Western Kentucky University

TopSCHOLAR®

2020

Mathematical Modeling of Diabetic Foot Ulcers Using Optimal Design and Mixed-Modeling Techniques

Michael Belcher
Western Kentucky University, m.belcher16@yahoo.com

Follow this and additional works at: https://digitalcommons.wku.edu/stu_hon_theses

 Part of the Other Applied Mathematics Commons

MATHEMATICAL MODELING OF DIABETIC FOOT ULCERS USING OPTIMAL
DESIGN AND MIXED-MODELING TECHNIQUES

A Capstone Project Presented in Partial Fulfillment

of the Requirements for the Degree Bachelor of Science

with Mahurin Honors College Graduate Distinction

at Western Kentucky University

By

Michael Belcher

May 2020

*****

CE/T Committee:

Dr. Richard Schugart, Mentor

Dr. Ozkan Ozer

Ms. Cheryl Kirby-Stokes

ABSTRACT

A mathematical model for the healing response of diabetic foot ulcers was developed using averaged data (Krishna *et al*., 2015). The model contains four major factors in the healing of wounds using four separate differential equations with 12 parameters. The four differential equations describe the interactions between matrix metalloproteinases (MMP-1), tissue inhibitors of matrix metalloproteinases (TIMP-1), the extracellular matrix (ECM) of the skin, and the fibroblasts, which produce these proteins. Recently, our research group obtained the individual patient data that comprised the averaged data. The research group has since taken several approaches to analyze the model with the individual patient data. One approach was to introduce mixed modeling techniques on certain parameters in that model. Mixed effects modeling is an analytical tool useful for the repeated measurement of data with subjects, patients, etc., that have random affects that deviate from a specified norm. This is accomplished by taking a parameter that is shared across all data sets and splitting it into a fixed variable and a random variable. Then all data sets are modeled so that the fixed variable is the same for all patients and the random variable is a unique modifier that accounts for the differences across patients. Another approach has been to use an optimal design technique to identify which times are ideal for data gathering for the model. For this project a Standard Error (SE) optimal design method was chosen with the goal of minimizing the sum of squared normalized standard errors. Our project worked to combine these techniques by first introducing mixed modeling parameter values into an SE-optimal design algorithm and then comparing the results collected to the standard algorithm.

Other optimal design techniques were used with and without these mixed modeling parameters to see if a certain technique was better than the others. Finally, we worked toward improving our estimates for mixed modeling parameters by attempting to implement in MONOLIX. Our focus for this part was to develop test cases that could be implemented in MONOLIX.

First and foremost, I would like to dedicate this work to my Lord and Savior, Jesus Christ, without whom none of this would be possible or be of any importance. I would also like to mention my wife, Erinly, for putting up with me being stressed and always supporting me in all that I do.

ACKNOWLEDGEMENTS

I would like to thank the WKU Mathematics Department for years of mentorship and education that lead to this work. Without many of the professors in the Math department and the opportunities and leadership that they have presented to me, this work would not have happened. Finally, I would like to thank Dr. Richard Schugart for allowing me to be a part of the project as well as being a mentor for me personally.

VITA

EDUCATION

Western Kentucky University, Bowling Green, KY   May 2020

    B.S. in Mathematics – Mahurin Honors College Graduate

    Honors Capstone: *Mathematical Modeling of Diabetic Foot Ulcers Using Optimal Design and Mixed-Modeling Techniques*

Gatton Academy of Mathematics and Science in Kentucky,   May 2016

    Bowling Green, KY

Barren County High School, Glasgow, KY   May 2016

PUBLICATIONS

Atici, Ferhan & Atıcı, Mustafa & Belcher, Michael & Marshall, Dana. (2017). **A New Approach for Modeling with Discrete Fractional Equations**. *Fundamenta Informaticae*. DOI: 151. 313-324. 10.3233/FI-2017-1494.

PRESENTATIONS

KYMAA Spring 2019 Conference, Centre College, Danville, KY   March 2019

NIMBioS Undergraduate Research Conference,   November 2019

    University of Tennessee, Knoxville, TN

39th Annual WKU Math Symposium,   November 2019

    Western Kentucky University, Bowling Green, KY

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

The wound-healing process is complex and requires the proper combination of biological healing factors to be present in the correct ratios for the healing process to progress. Proper healing occurs in four distinct stages: coagulation and hemostasis, inflammation, proliferation, and wound remodeling (Velnar et al., 2009). Coagulation and hemostasis begin the instant a wound is formed with the primary function being to create a blood clot to prevent extreme blood loss that could lead to further heal risks. For this to occur, an interconnected matrix of cells must be created which is permeable enough to allow cells through that are needed for future healing while also preventing excessive bleeding and further damage. This process begins immediately after trauma by causing an involuntary reaction from the neuronal reflex that leads to the contraction of blood vessels and vascular smooth muscle cells in the muscle layer. After this constriction is ended, "blood spills into the site of injury, the blood components and platelets come in contact with exposed collagen and other extracellular matrix components" (Velnar et al., 2009).   This creates a blood clot that is composed of fibronectin, fibrin, vitronectin, and thrombospondin. These platelet clots contain a series of growth factors that are important for later stages of the healing process. Without these clots, platelet derived growth factor (PDGF), transforming growth factor-β (TGF- β), epidermal growth factor (EGF), and insulin-like growth factor (IGF-1) would not be present in the healing site. These promote the healing process by recruiting additional healing factors into the site and activating them to begin the process of healing. In addition to the introduction of growth factors, platelet clots also contain vasoactive

amines that are stored and lead to "fluid extravasation in the tissue that results in oedema, which, in turn, potentiates itself during the following inflammation phase" (Velnar et al., 2009).As coagulation and hemostasis nears completion, the inflammatory phase begins. The goal of this phase is to purge any outside bacteria or infection-causing agents that have entered the body during or after the formation of the wound. In the early stages of inflammation, neutrophils enter and remove any of the outside bacteria that may cause infection or hinder the healing process. The neutrophils enter the wound site around 24-36 hours after a wound has formed. The neutrophils stick to the capillaries surrounding the wound and use the flow of blood to move around the wound site gathering up bacteria and any dangerous materials that have entered into the site. Once the site has been cleared the neutrophils are purged and are replaced by macrophages that continue the cleansing of the site. Macrophages have a longer lifespan than the fast acting neutrophils and can work at lower pH levels as the wound healing progresses. As the healing process continues, macrophages continue to clean the site as well as release growth factors that recruit cells like fibroblasts, keratinocytes, and endothelial cells. These cells are so important to the healing process that the lack of macrophages in the wound site "… causes severe healing disturbances due to poor wound debridement, delayed fibroblast proliferation and maturation, as well as delayed angiogenesis, resulting in inadequate fibrosis and a more weakly repaired wound" (Velnar et al., 2009). Finally, the last step in the inflammation stage is the entrance of lymphocytes and immunoglobulin G (IgG) breakdown products that are crucial for later healing.

