

Fitting PK Models with SAS NLMIXED Procedure

Halimu Haridona, PPD Inc., Beijing

ABSTRACT

Pharmacokinetic (PK) models are important for new drug development. Statistical models with both fixed and random effects can be fitted by nonlinear mixed models for PK analysis. SAS/STAT NLMIXED procedure fits these models using likelihood-based methods. This paper uses a PK data example to demonstrate how to establish PK model and how to implement it using the SAS PROC NLMIXED.

INTRODUCTION

Pharmacometrics is the science of developing and applying mathematical and statistical methods to characterize, understand, and predict a drug's pharmacokinetic (PK) and pharmacodynamic (PD). It is an important tool for new drug development.

As an important component of the pharmacometrics, pharmacokinetic (PK) studies the kinetics of drug absorption, distribution, and elimination, i.e., it explores drug metabolism and excretion. Refer to the figure 1, it reflects the change of drug plasma concentration along with time after a single dose administration. The concentration is getting lower and lower along with the time. For the PK profile of a repeated dose administration, refer to figure 2, the concentration present as waviness along with time. Whatever the profile is, at a particular time point, a certain number of subjects' concentration can be supposed to be followed a normal distribution. This feature of the PK profile can help us to establish the statistical models.

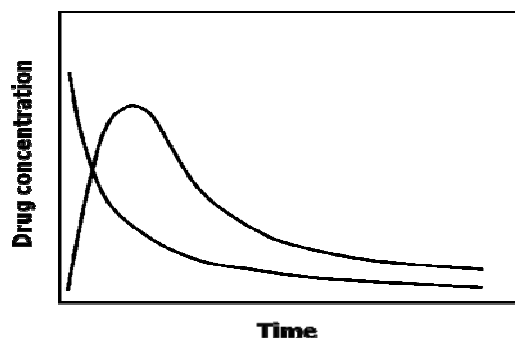


Figure 1 Curve of Drug Plasma Concentration on the Time Axis for Single Dose Administration

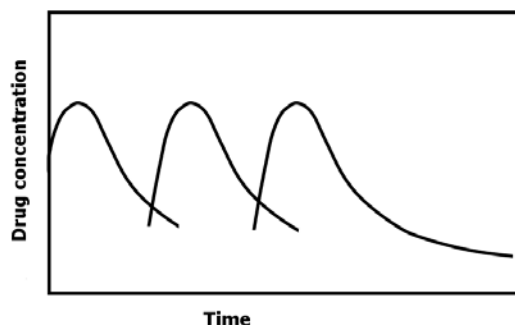


Figure 2 Curve of Drug Plasma Concentration on the Time Axis for Repeated Dose Administration

The PK modeling is to fit the nonlinear variation trend of the drug concentration using mathematical and statistical methods, and to consider both of the fixed and random effects in the model. PK models aid us in understanding the influence factors of drug absorption and elimination. The SAS/STAT NLMIXED procedure can fit these models using likelihood-based methods.

ONE-COMPARTMENT PHARMACOKINETIC MODEL

The model we used in this paper is a one-compartment model with intravenous administration and first-order elimination, which was introduced by Roman C. Littell etc. (2006).

The one-compartment model and the first-order elimination are the basic assumption of PK model we used in this paper. Regarding to the one-compartment model, if a drug rapidly equilibrates with the tissue compartment after the administration, then, for practical purposes, we can use the much simpler one-compartment model which uses only one volume term in our mathematic models, the apparent volume of distribution. The definition of first-order elimination is: elimination of a constant fraction per time unit of the drug quantity present in the organism. The elimination is proportional to the drug concentration.

PHARMACOKINETIC MODEL

The mean concentration for subject i at time t due to a single dose d_{ij} given at time t_{ij} ($t > t_{ij}$) is modeled as

$$(d_{ij}/v_i) \exp[-(t - t_{ij}) c_i/v_i]$$

Here c_i is the total clearance of drug in (liter/hour)/kg, and v_i is the apparent volume of distribution of drug in liter/kg. The specific models for c_i and v_i in terms of fixed and random effects parameters are as follows:

$$c_i = \beta_1 w_i \exp(u_{1i})$$

$$v_i = \beta_2 w_i (1 + \beta_3 \delta_i) \exp(u_{2i})$$

where

β_1 , β_2 , and β_3 are the fixed effects parameters to be estimated

u_{1i} and u_{2i} are the i^{th} subject's random effects having mean 0 and variance-covariance matrix to be estimated

w_i is the i^{th} subject's birth weight

δ_i is the influence factor, it equals 1 if the i^{th} subject's 5-minute Apgar score is less than 5 and 0 otherwise in this case

Because each subject receives multiple doses at different time points, the concentration at any particular time is computed as the sum of the contributions from all doses received prior to that time. Therefore, the model should be written as follow

$$f(d_{ij}, t_{ij}, c_i, v_i) = d_{ij}/v_i + f(d_{ij-1}, t_{ij-1}, c_i, v_i) \exp[-(t_{ij} - t_{ij-1}) c_i/v_i]$$

where $f(d_{ij}, t_{ij}, c_i, v_i)$ is the concentration for subject i immediately after receiving dose d_{ij} at time t_{ij} .

