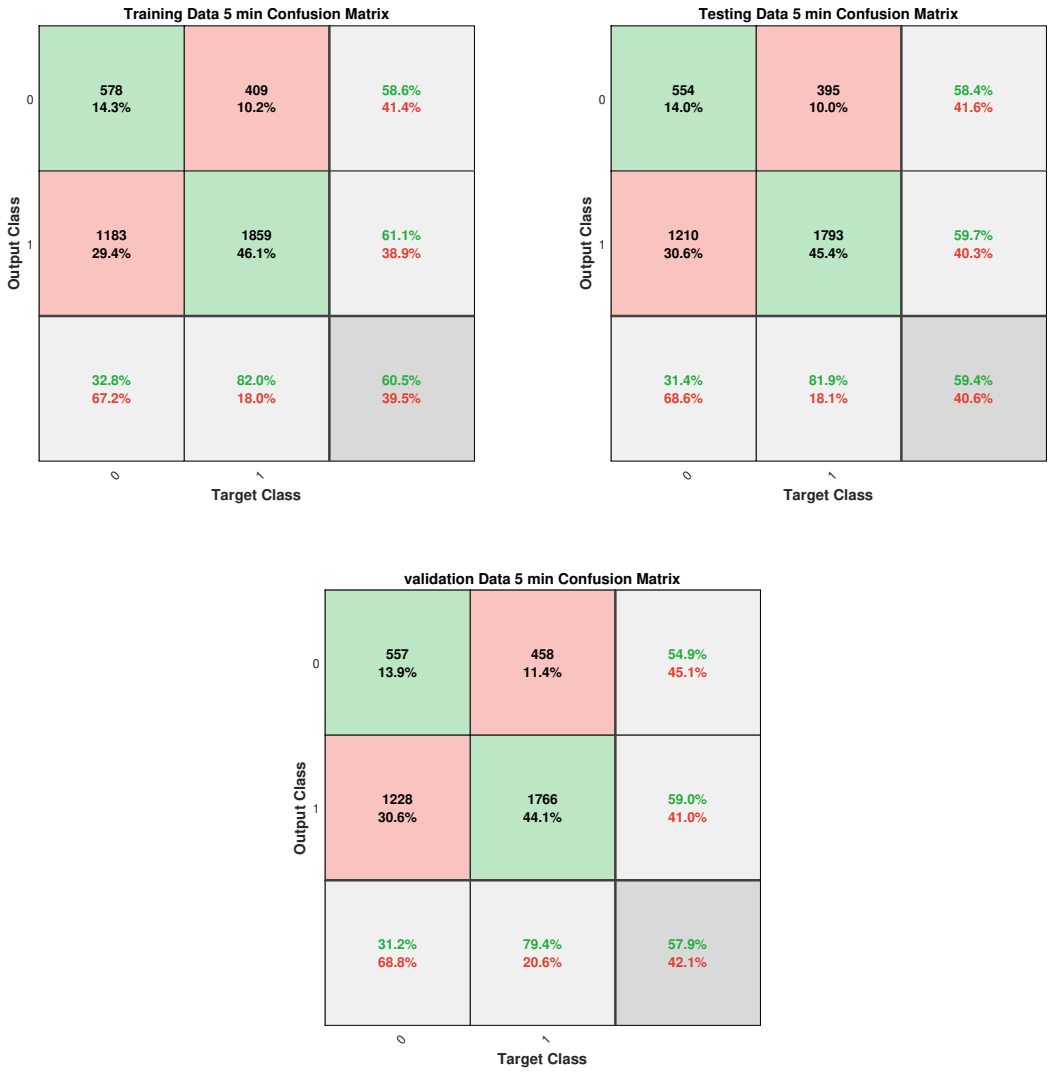

Figure 41: Confusion matrix for univariate logistic regression ($\geq 3\%$) of start to 5 minutes data.

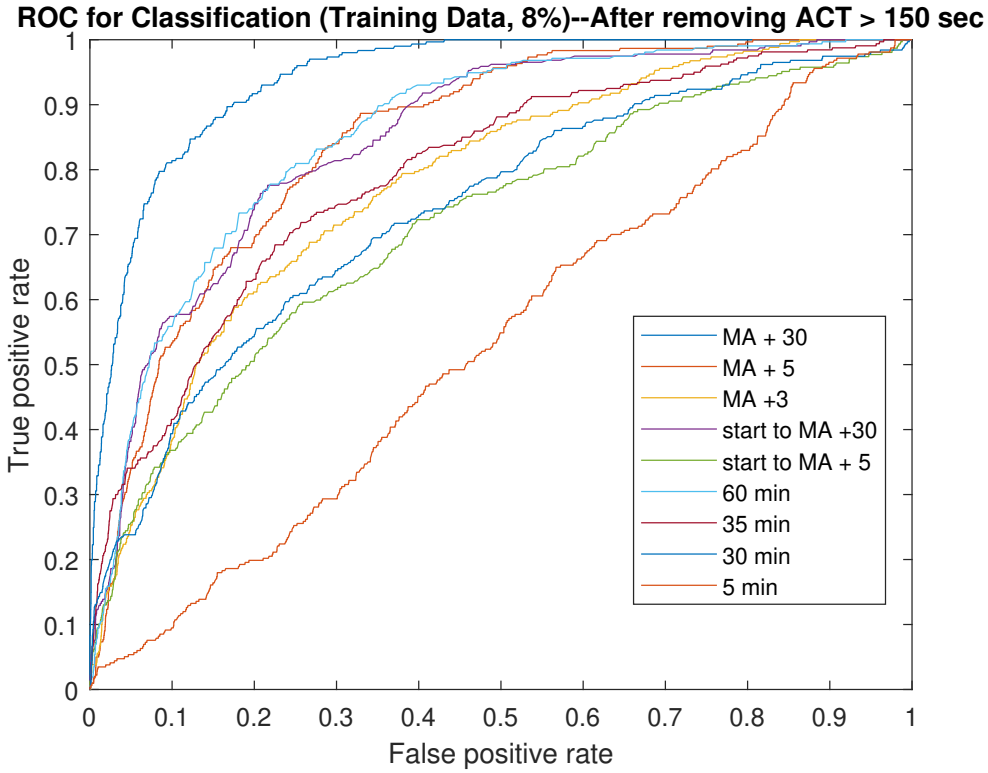

Figure 42: ROC curve for training data at various time intervals in TEG tracing ($\geq 8\%$).

6.5 ROC curves for $3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$

Supporting ROC curves for the prediction of $3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$:

6.6 Confusion matrices for $3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$

Supporting confusion matrices for the prediction of $3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$:

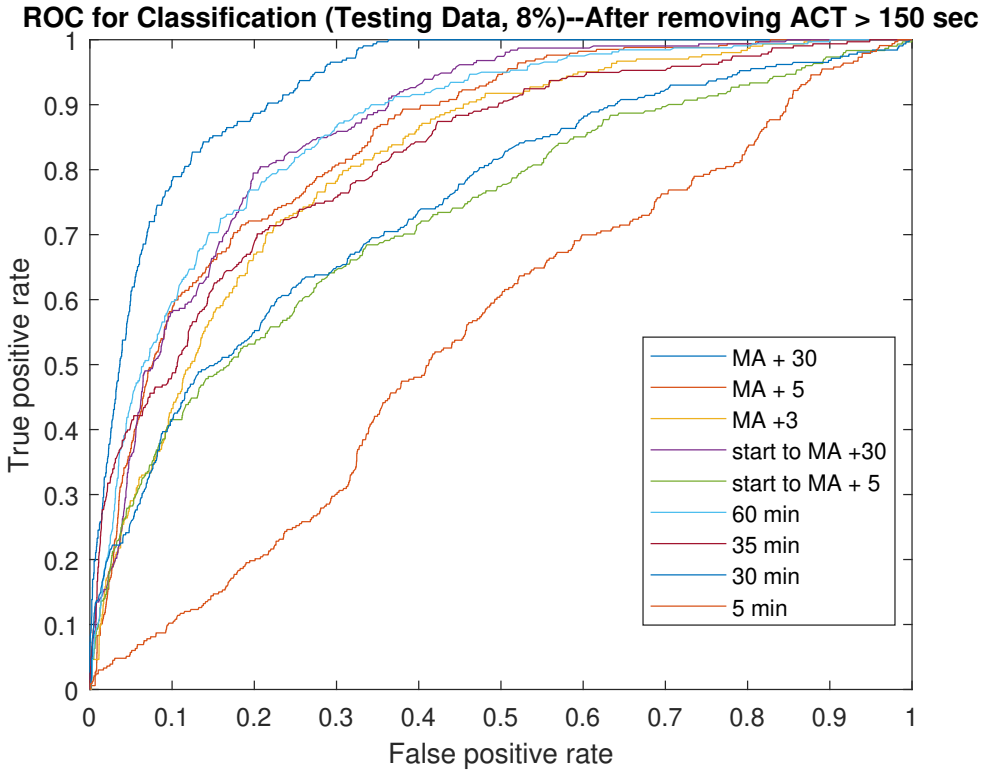

Figure 43: ROC curve for testing data at various time intervals in TEG tracing ($\geq 8\%$).

6.7 Confusion matrices for Multivariate Regression

Supporting confusion matrices for the multivariate LY30 prediction:

Figure 44: Confusion matrix for univariate logistic regression ($\geq 8\%$) of MA+30 data.

Figure 45: Confusion matrix for univariate logistic regression ($\geq 8\%$) of MA+5 data.

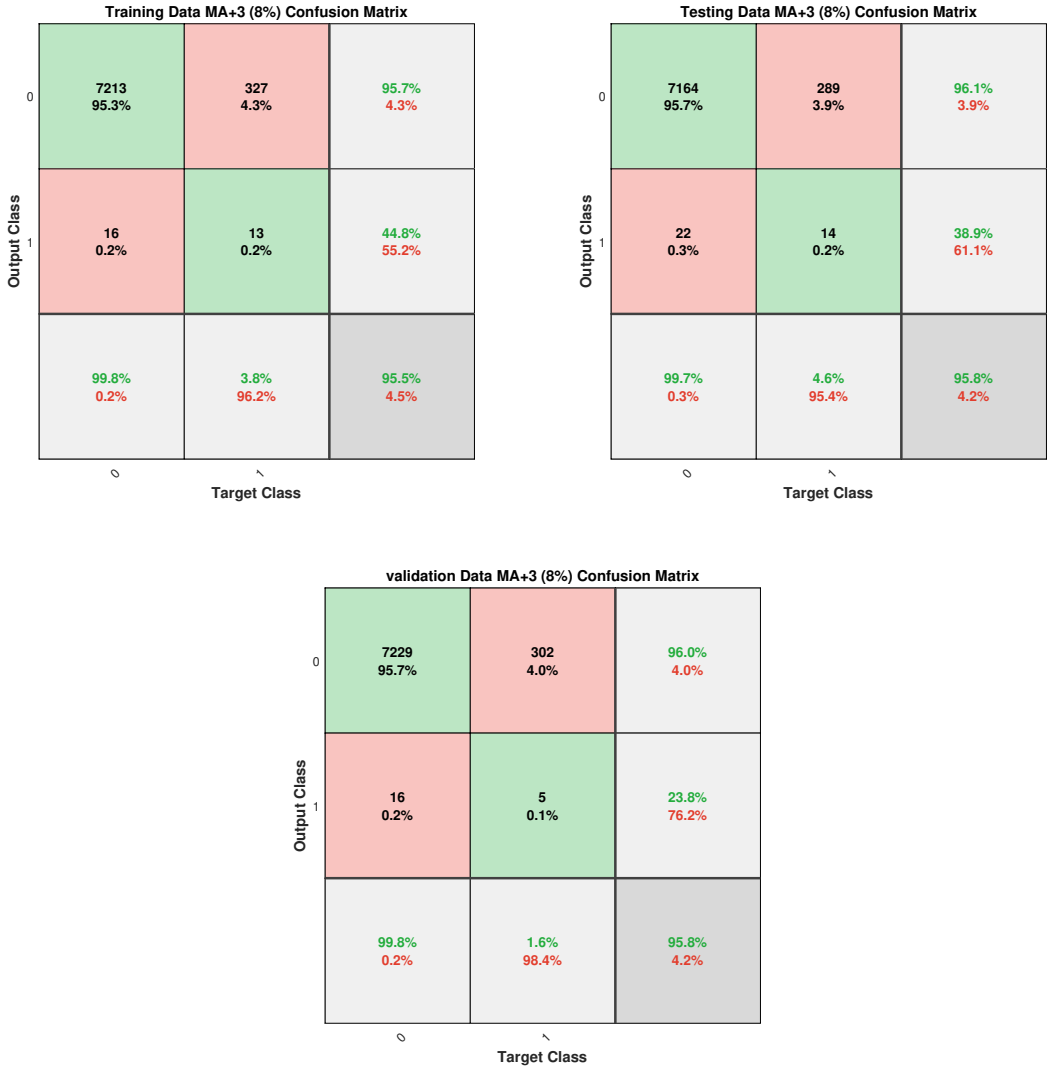

Figure 46: Confusion matrix for univariate logistic regression ($\geq 8\%$) of MA+3 data.

Figure 47: Confusion matrix for univariate logistic regression ($\geq$ 8%) of start to MA+30 data.

Figure 48: Confusion matrix for univariate logistic regression ($\geq 8\%$) of start to MA+5 data.

Figure 49: Confusion matrix for univariate logistic regression ($\geq 8\%$) of start to 60 minutes data.

Figure 50: Confusion matrix for univariate logistic regression ($\geq 8\%$) of start to 35 minutes data.

Figure 51: Confusion matrix for univariate logistic regression ($\geq 8\%$) of start to 30 minutes data.

