Estimation of pharmacokinetic parameters in random effects models

Jørn Attermann
Department of Biostatistics
University of Aarhus

Updates

- 19 July 2002: ”FOCE” removed from ”A collection of software” -> SAS NLMIXED.

Disposition

1. Introduction: A simple PK model
2. A brief review of models and inferential methods
    - Linear mixed models
    - Hierarchical nonlinear models
3. Example: A simple PK model analyzed by four popular software packages

Example: One-compartment model

- Theophylline data analyzed by Pinheiro & Bates (1995)
- Broncho-dilating drug
- Serum concentrations of theophylline measured 11 times over a 25 hour period in each of 12 subjects after oral administration
- Variables considered: subject, time, concentration, dose
- PK model: one-compartment model with first-order absorption

Thoretical background for the PK model

$$X_a \xrightarrow{\;k_a\;} \boxed{X} \xrightarrow{\;K\;}$$

$$\left.\begin{aligned}
&\frac{dX(t)}{dt} = k_a X_a(t) - KX(t) \\
&\frac{dX_a(t)}{dt} = -k_a X_a(t) \\
&X(0) = 0
\end{aligned}\right\} \Rightarrow \mathrm{X}_t = \frac{k_a D}{k_a - K}\left(e^{-Kt} - e^{-k_a t}\right)$$

where X_a and X is the amount of drug in the absorption and central compartment, respectively, k_a and K are rate constants, and D is the dose given at t = 0. It is assumed that all the drug is absorbed.

Let C_t denote the concentration in the central compartment at time t, and let V be the volume. Then

$$C_t = \frac{k_a D}{V\left(k_a - K\right)}\left(e^{-Kt} - e^{-k_a t}\right).$$

Since the clearance is $Cl = KV$ we can also write

$$C_t = \frac{k_a K D}{Cl\left(k_a - K\right)}\left(e^{-Kt} - e^{-k_a t}\right).$$

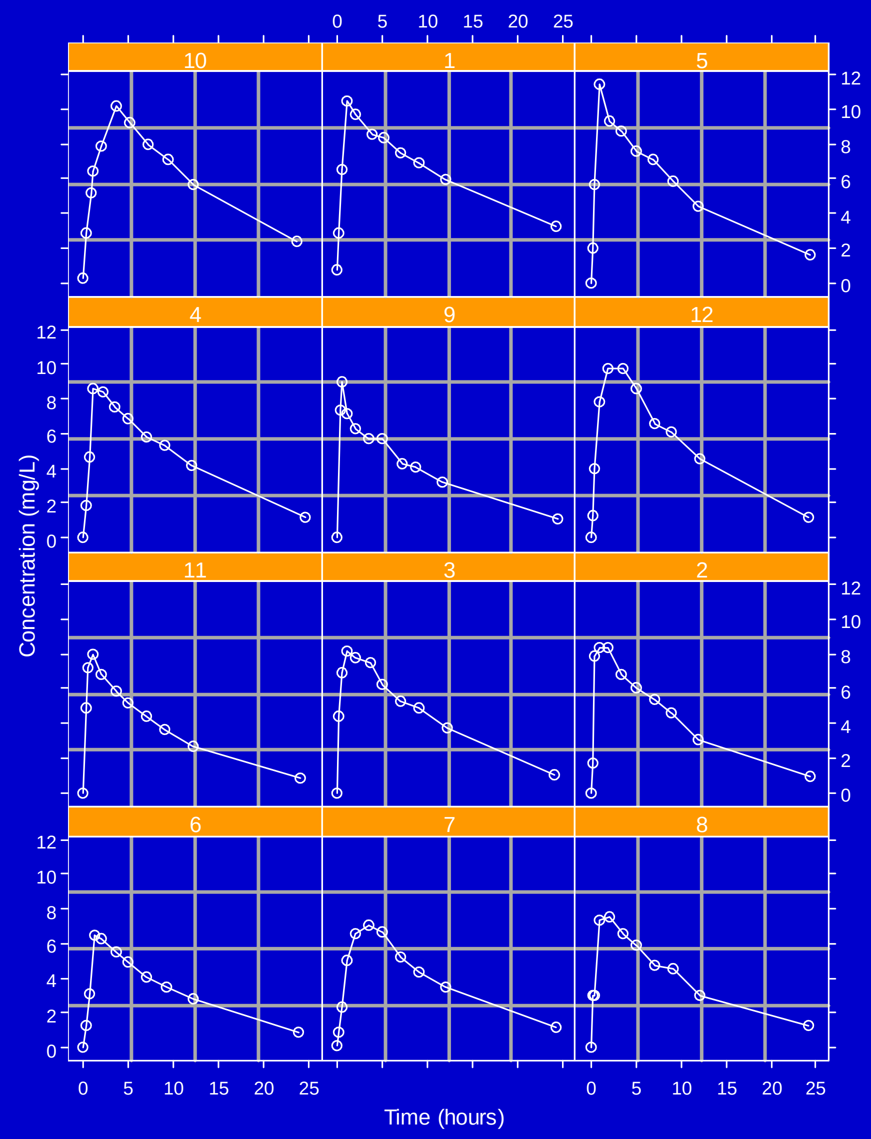
Another parameterization, utilized by PKBugs, is

