Cell Proliferation Kinetics for *In Vitro* and *In Vivo* Models

Andrea Mazzocchi

May 1, 2018

A. Introduction

Cancer can broadly be defined as a disease in which cells become mutated and are given the ability to proliferate uncontrollably. These mutations can vary within each type of cancer and lead to very different cancer growth, malignancy, and response to drug. The complexity of cancer is vast and researchers are regularly trying to determine pathways and mechanisms related to behavior. However, cancer is both a biologist's and clinician's focus. With over 1.6 million new cases of cancer each year, diagnostics and predictive disease modeling have become the center of translational medicine with the intentions of bringing cancer biologists' understanding to tools useful for clinicians[1]. One such avenue for translational medicine is through the use of numerical methods to predict the outcomes of treatment. These methods have not yet been heavily implemented into the space of cancer outside of modeling basic growth through agent-based models[2, 3]. Although such models are useful there is still a gap between computer-based models and clinical application. Here, I am **proposing** a simple ordinary differential equation-based kinetics model to predict growth, malignancy, and treatment response of hepatocellular carcinomas based on cancer related gene mutation and extracellular matrix. The **long-term goal** of this project is to create a predictive model in which known biological measures in controlled microenvironments with well characterized cellular mutations will be used. The model will then be representative of cell specific mutations. Once these models are created, we will be able to take tissues of known origin but unknown genotype and compare their growth patterns to our models. By linking the models to patient samples, we will be able to better predict patient response to drug treatment and give the best treatments first. It is **hypothesized** that the use of numerical methods will allow us to create a model from which we can predict the treatment outcomes of cancer cells based on change in cellular populations untreated and treated conditions. The **sub-hypothesis** is that *in vitro* tissue engineered models will yield quantitative model equal to *in vivo* models which will validate the *in vitro* models and further will be able to predict cancer related genetic mutations and best treatments for patient derived hepatocellular carcinoma samples. To **test this hypothesis**, a 3D tissue engineered cancer model will be employed to yield data to create an appropriate model utilizing ordinary differential equations for prediction of growth, malignancy, and treatment response.

B. Numerical Problem & Significance

The **numerical problem** presented is that we are currently unable to predict through mathematical models how cells cancer cells will behave which has limited our ability to bring best treatments to patients. The proposed work is **significant** in that it will allow for quantitative models to be used to determine growth, malignancy, and treatment outcomes of cancer cells of unknown mutational background.

C. Biological Background

Cancer and Hepatocellular Carcinoma.

Cancer is an interesting disease in that the cancerous cells composing solid tumors have many abilities. Historically known as the hallmarks of cancer, they have the ability to sustain proliferative signals, evade growth suppressors, activation invasion, enable replicative immortality, induce angiogenesis, and resist cell death[4]. The methods in which cancer cells are able to carry out such hallmarks are directly related to their genetic mutations. Either hereditary or acquired over time, cellular mutations drive cell specific phenotype and behavior directly and significantly impacts drug response and efficacy. The cancer microenvironment is additionally very supportive of tumor behavior; cancer cells will recruit stromal cells to create more advantageous extracellular matrices and drive mutational changes to create drug resistance. The dynamic ability of cancer and its surrounding environment makes it a challenging disease to treat. Hepatocellular carcinoma (HCC) specifically is defined as

cancer originating within the liver from mutated hepatocytes (cells primarily responsible for carrying out liver function)[5]. Liver is of specific interests as liver cancer incidence and death rates continue to rise despite advances in diagnostics and treatment; it currently stands as the 6th most common cancer worldwide. With just an 18% 5-year survival rate, each year there are approximately 40,000 new cases of primary liver cancer in the US[1]. Additionally, advanced primary colorectal, pancreatic, breast, and lung cancer often spread to the liver leading to the need for resection and yield low survival rates[5]. Thus, this area of cancer research has become of interest as the liver microenvironment is advantageous to primary and secondary cancer types.