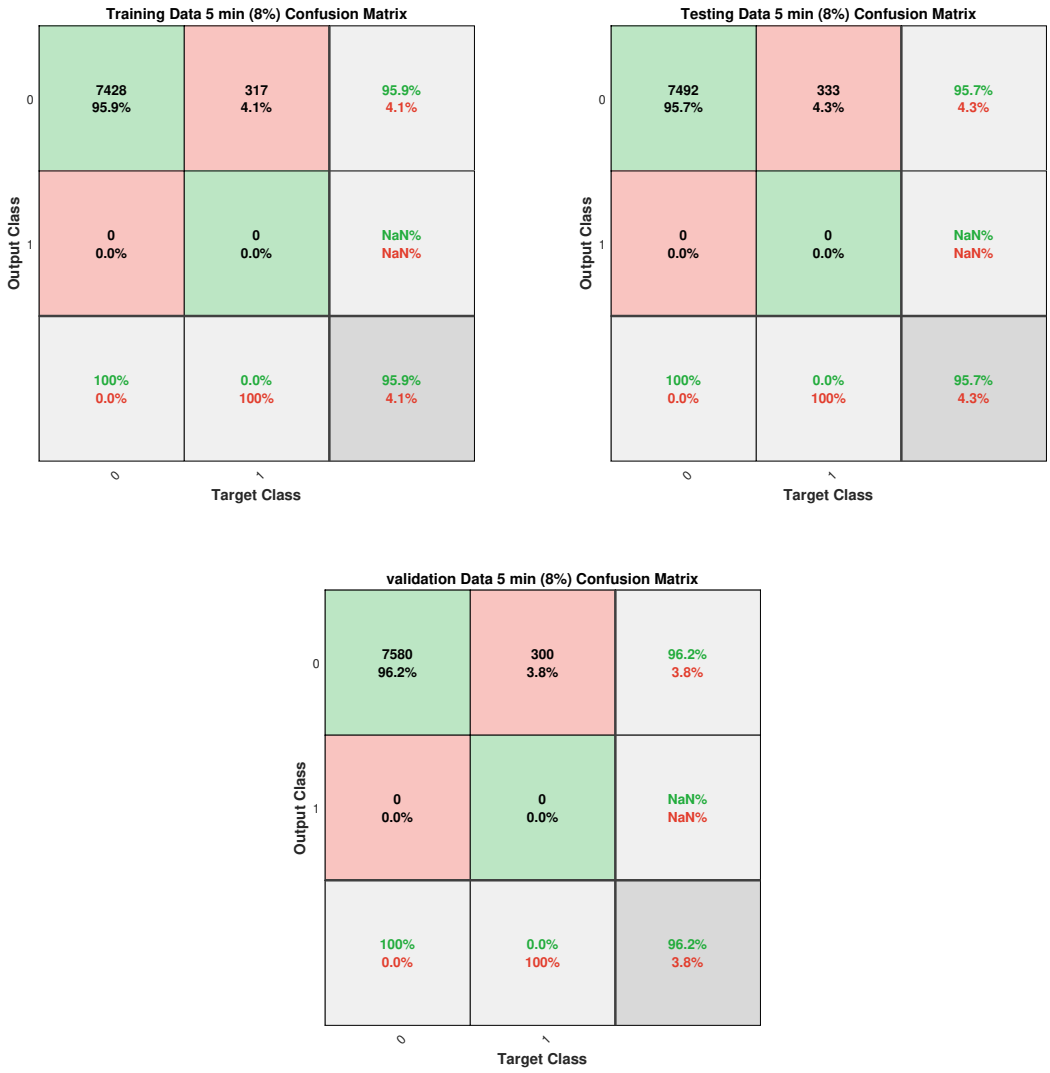

Figure 52: Confusion matrix for univariate logistic regression ($\geq 8\%$) of start to 5 minutes data.

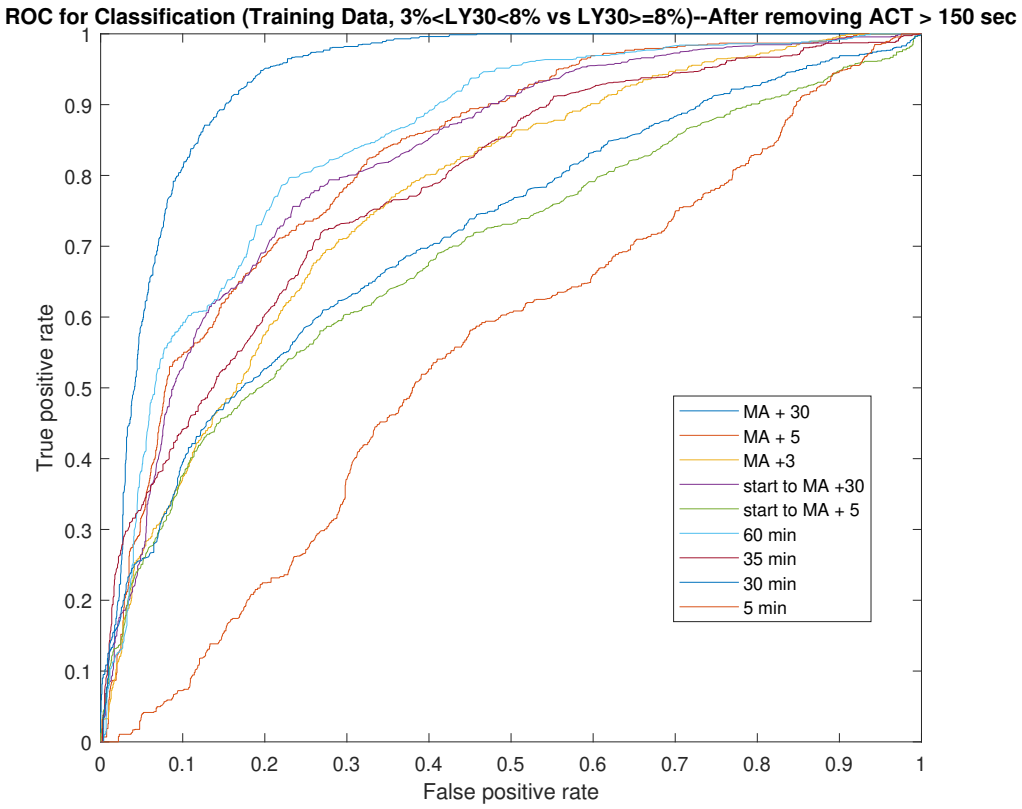

Figure 53: ROC curve for training data at various time intervals in TEG tracing ($3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$).

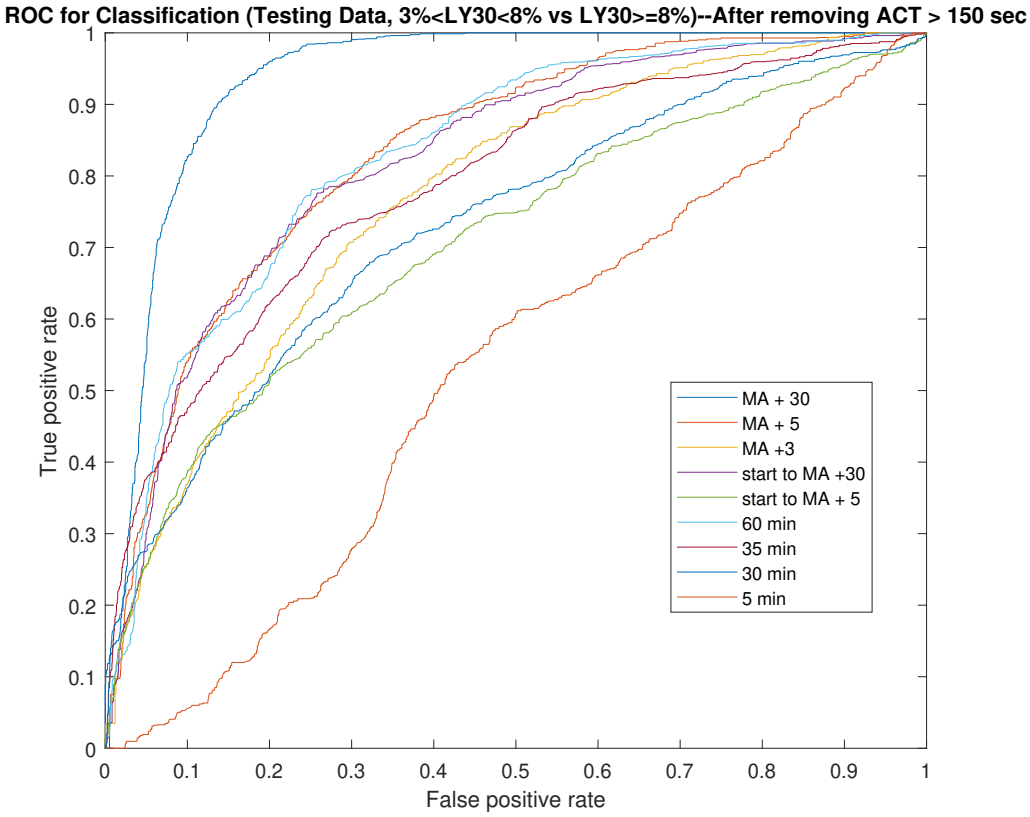

Figure 54: ROC curve for testing data at various time intervals in TEG tracing ($3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$).

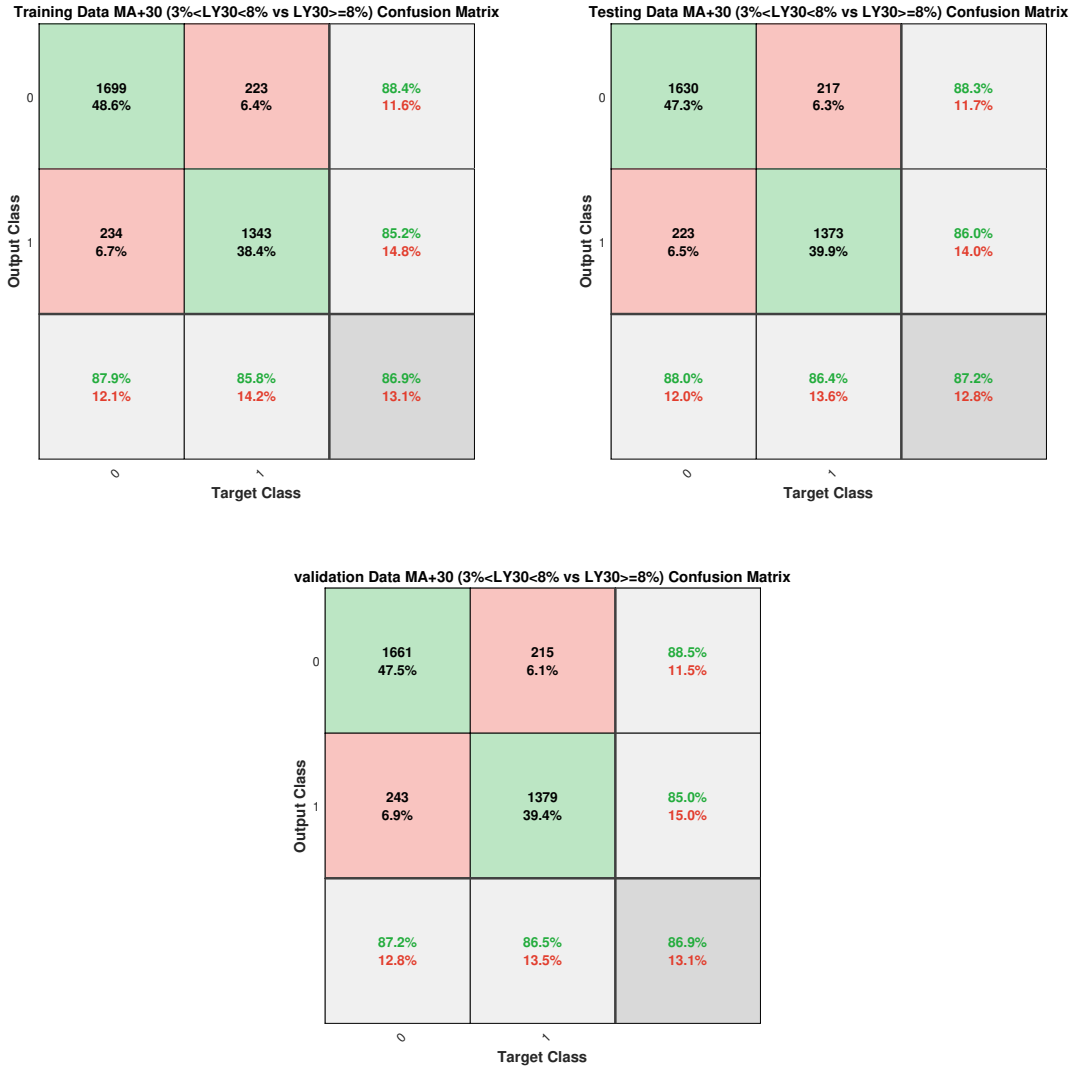

Figure 55: Confusion matrix for univariate logistic regression ($3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$) of MA+30 data.