$$C_t = \frac{D\left(k_a^* + Cl/V\right)}{V k_a^*}\left(e^{-\frac{Cl}{V}t} - e^{-\left(k_a^* + \frac{Cl}{V}\right)t}\right),$$

where

$$k_a^* = k_a - K = k_a - Cl/V.$$

Observed data



The Linear Mixed Model (LMM)

Laird & Ware (1982)

$$y_i = X_i \beta + Z_i b_i + e_i, \qquad i = 1, \ldots, m$$

where

y_i: vector of responses for *i*'th individual

β: vector of fixed effects parameters

X_i: design matrix for fixed effects, *i*'th individual

b_i: vector of random effects for *i*'th individual

Z_i: design matrix for random effects, *i*'th individual

e_i: vector of intraindividual errors

$$\begin{pmatrix} b_i \\ e_i \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} G(w) & 0 \\ 0 & R_i(w) \end{pmatrix} \right), \text{ independent}$$

w: a vector of random effects parameters.

Estimation

Since

$$y_i \sim N\left(X_i\beta,\ V_i\left(w\right)\right),$$

where

$$V_i = Z_i G Z_i^{'} + R_i,$$

the ML estimator of β, for known w, is

$$\hat{\beta} = \sum_{i=1}^{m}\left(X_i^{'} V_i^{-1}\left(w\right) X_i\right)^{-1} \sum_{i=1}^{m} X_i^{'} V_i^{-1}\left(w\right) y_i .$$

The best linear unbiased predictor (BLUP) for b_i is

$$\tilde{b}_i\left(w\right) = G\left(w\right) Z_i^{'} V_i^{-1}\left(w\right)\left(y_i - X_i\hat{\beta}\left(w\right)\right).$$

ML estimation of w involves joint maximization with respect to β and w of the likelihood

$$L = \prod_{i=1}^{m} N\left(X_i \beta,\ V_i\right).$$

REML estimation of w is done by maximizing the restricted log-likelihood function

$$\log\left(L_R\right) = \log\left(L\right) + \frac{1}{2} p \log\left(2\pi\right) - \frac{1}{2} \sum_{i=1}^{m} \log\left|X_i' V_i^{-1} X_i\right|.$$

Hierarchical nonlinear models

Let the joint density function of y_i and b_i be

$$f_i\left(y_i, b_i; \phi, \xi\right) = p_i\left(y_i \mid b_i; \phi\right)\, q_i\left(b_i; \xi\right), \quad i = 1, \ldots, m$$

where

y_i: observed data vector for i'th subject, independent across subjects

b_i: latent random effects vector for i'th subject, independent across subjects

ϕ: vector of unknown fixed-effects and intraindividual variance parameters

ξ: vector of unknown interindividual random-effects parameters.

