Nonlinear Pharmacokinetic Models Based on Statistical Regression Methods Using Secondary Data

Ishmael Ortiz Stewart, Professor Juan Valera
Polytechnic University of Puerto Rico
Undergraduate Research Program for Honor and Outstanding Students (URP-HOS-2022-2023)
2022 - 2023

HSI STEM TITLE III
POLYTECHNIC UNIVERSITY OF PUERTO RICO
A MULTIFACETED APPROACH TO STUDENT CENTERED STEM EDUCATION
P031S210139

Abstract:

- This research is about creating nonlinear pharmacokinetic models to better understand the behaviors that take place. From there, see patterns, relations between behaviors, and statistical techniques and, as an end goal, not have to rely on living test subjects to understand what the behaviors will be.

Introduction:

- Pharmacokinetic models mathematically model the drug disposition in the body in a physiological compartment and predict specific tissue concentrations in biological systems [1]. A technique used in pharmaceutical research to predict the compartments of absorption, distribution, metabolism, and excretion [2].
- The problem is that there is a lack of integration of physiological knowledge and physicochemical processes which limits the effective creation of new models and is full of risk in studying with living subjects. Proper understanding of these processes will lead to overall better pharmacokinetic models.
- Through a virtual setting, we do the black box method (focuses on input and output data), and we apply secondary data from pharmacokinetic databases to make non-linear drug concentration vs. time models.

Objectives:

- Collect data through pharmacokinetic databases to obtain drug concentration vs. time and find the best parameters to configure complex models of pharmacokinetic phenomena and simulate them.
- Apply second-order polynomials and exponential-order regression methods (#1 & #2) to calculate the goodness of fit and identify trends.
  - Second Order Polynomial: $Y = b_0 + a_1 x + a_2 x^2$
  - Exponential-order regression: $y_t = A \,.\, e^{-B.t}$

$$y_t = A \,.\, e^{-B.t} + C \,.\, e^{-D.t}$$

- Compare models with estimations and see the errors.

Equipment and Materials:

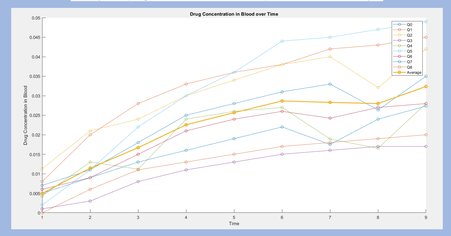
- Computer, MATLAB, and a pharmacokinetic data from a reliable pharmacokinetic database.

Methodology:

1. Find pharmacokinetic data from secondary databases. Drug concentration vs. time data.
2. Clean the data and eliminate unnecessary parameters.
3. Use imputation for missing variables (NANs). If necessary.
4. Through MATLAB, make a code to read and make drug concentration vs. time graphs by applying statistical techniques and calculating the goodness of fit (trendline). These will be the modules.
5. Validate the modules made with the parameters obtained and do re-estimations if needed.

Coding (Visualize data):

```
clear, clc

% data
time = [1, 2, 3, 4, 5, 6, 7, 8, 9];
Q0 = [0.005, 0.009, 0.013, 0.016, 0.019, 0.022, 0.0175, 0.024, 0.027346];
Q1 = [0.008, 0.02, 0.028, 0.033, 0.036, 0.038, 0.042, 0.045, 0.045];
Q2 = [0.011333, 0.021, 0.024, 0.03, 0.034, 0.038, 0.04, 0.032125, 0.042];
Q3 = [0.001, 0.003, 0.008, 0.011, 0.013, 0.015, 0.016, 0.017, 0.017];
Q4 = [0.0043, 0.013, 0.011372, 0.024, 0.026, 0.027, 0.018906, 0.018516, 0.028];
Q5 = [0.002, 0.011, 0.022, 0.03, 0.036, 0.044, 0.045, 0.047, 0.049];
Q6 = [0.006, 0.009, 0.015, 0.021, 0.024, 0.026, 0.02425, 0.027, 0.028];
Q7 = [0.007, 0.011, 0.018, 0.025, 0.028, 0.031, 0.033, 0.026359, 0.035];
Q8 = [0, 0.006, 0.011, 0.013, 0.015, 0.017, 0.018, 0.019, 0.02];
% Calculate the average concentration for each time point
Q_avg = mean([Q0; Q1; Q2; Q3; Q4; Q5; Q6; Q7; Q8]);
% Plot the individual cases and average concentration
figure;
hold on;
plot(time, Q0, '-o');
plot(time, Q1, '-o');
plot(time, Q2, '-o');
plot(time, Q3, '-o');
plot(time, Q4, '-o');
plot(time, Q5, '-o');
plot(time, Q6, '-o');
plot(time, Q7, '-o');
plot(time, Q8, '-o');
plot(time, Q_avg, '-o', 'LineWidth', 2);
hold off;
xlabel('Time');
ylabel('Drug Concentration in Blood');
legend('Q0','Q1','Q2','Q3','Q4','Q5','Q6','Q7','Q8','Average');
title('Drug Concentration in Blood over Time');
```