Once the inflammation phase has begun to subside and the wound site has been thoroughly "cleaned" by neutrophils and macrophages, proliferation begins. This begins

the tissue repair stages of wound healing. This stage begins with what is known as a "fibroblast migration" to the wound site that essentially floods the wound site with fibroblasts and myofibroblasts that are going to be in charge of synthesizing new extracellular matrix in the wound (Velnar et al., 2009). These fibroblasts produce collagen, which is vital for all stages of the healing process as it is the main structural protein in the skin as well as in other tissues. While the production of the extracellular matrix is obviously the main priority during this stage, blood vessels and other supportive tissue is also developed and controlled by cell migration to the growth factors that have been brought to the site during the inflammation phase. This ensures that any new extracellular matrix that is created will have adequate numbers of capillaries and blood flow present for continued health and re-growth. Endothelial cells are then "moved" into place creating the matrix using three processes that work together: protrusion, adhesion, and traction. Protrusion refers to filaments that "protrude" from the cell allowing for cell-cell movement along the extracellular matrix. These filaments act as a sort of anchor point for the cell and allow strong connections that can be manipulated to provide movement. Adhesion is controlled by integrin, extracellular receptors that allow for the adhesion between two cells, which are present in all cells in the extracellular matrix. The proper regulation of adhesion is vital to the movement of these cells to the correct place in the matrix for optimal migration rates that ebb and flow during the healing process adhesion is vital. If there is not adequate cell adhesion, then the site will be flooded with too many cells and healing is hindered. At the same time, if cells adhere too much to one another, then endothelial cells struggle to get to the site, which delays or even stops the healing process. Finally, traction is the pulling motion that is created between cells

through their integrin connections. This pulling is what actually facilitates cell movement, which is controlled by the adhesion of the cell and is made possible because of the protrusion from the cell. Without proper protrusion and cell adhesion, wound healing is hindered. When these three processes work together properly, endothelial cells are transported to the wound site and production of the extracellular matrix continues.

In the remodeling phase of the process, scar tissue and epithelium, the thin outside layer of the skin, are produced to complete healing of the wound. This process is actually begun quite soon after injury and can take up at least a year or two for the entire process to be complete, and depending on the severity of the wound, could take even longer (Velnar et al., 2009). In the wound-healing process, this phase actually overlaps with nearly all of the other phases. For the healing to progress properly through remodeling, there must be control of the degradation of old tissue as well as synthesis of new tissue. The main agent responsible for the breakdown of the extracellular matrix are matrix metalloproteinases (MMPs), more specifically for this project MMP-1, and is controlled by its inhibitor TIMP-1. If degradation is too prolific because of an overabundance of MMP-1, then the wound may never be able to heal properly. At the same time, if old tissue is not broken down quickly enough to be replaced with new matrix then the new site will not be strong enough to withstand potential further injury. Having the degradation and synthesis in equilibrium is essential for the long-term health of the wound site. As new extracellular matrix is produced, there is an abundance of collagen being infused into the new skin being formed, which can almost entirely recreate the strength of the skin present before the wound. During this remodeling phase the collagen present is transferred from being an unorganized mass present in the site to an organized

matrix that interconnects building strength and structure at the wound site. As structures are created, the underlying tissues shrink and contract closing the site as the healing continues. This closing happens in direct reaction to the creation of new matrix as scar tissue is replaced with properly created extracellular matrix and as the blood vessels and capillaries are extended into the newly created matrix. Over time the wound closes itself and proper healing is complete.

Because of the complexity of this process, it is difficult to analyze all phases of the healing process at one time. However, a few biological markers have been identified that influence all stages of the wound-healing process. Proper wound healing requires matrix metalloproteinases (MMP-1) and its corresponding inhibitor (TIMP-1) to remain in proper ratios and be present at appropriate times for the wound to heal in a normal and sufficient manner. The proportions of these proteins in the wound must be precise and, if there is too much or too little of one specific factor, the healing process will be hindered and can directly lead to the wound becoming chronic (Muller et al., 2008). The impact of these factors directly impacts the success or failure of the wound-healing process. Wounds that follow this process are known as acute healers and wounds that, for any of a variety of reasons, do not follow this process are called chronic wounds. There are a number of issues that could be present to cause a wound to become chronic as well as a number of treatments and management courses that can be taken. This project looks specifically at patients who have diabetes as they are much more likely to develop chronic wounds than the general population. People with diabetes are more likely to have complications "caused by several intrinsic factors (neuropathy, vascular problems,

other complicating systemic effects due to diabetes) and extrinsic factors (wound infection, callus formation, and excessive pressure to the site)" (Falagna, 2005).

To further evaluate the role of MMPs & TIMPs in the healing of chronic wounds, 16 separate patients with diabetic foot ulcers who were monitored over a twelve-week period or until the wound has fully healed (Muller et al., 2008). They measured a variety of wound proteins over this time, which included MMP-1 and TIMP-1. Cross-sectional areas of the wound were also measured. To further study the influences on the healing response, patients were subdivided into two categories – "good" healers and "poor" healers. A "good" healer is defined to be a patient with at least 82% wound closure at the 4-week point, while a "poor" healer is not. This definition was based on another study (Sheehan et al., 2003) that looked at four-week healing rates of chronic wounds that healed within 12 weeks. To further analyze the impact of these factors on the healing response of the wound, a mathematical model was formulated using the median patient data of the "good" and "poor" healers (Krishna et al., 2015).

In this model, four distinct state variables are used to describe the healing process through corresponding differential equations containing a total of twelve parameters. Specifically, these equations look at the evolution of MMP-1, TIMP-1, extracellular matrix, and fibroblast cell count which are labeled M, T, E, and $\tilde{f}$, respectively, and are shown below in Equation (1) – (4). There are twelve total parameters, $k_1 - k_{11}$ and the initial fibroblast parameter, f_i.

$$\frac{dM}{dt} = \frac{k_1 M^\alpha (\tilde{f} + f_i)}{k_2^\alpha + M^\alpha} - k_3 M - k_4 MT \tag{1}$$

$$\frac{dT}{dt} = \frac{k_5 T^\beta (\tilde{f} + f_i) M}{k_6^\beta + T^\beta} - k_7 T - k_4 MT \tag{2}$$

$$\frac{dE}{dt} = k_8 (\tilde{f} + f_i)(1 - E) - k_9 ME - k_{10} E \tag{3}$$

$$\frac{d\tilde{f}}{dt} = k_{11} (\tilde{f} + f_i)[1 - (\tilde{f} + f_i)] \tag{4}$$

Equation 1-4: Mathematical Model of Wound Healing

All equations are non-dimensionalized through rescaling. Both M and T and their corresponding data with MMP-1 and TIMP-1 are scaled by an average initial value of TIMP-1. We use the data of the wound cross-sectional area for the ECM by setting $E = \frac{A_{max} - A}{A_{max}}$, where A_{max} is the largest cross-sectional are for all patients. While there are no data for f, we scale f by its carrying capacity. While the original work was with median data (Krishna et al., 2015), all subsequent work (French, 2017; Karimli, 2019; Prassad, 2017; Alotaibi, 2019), including this thesis, is with the individual patient data.

Using multiple computational techniques is the basis for this work. One technique is called mixed effects modeling. Mixed effects modeling is an analytical tool useful for the repeated measurement of data with subjects, patients, etc. that have random effects that deviate from a specified norm. Modern uses for mixed modeling effects can remedy a variety of issues that analysts have run into in the past including, but not limited to,

(a) deficiencies in statistical power related to the problems posed by repeated observations, (b) the lack of a flexible method of dealing with missing data, (c) disparate methods for treating continuous and categorical responses, as well as (d)

unprincipled methods of modeling heteroskedasticity and non-spherical error

variance (for either participants or items) (Baayen at al., 2008).