Inference based on linearization

First-order linearization (Beal & Sheiner, 1982)

Assume that

$$y_i \mid b_i \sim N\left[m_i\left(\phi, b_i\right), R_i\left(\phi\right)\right],$$

$$b_i \sim N\left[0, G\left(\xi\right)\right].$$

Let an initial estimate $\hat{\phi}$ of ϕ be given. A first-order Taylor expansion of m_i about $\phi = \hat{\phi}$ and $b_i = 0$ yields

$$m_i\left(\phi, b_i\right) \approx m_i\left(\hat{\phi}, 0\right) + X_i\left(\hat{\phi}\right)\left(\phi - \hat{\phi}\right) + Z_i\left(\hat{\phi}\right) b_i$$

where

$$X_i = \frac{\partial m_i}{\partial \phi}\left(\hat{\phi}, 0\right),$$

$$Z_i = \frac{\partial m_i}{\partial b_i}\left(\hat{\phi}, 0\right).$$

Now, the adjusted observations

$$v_i = y_i - m_i\left(\hat{\phi}, 0\right) + X_i\left(\hat{\phi}\right)\hat{\phi}$$

are approximately independently normally distributed

$$\sim N\left[X_i\left(\hat{\phi}\right)\phi,\ Z_i\left(\hat{\phi}\right)G(\xi)Z_i\left(\hat{\phi}\right)' + R_i(\phi)\right].$$

Thus, the adjusted observations follow (approximately) a LMM and the unknown parameters ϕ and ξ can be iteratively reestimated by the techniques given above.

Conditional first-order linearization (Lindstrom & Bates, 1990; Wolfinger, 1993)

Roughly the first-order method of Beal & Sheiner (1982), except that the Taylor expansion is about the (approximate) BLUP, $\tilde{b}_i$, of b_i rather than about $b_i = 0$.

Inference based on integral approximations

<u>Adaptive Gaussian Quadrature (Pinheiro & Bates, 1995)</u>

Gauss-Hermite: For a smooth function *f*, and fixed *n*,

$$\int_{-\infty}^{\infty} f(z)\,dz = \int_{-\infty}^{\infty} e^{-z^2}\left[e^{z^2} f(z)\right]dz = \sum_{k=1}^{n} w(z_k)\, e^{z_k^2} f(z_k) + R_n$$

where z_k is the *k*'th root of the Hermite polynomial H_n, the weights are given by

$$w(z_k) = \frac{2^{n-1} n! \sqrt{\pi}}{n^2 \left[H_{n-1}(z_k)\right]^2},$$

and R_n is a residual term.

Let f be the joint density function of y_i and $b_{i.}$ Let $\tilde{b}_i$ denote the empirical Bayes' estimator of b_i, i.e. the vector that maximizes f in b_i, and let Γ_i denote the final Hessian from this maximization. Then the contribution of the i'th subject to the likelihood is

$$L_i\left(\phi,\xi\right)=\int f\left(y_i,b_i;\phi,\xi\right)\,db_i=\int p\left(y_i\mid b_i;\phi\right)\,q\left(b_i;\xi\right)db_i$$

$$\approx 2^{r/2}\left|\Gamma\left(\phi,\xi\right)\right|^{-1/2}$$

$$\times\sum_{j_1=1}^{p}\cdots\sum_{j_r=1}^{p}\left[p\left(y_i\mid a_{j_1,\ldots,j_r};\phi\right)q\left(a_{j_1,\ldots,j_r};\xi\right)\prod_{k=1}^{r}w_{j_k}e^{z_{j_k}^2}\right]$$

where $z_{j_1,\ldots,j_r}$ is the vector with elements $z_{j_1},\ldots,z_{j_r}$, r is the dimension of b_i, and

$$a_{j_1,\ldots j_r}=\tilde{b}_i+2^{1/2}\Gamma\left(\phi,\xi\right)^{-1/2}z_{j_1,\ldots j_r}.$$

Inference by Markov chain Monte Carlo (MCMC)

Assume the nonlinear hierarchical model with a hyperprior distribution of the parameters.
Inference by Gibbs sampling.