It is becoming well understood that cancer populations are not homogenous but rather widely heterogenous making the disease less predictable from patient to patient. Researchers have identified two broad classes of cancer cells play a substantial role in tumor behavior and may further be used to predict treatment outcomes[6]. The broadest and most commonly found cancer cell within a heterogenous population is the "proliferating" cancer cell. This type of cell is proliferating (dividing) uncontrollably as its proliferation pathway is no longer regulated. These cells spend little time in the G-phase of the cell cycle and can easily be caught in the M-phase of the cell cycle making them targets for many standard chemotherapies that attack DNA replication and cell proliferation. The second type of cancer cell within the heterogenous population is the "quiescent" cancer cell[7]. This cell has the same phenotype and genotype of the proliferating but is uniquely "stuck" in the G-phase of the cell cycle[7]. This cell is not proliferating but remains alive and protected from chemotherapies and treatments targeting proliferation. These cells can later become active again and proliferate – they have been directly related to cancer recurrence and metastases. It is important to understand both of these cell types as they each behave differently and must be appropriately considered within the mathematical models.

In Vitro Tissue Engineered Model

To create the quantitative model, tissue engineered constructs will be used to grow liver cancer cell lines with known, well-established mutational backgrounds. The cells will be grown in a 3D culture construct made of common liver specific extracellular matrices and grown over the course of 3 days to observe "normal" behavior which will act as controls in comparison to chemotherapy treatments. Chemotherapy treatments will also be administered over the course of three days. At 12-hour intervals, samples will be sacrificed and cells within them will be quantified and counted as either proliferating or quiescent. These studies will be done for a large number of liver cancer cell lines to collect a large amount of data with each condition and time point having a sample size of n=3 or greater. The data collected will be directly used to find coefficients within the kinetics model described below and further output cell proliferation ratios that can be used. It is intended that with enough data, the models will be refined enough that clearly defined cell proliferation ratios for "proliferating" and "quiescent" cells over 3 days can be established for each major cancer cell mutation. Patient cells or cells of unknown mutational background will then be used within the culture system and compared to the proliferation ratios to determine which mutation they have and further how they should respond to a wide variety of chemotherapy treatments. Collectively, many models will be produced in which we can compare patient data to and determine mutational background and treatment response. In addition to *in vitro* modeling, tissue constructs will also be surgically placed into mice (*in vivo*) and observed over time as controls and under chemotherapy treatment to determine if behavior is similar to that of the *in vitro* models.

D. Approach

Kinetics Model

Considering the two cancer cell types outlined above, it is important to consider the fates in which they can have to create kinetics models appropriate for predicting tumor growth, malignancy, and drug response. Many models have been developed considering cancer stem cells, differentiated cancer cells, and each of their roles in the heterogenous cell population[8]. Here, we are modifying a previously used kinetics model which considers proliferating cell, cancer stem cells, and terminally differentiated cells and looking at only quiescent cells and proliferating cells[8]. This modification is

made because it is believed that these two cell types play the most substantial role in solid tumor response to chemotherapy. For the kinetics models in general, proliferative cells have many fates, they can divide into two proliferating cells, divide into two quiescent cells, divide into one proliferating cell and one quiescent cell, or die. Quiescent cells have limited fate pathways and for our purposes they will be considered alive but not proliferating or dying. Related to each cell fate is a rate constant (k_1 through k_5) from which we can derive two rate equation dN_p/dt and dN_q/dt representing the rate of change in proliferating cells over time and the rate of change in quiescent cells over time (**Figure 1**). Within the kinetics models, N_p represents the number of proliferating cells and N_q represents the number of quiescent cells. As for rate constants, k_1 is doubling rate of proliferative cells, this value would vary for each cell mutation type and is known for each cell line (in literature). K_2 is rate at which a proliferating cell divides to produce a proliferating cell and a quiescent cell. K_3 is the rate at which a proliferating cell divides to produce two quiescent cells. Lastly k_4 and k_5 are the rates in which the proliferating and the quiescent cells are dying. It is important to understand that for each forward reaction which produces products (proliferating cells, quiescent cells, dead cells) loses the reactant as it no longer exists. This is indicated via subtracting the rates constants multiplied by the reactants in each of the overall rate equations (**Figure 1**). For this model and the purposes of the class, k_1 is considered to be known and k_2 and k_3 are assumed to be related k_1. This reduces the model to two unknowns for each time point, k_4 and k_5 which will be found at each time point and the average will be used for the ordinary differential equations model (ODE).

