Evaluation of ChatGPT and Gemini Large Language Model for Pharmacokinetic Data Analysis and NONMEM

Abstract

Purpose

To systematically assess the ChatGPT large language model on diverse tasks relevant to pharmacokinetic data analysis and NONMEM coding in pharmacometrics and clinical pharmacology settings.

Method

ChatGPT-4 was evaluated with prototypical tasks related to non-compartmental analysis, pharmacokinetic word problems, and real-world applications of NONMEM.

Results

ChatGPT showed satisfactory performance on the R code generation tasks and provided accurate information on the principles and methods underlying pharmacokinetic data analysis. However, ChatGPT had high error rates in numerical calculations involving exponential functions. ChatGPT was able to generate applicable code blocks for the NONMEM control stream from the lay language prompts for the code tasks. The control streams contained focal structural and NONMEM syntax errors that required revision before they could be executed without errors and warnings. The outputs generated by ChatGPT were not reproducible.

Conclusion

LLMs may be useful in pharmacometrics for efficiently generating an initial coding template for modeling projects. However, the output can contain errors that require correction.

Large language models (LLM): such as ChatGPT from OpenAl, Bard from Google, and others are an emerging artificial intelligence technology that has engendered great public interest. LLM are deep learning neural networks trained on a large body of text and other information that are capable of two-way interactions with users in a manner that approximates the complexity and nuance of human conversations.

Pharmacometrics: approaches enable analysis of the time courses and variability of drug concentrations and are leveraged to inform dosage recommendations and therapeutic strategies. NONMEM is a software package for implementing nonlinear mixed effects regression methods widely used by pharmacometricians.

Objective

Case 1: PK word problems involving PK calculations related to intravenous and extravascular bolus dosing and multiple oral dosing from textbook Clinical Pharmacokinetics: Concepts and Applications, 3rd Edition, by Rowland and Tozer

Case 2: NCA analysis consisted of generating R code for semi-logarithmic graphing of concentration-time profiles and calculating the AUC and the AUMC from time zero to infinity

Case 3: NONMEM code for a linear pharmacokinetic (PK) model with oral administration

Case 4: NONMEM code for one compartment model with two parallel first-order absorptions and linear elimination

Euibeom Shin¹, Yifan Yu¹, Robert Bies¹, Murali Ramanathan¹

¹ Department of Pharmaceutical Sciences, State University of New York, Buffalo, NY, USA

Case 1 – PK word problems

ChatGPT's performance in solving pharmacokinetics word problems varied across three case studies. In Case Study 3A, which involved a five-part IV bolus dosing problem, it achieved an 88% accuracy rate. In Case Study 3B, focused on bioavailability calculations after subcutaneous or oral dosing, the accuracy was higher at 94.3%. However, in Case Study 3C, involving a three-part problem on multiple oral dosing, ChatGPT's accuracy was 80%.

Problem Statement Accuracy **Case 1A.** When 100 mg of a drug was given as an i.y. bolus, the following plasma concentration-time relationship (C in mg/L and t in hours) was observed, $C = 7.14e^{-0.051t}$

U = 7.140	
a. Volume of distribution.	10/10
b. Elimination half-life.	10/10
c. Total AUC.	10/10
d. Total clearance.	9/10
 e. Plasma concentration 70 min after a 250-mg i.v. bolus dose. 	6/10

Case 1B. The information in Table 6-6 on cidofovir (Vistide), an antiviral agent, was obtained by Wachsman et al. (Wachsman M, Petty PG, Cundy KC, et al. Pharmacokinetics, safety and bioavailability of HPMPC [Cidofovir] in human immunodeficiency virusinfected subjects. Antiviral Res 1996;29:153-161). Use mean values to answer the following questions.

TABLE 6-6 Area, Half-life, and Amount Excreted Unchanged of Cidofovir after Intravenous,