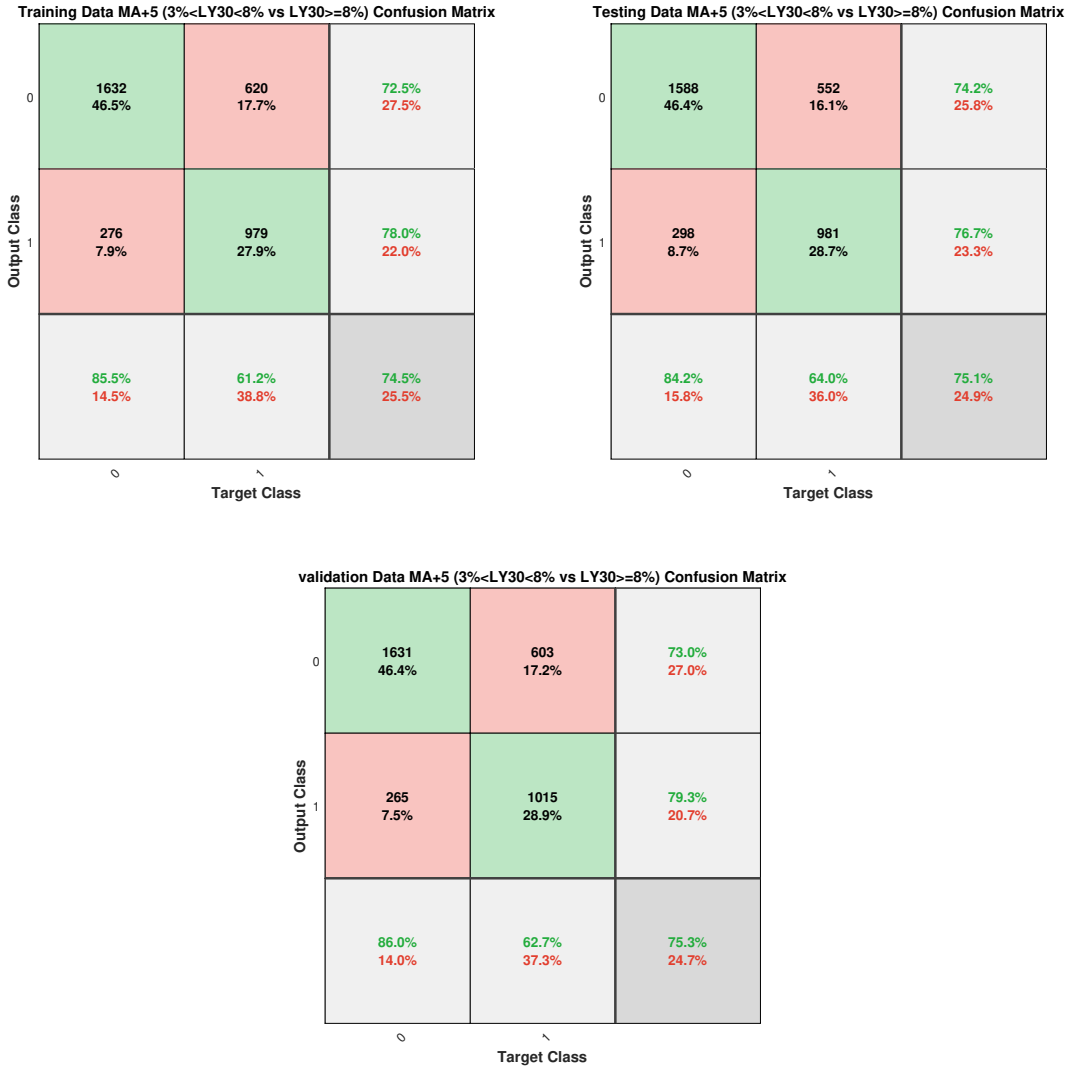

Figure 56: Confusion matrix for univariate logistic regression ($3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$) of MA+5 data.

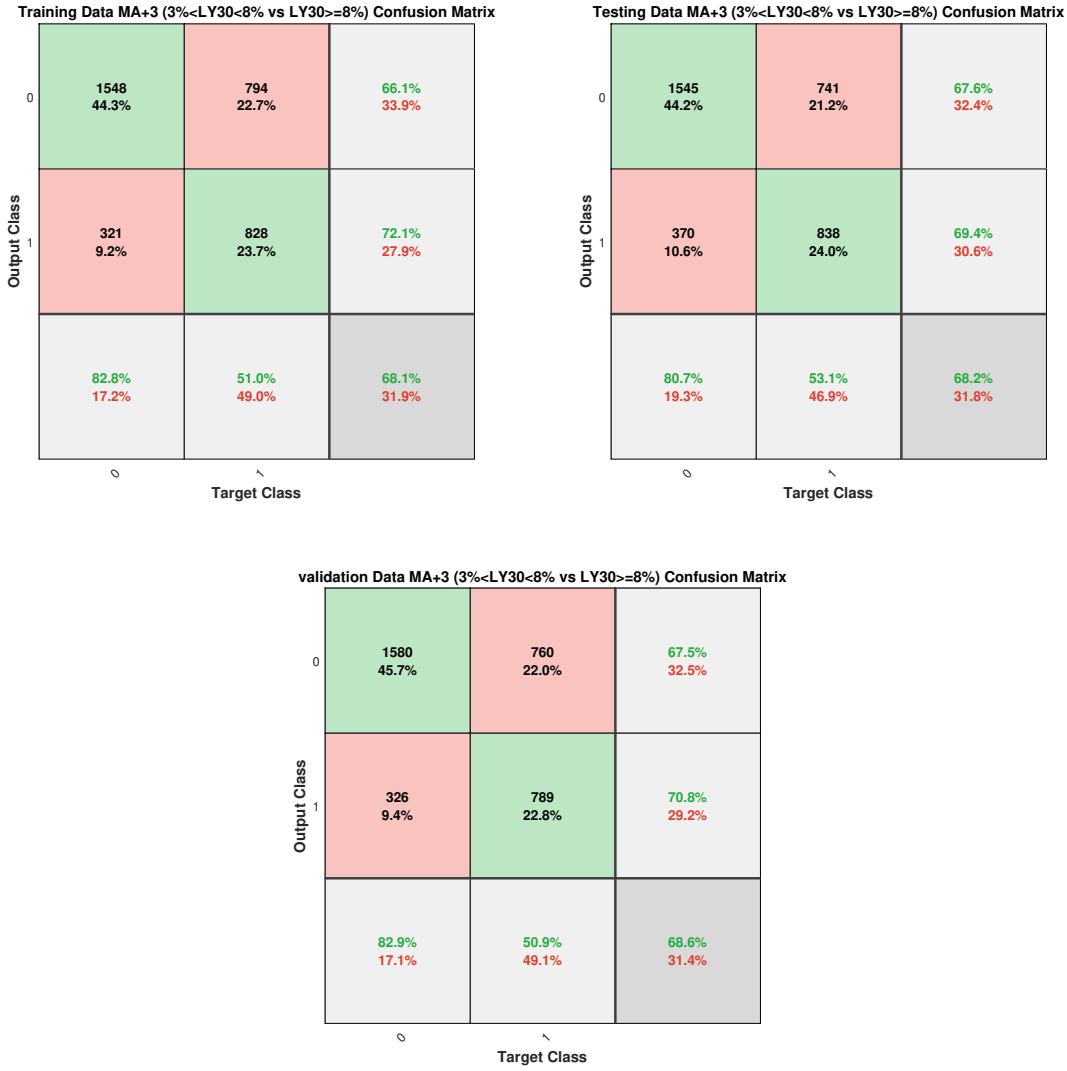

Figure 57: Confusion matrix for univariate logistic regression ($3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$) of MA+3 data.

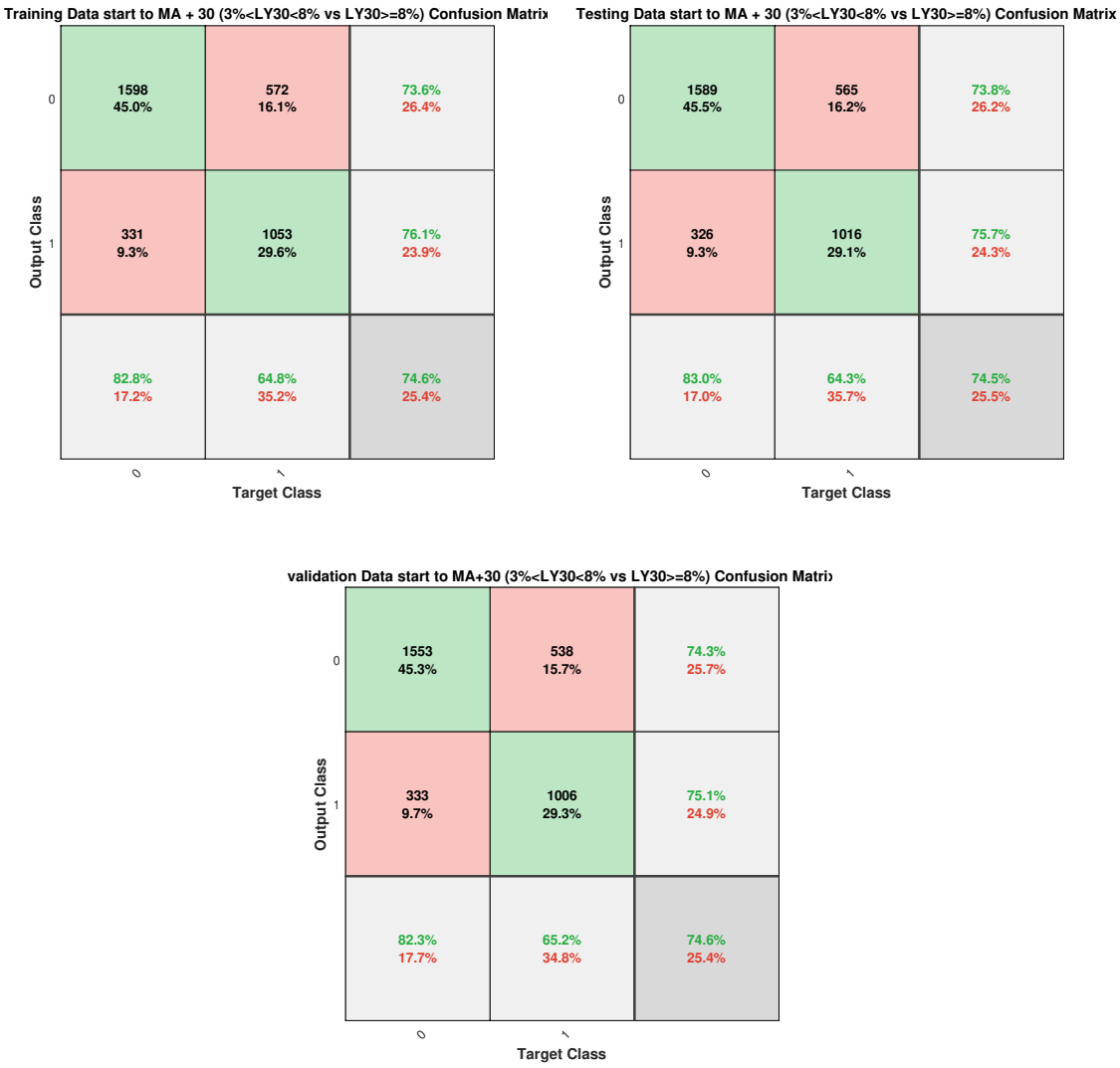

Figure 58: Confusion matrix for univariate logistic regression ($3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$) of start to MA+30 data.

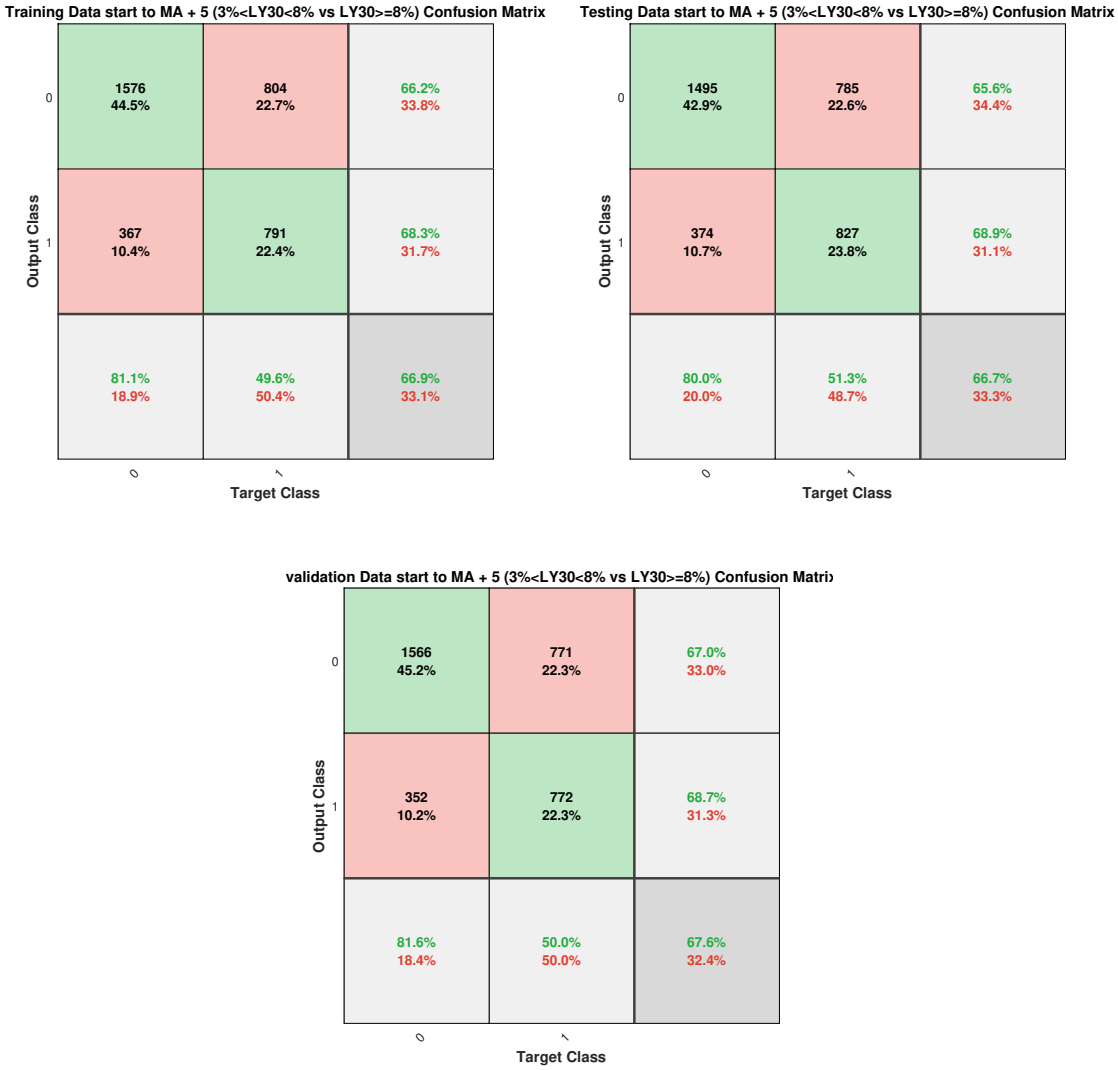

Figure 59: Confusion matrix for data of univariate logistic regression ($3\% \leq \text{LY30} < 8\%$ vs. $\text{LY30} \geq 8\%$) of start to MA+5 data.