A collection of software

- NONMEM:
  - ”Gold standard” in PK modeling,
  - FO and FOCE,
  - Developed by Beal & Sheiner
- nlme in R (or S-Plus):
  - FOCE
  - Developed by J. Pinheiro, D. Bates and M. Lindstrom
- SAS NLMIXED:
  - FO, adaptive Gaussian quadrature

- PKBugs:
  - specialized for PK modeling with 32 different compartmental models
  - MCMC
  - developed by D.J. Lunn, J. Wakefield, A. Thomas, N. Best, and D. Spiegelhalter

Example, revisited

Statistical model, first run

Let y_i denote the 11-dimensional vector of observed concentrations for the i'th subject, $i = 1, \ldots, 12$. Let the conditional mean concentration be given by

$$m_{it}\left(\phi, b_i\right) = \frac{k_{a,i} D_i K_i}{Cl_i}\left(e^{-K_i t} - e^{-k_{a,i} t}\right)$$

where

$$Cl_i = \exp\left(\theta_1 + b_{i1}\right)$$

$$k_{a,i} = \exp\left(\theta_2 + b_{i2}\right)$$

$$K_i = \exp\left(\theta_3 + b_{i3}\right)$$

and

$$\phi = \left(\theta_1, \theta_2, \theta_3, \sigma^2\right)'.$$

We assume that

$$y_{it} \mid b_i \sim N\left(m_{it},\ \sigma^2\right)$$

and that $\left\{y_{it}\right\}_{t=t_1,\ldots t_{11}}$ are conditionally independent given b_i for every i. Finally, we assume that the latent random vectors b_i are independent with

$$b_i \sim N\left(0,\ \operatorname{diag}\left(\xi_1, \xi_2, \xi_3\right)\right).$$

PKBugs uses the alternative parameterization as mentioned above.

The analyses were performed on a Dell Inspiron 8100 with a 1 GHz Intel processor and 512 MB RAM.

Implementation in NONMEM

```
$PROBLEM Theophylline data, three interindiv var comp
$INPUT ID AMT TIME DV MDV
$DATA PKDATA
$SUBROUTINES ADVAN2
$PK
      CL=EXP(THETA(1)+ETA(1))
      KA=EXP(THETA(2)+ETA(2))
      K=EXP(THETA(3)+ETA(3))
      S2=CL/K
$ERROR
      Y = F+EPS(1)
$THETA 1 1.1 1.2
$OMEGA 0.2 0.2 0.2
$SIGMA 0.5
$ESTIMATION MAXEVAL=2000 PRINT=5 SIGDIGITS=4
$COVARIANCE
$SCAT    PRED VS TIME
$SCAT    RES VS TIME
$SCAT    PRED VS DV  UNIT
```