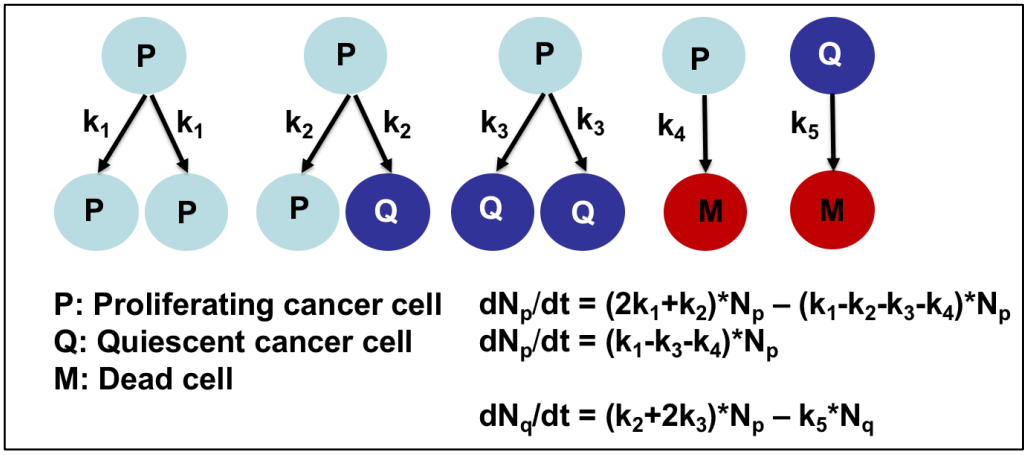

Figure 1, Kinetics Model for Proliferating and Quiescent Cancer Cells. Rates (k_1 through k_5) are shown with appropriate cellular fates. Arrows indicate the rate at which the top row of cells are able to produce the bottom row of cells. It is shown that proliferating cells have four fates whereas for the scope of this project quiescent cells have just one fate. Cumulative rate equations are given as dN_p/dt and dN_q/dt.

Ordinary Differential Equations

Ordinary differential equations are broadly defined as initial value problems that can be solved using numerical methods. Such equations require knowing the derivatives of the functions of interest as well as initial values of the function as well as the derivative. This method of modeling was specifically selected for the application at hand because it can appropriately utilize rate equation (kinetics equations) that are changing over time. In application, ODEs are commonly used to approximate solutions that can be challenging to numerically solve. For this application, the equations can be solved exactly, however, when we are unable are only looking at cells at two time points, we will not be able to solve the equations exactly and will rely on our well-developed ODE models to predict behavior. For this application specifically, a system of differential equations was being solved involving N_p and N_q. This modified and slightly more complex use of ODE45 involves evaluating two linear first order differential equations to subsequently model both of them. Both equations are continuous and accurate results are derived. Because two equations were to be used due to the nature of the problem, the use of ODEs seemed most appropriate. Such equations require only the derivatives and initial values but can also use a system of derivatives intertwined within each other to create reliable models.

MATLAB Implementation

MATLAB was used to implement the kinetic equations discussed above. In implementing MATLAB, the following approach was taken (**Appendix A**). First, determine the coefficients (k_4 and k_5) by simply solving the equations at each time point because dN_p/dt, dN_q/dt, N_p, and N_q were known at each time point. Then take an average of both k_4 and k_5 to use within the ODE system. Additionally, regression modeling was done using R-squared to determine how well the coefficients were representing the actual data. Although only one model is used here, in future practice there will be many k_4 and k_5 from which the average will be taken (many sets of data). ODE modeling is next done using the ODE45 script in MATLAB. This was selected over ODE23 because although it can require more work per each step, it is more accurate for data with smooth solutions. It is believed that the model should be a smooth slowly increasing equation and thus ODE45 would be more appropriate. The data being used in the model is simple and both ODE45 and ODE23 yield very similar results, looking forward however, it is more advantageous to use the more accurate model. Additionally, only a small amount of computing power is required and the complexities of ODE45 do not hinder processing. For this specific ODE model, two equations were used within the MATLAB code. These were implemented in a function (**Appendix B**) that was called by the ODE45 script within a separate MATLAB file (**Appendix A**). After the ODE analysis was performed, a comparison of the model with the actual data was done. When this model is implemented for prediction of behavior this type of comparison cannot be made but is extremely valuable when developing the model to ensure that it is appropriately representing the actual data rather than diverging. Lastly, usefully outputs were required to yield a quantitative model to be used for prediction of growth, malignancy, and drug treatment. The output metric determined to be most useful at this time was entitled “proliferation ratios”. The first ratio is that of the proliferating cells at time 72hrs over number of proliferating cells at time 0hrs and the second ratios is that of the quiescent cells at time 72hrs over number of quiescent cells at time 0hrs. After a large amount of data processing, we will have these values for many different cancer cell mutations and the be able to grow cells with unknown mutational background and map them to the models developed here. We will then be able to extrapolate their expected response to chemotherapy based on what we have been able to determine with our drug treated conditions.