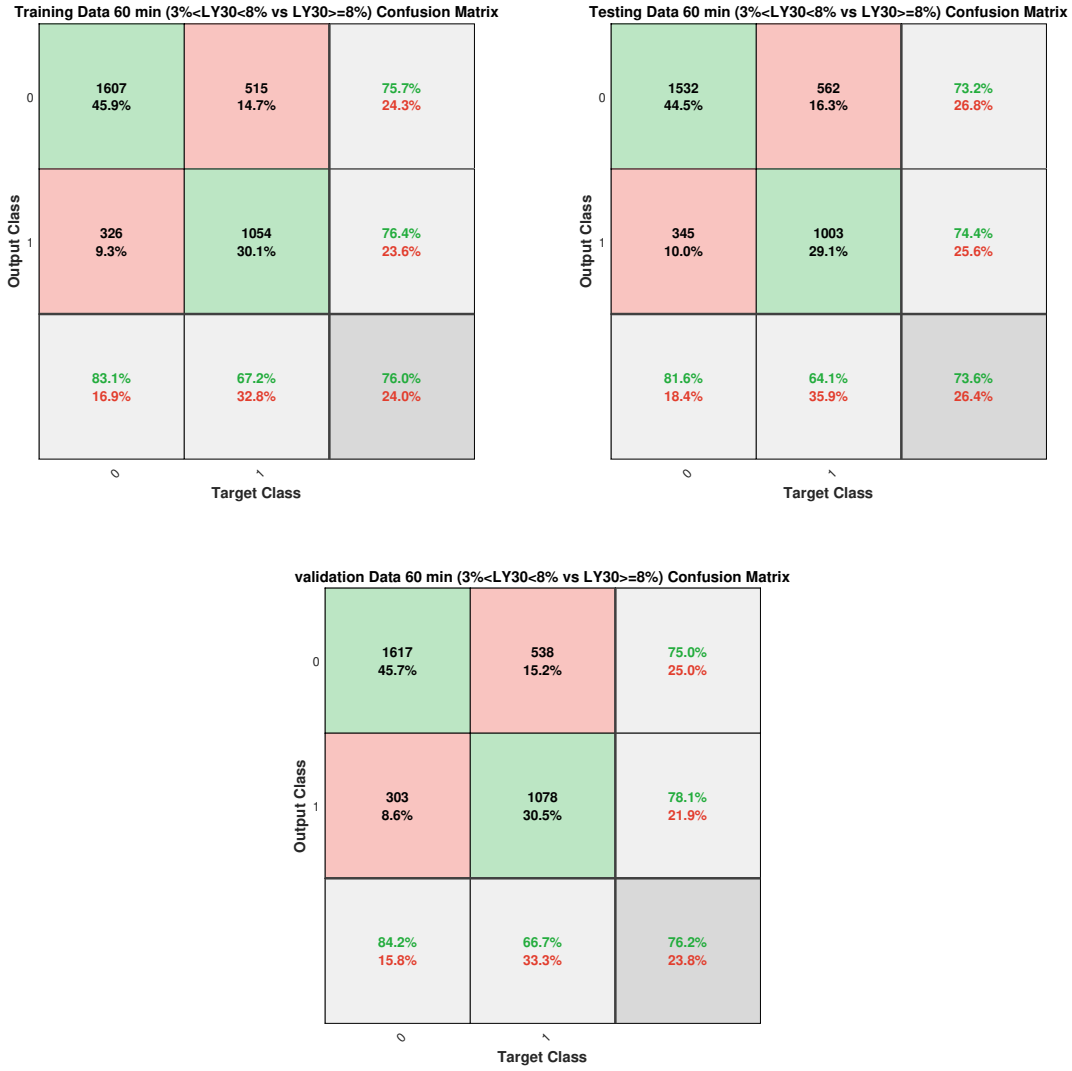

Figure 60: Confusion matrix for univariate logistic regression ($3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$) of start to 60 minutes data.

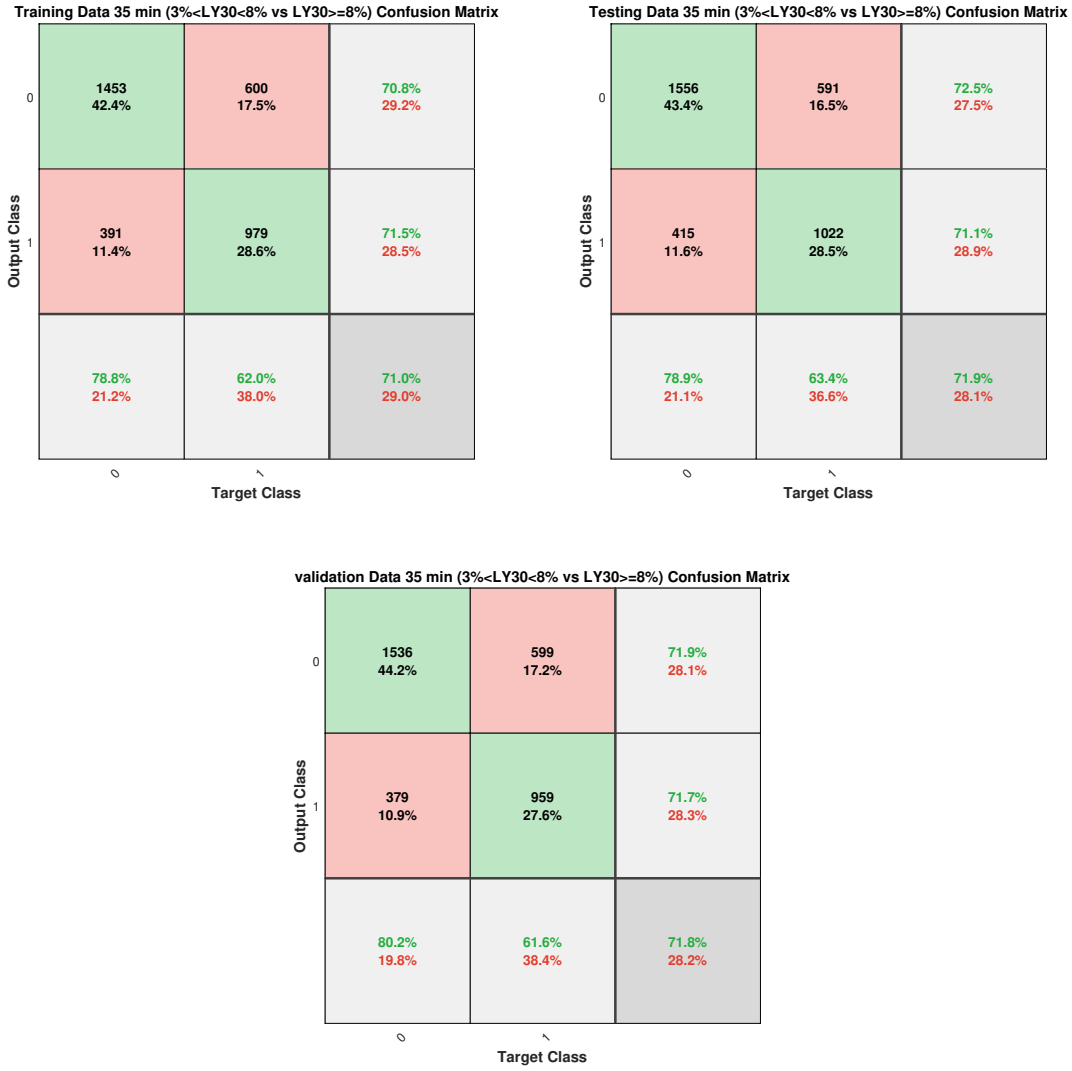

Figure 61: Confusion matrix for univariate logistic regression ($3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$) of start to 35 minutes data.

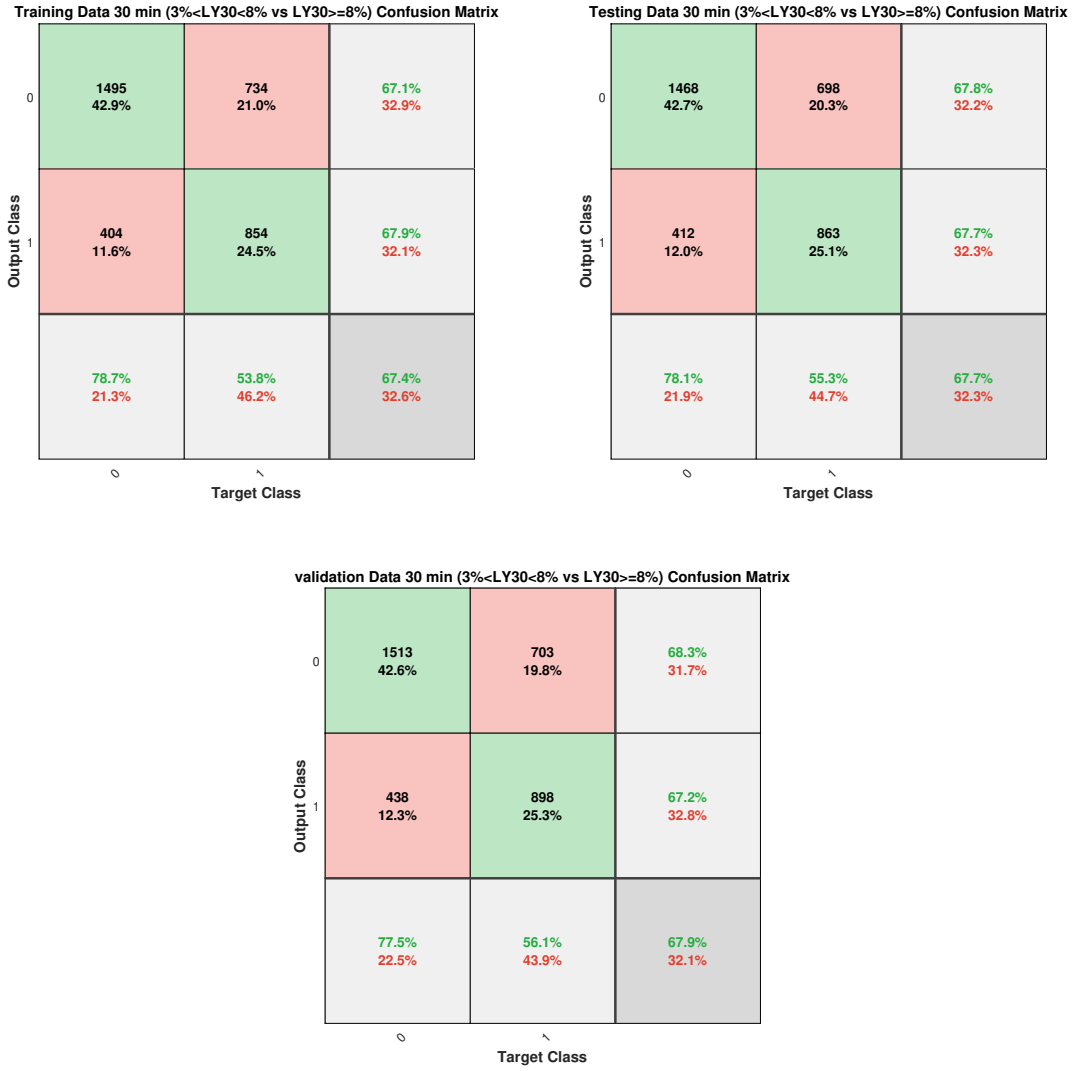

Figure 62: Confusion matrix for univariate logistic regression ($3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$) of start to 30 minutes data.