Previously, French (2017) applied the technique of mixed modeling effects to the

parameters present in the model using an exponential function of a sum of two

parameters (Equation 5). The parameter β_i is a fixed effect that is the same value for all

13 patients (Three patients were unable to be curve fit due to issues with their data.),

while the random variable, $\Phi_{i,j}$, is a parameter that measures the variability that occurs for

each of the patients. This technique can be applied to any parameter.

$$k_{i,j} = e^{(\beta_i + \Phi_{i,j})}, \ i = 1, \dots, 12 , \ j = 1, \dots , 13 \tag{5}$$

Equation 5: Mixed Modeling of Parameters (French, 2017)

A second set of techniques is optimal design methods. Optimal Design is a method

that has historically been used to identify optimal sampling times and distributions in

relation to the cost of gathering data in effort to minimize the error of the parameter

estimates. In medicine, where data collection can be both sparse and expensive,

identifying optimal times for collecting data for individual patients is essential.

There are many different optimal design techniques that can be used, many of which

were explored in Banks et al. (2011). These differing methods all work by minimizing a

specific characteristic of the model and data set. For example, the Standard Error (SE)

optimal design minimizes the standard error between the model and the original data.

The determinant (D) optimal-design technique works by minimizing the determinant of

the Fisher Information Matrix (FIM), while the eigenvalue (E) optimal-design technique

works by minimizing the minimum eigenvalue of the FIM. The Fisher Information Matrix measures the information that an observable variable contains regarding an unknown parameter assuming a known distribution (Frieden, 2004).

This project seeks to combine a series of previously-developed methods to analyze this mathematical model in wound healing. The techniques of mixed-modeling & optimal design were previously used separately (French, 2017; Karimli, 2019; Alotaibi, 2019; Prassad, 2017) and part of this project aims to see how these techniques may work together to create a better understanding of the model. Using parameters that were identified through mixed-modeling techniques can be put through the optimization routines that can then be used to identify specific days that data should be collected for optimal modeling accuracy.

MIXED MODELING AND OPTIMAL DESIGN

METHODS

The first step in working with the wound-healing data is to create and implement an algorithm that would use the model from previous work (Krishna et al., 2015) and adapt it to a given data set. The first previously-developed method (Prassad, 2017) is mixed effect modeling and used with the following procedure in Alotaib (2019).

Procedure 1: Mixed Modeling Effects

1. Curve fit the individual patient data using the GlobalSearch algorithm (as outlined in Krishna et al., 2015).
2. Identify a subset of parameters that are divided into fixed and random effects using $k_i = k_{i\,random} + k_{i\,fixed}$ for the ith parameter. The subset of parameters used in this project are $k_1, k_5, k_8, and\ k_{11}$.
3. Fix the other parameters to values from the curve fits in Step 1.
4. Recurve fit all patients simultaneously to find fixed and random effect values $k_1, k_5, k_8, and\ k_{11}$ for all patients using the GlobalSearch algorithm.
5. Combine the random & fixed effect measurement to obtain a mixed modeling effect estimate for $k_1, k_5, k_8, and\ k_{11}$.

The GlobalSearch algorithm in MATLAB was unable to curve fit all parameters simultaneously for the mixed effects model. A subset of parameters was then used. It was decided that parameters *k₁, k₅, k₈, and k₁₁*, which are known as the growth parameters, would be the ones chosen. This work was completed in Alotaibi (2019) and the mixed-effect parameter values are given in Table 1, where $k_{i\,fixed}$ represents the random effect for the *i*-the parameter, and $k_{i\,combined} = k_{i\,fixed} + k_{i\,random}$ for the *i*-the parameter. $k_{i\,combined}$ and the fixed values from the individual curve fits for the

remaining parameters are used for the rest of this work. The next step is to use the

parameter values to run an optimal-design algorithm that would result in the optimal

times for data collection for each patient. The procedure is as follows:

Procedure 2: Optimal Design Code

1. Define a final time T and the number of optimal time values for collecting data.
2. Choose an initial guess for the optimal time points and define the error tolerance for the GlobalSearch algorithm.
3. Define an optimal design method.
4. After running the optimal design method, return the time values in both days and weeks and the least-squares minimum value J.

This process was initially done using an SE optimal-design technique, where the J-

min value represented the Standard Error of the FIM. Likewise, we also used E-optimal

design which requires $J = \min \{eig(FIM)\}$ which minimizes the minimum eigenvalue of

the FIM. $FIM = X^T X$ where X is a sensitivity matrix containing partial derivatives of the

state variables with respect to the parameters. For the D-optimal design algorithm, $J =$

$\min \{det(FIM)\}$, which minimizes the determinant of the FIM. Then we compare the

results from each technique to see if one technique is better or worse than the others. This

process outputs four individual graphs, one for each of the state variables in the model,

which shows the original data, the model that was created, and final the optimal design

time points. In addition to the graphs, there is a list of data points that gives the exact

values of the optimal design points. These values are given first in weeks and then in

days; this allows the user to make the best decision possible when collecting data

considering the data collection process may be complicated.

p	k1fixed	k1random	k1combined	k5fixed	k5random	k5combined
p1	2.2446e-09	1.1460e-09	0.000418	0.0768	66.9919	67.0687
p2	2.2446e-09	2.1011e-05	0.014434	0.0768	1.9117	1.9885
p3	2.2446e-09	5.4684e-08	0.002111	0.0768	7.2137	7.2905
p4	2.2446e-09	3.1597e-05	0.021567	0.0768	1.5742e-05	0.087407
p5	2.2446e-09	0.0294	0.029677	0.0768	0.5163	0.5931
p6	2.2446e-09	0.0375	0.037777	0.0768	1.8434	1.9202
p7	2.2446e-09	8.5962	8.59648	0.0768	2.8857	2.9625
p8	2.2446e-09	0.0129	0.013177	0.0768	5.6119	5.6887
p9	2.2446e-09	0.1639	0.164177	0.0768	3.1099	3.1867
p10	2.2446e-09	0.0040	0.004277	0.0768	2.1522	2.229
p11	2.2446e-09	0.0471	0.047377	0.0768	1.5168	1.5936
p12	2.2446e-09	0.0704	0.070677	0.0768	0.4363	0.5131
p13	2.2446e-09	5.0880e-05	0.03456	0.0768	0.0668	0.1436
p14	2.2446e-09	0.0104	0.010677	0.0768	1.5696	1.6464
p15	2.2446e-09	0.5023	0.502577	0.0768	2.5438	2.6206

p	k8fixed	k8random	k8combined	k11fixed	k11random	k11combined
p1	1.1123	7.6166	8.7289	8.2047e-09	9.6692	9.67021
p2	1.1123	4.7803e-04	1.19985	8.2047e-09	2.1153e-04	0.039756
p3	1.1123	4.3412	5.4535	8.2047e-09	10.2714	10.2724
p4	1.1123	4.5194	5.6317	8.2047e-09	2.8459	2.84691
p5	1.1123	5.4848	6.5971	8.2047e-09	6.3479	6.34891
p6	1.1123	13.3527	14.465	8.2047e-09	0.7030	0.704013
p7	1.1123	1.8960	3.0083	8.2047e-09	0.0116	0.012613
p8	1.1123	3.4375	4.5498	8.2047e-09	0.0138	0.014813
p9	1.1123	4.1542	5.2665	8.2047e-09	14.3361	14.3371
p10	1.1123	3.8261	4.9384	8.2047e-09	3.7911	3.79211
p11	1.1123	0.4480	1.5603	8.2047e-09	0.0635	0.064513
p12	1.1123	3.8193	4.9316	8.2047e-09	0.7816	0.782613
p13	1.1123	11.9934	13.1057	8.2047e-09	3.0884	3.08941
p14	1.1123	5.4312	6.5435	8.2047e-09	0.4179	0.418913
p15	1.1123	4.6130	5.7253	8.2047e-09	15.8611	15.8621

Table 1: Mixed-Modeled Parameters

Finally, after running each optimal design algorithm with the individual curve-fitted parameter values, the optimal design algorithms were run with the parameters given in Table 1. All told, the optimal design algorithms were run 6 times total: 3 with the individual curve fits and 3 with the mixed modeling parameters.