F. Preliminary Results

MATLAB code was written and self-created (mock) data was inputted to test the ability of the code to produce the desired outputs. Using the code shown in **Appendices A and B** and using three different data sets (untreated/control cells, G-phase drug toxicity treatment, and M-phase drug toxicity treatment) the desired outputs were produced. Using a simple for loop and basic calculations, the k_4 and k_5 rate constants were calculated. Maximum error between the ODE45 model and the actual data was produced, as well as plots showing the actual data in comparison to what was produced by the ODE45 (**Figure 2**). The population ratios for each condition were also calculated using data from the ODE model rather than the actual data. Although mock data was used for each of the models, proof of concept was shown. Conclusively, utilizing numerical methods in MATLAB, cancer cell lines can be modeled and compared to treatment conditions to create a potentially useful predictive tool for future precision medicine applications.

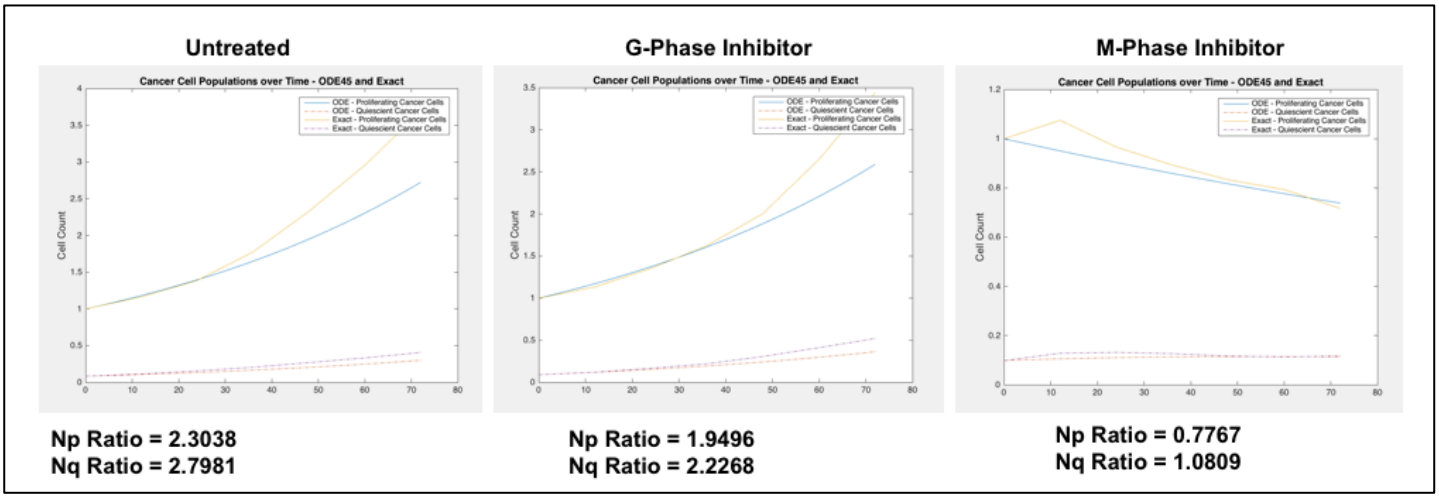

Figure 2, Preliminary data from MATLAB. Panel shows mock data for three conditions (L to R: untreated/control, G-phase inhibitor, M-phase inhibitor). Each shows ODE produced proliferating cell model in blue and quiescent cell model in red dotted lines and actual data show in yellow for proliferating cells and purple dotted lines for quiescent. Ratios were calculated using ODE model data (not actual data).