Subcutaneous, and Oral Administration								
Dose	1mg/kg	3 mg/kg		10 mg/kg				
Route of administration	i.v.	i.v.	s.c.	i.v.	p.o.			
AUC(0 — ∞) (µg-hr/mL)	7.7 ± 2.6	17.9 ± 2.3	17.8 ± 1.5	64.2 ± 8.8	< 3.5			
Half-life observed (hr)	1.7 ± 0.5	2.5 ± 0.9	2.3 ± 0.7	2.9 ± 0.7	*			
Amount excreted unchanged,	0.92 ± 0.20	2.38 ± 0.36	2.47 ± 0.25	10.5 ± 1.6	0.24 ± 0.03			
$Ae(\infty)(mg/kg)$								
p.o., oral								
* Concentration too low to obtain a proper estimate.								
a. Calculate the bioavailability (extent of systemic								
absorption) of cidofovir when given by the subcutaneous								
(s.c.) route using:								
1.) The plasma data.				10/10				
2.) The urine data.		10/1	10					
b. Calculate the bioavailability of cidofovir when given orally								
using:								
1.) The plasma data.								
2.) The urine data. 8/10				0				
c. Estimate the fraction excreted unchanged after all three in desce								
$i_{\rm N}$ doses.								
1.) 1 mg/kg								
2.) 3 mg/kg 10/10								
3.) 10 mg/kg 10/10								

Case 1C... Table 11-8 lists a typical plasma concentration-time profile obtained following an oral 500-mg dose of a drug. The AUC is 80.6 mg-hr/L, and the terminal half-life is 5 hr. TABLE 11-8 Plasma Concentration-Time Profile of a Drug After a Single 500-mg Oral Dose Time (hr) Drug Concentration (mg/L) 0 2.3 4.7 5.2 4 2.8 0.6 0.14 0.03 a. What oral dosing rate of drug is needed to maintain an average plateau concentration of 10 mg/L? 8/10 b. The decision has been made to give the drug once every

8/10

0/10



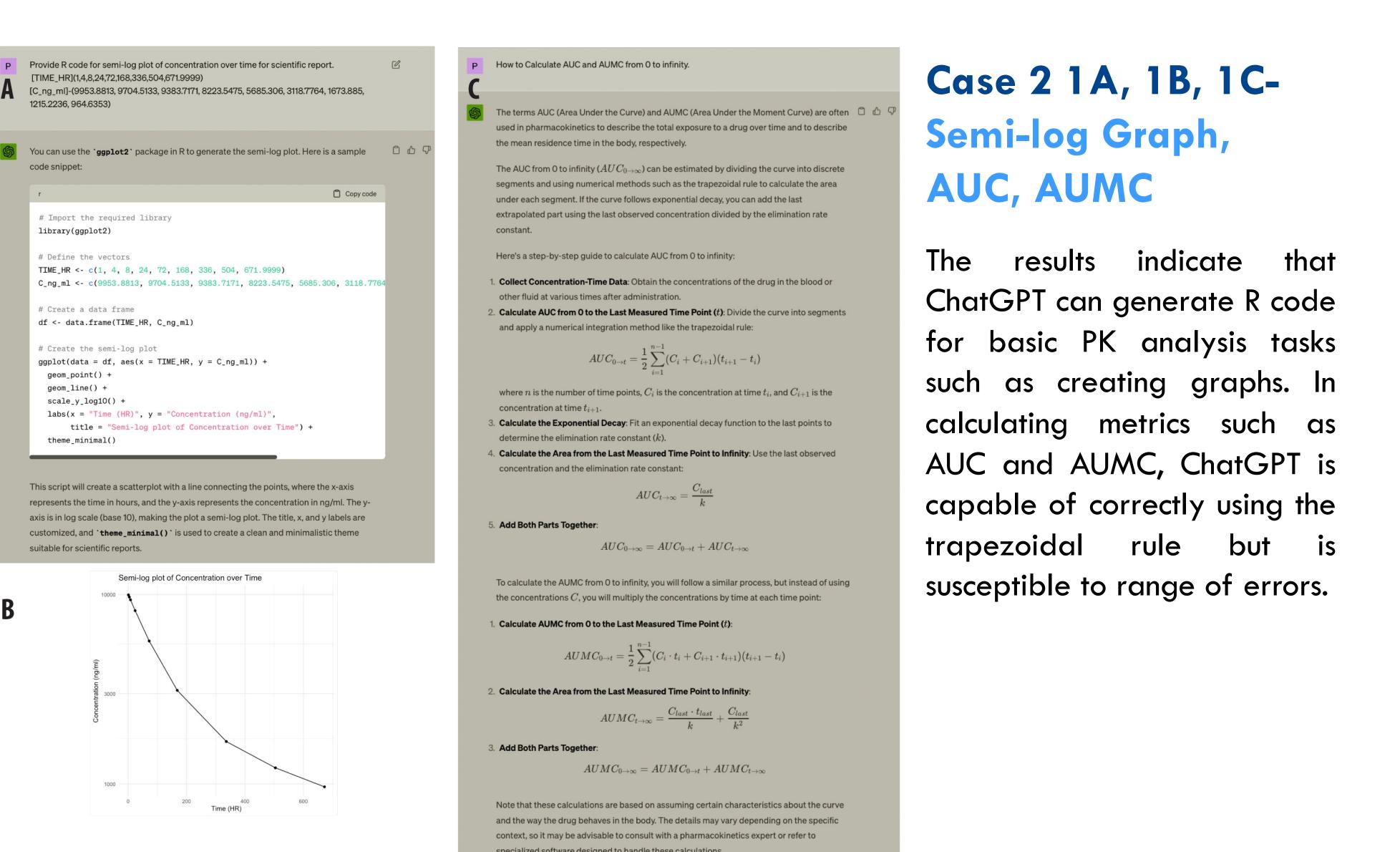
The unit dose strength of product needed?