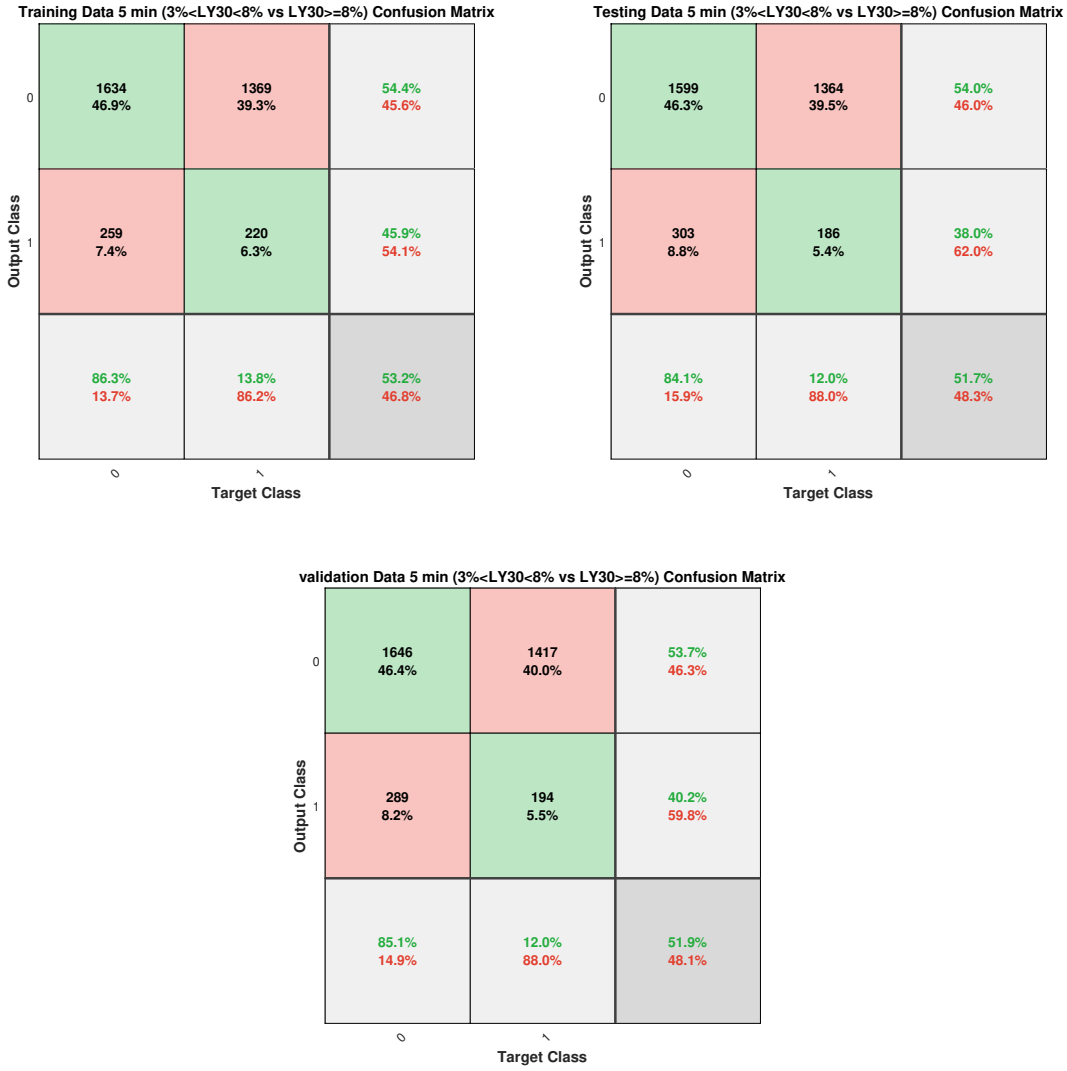

Figure 63: Confusion matrix for univariate logistic regression ($3\% \leq$ LY30 $< 8\%$ vs. LY30 $\geq 8\%$) of start to 5 minutes data.