DATA

We will take the phenobarbital data from Grasela and Donn (1985) as an example to demonstrate the PK model implementing by SAS PROC NL MIXED. The data are collected from 59 newborn infants treated with phenobarbital. Refer to the appendix 1 for the complete data. There are eight variables in the data and the names and descriptions as below in table 1:

Table 1 Descriptions of Variables

Variables	Descriptions
INDIV	Subject ID
NEWSUB	Equal to 1 if the current observation is from a new subject, 0 otherwise
WEIGHT	Birth weight (kg)
APGARLOW	Equal to 1 if 5-minute Apgar score is less than 5, 0 otherwise
TIME	Time of administration
TLAG	The previous time value (if NEWSUB=0) or 0.0 (if NEWSUB=1)
DOSE	Dose amount ($\mu\text{g/kg}$)
CONC	Plasma concentration of phenobarbital (in $\mu\text{g/kg}$)

SAS PROGRAMS

The SAS PROC NLMIXED programs for the previously described PK model are as follows.

```
proc nlmixed data=pheno qpoints=1;
* parameters and starting values;
  parms beta1=0.5 beta2=1 beta3=0.15
        s2u1=0.05 s2u2=0.03 cu12=0.015 s2e=0.01;
* boundary constraints;
  bounds beta1>=0, beta2>=0, s2u1>=0, s2u2>=0, s2e>0;
* compute key terms in the model;
  clear = beta1/100*weight*exp(u1);
  vol = beta2*weight*(1+beta3*apgarlow)*exp(u2);
  eterm = exp(-(time-tlag)*clear/vol);
* compute predicted concentration value;
  if (newsb = 1) then pred = dose/vol;
  else pred = dose/vol + zlag(pred)*eterm;
* variance of the predicted concentration value;
  var = s2e*(pred**2);
* model specification;
  model conc ~ normal(pred,var);
* random effects specification;
  random u1 u2 ~ normal([0,0],[s2u1,cu12,s2u2])
  subject=indiv out=ebayes;
* compute an additional estimate of the correlation;
  estimate 'corr(u1,u2)' cu12/sqrt(s2u1*s2u2);
* output subject-specific predictions of clearance and volume;
  predict clear out=clear;
  predict vol out=vol;
run;
```

The option QPOINTS= specifies the number of quadrature points to be used during evaluation of integrals. We specify QPOINTS=1 in order to fit a single adaptive Gaussian quadrature point. The PARMS statement lists names of parameters and specifies initial values. You can specify the parameters and values either directly in a list or provide the name of a SAS data set that contains them using the DATA= option. While the PARMS statement is not required, you are encouraged to use it to provide PROC NLMIXED with accurate starting values. Parameters not listed in the PARMS statement are assigned an initial value of 1. Starting values here were chosen by trial and error. Boundary constraints are specified with a BOUNDS statement. We compute the clearance, distribution volume, exponential terms, and the predicted concentration values using the previously described models and parameters. The β_1 is divided by 100 to improve numerical stability. Predicted values are computed depending on whether this is a new subject or not. For the second and subsequent observations for each subject, the ZLAG function is used to model the current value of PRED as a function of the previous value of PRED in time. You should always use ZLAG instead of LAG when programming in PROC NLMIXED. The variance of predicted value goes up with the square of the predicted value. The MODEL statement is the mechanism for specifying the conditional distribution of the dependent variable. The RANDOM statement defines the random effects and their distribution. The random effects must be represented by symbols that appear in your SAS programming statements. They typically influence the mean value of the distribution specified in the MODEL statement. The ESTIMATE statement enables you to compute an additional estimate that is a function of the parameter values. The PREDICT statement enables you to construct predictions of an expression across all of the observations in the input data set.

RESULTS

SAS outputs of PROC NLMIXED for previous model:

Specifications	
Data Set	WORK.PHENO
Dependent Variable	conc
Distribution for Dependent Variable	Normal
Random Effects	u1 u2
Distribution for Random Effects	Normal
Subject Variable	indiv
Optimization Technique	Dual Quasi-Newton
Integration Method	Adaptive Gaussian Quadrature

The "Specifications" table lists some basic information about the nonlinear mixed model that you have specified.

Dimensions	
Observations Used	155
Observations Not Used	589
Total Observations	744
Subjects	59
Max Obs Per Subject	6
Parameters	7
Quadrature Points	1

The "Dimensions" table lists various counts related to the model, including the number of observations, subjects, and parameters. You can use these numbers to check that you have specified your input information correctly.

Parameters							
beta1	beta2	Beta3	s2u1	s2u2	cu12	s2e	NegLogLike
0.5	1	0.15	0.05	0.03	0.015	0.01	437.759491

The "Parameters" table lists the parameters to be estimated, their starting values, and the negative log likelihood evaluated at these starting values.