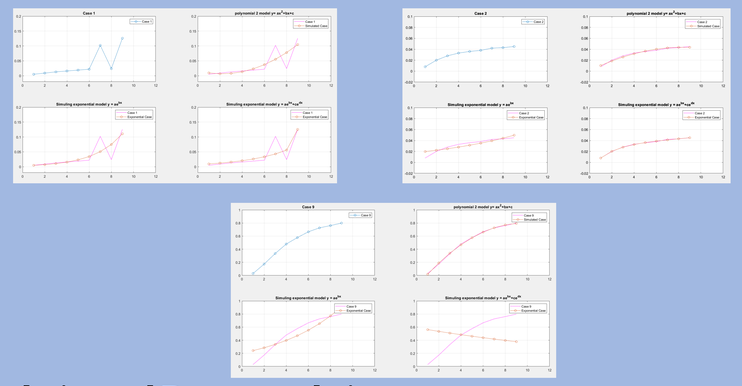
Coding (Simulate data separately):

```
clear, clc

%read excel data

T = readtable('data.xlsx')
Q0 = T.time;
Q0_mean = mean(Q0);
Q0_sd = std(Q0);

%Subplot Case 1 data
subplot(2,2,1)

plot(t,Q0, '-o');
xlim([0 12])
ylim([-0.02 0.2])
grid on
legend('Case 1')
title('Case 1')

% Simuling the polynomial second order model
[p,S]=polyfit(t,Q0,2)

Q01 = polyval(p,t)
subplot(2,2,2);

plot(t,Q0,'m', t,Q01,'--o')
format long
a=p(1)
b=p(2)
c=p(3)

xlim([0 12])
ylim([-0.02 0.2])
grid on
legend('Case 1','Simulated Case')
title('polynomial 2 model y= ax^2+bx+c')

Q01_mean = mean(Q01);
Q01_sd = std(Q01);
```

```
%percent error

%error_percent = ((Q01-Q0)/Q01)*100
error_percent = abs(Q01-Q0)./Q01*100

R2_1 = ((Q0 - Q0_mean)'*(Q01 - Q01_mean))/((length(Q0)-1)*Q0_sd*Q01_sd);
AdjR2_1 = 1 - (1 - R2_1)*(length(Q0) - 1)/(length(Q0) - 1 - 1);
Dev_Index1 = sqrt(sum((Q0 - Q01).^2)/length(Q0));
Chi_Index1 = sqrt(sum((Q0 - Q01).^2./abs(Q01))/length(Q0));

%Estimating parameters of exponential 1 model
model1 = fit(t,Q0,'exp1');
a = model1.a;
b = model1.b;

%Simuling the exponential 1 model
Q02 = a*exp(b*t);
subplot(2,2,3);

%plot(model1,'b',t,Q0,'m')
plot(t, Q0,'m',t,Q02,'-o')
xlim([0 12])
ylim([-0.02 0.2])
grid on
legend('Case 1', 'Exponential Case')
title('Simuling exponential model y = ae^{bx}')

Q02_mean = mean(Q02);
Q02_sd = std(Q02);

R2_2 = ((Q0 - Q0_mean)'*(Q02 - Q02_mean))/((length(Q0)-1)*Q0_sd*Q02_sd);
AdjR2_2 = 1 - (1 - R2_2)*(length(Q0) - 1)/(length(Q0) - 1 - 1);
Dev_Index2 = sqrt(sum((Q0 - Q02).^2)/length(Q0));
Chi_Index2 = sqrt(sum((Q0 - Q02).^2./abs(Q02))/length(Q0));

%Estimating parameters the exponential 2 model
model2=fit(t,Q0,'exp2');
c = model2.a;
b = model2.b;
a = model2.c;
d = model2.d;
```

```
%Simuling the exponential 2 model
Q03 = a*exp(b*t) + c*exp(d*t);
subplot(2,2,4);

%plot(model2,'b',t,Q0,'m')
plot(t, Q0,'m',t,Q03,'-o')
xlim([0 12])
ylim([-0.02 0.2])
grid on
legend('Case 1', 'Exponential Case')
title('Simuling exponential model y = ae^{bx}+ce^{dx}')

Q03_mean = mean(Q03);
Q03_sd = std(Q03);

R2_3 = ((Q0 - Q0_mean)'*(Q03 - Q03_mean))/((length(Q0)-1)*Q0_sd*Q03_sd);
AdjR2_3 = 1 - (1 - R2_3)*(length(Q0) - 1)/(length(Q0) - 1 - 1);
Dev_Index3 = sqrt(sum((Q0 - Q03).^2)/length(Q0));
Chi_Index3 = sqrt(sum((Q0 - Q03).^2./abs(Q03))/length(Q0));

RSqu_Index = [R2_1 R2_2 R2_3]
AdjR2_3_Index = [AdjR2_1 AdjR2_2 AdjR2_3] %Adjusted r2
Std_Index = [Dev_Index1 Dev_Index2 Dev_Index3] %Standard Error of Estimatation
Chi_Index = [Chi_Index1 Chi_Index2 Chi_Index3];

error_percent_M1 = mean(100*abs(Q0 - Q01)./(Q0));
error_percent_M2 = mean(100*abs(Q0 - Q02)./(Q0));
error_percent_M3 = mean(100*abs(Q0 - Q03)./(Q0));
fprintf('Average Percentange Error for Polynomial  Regression: %6.2f%%\n',error_percent_M1);
fprintf('Average Percentange Error for Exponential1 Regression: %6.2f%%\n',error_percent_M2);
fprintf('Average Percentange Error for Exponential2 Regression: %6.2f%%\n',error_percent_M3);

e1 = mape(Q01,Q0)
e2 = mape(Q02,Q0)
e3 = mape(Q03,Q0)
```