RESULTS

The first series of numerical experiments in the project was done using the Standard

Error (SE) optimal design technique on the data collected for patients 1-15 (Muller et al.,

2008), excluding patients 10 and 13 because of inconsistent or incomplete data. This

series of data collection times provide a general basis that future results would be

compared to and also give an idea of which patients will be difficult or easy to model

data collection. An example of what ideal data collected can be seen for Patient 14

(Figure 1).
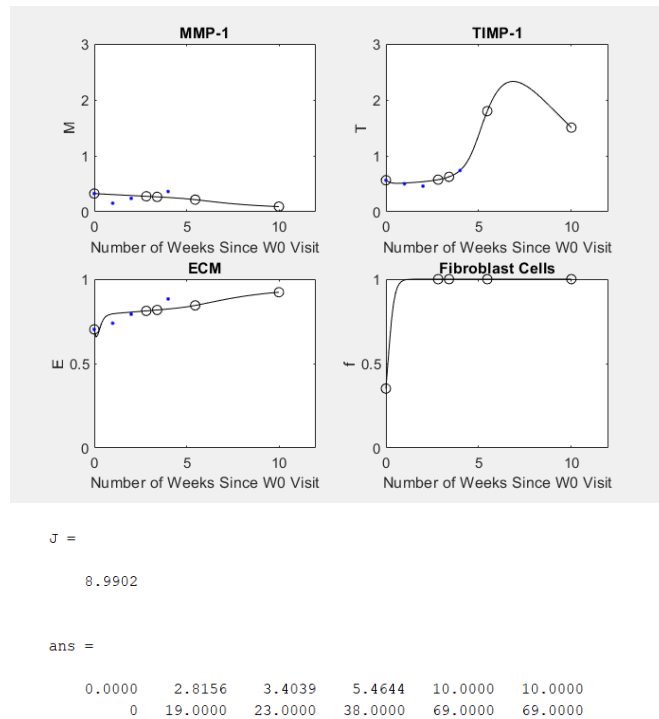
Figure 1: Data Collected for Patient 14 Using SE Optimization on Original Parameter Guesses

The four graphs in Figure 1 represent the four state variables in the mathematical

model. The graphs offer a visual comparison of the original data (small data points) and

the mathematical model (solid line). In addition to the graphs, there is a series of data that

represents the results using the optimization technique (circles). Below the graphs is the
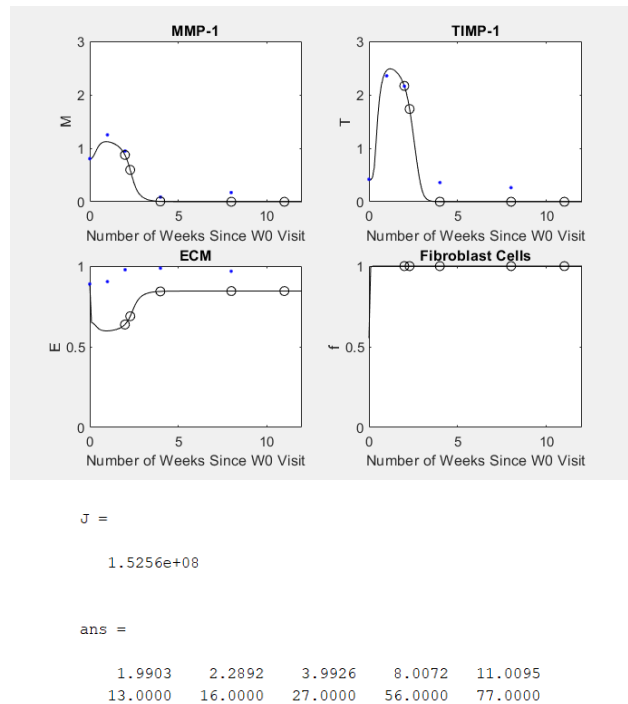
J value, which represents the current value of the Standard Error associated with the

model. When using SE optimization, this value is minimized and the final value is

displayed along with the graphs. In the case of Patient 14, the final *J* value is 8.9902.

Finally, there are two rows of data points that represent the optimal time measurements

for data collection that would have resulted in the best model of the original data. The

first row of data displayed represents these optimal data collection times in week and the

second row are those same collection times in days. The input data includes an initial

guess for the optimal time points of 6 data points that are evenly spread out from week 0

to week 12 and then these points are refined to the final values. For patient 14, it actually

turns out that only 5 times for optimal data collection are needed. There were two groups

that data collected fell into once it was collected. When the parameter values used

represented the data well, there were overall positive results using the optimal design

algorithm. This was the case for patients 2, 3, 4, 5, 6, 7, 11, 12, 14, and 15, but not for 1,

8, and 9.

Next, the mixed modeling parameters (Table 1) were introduced into the optimal-

design program and then the above process was then repeated to observe any changes in

the modeling accuracy and the optimal design data results. This series of computational

runs began to show a problem. While there were certain patients that showed decent

modeling results with the introduction of these new parameters (*i.e*, Patients 4, 14, 15),

most of the results illustrated issues with the mixed modeling parameters. The main issue

that appeared to be distorting the data was the Fisher Information Matrix (FIM)

approaching or becoming singular. Because the program works by minimizing a value

associated with the inverse of the FIM this raised a consistency issue. The program would begin as intended and then somewhere along the line it would fail.

The algorithm was then rerun using D-optimal and E-optimal design for both sets of parameters. Once again, this series of results fell into two groups. When using D- and E-optimal design, these two techniques produced similar results. For both optimal-design techniques patients 1, 2, 4, 5, 6, 7, 12, 14, and 15 displayed accurate fits (like Figure 3), while patients 3, 8, 9, and 11 did not. However, the high J-value for almost all patients (Table 3) using D-optimal and E-optimal design suggests all cases had difficulty with the FIM being singular or near singular. Yet, even with high J values, the algorithm still seems to produce reasonable results, such as having two time measurements between the extreme and inflection points of the MMP-1, TIMP-1, and ECM graphs (Figure 2).



```
J =

   1.5256e+08


ans =

    1.9903    2.2892    3.9926    8.0072   11.0095
   13.0000   16.0000   27.0000   56.0000   77.0000
```

Figure 2: Data Collected for Patient 3 Using E Optimization on Original Parameter Guesses