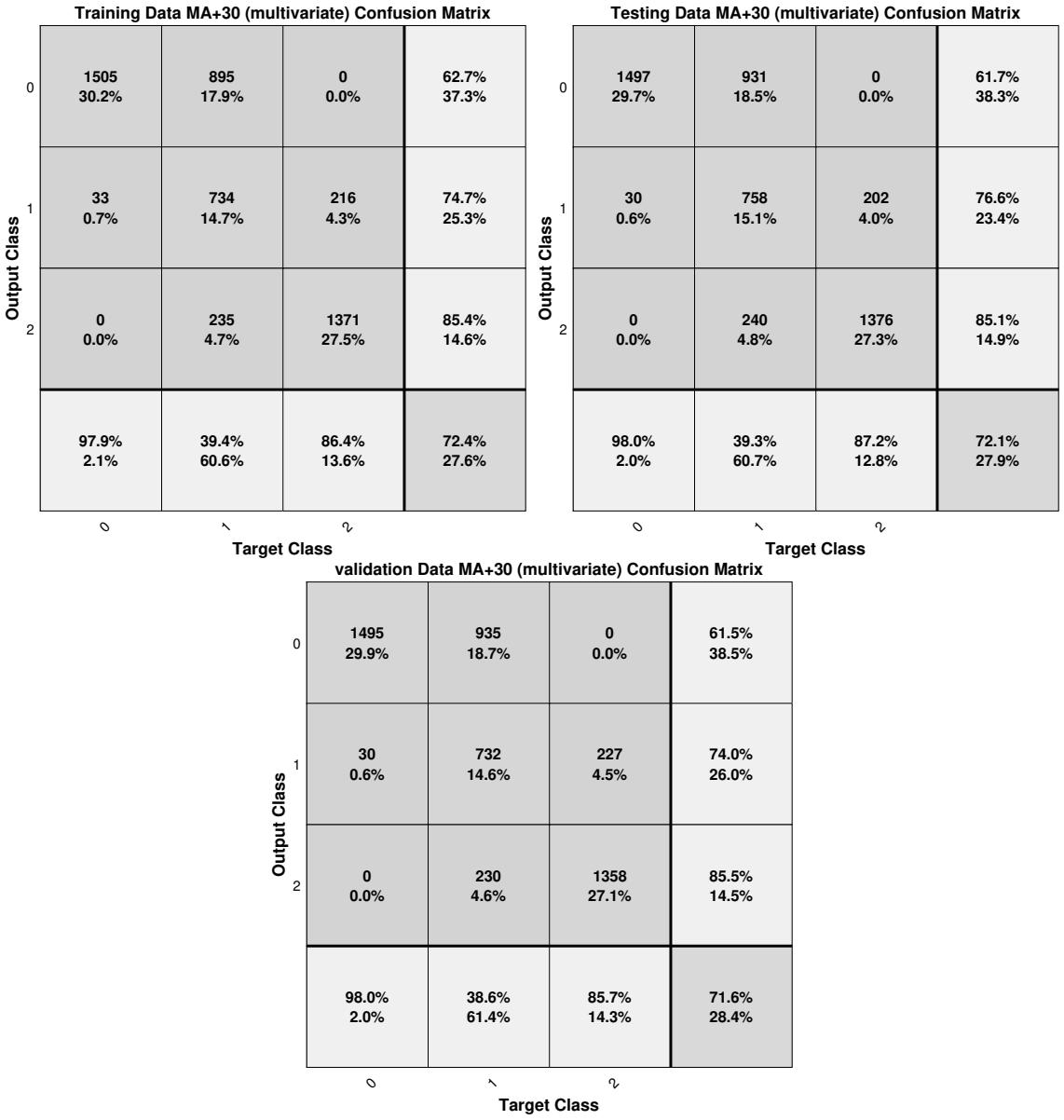

Figure 64: Confusion matrix for multivariate regression of MA+30 data

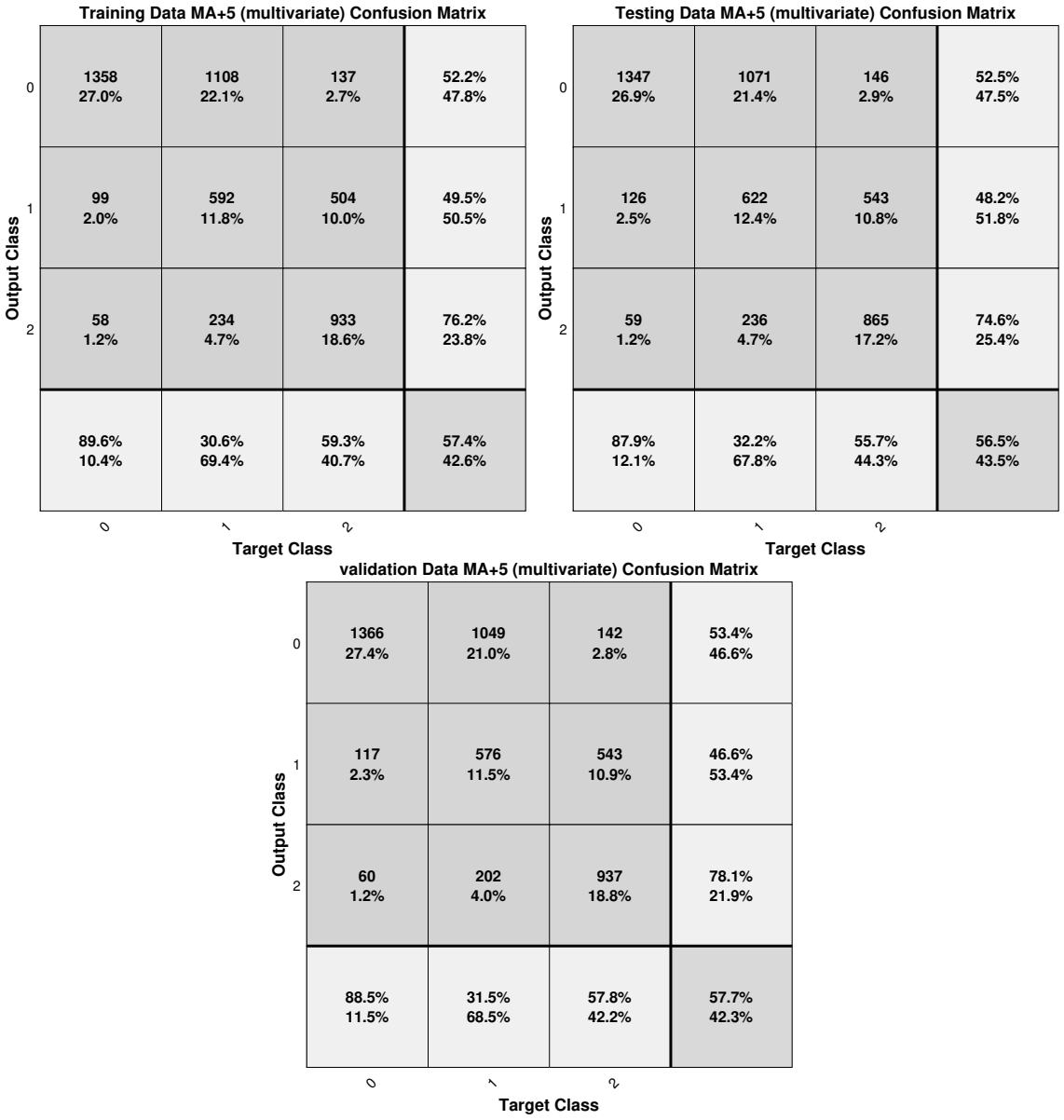

Figure 65: Confusion matrix for multivariate regression of MA+5 data

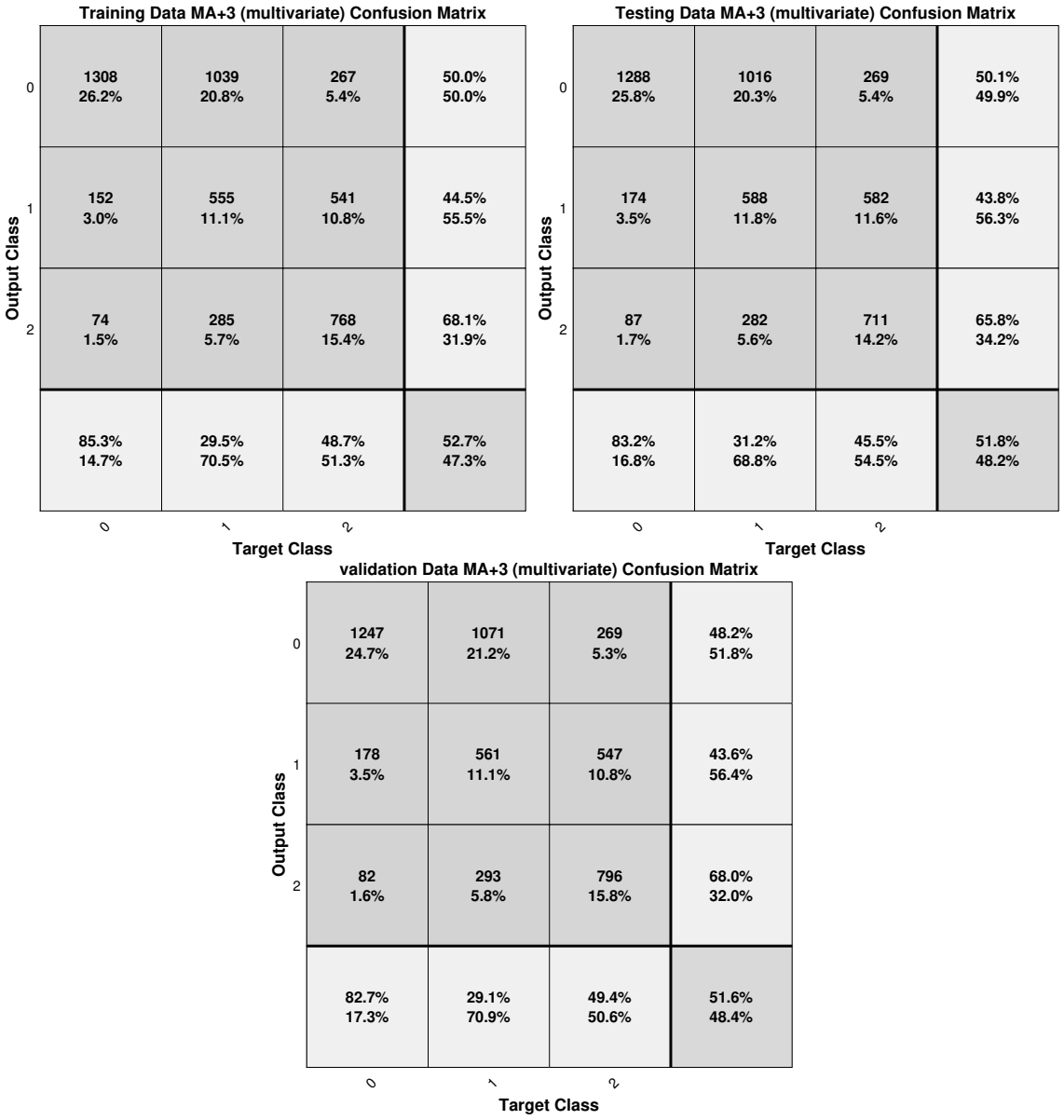

Figure 66: Confusion matrix for multivariate regression of MA+3 data

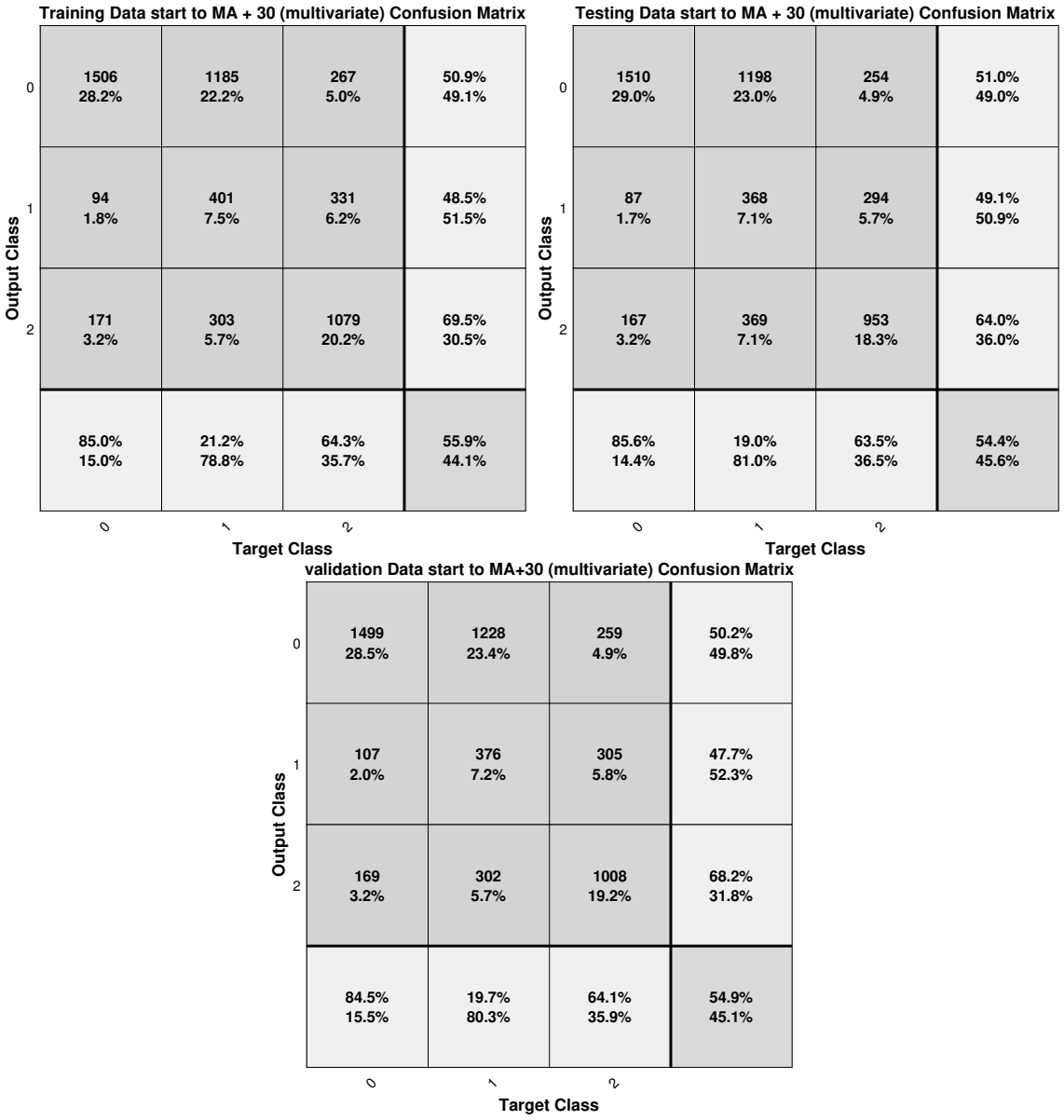

Figure 67: Confusion matrix for multivariate regression of start to MA+30 data

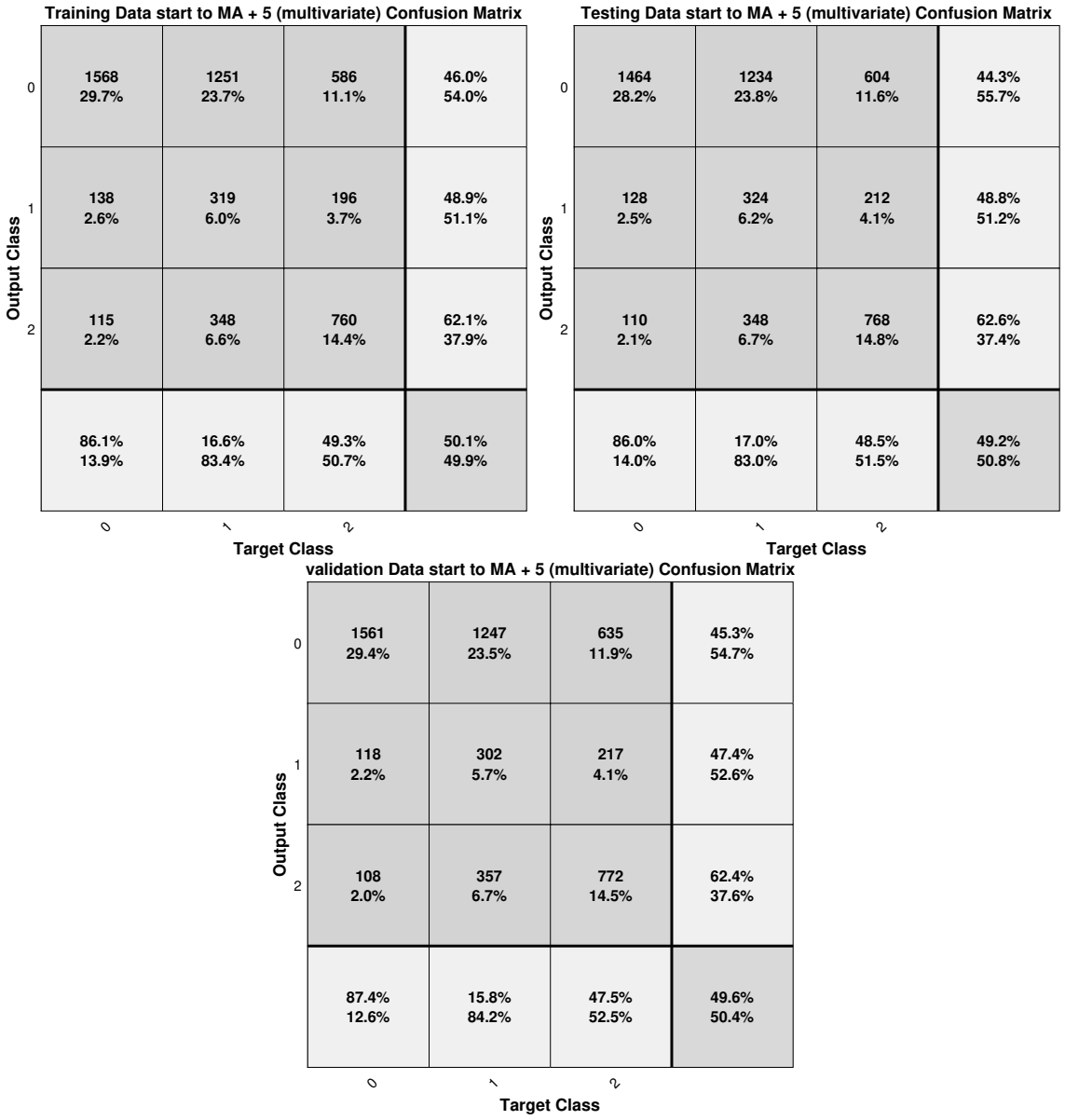

Figure 68: Confusion matrix for multivariate regression of start to MA+5 data

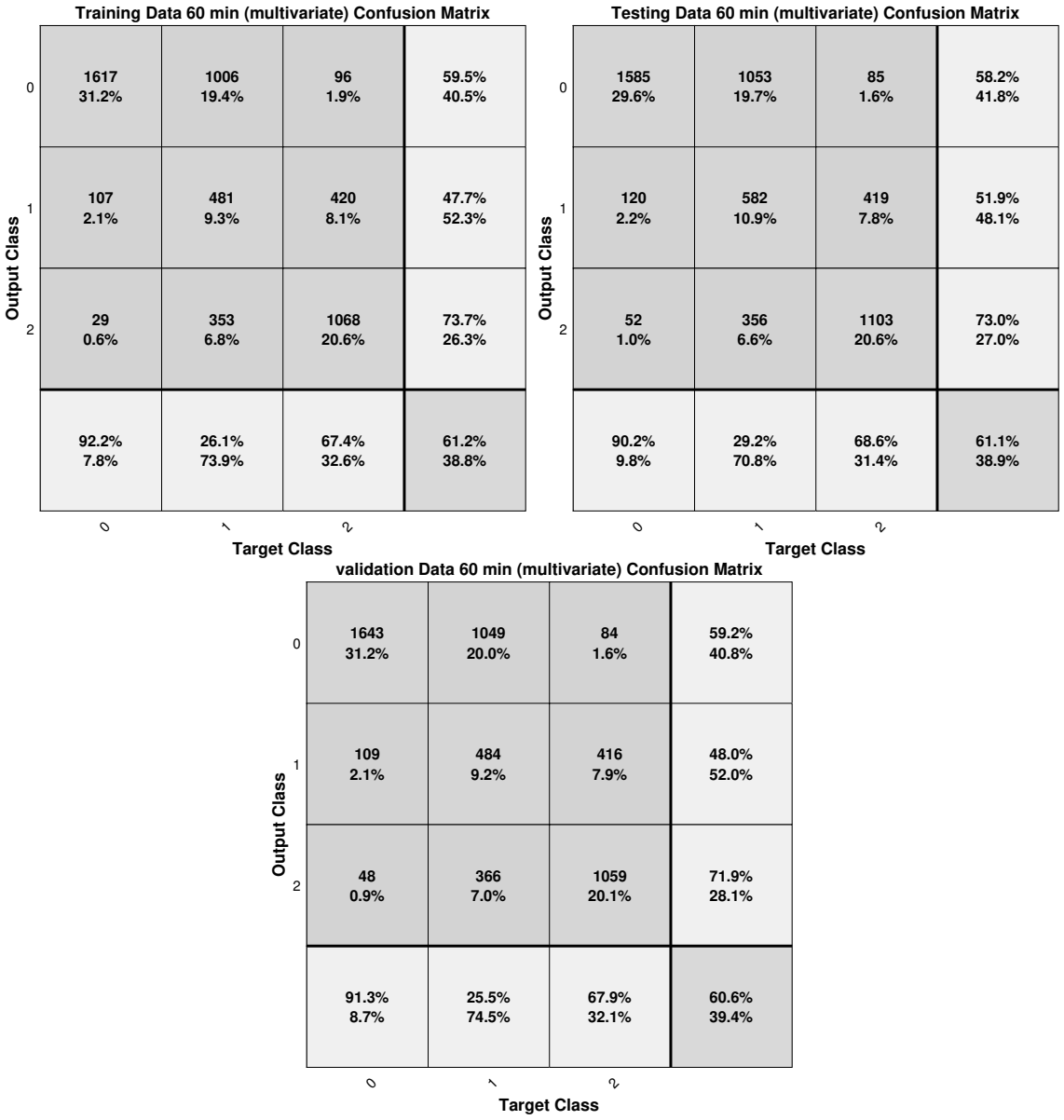

Figure 69: Confusion matrix for multivariate regression of start to 60 minutes data