Iteration History					
Iter	Calls	NegLogLike	Diff	MaxGrad	Slope
1	5	436.624676	1.134816	249.0696	-10645.5
2	9	435.107501	1.517175	225.3187	-371.212
3	11	434.015793	1.091707	121.1016	-48.9049
4	14	433.996886	0.018907	105.5664	-10.0008
5	16	433.941294	0.055592	90.84407	-2.21844
6	18	433.80331	0.137984	80.67595	-1.12135
7	20	433.764465	0.038845	89.06237	-1.32549
8	21	433.708496	0.055968	19.46526	-0.10437
9	23	433.705581	0.002916	5.191272	-0.00864
10	25	433.704789	0.000791	3.106978	-0.0016
11	27	433.704749	0.00004	0.296499	-0.00007
12	29	433.704746	3.366E-6	0.115745	-5.74E-6
13	31	433.704746	6.569E-8	0.001814	-9.36E-8

NOTE: GCONV convergence criterion satisfied.

The “Iterations History” table records the history of the minimization of the negative log likelihood. This model converges in 13 iterations.

Fit Statistics	
-2 Log Likelihood	867.4
AIC (smaller is better)	881.4
AICC (smaller is better)	882.2
BIC (smaller is better)	896.0

The “Fit Statistics” table lists the final maximized value of the log likelihood. The Fit Statistics can be used to compare the different models that you established in the model development stage and to select the most appropriate one.

Parameter Estimates									
Param	Estimate	Std Error	DF	t Value	Pr > t	Alpha	Lower	Upper	Gradient
beta1	0.4738	0.02202	57	21.52	<.0001	0.05	0.4297	0.5179	-0.00072
beta2	0.9756	0.02549	57	38.28	<.0001	0.05	0.9246	1.0267	0.000067
beta3	0.1380	0.07437	57	1.86	0.0687	0.05	-0.01094	0.2869	0.000049
s2u1	0.03707	0.01765	57	2.10	0.0402	0.05	0.001721	0.07243	0.000423
s2u2	0.02171	0.006737	57	3.22	0.0021	0.05	0.008221	0.03520	-0.00062
cu12	0.01903	0.009439	57	2.02	0.0485	0.05	0.000133	0.03793	0.000755
s2e	0.01320	0.002152	57	6.13	<.0001	0.05	0.008889	0.01751	0.001814

From the “Parameter Estimates” table we can see that the estimates of β_1 and β_2 are significantly different from zero, that means the clearance and volume have significantly influence on the drug plasma concentration. The estimate for Apgar score β_3 is not significantly larger than zero, so the Apgar score has no influence on the concentration. We can also use some other subject characteristic factors instead of Apgar score to test whether they have influence on the drug concentration or not. The estimates of variances and variance-covariance matrix of random effects are significantly larger than zero.

Additional Estimates								
Label	Estimate	Std Error	DF	t Value	Pr > t	Alpha	Lower	Upper
corr(u1,u2)	0.6709	0.2703	57	2.48	0.0160	0.05	0.1297	1.2121

From the “Additional Estimates” table we can see that the correlation between the random effects also appears to be significantly larger than zero.

CONCLUSIONS

Fitting a nonlinear mixed model is more challenging than fitting a linear one. The challenges come from at least three sources. First, the likelihood function for random effects of a nonlinear mixed model is an integral. The second one is the parameters with widely varying scales, which maybe need to be rescaled to get a numerical stability. The third one is the starting values for the parameters. Starting values far from their optimal estimated values can produce unpredictable results. Trial and error is the most commonly used way to assign the initial values for the parameters. If your nonlinear model has interpretable parameters, you can sometimes obtain good guesses for the parameters simply by visually inspecting a plot of the data.

However, the SAS PROC NLMIXED is a flexible and powerful tool to fit nonlinear mixed models. Its main computing components are a SAS engine for processing programming statements, and a numerical integrator. These enable you to fit many nonlinear mixed models using likelihood-based methods flexibly.

REFERENCES

Ramon C. Littell, George A Milliken, Walter W. Stroup, Russell D. Wolfinger, and Oliver Schabenberger (2006), SAS for Mixed Models, Second Edition, SAS Institute Inc, Cary, NC, USA.

Russell D. Wolfinger (1999), Fitting Nonlinear Mixed Models with the New NLMIXED Procedure, SAS Institute Inc, Cary, NC, USA.

Sheiner, L.B. and Beal, S.L. (1985), "Pharmacokinetic Parameter Estimates from Several Least Squares Procedures: Superiority of Extended Least Squares," Journal of Pharmacokinetics and Biopharmaceutics, 13, 185-201.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Halimu Haridona

Enterprise: PPD Inc.

Address: 8/F, Tower B, Central Point Plaza No 11, Dongzhimen South Ave. Dongcheng District

City, State ZIP: Beijing, 100007

Work Phone: +86 10-57636250

Fax: +86 10-57636251

E-mail: Halimu.Haridona@ppdi.com

Web: www.ppdi.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.