	SE-Optimal Design, Individual Parameters	SE-Optimal Design, Mixed-Model Parameters	D-Optimal Design, Individual Parameters	D-Optimal Design, Mixed-Model Parameters	E-Optimal Design, Individual Parameters	E-Optimal Design, Mixed-Model Parameters
Patient 1	3.954	0.1127	-2.9949 E 73	1.4816 E 45	7.6152 E 13	1.0016 E 11
Patient 2	66.8835	93.1693	2.1480 E 69	7.6401 E 57	7.4452 E 14	4.8460 E 19
Patient 3	4.821 E -04	2.9442	6.0800 E -20	1.2684 E 70	7.5256 E 08	1.5758 E 14
Patient 4	-2.6744 E 09	0.4150	-3.5005 E 33	-4.6888 E 69	3.9520 E 12	1.4439 E 15
Patient 5	1.1963	5.5534 E 03	1.3386 E 05	1.6645 E 41	1.3005 E 08	7.0634 E 11
Patient 6	192.5474	288.8716	-3.9193 E 67	1.6315 E 59	9.1881 E 24	1.9220 E 25
Patient 7	2.3674 E 03	7.0582 E 20	2.7343 E 42	7.2603 E 74	1.1613 E 12	3.9419 E 14
Patient 8	3.1383 E 09	-1.8754 E 21	-2.421 E 101	-1.6763 E 132	7.7328 E 26	4.9932 E 27
Patient 9	-3.4565 E 19	317.4927	-6.6248 E 65	1.8611 E 73	4.0195 E 25	3.4141 E 17
Patient 11	-3.4866 E 20	-1.4192 E 24	-1.9727 E 70	2.2910 E 78	9.8501 E 20	2.8563 E 21
Patient 12	-3.7106 E 19	592.6011	4.7766 E 66	-1.0669 E 75	2.5829 E 13	4.2474 E 27
Patient 14	8.9902	495.7772	3.1674 E 25	2.2224 E 57	8.4847 E 14	2.5079 E 22
Patient 15	13.7374	2.5742 E 04	2.8610 E 25	5.1925 E 33	7.4692 E 07	2.9750 E 10

Table 2: J-min Values for Each Patient

When comparing SE-optimal design, in which the *J*-min value is the final Standard Error value derived from the FIM, individual-parameter to mixed-modeling parameter results, the *J*-min value is fair-to-excellent for Patients 1, 2, 3, 5, 6, 14, and 15 for both sets of parameters. Only Patient 7 has a fair *J* value for the individual parameters, but not with the mixed-modeling parameters, while 9 and 11 are the opposite. Patient 4 has an excellent *J*-value for its mixed-modeling parameters, but not with regard to individual parameters. Patients 8 and 11 have poor *J* values with either technique. In all cases, the SE-optimal design performed better than D- and E-optimal design, which is consistent with another study (Banks *et al.,* 2011).
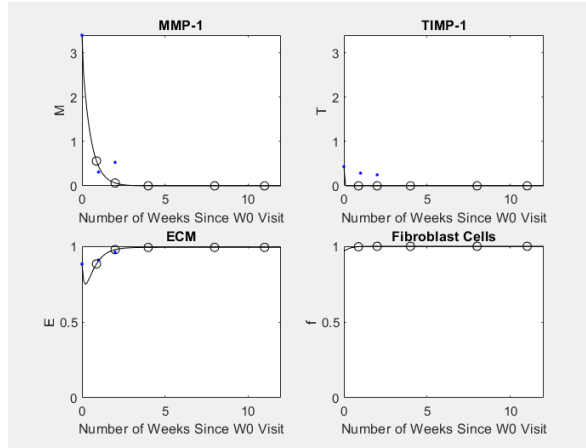
```
J =

   -4.6888e+69


ans =

       0.8756     1.9981     3.9944     7.9998    11.0100
       6.0000    13.0000    27.0000    55.0000    77.0000
```

Figure 3: Data Collected for Patient 4 Using D Optimization on Mixed-Modeling Parameters

The high *J*-values represent a problem that has occurred across all three optimal design techniques and seems to be an issue for both the individual parameters and the mixed-modeling parameters as the FIM becomes singular. Because this issue occurs after several iterations of the algorithms, it seems likely that there is a solution to this problem which would allow for the collection of relevant data that is not skewed by the issues currently present.

CONCLUSIONS

The first major conclusion that can be derived from the work is that SE-optimal design is an excellent tool for the project, especially when compared to the other optimal design techniques that were utilized in the project. The best indication of this are the *J*-

17

values obtained from using SE-optimal design are much more consistent and reasonable than the values obtained from the other techniques. In addition, the graphs that were obtained with these values also tell the same story. For each set of simulations, there were groups of patients that were identified as good and poor fits by looking at the graphs that represent each of the 4 state variables and in each case the SE-optimal design performed as well or better than the other two techniques. All in all, this supports previous work that had been done in regards to developing and testing SE-optimal design (Banks *et al.*, 2011). In this study, it was decided that while SE-optimal design is not always better than more traditional techniques, like D-optimal and E-optimal, it was at the very least on par and can sometimes lead to better results. Our results support the previous conclusions. Therefore, any future work will continue with SE-optimal design as the primary optimal design technique used.

To potentially improve our results, we will work to test and improve the estimates for mixed-modeling parameters. If the parameters that are being fed into the optimal design process are poor estimates, this can lead to poor results when using the optimal design algorithms. To do this, future work will be done exploring MONOLIX, as a possible mixed-modeling alternative to using MATLAB. It is hoped that this focus can further improve the results for both optimal design and the mixed modeling areas of the project.

FORMULATING A NEW MODEL AND MONOLIX TRANSITION

METHODS

MONOLIX is an advanced solver for non-linear mixed effects modeling (MONOLIX, 2016). It has been recommended by various experts in the field whom we have talked to as a software tool to analyze our model. MONOLIX is an advanced solver for non-linear mixed effects modeling. To begin with, we decided to work with a simpler model to test MONOLIX and see its effectiveness. From there, we would continue to increase the model complexity until either implementing MONOLIX using our model or the academic year ended.

First, we decided that the modeling of logistical growth would be a reasonable starting point. The logistical growth equation is easy to manipulate, has a well-known solution, and its numerical solution generally resembles the behavior that we expect to see in the fibroblast equation in our model. We started by creating a pseudo-random data set from the logistic growth model and take that data and try to fit it to confirm that we had a data set that would match the original logistic model. If we were able to do this for a simple version of the function as well as a more complex system of equations we would be able to easily move this data set, as well as our custom model, over to MONOLIX.

$$\frac{dY}{dt} = x * Y(t) * (1 - \frac{Y(t)}{w})$$

(6)

Equation 6: Basic Logistic Growth Model

Equation (6) is the logistic growth differential equation, where *x* represents the growth

rate, and *w* represents the carrying capacity. To create a data set, we took this model and

produced two data sets, which can represent two patients, by adding random "noise" for

integer time points between 0 and 5, inclusive. One data set is generated with a growth

rate of 1.5 and a carrying capacity of 150, while the other with a growth rate of 1 and a

carrying capacity of 100. Table 3 shows these two data sets that were created using the

single ODE with two sets of parameter values.