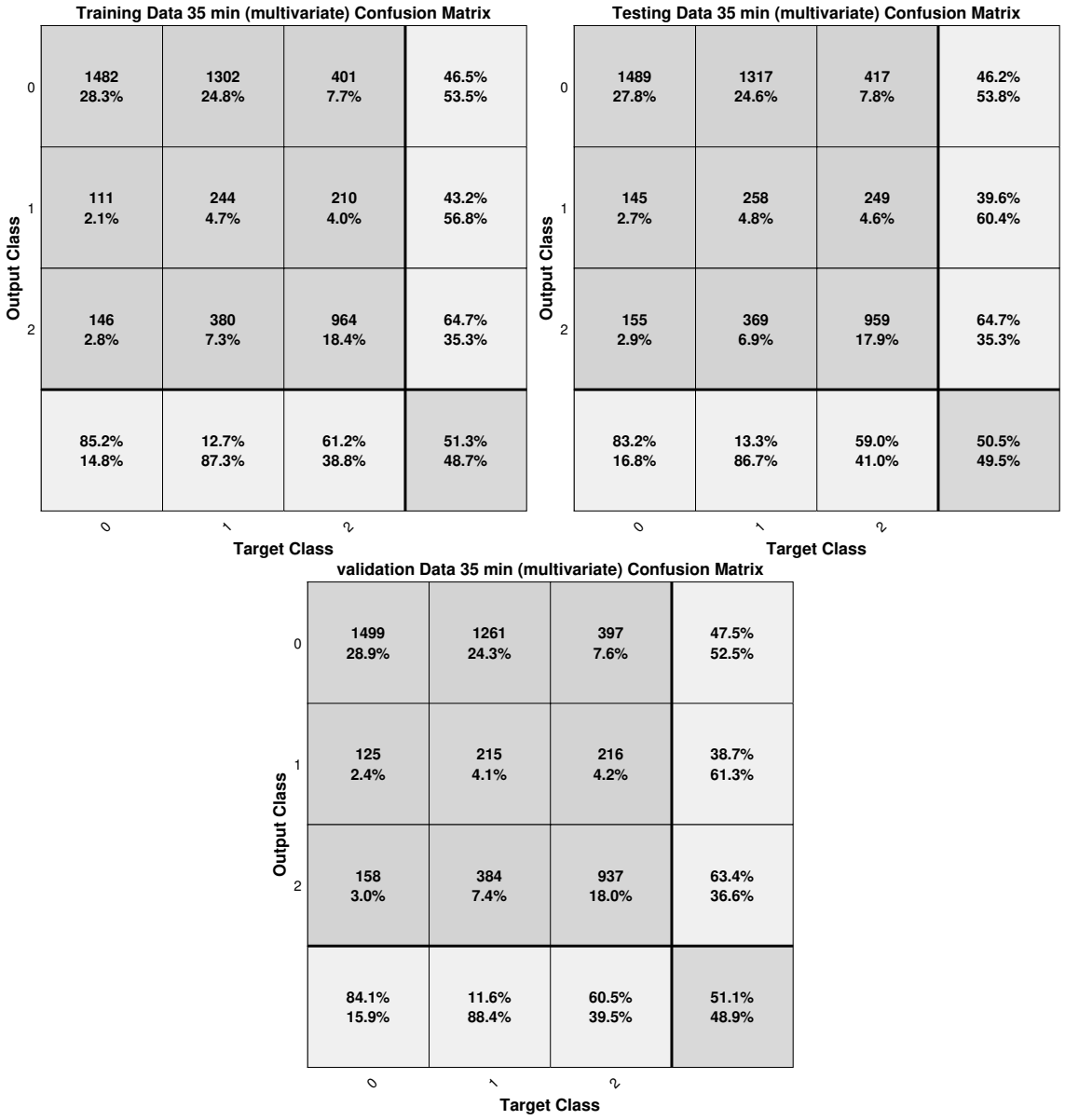

Figure 70: Confusion matrix for multivariate regression of start to 35 minutes data

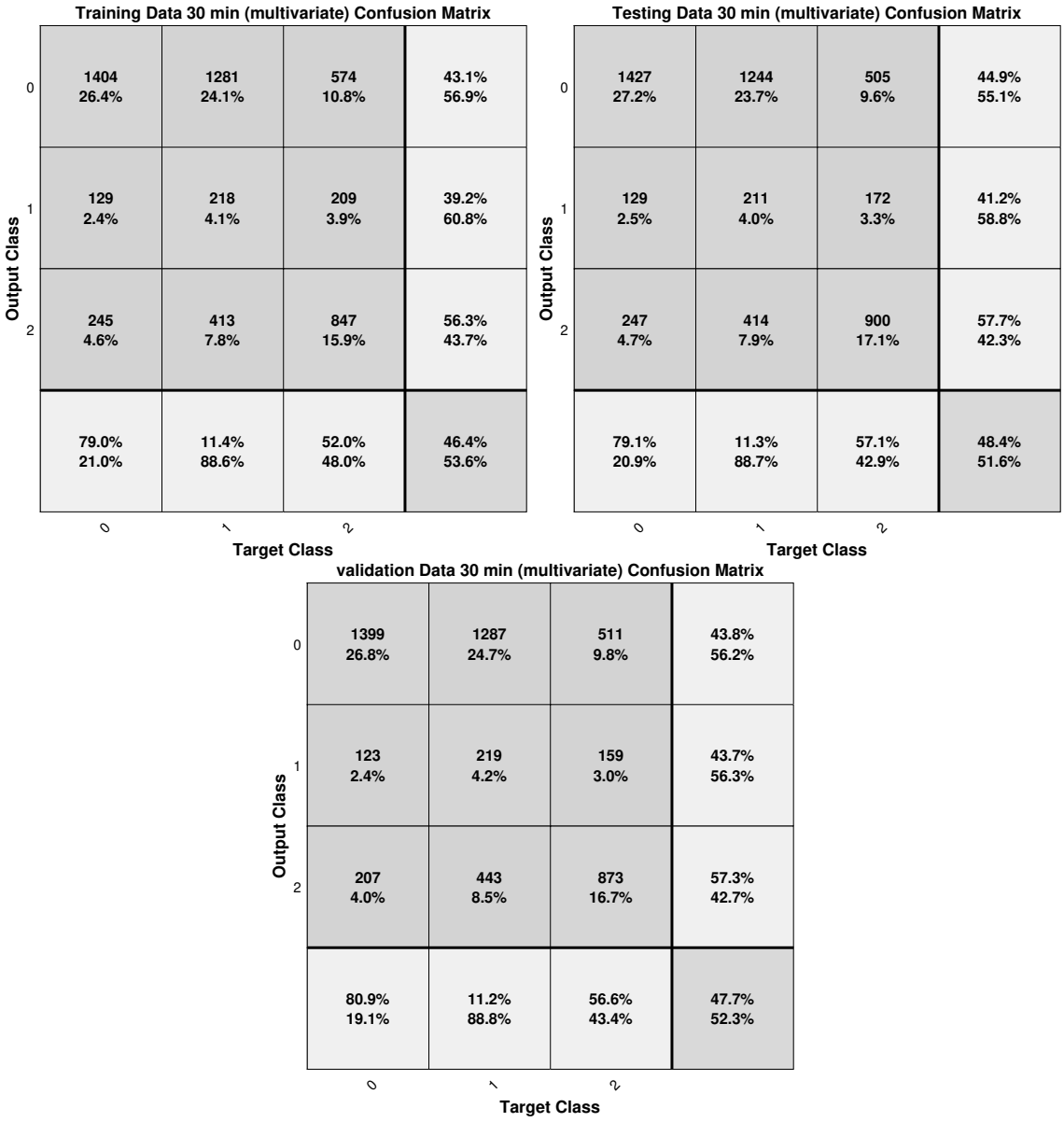

Figure 71: Confusion matrix for multivariate regression of start to 30 minutes data

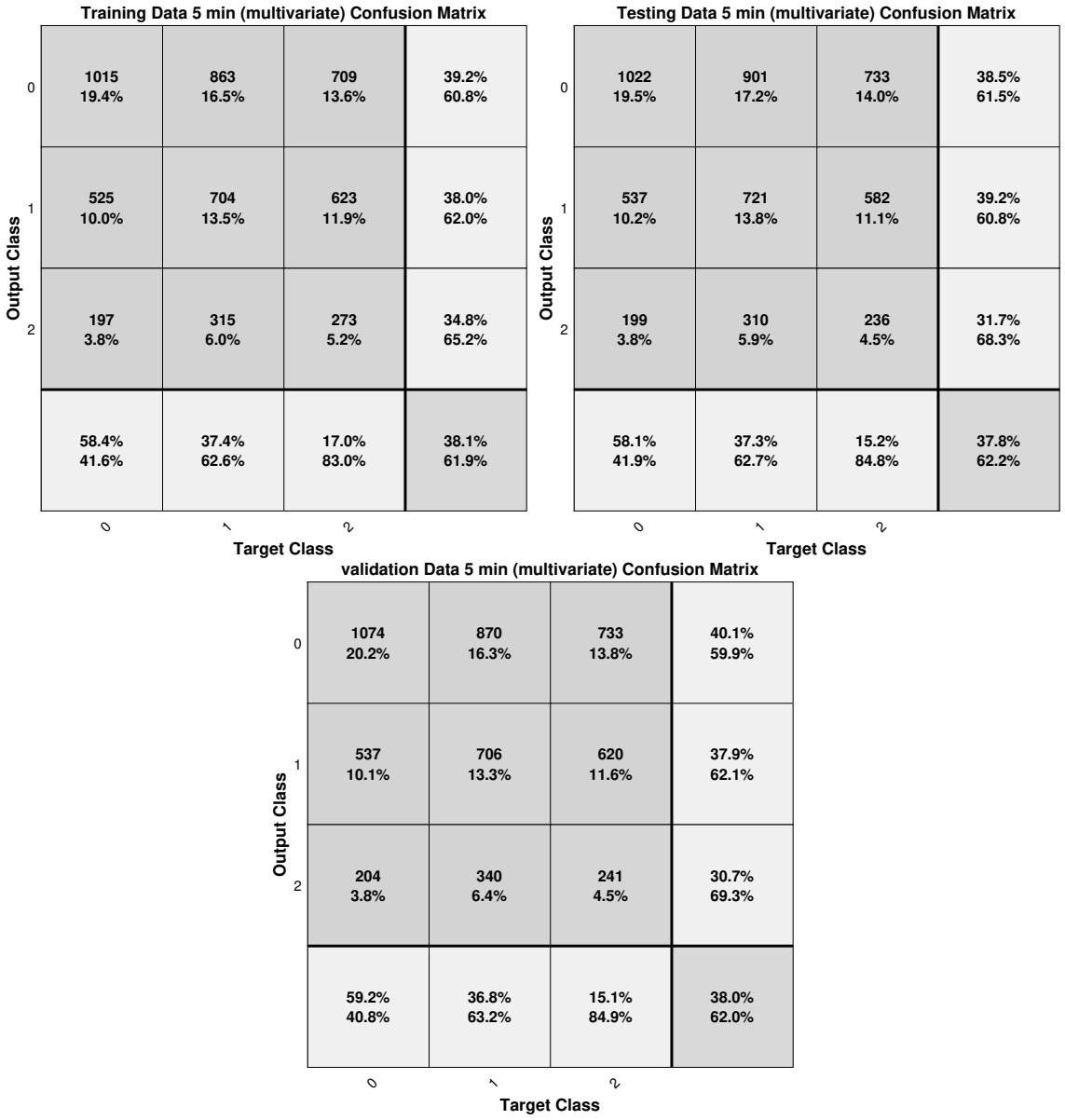

Figure 72: Confusion matrix for multivariate regression of start to 5 minutes data

Bibliography

[1] Khalid O Alfarouk, Christian-Martin Stock, Sophie Taylor, Megan Walsh, Abdel Khalig Muddathir, Daniel Verduzco, Adil HH Bashir, Osama Y Mohammed, Gamal O Elhassan, Salvador Harguindey, et al. Resistance to cancer chemotherapy: failure in drug response from adme to p-gp. *Cancer cell international*, 15(1):1–13, 2015.

[2] M Anand, K Rajagopal, and KR Rajagopal. A model for the formation, growth, and lysis of clots in quiescent plasma. a comparison between the effects of antithrombin iii deficiency and protein c deficiency. *Journal of theoretical biology*, 253(4):725–738, 2008.

[3] Jovan P Antovic and Margareta Blombäck. *Essential guide to blood coagulation*. John Wiley & Sons, 2013.

[4] KB Bischoff, RL Dedrick, DS Zaharko, and JA Longstreth. Methotrexate pharmacokinetics. *Journal of pharmaceutical sciences*, 60(8):1128–1133, 1971.

[5] Joshua B Brown, Francis X Guyette, Matthew D Neal, Jeffrey A Claridge, Brian J Daley, Brian G Harbrecht, Richard S Miller, Herb A Phelan, Peter W Adams, Barbara J Early, et al. Taking the blood bank to the field: the design and rationale of the prehospital air medical plasma (pamper) trial. *Prehospital Emergency Care*, 19(3):343–350, 2015.

[6] Joshua B Brown, Matthew D Neal, Francis X Guyette, Andrew B Peitzman, Timothy R Billiar, Brian S Zuckerbraun, and Jason L Sperry. Design of the study of tranexamic acid during air medical prehospital transport (staamp) trial: addressing the knowledge gaps. *Prehospital Emergency Care*, 19(1):79–86, 2015.

[7] Eduardo F Camacho and Carlos Bordons Alba. *Model predictive control*. Springer science & business media, 2013.

[8] GS Chatta, M Fakih, S Ramalingam, CP Belani, RK Ramanathan, W Zamboni, D Friedland, D Lis, S Tutchko, and M Egorin. Phase i pharmacokinetic (pk) study of daily imatinib in combination with docetaxel for patients with advanced solid tumors. *Journal of Clinical Oncology*, 22(14_suppl):2047–2047, 2004.

[9] 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*, 6(9):e1000950, 2010.

[10] Raheel Chaudhry, Syed Muhammad Usama, and Hani M Babiker. Physiology, coagulation pathways. 2018.

[11] Stephen J Clarke and Laurent P Rivory. Clinical pharmacokinetics of docetaxel. *Clinical pharmacokinetics*, 36(2):99–114, 1999.

[12] Meindert Danhof, Joost de Jongh, EC De Lange, Oscar Della Pasqua, Bart A Ploeger, and Rob A Voskuyl. Mechanism-based pharmacokinetic-pharmacodynamic modeling: biophase distribution, receptor theory, and dynamical systems analysis. *Annu. Rev. Pharmacol. Toxicol*, 47:21–21, 2007.

[13] Hartmut Derendorf and Bernd Meibohm. Modeling of pharmacokinetic/pharmacodynamic (pk/pd) relationships: concepts and perspectives. *Pharmaceutical research*, 16(2):176–185, 1999.

[14] Alberto Fernández, Salvador Garcia, Francisco Herrera, and Nitesh V Chawla. Smote for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. *Journal of artificial intelligence research*, 61:863–905, 2018.

[15] Kieran P Fitzmaurice, Michelle A Pressly, Gilles Clermont, Jonathan H Waters, and Robert S Parker. Dynamic modelling of obstetric patient coagulation from kaolin-activated thromboelastogram data. *IFAC-PapersOnLine*, 53(2):16329–16334, 2020.

[16] Jeffry Alan Florian Jr. *Modeling and dose schedule design for cycle-specific chemotherapeutics*. PhD thesis, University of Pittsburgh, 2008.