Implementation in SAS

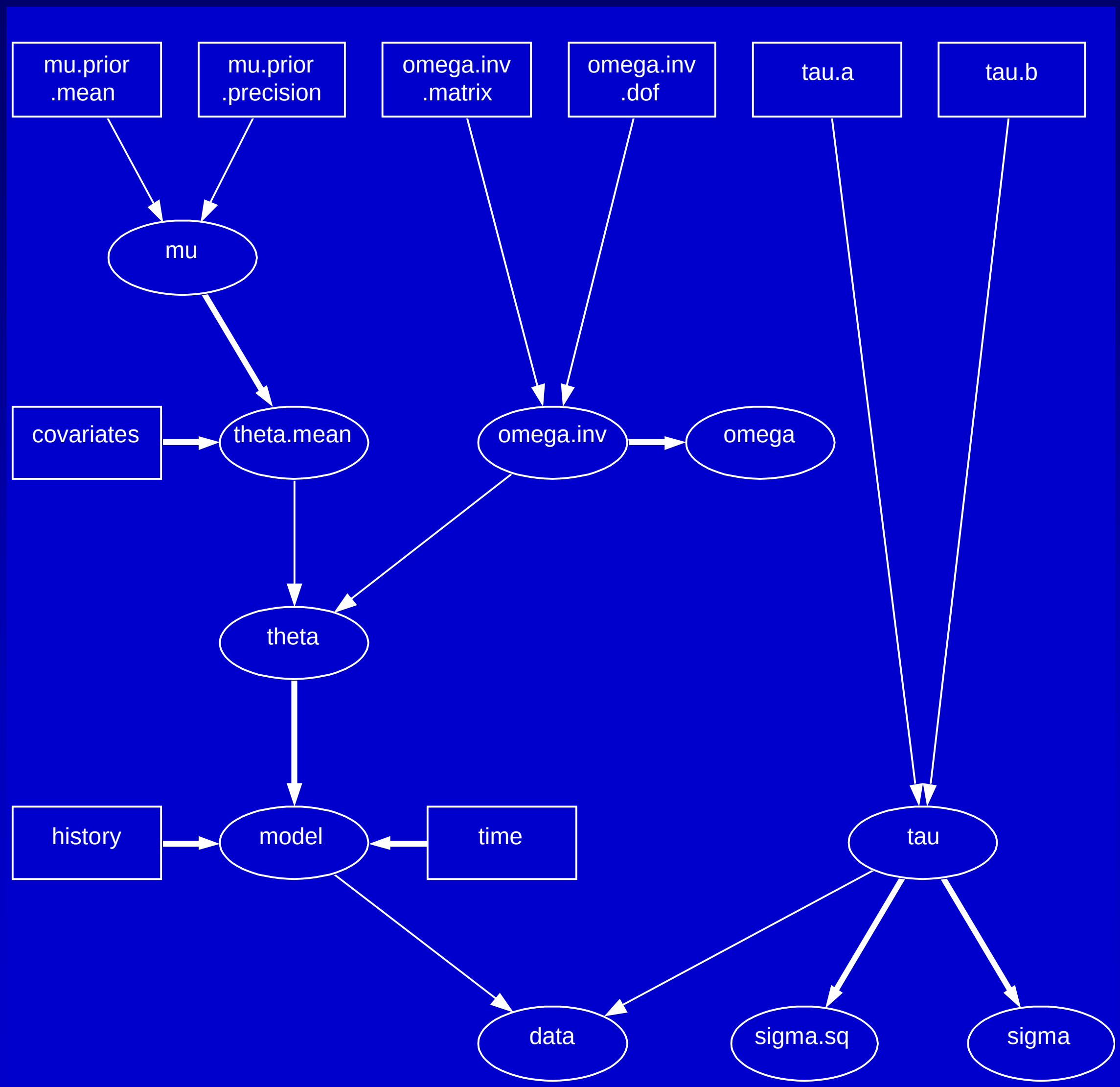
```
proc nlmixed data=theoph;
   parms theta1=-3.25 theta2=0.4
      theta3=-2.52        /* Fixed effects */
      sigma2=2            /* Intraindividual var */
      omega1=1 omega2=1
      omega3=1;           /* Interindividual var */
   Cl = exp(theta1 + b1); /* Assumes F=1 */
   ka = exp(theta2 + b2);
   K = exp(theta3 + b3);
   conc_m = dose * K * ka / Cl / (ka-K)
            * ( exp(-K*time)-exp(-ka*time) );
   model conc ~ normal(conc_m, sigma2);
   random b1 b2 b3 ~ normal([0,0,0],
      [omega1,0,omega2,0,0,omega3]) subject=subject;
   predict b1 out=b1;
   predict b2 out=b2;
   predict b3 out=b3;
   run;
```