	0	1	2	3	4	5
Solution Data for Patient 1	10	36.3740	88.3903	129.8110	144.9692	148.8474
Pseudo-Random Data for Patient 1	10	43.0890	76.3154	136.9834	161.2716	153.7364
Solution Data for Patient 2	5	12.5161	28.0005	51.3887	74.1841	88.6508
Pseudo-Random Data for Patient 2	5	22.8630	35.2693	48.3543	77.1228	80.7780

Table 3: Data for Patients 1 and 2 Created Using Logistic ODE

To test the data, the data was curve fit using MATLAB's *fminseearch*, a local

optimization routine, minimizing the sum of squares of the error. When running this

algorithm there are certain values that must be estimated and known. First, the initial

conditions for each patient are input into the algorithm. Then, the algorithm requires an

initial guess estimated for the parameter values. This allows the algorithm to differentiate

potential conflicts between local and global minimum. However, a variety of initial

guess estimates were tested and similar solutions were always obtained. While accurate

initial guesses are likely necessary for a more complex system, for these data sets they

were not. Overall, this supports the data being an accurate representation of the solutions

to Equation (6) for each patient as the returned parameter values were always close to the

parameters that were used to create the data.

Next, we used the data with a mixed-effects model. This is to confirm that we

have data to use with a mixed-effects model before using. To make this change, we need

six parameters with initial estimates for these parameters. The mixed-effects model is

given in Equations (7) and (8), where Y_1 is for Patient 1 and Y_2 is for Patient 2.

$$\frac{dY_1}{dt} = (k_1 + k_5) * Y_1(t) * (1 - \frac{Y_1(t)}{(k_2 + k_6)}) \quad\quad (7)$$

$$\frac{dY_2}{dt} = (k_3 + k_5) * Y_2(t) * (1 - \frac{Y_2(t)}{(k_4 + k_6)}) \quad\quad (8)$$

Equations 7-8: Models for Patients 1 and 2

To represent as a mixed-effects model, the carrying capacity and growth rates

have been split into the sum of two parameters giving two new parameters that are then

shared between the two patients. Here, k_1-k_4 are still unique to each patient and

represent the random variables, while k_5 and k_6 are shared by each and are thus the fixed

variables for the two patients.

We, then, decided to do a quick numerical study to see the effects of the data

becoming sparse. This was done by only using subsets of the data and then re-curve fit to

see if there was any change in the overall fit of the model. The original data set that used

integer time values from day 0 to day 5 with day 0 being the initial condition. We

comparted this set with data from days 0, 1, 3, and 5 as well as the data from days 0, 2,

and 4.

Once the data had been properly modeled using one simple ordinary differential

equation (ODE) for each patient, the next step was to create a system of equations for

each patient. This adds complexity to the model and is closer to what our model is for

diabetic ulcer healing. If we are able to correctly model a system of equations for each

patient we should be able to port that system of equations over to MONOLIX as well as

the full diabetic healing model. The system of equations that was decided to work on for

each patient is in Equations (9) and (10).

$$\frac{dY}{dt} = x * Y(t) * (1 - \frac{Y(t)}{w}) \tag{9}$$

$$\frac{dG}{dt} = z * Y(t) * G(t) \tag{10}$$

Equations 9-10: System of Equations Logistic Growth Model

This system of equations is more complex than the previous model as it

introduces another parameter z, which in this case is an interaction rate that, introduces a

relationship between the first ODE and the second. The second function being a

decaying interaction function is quite important as it closely resembles the biological

interaction relationships that are expected in Equations (1) – (4). In that model,

biological enzymes and their inhibitors interact and compete with one another and this

relationship is best replicated with a decaying interaction ODE as seen in our system. As

before, two sets of parameters were introduced to simulate two separate patients. The

first patient has a growth rate of 1.5, a carrying capacity of 150, and an interaction rate of

.002. The second has a growth rate of 1.0, a carrying capacity of 100, and an interaction

rate of .003. To create a new data set for each of the patient's data these parameters were

entered into the above ODEs and then solved. Time points where once again intervals of

1 starting at 0 and ending at 5. This data was collected, and then random noise was once again added to each data point. While the first data set for each patient is essentially the same as before the resulting data from the second ODE for each patient can be seen below.

	0	1	2	3	4	5
Solution Data for Patient 1	20	18.7860	15.6346	11.1873	7.3737	4.7390
Pseudo-Random Data for Patient 1	20	18.0399	17.2656	12.7851	7.6141	5.8815
Solution Data for Patient 2	40	39.0231	36.8080	32.7163	27.0588	21.1462
Pseudo-Random Data for Patient 2	40	38.6131	36.5597	35.6957	29.8769	23.9806

Table 4: Data for Patients 1 and 2 Created by Second ODE in System of Equations

The generated data was then used to refit equations (9) – (10). The same *fminsearch* MATLAB function that was used before was modified to model each of the patients' data sets separately and then solve for the individual parameters. This function produced two modeled graphs of the data fit, one for each ODE, as well as three individual parameters that corresponded to each patients' x, w, and z values from Equations (9) and (10). We again conclude that we can determine the model parameter values well from simulated data.

Next, we consider the system of ODEs using mixed-effects modeling. In this case, there were nine parameters between the two patients and the ODEs that were used to model each patient given in equations (11)-(14).

$$\frac{dY_1}{dt} = (k_1 + k_7) * Y_1(t) * (1 - \frac{Y_1(t)}{(k_2 + k_8)}) \tag{11}$$

$$\frac{dG_1}{dt} = (k_3 + k_9) * Y_1(t) * G_1(t) \tag{12}$$

$$\frac{dY_2}{dt} = (k_4 + k_7) * Y_2(t) * (1 - \frac{Y_2(t)}{(k_5 + k_8)}) \tag{13}$$

$$\frac{dG_2}{dt} = (k_6 + k_9) * Y_2(t) * G_2(t) \tag{14}$$

Equations 11-14: System of Equations Models for Patients 1 and 2

Here, Equations (11) and (12) are for Patient 1 and Equations (13) and (14) are for Patient 2. As before, each parameter has been split into the sum of two parameters, with one of those being shared with the other patient. Once again a variety of initial parameter guesses were used to see if being off on the parameters guess would produce bad results from fitting. And as before there were no issues found with trying a variety of parameters guesses; the only thing that it seemed to effect by the various inputs where how long the program would take to reach a solution. This shows that the parameters in the model are identifiable and able to be used to create a pseudo-random data set, and that this data set is able to be modeled back to estimate the original parameter values. This means that the next step, the transition to MONOLIX, is ready.

When looking at moving this data over to MONOLIX the first place to start is to use the model library that comes preloaded with the software, specifically the PKPD model library. PKPD stands for pharmacokinetic/pharmacodynamic modeling, which is typically used for modeling the intensity of medical pharmaceutical and its relationship to dosing times and regiment. The models present in the PKPD library cover a variety of dosing timings, models, and data collection methods. The issue that became apparent with this was that none of the preloaded models worked in the way that we were hoping

and we did not figure out how to use MONOLIX for the above work at the time of this writing. While the originally planned stopping spot for this section of the project was to complete the transfer of the model to MONOLIX, this was where the work for this project ended.

RESULTS

First, we generated pseudo-randomized data for each of our imaginary patients. We solved Equation (6) using pre-selected parameter values at integer time values. Then we added pseudo-random noise to the numerical solutions at the integer time values to generate our data. Each of the below results was also replicated using a variety of initial parameter guesses as well as with multiple replications for each step.