[17] Nazzareno Galiè, Paul A Corris, Adaani Frost, Reda E Girgis, John Granton, Zhi Cheng Jing, Walter Klepetko, Michael D McGoon, Vallerie V McLaughlin, Ioana R Preston, et al. Updated treatment algorithm of pulmonary arterial hypertension. *Journal of the American College of Cardiology*, 62(25S):D60–D72, 2013.

[18] Felice Galluccio, Tolga Ergonenc, Alvaro Garcia Martos, Abdallah El-Sayed Allam, Maria Pérez-Herrero, Ricardo Aguilar, Giacomo Emmi, Michele Spinicci, Ignacio Terrancle Juan, and Mario Fajardo-Pérez. Treatment algorithm for covid-19: a multidisciplinary point of view. *Clinical Rheumatology*, 39(7):2077–2084, 2020.

[19] Eduardo Gonzalez, Ernest E Moore, Hunter B Moore, Michael P Chapman, Theresa L Chin, Arsen Ghasabyan, Max V Wohlauer, Carlton C Barnett, Denis D Bensard, Walter L Biffl, et al. Goal-directed hemostatic resuscitation of trauma-induced coagulopathy: a pragmatic randomized clinical trial comparing a viscoelastic assay to conventional coagulation assays. *Annals of surgery*, 263(6):1051, 2016.

[20] Jamie T Griffin, T Deirdre Hollingsworth, Lucy C Okell, Thomas S Churcher, Michael White, Wes Hinsley, Teun Bousema, Chris J Drakeley, Neil M Ferguson, María-Gloria Basáñez, et al. Reducing plasmodium falciparum malaria transmission in africa: a model-based evaluation of intervention strategies. *PLoS medicine*, 7(8):e1000324, 2010.

[21] JB Gross. American society of anesthesiologists task force on perioperative management: Practice guidelines for the perioperative management of patients with obstructive sleep apnea: a report by the american society of anesthesiologists task force on perioperative management of patients with obstructive sleep apnea. *Anesthesiology*, 104:1081–1093, 2006.

[22] Gaorav P Gupta and Joan Massagué. Cancer metastasis: building a framework. *Cell*, 127(4):679–695, 2006.

[23] Howard Gurney. How to calculate the dose of chemotherapy. *British journal of cancer*, 86(8):1297–1302, 2002.

[24] Haemooscope. TEG 5000 Thromboelastograph hemostasis system user manual, teg analytical software: version 4.2.3, 2007.

[25] John M Harrold. *Model-based design of cancer chemotherapy treatment schedules*. PhD thesis, University of Pittsburgh, 2005.

[26] Michael A Hillman, Russell A Wilke, Steven H Yale, Humberto J Vidaillet, Michael D Caldwell, Ingrid Glurich, Richard L Berg, John Schmelzer, and James K Burmester. A prospective, randomized pilot trial of model-based warfarin dose initiation using cyp2c9 genotype and clinical data. *Clinical medicine & research*, 3(3):137–145, 2005.

[27] Taizo Hirata, Shinji Ozaki, Masahiro Tabata, Takayuki Iwamoto, Shiro Hinotsu, Akinobu Hamada, Takayuki Motoki, Tomohiro Nogami, Tadahiko Shien, Naruto Taira, et al. A multicenter study of docetaxel at a dose of 100 mg/m2 in japanese patients with advanced or recurrent breast cancer. *Internal Medicine*, pages 5089–20, 2021.

[28] Thang Ho. *A model-based clinically-relevant chemotherapy scheduling algorithm for anticancer agents*. PhD thesis, University of Pittsburgh, 2014.

[29] Thang Ho, Gilles Clermont, and Robert S Parker. A model of neutrophil dynamics in response to inflammatory and cancer chemotherapy challenges. *Computers & chemical engineering*, 51:187–196, 2013.

[30] 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.

[31] Arti Hurria, Mark T Fleming, Sharyn D Baker, Wm Kevin Kelly, Katie Cutchall, Katherine Panageas, James Caravelli, Henry Yeung, Mark G Kris, Jorge Gomez, et al. Pharmacokinetics and toxicity of weekly docetaxel in older patients. *Clinical cancer research*, 12(20):6100–6105, 2006.

[32] Celestine Iwendi, Ali Kashif Bashir, Atharva Peshkar, R Sujatha, Jyotir Moy Chatterjee, Swetha Pasupuleti, Rishita Mishra, Sofia Pillai, and Ohyun Jo. Covid-19 patient health prediction using boosted random forest algorithm. *Frontiers in public health*, 8:357, 2020.

[33] Sara S Johnson, Andrea L Paiva, Carol O Cummins, Janet L Johnson, Sharon J Dyment, Julie A Wright, James O Prochaska, Janice M Prochaska, and Karen Sherman. Transtheoretical model-based multiple behavior intervention for weight management: effectiveness on a population basis. *Preventive medicine*, 46(3):238–246, 2008.

[34] Kenneth C Jones and Kenneth G Mann. A model for the tissue factor pathway to thrombin. ii. a mathematical simulation. *Journal of Biological Chemistry*, 269(37):23367–23373, 1994.

[35] William J Jusko. Moving from basic toward systems pharmacodynamic models. *Journal of pharmaceutical sciences*, 102(9):2930–2940, 2013.

[36] William J Jusko and Hui C Ko. Physiologic indirect response models characterize diverse types of pharmacodynamic effects. *Clinical Pharmacology & Therapeutics*, 56(4):406–419, 1994.

[37] U Ledzewicz and H Schättler. Optimal Controls for a Model with Pharmacokinetics Maximizing Bone Marrow in Cancer Chemotherapy. *Math. Biosci.*, 206:320–342, 2007.

[38] Gerhard Levy. Kinetics of pharmacologic effects. *Clinical Pharmacology & Therapeutics*, 7(3):362–372, 1966.

[39] Joseph T Liparulo, Timothy D Knab, and Robert S Parker. Toxicity-centric cancer chemotherapy treatment design. *IFAC-PapersOnLine*, 53(2):16353–16358, 2020.

[40] Marc Maegele, Rolf Lefering, Nedim Yucel, Thorsten Tjardes, Dieter Rixen, Thomas Paffrath, Christian Simanski, Edmund Neugebauer, Bertil Bouillon, AG Polytrauma of the German Trauma Society (DGU, et al. Early coagulopathy in multiple injury: an analysis from the german trauma registry on 8724 patients. *Injury*, 38(3):298–304, 2007.

[41] Rory Martin and KL Teo. *Optimal control of drug administration in cancer chemotherapy*. World Scientific, 1994.

[42] Hunter B Moore, Ernest E Moore, Theresa L Chin, Eduardo Gonzalez, Michael P Chapman, Carson B Walker, Angela Sauaia, and Anirban Banerjee. Activated clotting time of thrombelastography (t-act) predicts early postinjury blood component transfusion beyond plasma. *Surgery*, 156(3):564–569, 2014.

[43] Larry Norton. A gompertzian model of human breast cancer growth. *Cancer research*, 48(24_Part_1):7067–7071, 1988.

[44] Larry Norton and Joan Massagué. Is cancer a disease of self-seeding? *Nature medicine*, 12(8):875–878, 2006.

[45] Marina Panova-Noeva, Lisa Eggebrecht, Jürgen H Prochaska, and Philipp S Wild. Potential of multidimensional, large-scale biodatabases to elucidate coagulation and platelet pathways as an approach towards precision medicine in thrombotic disease. *Hämostaseologie*, 39(02):152–163, 2019.

[46] Robert S Parker, Francis J Doyle, and Nicholas A Peppas. A model-based algorithm for blood glucose control in type i diabetic patients. *IEEE Transactions on biomedical engineering*, 46(2):148–157, 1999.

[47] Matthew E Pollard, Alan J Moskowitz, Michael A Diefenbach, and Simon J Hall. Cost-effectiveness analysis of treatments for metastatic castration resistant prostate cancer. *Asian journal of urology*, 4(1):37–43, 2017.

[48] Michelle A Pressly. *Dynamic Modeling of Coagulation Advances the Bloody Fight to Improve Patient Outcomes*. PhD thesis, University of Pittsburgh, 2020.

[49] Michelle A Pressly, Matthew D Neal, Gilles Clermont, and Robert S Parker. Dynamic modeling of thromboelastography to inform state of coagulopathy in trauma patients. In *Foundations of Computer Aided Process Operations and Chemical Process Control Conference*, volume 105, 2017.

[50] 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.

[51] E Salminen, M Bergman, S Huhtala, and E Ekholm. Docetaxel: standard recommended dose of 100 mg/m2 is effective but not feasible for some metastatic breast cancer patients heavily pretreated with chemotherapy—a phase ii single-center study. *Journal of clinical oncology*, 17(4):1127–1127, 1999.

[52] Ajith Sankarankutty, Bartolomeu Nascimento, Luis Teodoro da Luz, and Sandro Rizoli. Teg® and rotem® in trauma: similar test but different results? In *World Journal of Emergency Surgery*, volume 7, pages 1–8. Springer, 2012.

[53] Amarnath Sharma, William F Ebling, and William J Jusko. Precursor-dependent indirect pharmacodynamic response model for tolerance and rebound phenomena. *Journal of pharmaceutical sciences*, 87(12):1577–1584, 1998.

[54] Lewis B Sheiner, Donald R Stanski, Samuel Vozeh, Ronald D Miller, and Jay Ham. Simultaneous modeling of pharmacokinetics and pharmacodynamics: application to d-tubocurarine. *Clinical Pharmacology & Therapeutics*, 25(3):358–371, 1979.

[55] W.S. Snyder, M.J. Cook, E.S. Nasset, L.R. Karhausen, G.P. Howells, and I.H. Tipton. *Report of the Task Group on Reference Man: A Report Prepared by a Task Group of Committee 2 of the International Commission on Radiological Protection*. Pergamon Press, Oxford, 1975.

[56] DB Solit, M Egorin, G Valentin, A Delacruz, Q Ye, L Schwartz, S Larson, N Rosen, and HI Scher. Phase 1 pharmacokinetic and pharmacodynamic trial of docetaxel and 17aag (17-allylamino-17-demethoxygeldanamycin). *Journal of Clinical Oncology*, 22(14_suppl):3032–3032, 2004.

[57] Jason L Sperry, Francis X Guyette, Joshua B Brown, Mark H Yazer, Darrell J Triulzi, Barbara J Early-Young, Peter W Adams, Brian J Daley, Richard S Miller, Brian G Harbrecht, et al. Prehospital plasma during air medical transport in trauma patients at risk for hemorrhagic shock. *New England Journal of Medicine*, 379(4):315–326, 2018.

[58] W Street. Cancer facts & figures 2022. *American Cancer Society: Atlanta, GA, USA*, 2022.

[59] S Strychor, JL Eiseman, RA Parise, MJ Egorin, E Joseph, and WC Zamboni. Plasma, tumor, and tissue disposition of docetaxel in scid mice bearing skov-3 human ovarian xenografts. In *Proc AACR*, volume 2005, page 4165, 2005.

[60] GW Swan. Cancer chemotherapy: Optimal control using the verhulst-pearl equation. *Bulletin of mathematical biology*, 48(3-4):381–404, 1986.

[61] US Food and Drug Administration. Highlights of docetaxel prescribing information, 2012.

[62] Oksana Volod, Connor M Bunch, Nuha Zackariya, Ernest E Moore, Hunter B Moore, Hau C Kwaan, Matthew D Neal, Mahmoud D Al-Fadhl, Shivani S Patel, Grant Wiarda, et al. Viscoelastic hemostatic assays: A primer on legacy and new generation devices. *Journal of Clinical Medicine*, 11(3):860, 2022.

[63] Mark Walsh, Scott Thomas, Hau Kwaan, John Aversa, Stephen Anderson, Rajalakshmy Sundararajan, Donald Zimmer, Connor Bunch, John Stillson, Dominik Draxler, et al. Modern methods for monitoring hemorrhagic resuscitation in the united states: Why the delay? *Journal of Trauma and Acute Care Surgery*, 89(6):1018–1022, 2020.

[64] Sholom M Weiss, Casimir A Kulikowski, Saul Amarel, and Aran Safir. A model-based method for computer-aided medical decision-making. *Artificial intelligence*, 11(1-2):145–172, 1978.

[65] Joseph Willis, Caleb Carroll, Virginia Planz, and Samuel J Galgano. Thromboelastography: a review for radiologists and implications on periprocedural bleeding risk. *Abdominal Radiology*, pages 1–7, 2022.

[66] Han Wu, Shengqi Yang, Zhangqin Huang, Jian He, and Xiaoyi Wang. Type 2 diabetes mellitus prediction model based on data mining. *Informatics in Medicine Unlocked*, 10:100–107, 2018.

[67] Shinji Yamazaki, Judith Skaptason, David Romero, Sylvia Vekich, Hannah M Jones, Weiwei Tan, Keith D Wilner, and Tatiana Koudriakova. Prediction of oral pharmacokinetics of cmet kinase inhibitors in humans: physiologically based pharmacokinetic model versus traditional one-compartment model. *Drug metabolism and disposition*, 39(3):383–393, 2011.

[68] William C Zamboni, Sandra Strychor, Erin Joseph, Robert A Parise, Merrill J Egorin, and Julie L Eiseman. Tumor, tissue, and plasma pharmacokinetic studies and antitumor response studies of docetaxel in combination with 9-nitrocamptothecin in mice bearing skov-3 human ovarian xenografts. *Cancer chemotherapy and pharmacology*, 62(3):417–426, 2008.

[69] Huixi Zou, Parikshit Banerjee, Sharon Shui Yee Leung, and Xiaoyu Yan. Application of pharmacokinetic-pharmacodynamic modeling in drug delivery: development and challenges. *Frontiers in pharmacology*, 11:997, 2020.