The plateau trough concentration expected.

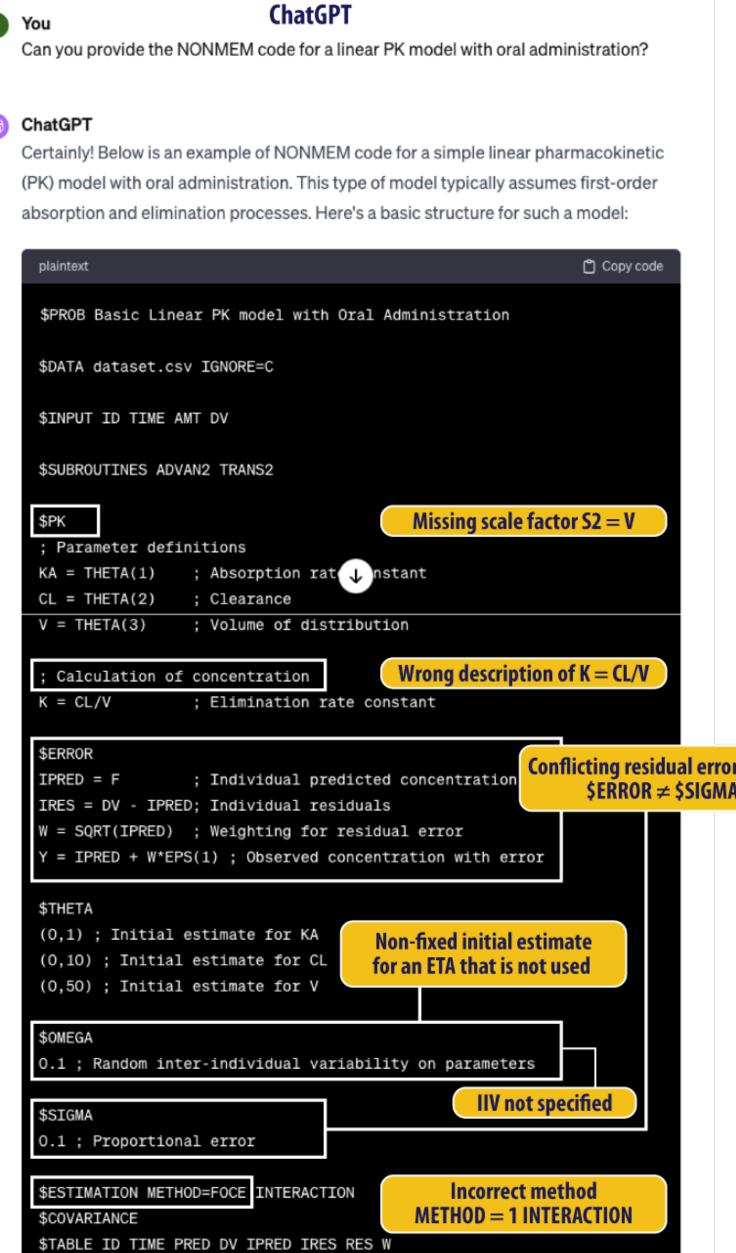
12 hr. What is:

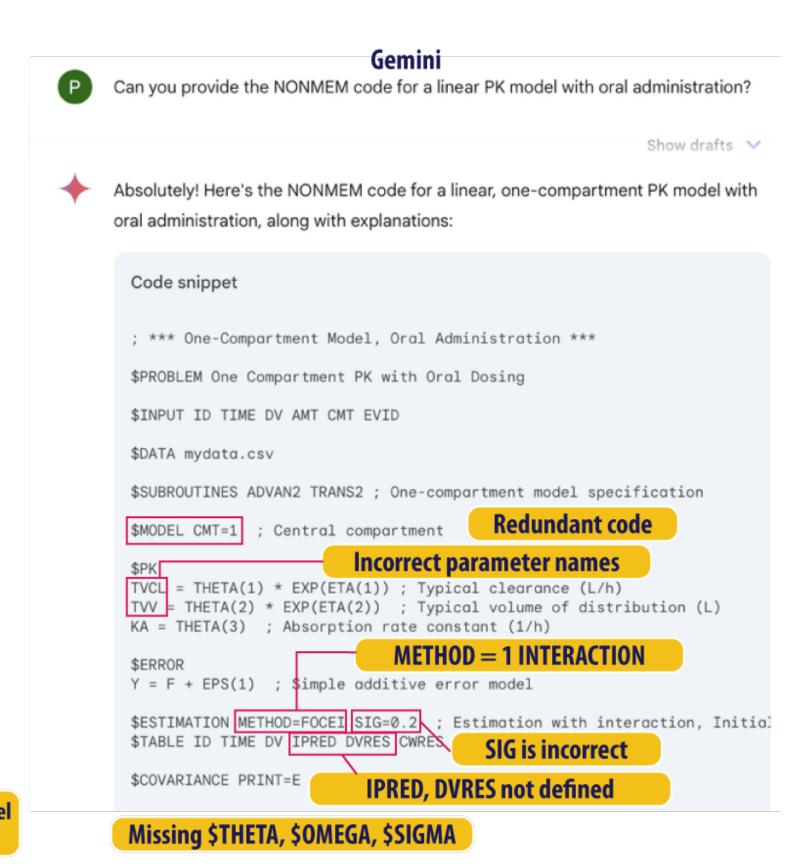


Results



Case 3 – NONMEM, Linear PK Compartmental Model with Oral Absorption





• Correct Structure and Components: ChatGPT generated a NONMEM code with appropriate sections for data input, correctly selecting subroutines (ADVAN2 and TRANS2) suitable for a one-compartment model with first-order absorption and elimination.

• Errors in Residual Error Model and Estimation Method: The code contained errors, including an incorrect residual error model and an inappropriate choice of the FOCE INTERACTION estimation method, which led to errors during execution.

Case 4 – NONMEM, one compartment model with two parallel first-order absorptions and linear elimination