Next, we used this random data to see how well we can re-curve fit both data sets in MATLAB using the built-in function, *fminsearch* (MATLAB, 2019) minimizing the sum of squared differences between the model and the data. Patient 1's data was created using parameter values of 1.5 and 150, while Patient 2's data was created using parameter values of 1 and 100. Figures 4 and 5 show that that the data was well fit using the MATLAB routine. In addition to these graphs, the values in Table 5 also show that the routine is able to identify and estimate the original parameter values that were used to create the data.
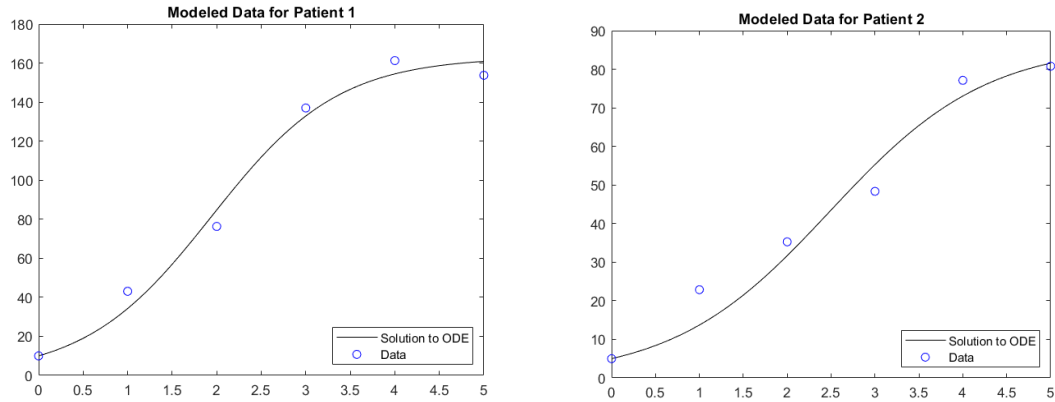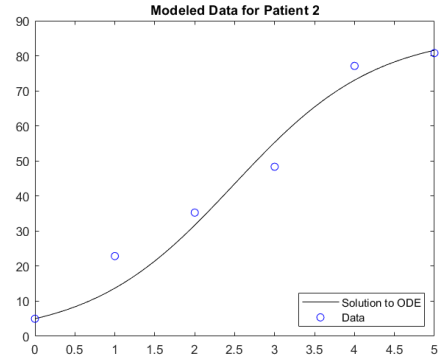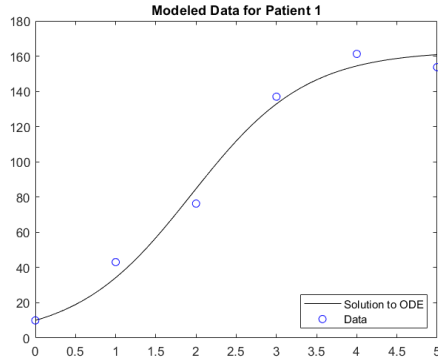
Figures 4-5: Modeled Data for Patients 1 and 2

	Growth Rate	Carrying Capacity
Parameters Used to Create Patient 1 Data	1.5	150
Parameter Values Returned for Patient 1	1.4035	163.0693
Parameters Used to Create Patient 2 Data	1.0	100
Parameter Values Returned for Patient 2	1.1215	86.4586

Table 5: Parameter Values for Patients 1 and 2

Next, the program was modified for mixed-modeling parameters. To maintain consistency and to be able to accurately compare results the same data was once again used. Figures 6 and 7 as well as the data in Table show that we were able to accurately able to model the data and estimate the original parameter values.
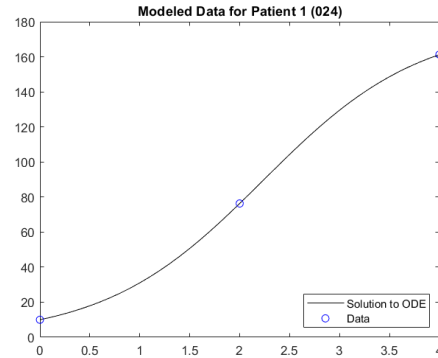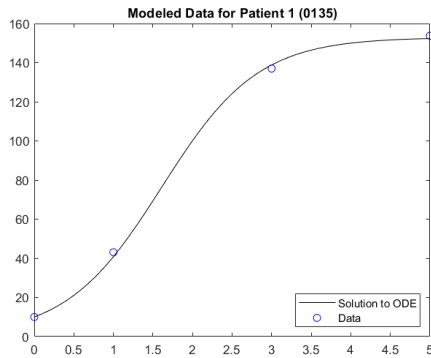
Figures 6-7: Mixed-Modeling Data for Patients 1 and 2

	Growth Rate (Fixed)	Growth Rate (Random)	Carrying Capacity (Fixed)	Carrying Capacity (Random)
Parameter Values Returned for Patient 1	1.3397	0.0638	76.1868	86.8826
Parameter Values Returned for Patient 2	1.3397	-0.2182	76.1868	10.2718

Table 6: Mixed-Modeling Parameter Values for Patients 1 and 2

Before moving on to the system of equations, the numerical study with fewer data points brought some interesting results. This study was done to test the effectiveness of our current technique when looking at data sets that were incomplete or having less and less data.

Figures 8-9: Mixed-Modeling Data for Patient 1 With Less Data

Figure 8 shows the model using mixed-modeling parameters with data from days 0, 1, 3, and 5, while Figure 9 shows the model using mixed-modeling parameters with data from days 0, 2, and 4. Both figures represent data collected for Patient 1, however the data collected for Patient 2 mirrored the above graphs in both instances. When comparing these graphs to those generated when using the individual parameters, it can be seen that these are slightly better. While the data set is quite small, and this is not surprising as having more parameters allows for a better fit. In addition, the parameter values that are returned in both of these instances shows that mixed-modeling parameters allows for accurate finding of the original parameter values even with less data points (Table 7).

	Growth Rate (Fixed)	Growth Rate (Random)	Carrying Capacity (Fixed)	Carrying Capacity (Random)
Parameter Values Returned for Patient 1 (0135)	1.5084	0.1417	70.0567	82.8408
Parameter Values Returned for Patient 2 (0135)	1.5084	-0.4726	70.0567	17.5804
Parameter Values Returned for Patient 1 (024)	1.1265	0.1399	91.6408	86.7694
Parameter Values Returned for Patient 2 (024)	1.1265	0.0794	91.6408	-4.3134

**Table 7: Mixed-Modeling Parameter Values with Less Data**

Next, we formulate a system of equations, which adds a level of complexity and more closely aligns to model Equations (1) – (4), which is our system of differential

equations. Rather than modify the first equation, we simply developed a second equation (Equation (10)), which couples the two equations.

Mirroring the previous work, the next step was to take these data sets and then to see if they could be modeled to identify the original parameters that were used to create the data. The previous MATLAB program that was used with the previous data is once again adapted to model the new system of equations. Because the new model is made up of two ODEs, the second of which is actually dependent on the solution of the first, this is a bit more complex. However, the same concept of modeling the data, minimizing the sum squared difference, and then outputting the model alongside the original data still stands. The data that was generated for this step can be seen above in Methods.