G. Numerical Method Limitations & Future Directions

This study is currently limited by the assumptions being made within the kinetics model and the ability of the ODE to predict behavior. Assumptions are being made for k_1 through k_3 although it would be ideal if only k_1 was known. The model is also being produced based on calculating the dN_p/dt, dN_q/dt, and knowing the cell counts at each time point. It would be more idea to require less information to create the model. These specific limitations can be addressed through the use of more robust numerical methods and an overall improved kinetics model that may be further simplified if possible. Going forward, it is intended that I collect more data to create predictive ODE models that are not utilizing just one set of data but rather many sets of data to act as a representation of cancer cell behavior for each mutation, rather than for a sample size of 1. Data collection will be carried out over the period of the next two years. Consideration of the extracellular matrix components will also be incorporated into the model which will add complexity to the data and will require modification of the ODEs used to predict behavior. This future direction will require modifications of both the tissue engineered model as well as the mathematical model.

H. Summary of Resources Used

Outside of the appropriately cited resources within the project document, the Burden and Faires Numerical Analysis textbook, class notes, and MathWorks website were each used. The Burden and Faires textbook was primarily used to better understand ODEs and how my work can appropriately intersect with the models they can produce. Class notes from this course as well as my previously taken numerical methods course (undergraduate course, 2014) were additionally used to help derive MATLAB code and implement kinetics models. The MathWorks website was heavily utilized for code troubleshooting and determine the best methods for extracting the data and models wanted for this specific application.

References

1. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2018.* CA Cancer J Clin, 2018. **68**(1): p. 7-30.
2. Barbolosi, D., et al., *Mathematical and numerical analysis for a model of growing metastatic tumors.* Math Biosci, 2009. **218**(1): p. 1-14.
3. Deasy, B.M., et al., *Modeling stem cell population growth: incorporating terms for proliferative heterogeneity.* Stem Cells, 2003. **21**(5): p. 536-45.
4. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
5. Sia, D., et al., *Liver Cancer Cell of Origin, Molecular Class, and Effects on Patient Prognosis.* Gastroenterology, 2017. **152**(4): p. 745-761.
6. Arnold, D., et al., *Prognostic and predictive value of primary tumour side in patients with RAS wild-type metastatic colorectal cancer treated with chemotherapy and EGFR directed antibodies in six randomized trials.* Ann Oncol, 2017. **28**(8): p. 1713-1729.
7. Sell, S. and H.L. Leffert, *Liver cancer stem cells.* J Clin Oncol, 2008. **26**(17): p. 2800-5.
8. Molina-Pena, R. and M.M. Alvarez, *A simple mathematical model based on the cancer stem cell hypothesis suggests kinetic commonalities in solid tumor growth.* PLoS One, 2012. **7**(2): p. e26233.