Implementation in R (nlme)

```
theo <- read.table("data.txt", header=TRUE)

library(nlme)

library(lattice)

theo <- groupedData(formula = Conc ~ Time | Subject,
            data = theo,
            labels = list(x="Time (hours)",
                          y="Concentration (mg/L)"))

data(theo)

plot(theo)
```

```
conc.m <- function(Dose, lCl, lka, lK, Time) {
      Cl <- exp(lCl)
      ka <- exp(lka)
      K <- exp(lK)
      ka * Dose * K / Cl / (ka-K)
            * (exp(-K * Time) - exp(-ka * Time))
}

theo.nlme <- nlme(Conc ~ conc.m(Dose, lCl, lka,
                                    lK, Time),
            fixed = lCl + lka + lK ~1,
            data = theo,
            random = lCl + lka + lK ~ 1,
            start = list(fixed = c(-3, 0.5, -2)))


summary(theo.nlme)

plot(theo.nlme)

plot(augPred(theo.nlme))
```

Implementation in PKBugs



The parameters τ, μ, and Ω^{-1} are updated using Gibbs sampling.
Θ is sampled using a Metropolis-Hastings algorithm.

Results of first run

	Time (sec)	Neg. loglike	lCl	lk_a	lK	σ^2	ξ_1	ξ_2	ξ_3
NONMEM	≈ 0	182.73	-3.29 0.085	0.961 0.339	-2.57 0.062	0.489 0.170	0.0365 0.027	0.741 0.422	0.012 0.014
SAS	600	177.74	-3.23 0.059	0.478 0.198	-2.46 0.051	0.504 0.069	0.028 0.012	0.427 0.195	7E-13 1E-5
nlme	∞								
PKBugs	160		-3.22 0.087	0.456	-2.45				

The seven rightmost columns contain parameter estimates and standard errors. PKBugs performed 100000 iterations.

Statistical model, second run

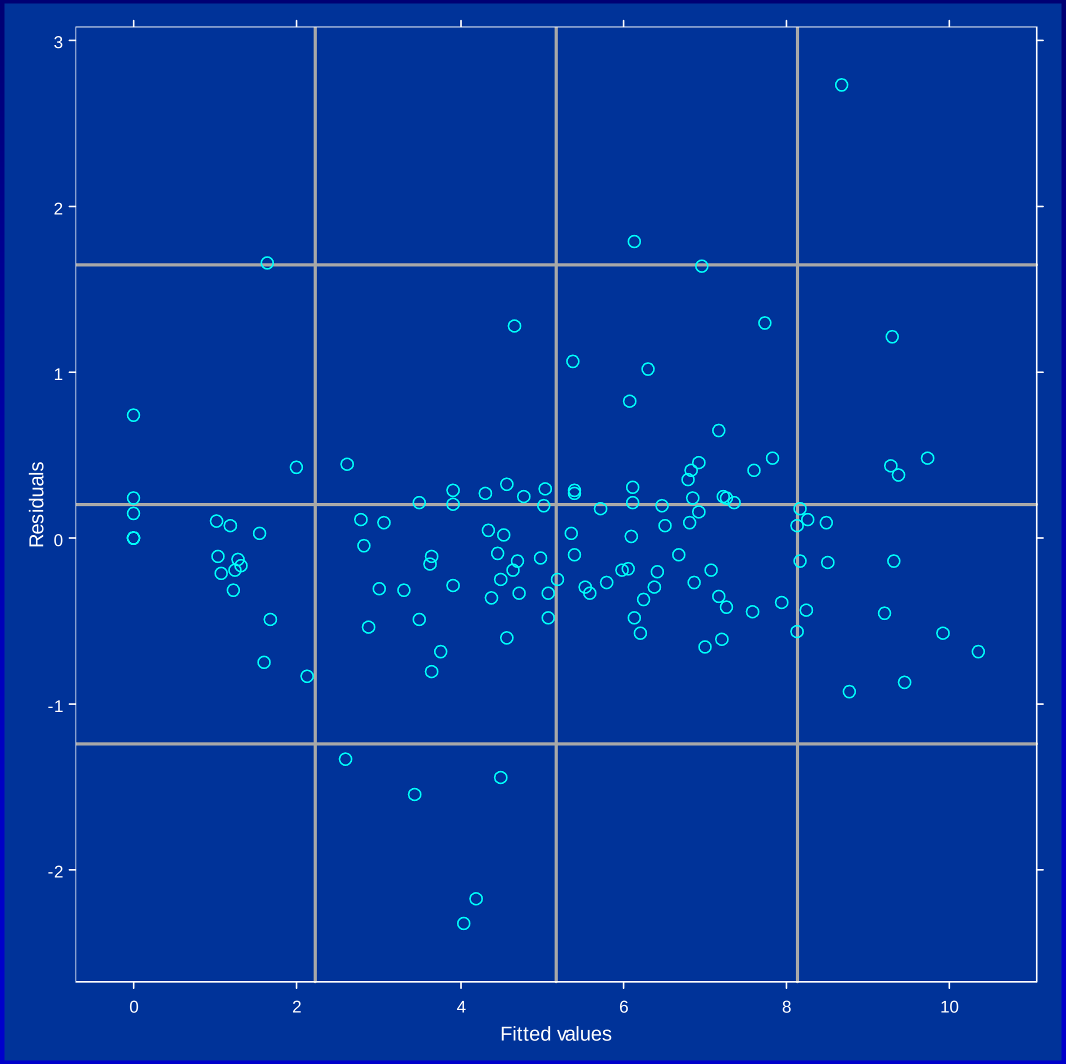
Remove the parameter ξ_3.