	ChatGPT		Gemini	
P	You Provide the NONMEM code for a one compartment model with two parallel first-order absorptions and linear elimination.	P	Provide NONMEM code for a one-compartment model with two parallel absorptions and linear elimination.	
			Show drafts 🗸	
9	ChatGPT	-	Absolutely! Here's a NONMEM code for the model you described, along with	
	Creating a NONMEM (Nonlinear Mixed Effects Modeling) code for a pharmacokinetic model involves		explanations:	
	specifying the structural model, which in this case is a one-compartment model with two parallel first- order absorption processes and linear elimination. This model is useful for drugs administered via			
	routes that lead to absorption through two parallel pathways, such as oral and sublingual, or for		NONMEM Code	
	formulations that result in two distinct absorption phases.		Onde entire at	
	Below is a simplified example of NONMEM code for this type of model. Please note that the actual		Code snippet	
	implementation may vary depending on the specifics of your study, such as the parameterization of the		<pre>\$PROBLEM One Compartment Model - Parallel Absorption - Linear Eliminatic</pre>	
	model, the data being analyzed, and any additional complexities like covariate models or inter-		\$INPUT ID TIME DV AMT	
	individual variability components you wish to include.		<pre>\$DATA DataFile.csv IGNORE=#</pre>	
	nmtran 🗋 Copy code		\$SUBROUTINE ADVAN6 TOL=5 ; 2 Depot compartments for parallel absorption	
	\$PROBLEM One Compartment Model with Two Parallel Absorptions and Linear Elimination		\$MODEL	
	\$INPUT ID TIME AMT DV CMT EVID MDV \$DATA dataset.csv IGNORE=CMT(3) Did not exclude column names		; Compartment amounts Incorrect format	
	\$DATA dataset.csv IGNORE=CMT(3) Did not exclude column names		COMP=(CENT, DEPOT1, DEPOT2)	
	\$SUBROUTINE ADVAN2 TRANS2 Incorrect subroutine		; Pharmacokinetic parameters ; F1/F2 = Fractions of dose absorbed through DEPOT1/DEPOT2	
	\$MODEL Incorrectly identified central compartment		; F1+F2 must sum to 1 ; KA1/KA2 = Absorption rate constants from DEPOT1/DEPOT2	
	COMP=(DEPOT1, DEFDOSE) COMP=(DEPOT2) as default dosing compartment		; V = Volume of distribution ; CL = Clearance	
	COMP=(CENTRAL, DEFDOSE)			
	\$PK		<pre>\$PK F1 = 0.6 ; Fraction in first absorption pathway</pre>	
	; Model parameters		F2 = 0.4 ; Fraction in second absorption pathway KA1 = 0.8	
	TVKA1 = THETA(1) ; Absorption rate constant for the first pathway TVKA2 = THETA(2) ; Absorption rate constant for the second pathway		KA2 = 0.2 V = THETA(1) SDES & STHETA missing	
	TVCL = THETA(3) ; Clearance		CL = THETA(2)	
	TVV = THETA(4) ; Volume of distribution		S1 = CENT/V \$PK missing scale factor S3 = V	
	; Inter-individual variability ETA(1) : TTV for KA1 Unnecessary and incorrect		\$ERROR X = S1*TPRED	
	ETA(1); IIV for KA1 ETA(2); IIV for KA2 inter-individual variability			
	ETA(3) ; IIV for CL		<pre>\$EST METHOD=1 INTER MAXEVAL=9999 SIG=3 PRINT=1 NOABORT \$OMEGA BLOCK(2) ; Inter-individual variability on V and CL</pre>	
	ETA(4) ; IIV for V		\$SIGMA 1 ; Proportional error Missing initial estimates	
	; Define actual parameters			
	KA1 = TVKA1 * EXP(ETA(1)) KA2 = TVKA2 * EXP(ETA(2)) Actual parameters and			
	CL = TVCL * EXP(ETA(3)) \$0MEGA conflict	•	Incorrect Subroutines: ChatGPT correctly	
	V = TVV * EXP(ETA(4))	- 1		
	\$DES	structured the model but chose the wrong		
	DADT(1) = -KA1*A(1) ; Change in amount in DEPOT1 over time DADT(2) = -KA2*A(2) ; Change in amount in DEPOT2 over time	subroutines (ADVAN2 and TRANS2), leading to		
	DADT(3) = KA1*A(1) + KA2*A(2) - (CL/V)*A(3) ; Change in amount in CENTRAL compartment	subrounnes (ADVANZ und IKANSZ), leuding to		
	SERROR Extra IIV = ETA(1) code block	errors.		
	IPRED = A(3)/V ; Predicted concentration IIV = ETA(1) Wrong proportional error			
	W = SQRT(SIGMA(1)**2 + (SIGMA(2)**2)*IPRED**2)	•	Parameter Errors: While key parameters were	
	Y = IPRED*(1+W*EPS(1)) ; Observation model with proportional error			
	\$THETA	defined, it failed to properly distribute the dose		
	(0,1) ; Guess for KA1 (0,1) ; Guess for KA2	between absorption pathways and included		
	(0,20) ; Guess for CL	between absorption pathways and included		
	(0,70) ; Guess for V	U	nnecessary code.	
	\$OMEGA 0.1 ; IIV block	; IIV block		
	\$SIGMA 0.1 ; Residual variability block \$SIGMA only includes a single	•	Repeated Mistakes: incorrect formulas for the	
	\$ESTIMATION METHOD=FOCE INTERACTION		error model and estimation method, requiring	
	\$COVARIANCE	5	ind model and estimation memory requiring	
	\$TABLE ID TIME AMT DV PRED RES WRES ETA(1) ETA(2) ETA(3) ETA(4) ONEHEADER NOPRINT FIL	fi	ixes for the code to run successfully.	

Conclusion

Our results suggest that ChatGPT could be a useful productivity aid for knowledge encapsulation, and programming tasks. However, the high rate of errors in arithmetic calculations and NONMEM coding could limit its utility for more complex tasks and diverse data analysis scenarios in PK and pharmacometrics.

The utility of ChatGPT for NONMEM coding might be metaphorically viewed as a cup that is currently half full.

References

Rowland M, Tozer TN. Clinical pharmacokinetics : concepts and applications. 3rd ed. Baltimore: Williams & Wilkins; 1995. xiv, 601 p. p.

Funding Support

This is unfunded research.

School of Pharmacy and Pharmaceutical Sciences University at Buffalo, State University of New York Buffalo, 14215