Appendix A
cancer_methods.m

```
% AR Mazzocchi
% Methods for Determining Cancer Ratios for Prediction
% Updated 23-April-2018

% No Treatment
%tspan = [0 72]; % hours
%np = [670 780 930 1190 1560 1980 2470]; % number of cells
%dnp = [0 110 150 260 370 420 490]; % rate of change
%nq = [56 77 105 138 180 223 272]; % number of cells
%dnq = [0 21 27 33 42 43 49]; % rate of change


% Treatment 1 - G phase inhibitor
%tspan = [0 72]; % hours
%np = [600 680 812 974 1203 1590 2064]; % number of cells
%dnp = [0 80 132 162 229 387 474]; % rate of change
%nq = [54 72 100 132 184 247 312]; % number of cells
%dnq = [0 18 28 32 52 63 65]; % rate of change

% Treatment 2 - M phase inhibitor
tspan = [0 72]; % hours
np = [600 645 580 536 500 476 430]; % number of cells
dnp = [0 45 -65 -44 -36 -24 -46]; % rate of change
nq = [59 77 79 76 70 68 70]; % number of cells
dnq = [0 18 2 -3 -6 -2 2]; % rate of change

%% Function
%cancer_methods(@(z)tspan,np,nq,dnp,dnq)
%function z = cancer_methods(tspan, np, nq, dnp, dnq)

%% Initial Conditions
tint = [tspan(1):12: tspan(2)]; % Works for this but needs adjusted
nq = [nq]/(np(1));
dnp = [dnp]/(np(1))/(tint(2));
dnq = [dnq]/(np(1))/(tint(2));
np = [np]/(np(1));
k1 = 0.125; % cells doubling per hour
k2 = 0.0075*k1;
k3 = 0.005*k1;
yzero = [np(1) nq(1)];

%% Calculating Constants and Determining Regression Error
i = 1;

for i = 1: length(np)
    k4(i) = ((k1 - k3)*np(i)-dnp(i))/(np(i));
    k5(i) = ((k2+2*k3)*np(i)-dnq(i))/(nq(i));
end

k4 = mean([k4]);
k5 = mean([k5]);

% Regression Error - k4
% Estimate
dnphat = (k1-k3-k4)*[np];
% Residuals
r = dnp-dnphat; % Actual minus when average k4 is used
% SSR
ssr = sum(r.^2);
```

```
% Coefficient of Determination
dnpbar = mean(dnp);
R2 = (1-ssr/sum((dnp-dnpbar).^2))*100;

% Plot
figure(1);
plot(np,dnp,'*')
hold on
plot(np,dnphat,'r')
title(['Linear Model, R2=' num2str(R2)])

% Regression Error - k5
% Estimate
dnqhat = (k2+2*k3)*[np]-k5*[nq];
% Residuals
r = dnq-dnqhat;
% SSR
ssr = sum(r.^2);
% Coefficient of Determination
dnqbar = mean(dnq);
R2 = (1-ssr/sum((dnq-dnqbar).^2))*100;

% plot

figure(2);
plot(nq,dnq,'*')
hold on
plot(np,dnqhat,'r')
title(['Linear Model, R2=' num2str(R2)])

%% ODE to Model Behavior

% ODE to Approximate Cancer Cell Behavior
[t,y]=ode45(@cancer_kinetics,tspan,yzero,[],k1,k2,k3,k4,k5);
npint = interp1(t,y(:,1),tint); % interpolate points for np at times matching exact
nqint = interp1(t,y(:,2),tint); % interpolates points for nq at times matching exact

% Error Calculations
error_np = np - npint(:,1)';
error_nq = nq - nqint(:,1)';
maxerr (1) = max(abs(error_np));
maxerr(2) = max(abs(error_nq));
maxerr

% Plot Cancer Cell Behavior from ODE45
figure(3);
plot(t,y(:,1),'-',t,y(:,2),'-.')
title('Cancer Cell Populations over Time - ODE45','FontSize',12)
ylabel('Cell Count','FontSize',12);
legend('Proliferating Cancer Cells','Quiescient Cancer Cells');

%% Exact Data vs ODE Model

% Plot Cancer Cell Behavior from Exact Data
figure(4);
plot([0:12:72],np,'-',tint,nq,'-.')
title('Cancer Cell Populations over Time - ACTUAL','FontSize',12)
ylabel('Cell Count','FontSize',12);
legend('Proliferating Cancer Cells','Quiescient Cancer Cells');

% Plot of Both ODE45 and Exact Data
figure(5);
plot(t,y(:,1),'-',t,y(:,2),'-.',[0:12:72],np,'-',[0:12:72],nq,'-.')
```

```
title('Cancer Cell Populations over Time - ODE45 and Exact','FontSize',12)
ylabel('Cell Count','FontSize',12);
legend('ODE - Proliferating Cancer Cells','ODE - Quiescient Cancer Cells','Exact -
Proliferating Cancer Cells','Exact - Quiescient Cancer Cells');

%% Outputs of Importance

% Consider ratios at day 1 (24 hrs), day 3 (72 hrs)
% From ODE:

np12 = interp1(t,y(:,1),tint(2));
nq12 = interp1(t,y(:,2),tint(2));

np72 = interp1(t,y(:,1),tint(7));
nq72 = interp1(t,y(:,2),tint(7));

np_ratio = np72/np12;
nq_ratio = nq72/nq12;

fprintf ('Np Ratio = %.4f\n', np_ratio);
fprintf ('Nq Ratio = %.4f\n', nq_ratio);
```

Appendix B
cancer_kinetics.m

```
% AR Mazzocchi
% Cancer Kinetics for ODE45
% Numerical Methods Project
% Updated 23-April-2018
% Notes:
% Cancer kinetics function that inputs k values from regression model into
% rate equations related to change in cancer cell populations


function dy = cancer_kinetics(t,y, k1,k2,k3,k4,k5)
% cancer_kinetics.m
% Contains equations for proliferating and quiescent cancer cells
% Variables
np = y(1);
nq = y(2);

% Equations
dy =[(k1-k3-k4)*np
    (k2+2*k3)*np-k5*nq];
```