Analysis and Result (Cases 1, 2, and 9):



Conclusion and Recommendations:

- Certain statistical techniques work better in specific data behaviors, while others don't.
- Imputation has proven to be effective in estimating missing variables (NANs), but in some cases, it gives negative variables, which is not possible. New methods of estimation and calculation should be explored.

Future Work:

- The investigation is very broad. Narrowing the search to focus on specific physiological compartments (ex. Absorption) would be a good start. That way more considerations in physiological knowledge and physicochemical processes are taken and could lead to better modules developed.
- Investigate which statistical technique benefit more in certain behaviors and try to find relations between behaviors and techniques.
- An additional pharmacokinetic database should be included.

Acknowledgments:

Thanks to the Undergraduate Research Program for Honor and Outstanding Students (URP-HOS-2022-2023) and my mentor Professor Juan Valera for supporting this study.

References:

[1] S. Gaynor, J., & W. Muir, W. (2009). 7 – Pharmacologic Principles and Pain: Pharmacokinetics and Pharmacodynamics. *Handbook of Veterinary Pain management, Second Edition*, 113-140. DOI: https://doi.org/10.1016/B978-032304679-4.10007-3
[2] Chen, A., Yarmush, M.L., & Maguire, T. (2014). Physiologically Based Pharmacokinetic Models: Integration of In Silico Approaches with Micro Cell Culture Analogues. *Curr Drug Metab*, *13*(6), 863-880. DOI: https://doi.org/10.2174%2F138920012800840419
[3] Nelson, E. (2019). Study: PKDB00194 Nelson1963 2019-09-30; Kinetics of the Metabolism of Acetaminophen by Humans. PK-DB DATA. DOI: https://pk-db.com/data?tab=studies