Results of second run

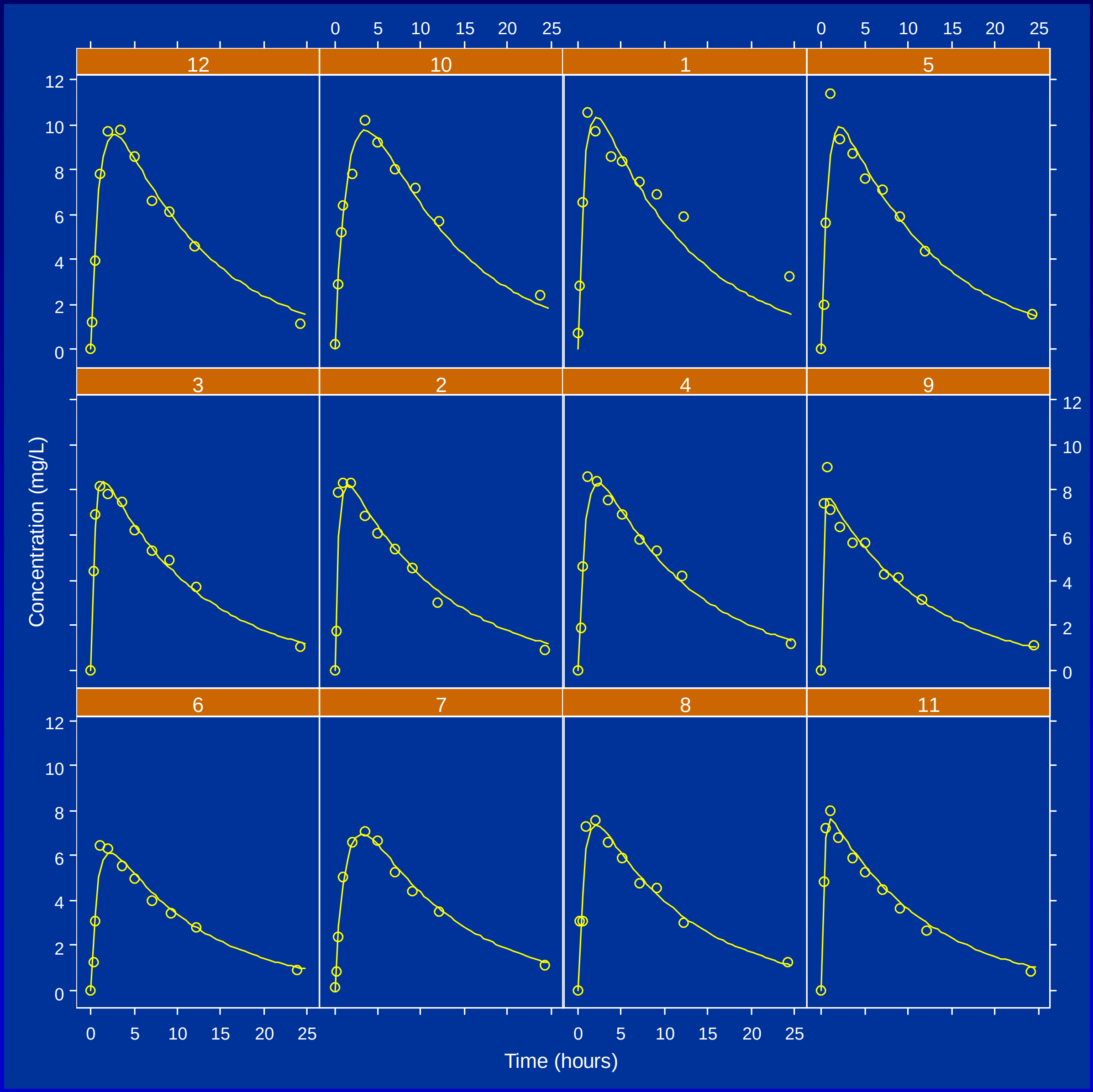
	Time (sec)	Neg. loglike	*lCl*	lk_a	*lK*	σ^2	ξ_1	ξ_2	ξ_3
NONMEM	≈ 0	181.31	-3,29 0.089	0.88 0.369	-2.57 0.064	0.543 0.159	0.029 0.015	0.657 0.419	0
SAS	≈ 0	177.75	-3.23 0.059	0.481 0.199	-2.45 0.051	0.502 0.068	0.028 0.012	0.433 0.200	0
nlme	≈ 0	177.02	-3.23 0.060	0.466 0.199	-2.46 0.052	0.503	0.028	0.414	0
PKBugs									

The seven rightmost columns contain parameter estimates and standard errors.

Residuals vs. fitted, second run



Fitted curves, second run



References

- Beal, S.L. and Sheiner, L.B. (1982) Estimating Population Kinetics. *CRC Crit. Rev. Biomed. Eng.*, 8, 195-222.
- Beal, S.L. and Sheiner, L.B. (1988) Heteroskedastic Nonlinear Regression. *Technometrics*, 30, 327-38.
- Davidian, M. and Giltinan, D.M. (1995) *Nonlinear Models for Repeated Measurement Data.* Chapman & Hall, New York.
- Laird, N.M and Ware, J.H. (1982) Random effects models for longitudinal data. *Biometrics*, 38, 963-74.
- Lindstrom, M.J. and Bates, D.M. (1990) Nonlinear Mixed Effects Models for Repeated Measures Data. *Biometrics*, 46, 673-87.
- McCullagh, P. and Nelder, J.A. (1989) *Generalized Linear Models* (2nd edn). Chapman & Hall, New York.
- Pinheiro, J.C. and Bates, D.M. (1995) Approximations to the Log-likelihood Function in the Nonlinear Mixed-effects Model. *Journal of Computational and Graphical Statistics*, 4, 12 -35.
- 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.
- Wolfinger R.D. (1993) Laplace's Approximation for Nonlinear Mixed Models. *Biometrika*, 80, 791-5.