Next, we re-curve fit all four equations. Again, there is excellent agreement between the model and generated data (Figures 10-11). The data for Patient 1 was created using parameter values of $k_1 = 1.5$, $k_2 = 150$, and $k_3 = -0.002$, while the data for Patient 2 was created using parameter values of $k_1 = 1$, $k_2 = 100$, and $k_3 = -0.003$. And since the program returned values of 1.4034, 163.0785, -0.0026, 1.12, 86.5608, and -.0022, we can reasonably identify the parameters.
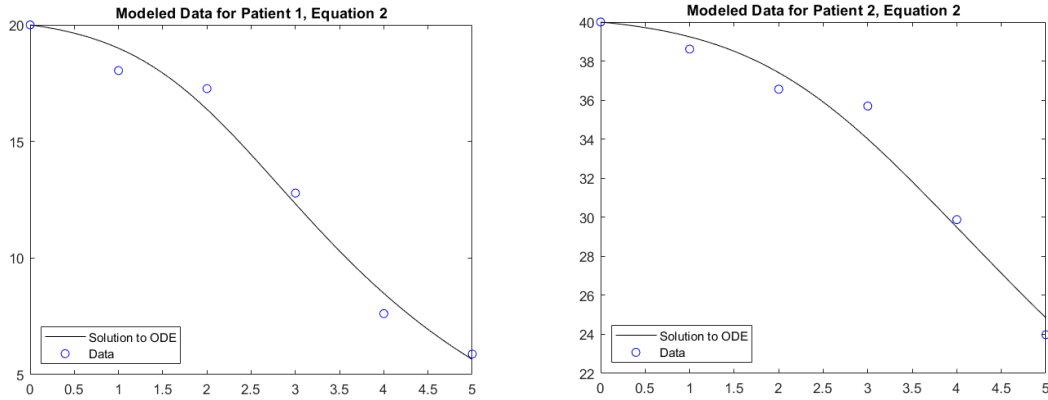
Figures 10-11: Data Fits for ODE 2 in System of Equations with Individual Parameters

The final addition of complexity to the model before the transition to MONOLIX was to introduce mixed modeled parameters into this system of equations. Once the necessary changes to the code were completed, the program was rerun and produced graphs and values that were remarkably similar to the above graphs.

	Growth Rate (Fixed)	Growth Rate (Random)	Carrying Capacity (Fixed)	Carrying Capacity (Random)	Interaction Rate (Fixed)	Interaction Rate (Random)
Parameter Values Returned for Patient 1	1.2641	0.1394	113.7930	49.2855	0.0024	0.0002
Parameter Values Returned for Patient 2	1.2641	-0.1440	113.7930	-27.2321	0.0024	-0.0002

Table 8: Mixed-Modeling Parameter Values for Patients 1 and 2, System of Equations

With the mixed modeling parameters, the final results for parameter values would nearly always be close to the initial guess, but if the fixed and random values were combined into a single parameter, that value was consistent across all of the initial guesses. Specifically, k_1 - k_6 are the random effects that are unique to each patient, and

k_7 - k_9 are the fixed effects that are shared between the two patients. While the graphs

obtained from modeling the data using the mixed-modeling parameters is nearly identical

to Figure (10) and (11), the parameter values returned, given in Table 8 show us that the

program was accurately able to estimate the parameters used to create the data. The

ability to provide clear and consistent answers across multiple data sets with a variety of

initial guesses tells us that the program is working properly to identify and accurately

predict the parameter values of the data that it is given.

Our results are consistent because the randomized data was reproducible, each set

was representative of the model, and each of the models were done with various initial

parameter guesses. The next step in the project was planned to be to now move this

model over to MONOLIX as a precursor to the diabetic healing model, but we have been

unable to replicate these types of results in MONOLIX at the time of this defense. If we

obtained reliable results using equations (1) – (4) in MONOLIX, we would then re-run

the SE-optimal design algorithm with the mixed-modeling parameters.

CONCLUSIONS

While the major goal, transitioning the new model into MONOLIX, of this

section of the project was unsuccessful because of technical, as well as situational, issues

there are still quite a few of results from the work done that allow for some general

conclusions. First, the model was able to be accurately solved and modeled using the

tools available in MATLAB. In addition is was shown that an interconnected system of

ordinary differential equations was able to be modeled and analyzed properly, which

points to the long-term viability of the full model. Finally, the small look at using various sizes of data sets was interesting as it showed that as less data was available the better mixed modeled parameters performed. Overall, this part of the project will serve as an excellent starting point for future work on an introduction of using tools like MONOLIX.

The next step for this project is to continue to work on integrating the current model into the MONOLIX software. With MONLIX's ability to identify and work easily with mixed-modeling parameters new and better estimates for the values may be found which can in turn improve the optimal design work with Equations (1) – (4). While this step has not been completed at the time of this defense, the work to accomplish this move has continued.

REFERENCES

Alotaibi, J. (2018). "Analysis of Wound Healing Process Using Parameter Estimation Techniques and Mixed Effects Modeling." *Masters Theses & Special Projects*.

Baayen, R.C., Davidson, D.J., Bates, D.M. (2008). "Mixed-effects modeling with crossed random effects for subjects and items." *Journal of Memory and Language*. 59.4.2008, 390-412. Retrieved from: https://doi.org/10.1016/j.jml.2007.12.005.

Bagiella, E., Sloan, R., & Heitjan, D. (2000). "Mixed-effects models in psychophysiology". *Psychophysiology, 37*(1), 13-20.

Banks, H.T., Holm, K., Kappel, F. (2011). "Comparison of Optimal Design Methods in Inverse Problems." *Inverse Probl*. 2011 Jul 1; 27(7): 075002. doi: 10.1088/0266-5611/27/7/075002

Comets, E., Lavenu, A., & Lavielle, M. (2017). Parameter estimation in nonlinear mixed effect models using saemix, an R implementation of the SAEM algorithm. *Journal of Statistical Software*, 80(3), 1 - 41.

Falagna, V. (2005), "Wound Healing and Its Impairment in the Diabetic Foot." *The Lancet*. 366.9498, 1736-743. Retrieved from: https://www.sciencedirect.com/science/article/pii/S0140673605677008

French, R. (2017), "Using Mixed Effects Modeling to Quantify Difference Between Patient Groups with Diabetic Foot Ulcers." *Honors College Capstone Experience/Thesis Projects*. Paper 723. https://digitalcommons.wku.edu/stu_hon_theses/723

Frieden, B. R. (2004) *Science from Fisher Information: A Unification*. Cambridge Univ. Press. ISBN 0-521-00911-1

Karimli, N. (2019). "Parameter Estimation and Optimal Design Techniques to Analyze a Mathematical Modeli in Wound Healing." *Masters Theses & Special Projects*. Paper 3114. http://digitalcommons.wku.edu/theses/3114

MATLAB and Global Optimization Toolbox Release R2019a, The MathWorks, Inc., Natick, Massachusetts, United States.

The MONOLIX software supported by the MONOLIX Group (2016). https://www.monolix.org/

Muller, M., Trocme, C., Lardy, B., Morel, F., Halimi, S., & Benhamou, P. Y. (2008). "Matrix metalloproteinases and diabetic foot ulcers: the ratio of MMP-1 to TIMP-1 is a predictor of wound healing." *Diabet Med*, 25(4), 419-426.

Sheehan P., Jones, P., Caselli, A., Giurini, J.M., Veves, A. (2003) "Percent change in wound area of diabetic foot ulcers over a 4-week period is a robust predictor of complete healing in a 12-week prospective trial." *Diabetes Care*, 26(6), 1879-1882.

Prassad, A. (2017). *SE-optimal design MATLAB code*. Center for Undergraduate Research in Mathematics Final Report.

Velnar, T., Bailey, T., & Smrkolj,V. (2009). The wound healing process: an overview of the cellular and molecular mechanisms. Journal of International Medical Research, 37(5). Retrieved from http://journals.sagepub.com/doi/pdf/10.1177/147323000903700531