

Kevin Freise, Chih-Wei Lin, and
 Mohammad Saleh

22S: 138 Final Project Report

12/10/2007

Instructor: Dr. Cowles

Introduction:

For our project we conducted a (hierarchal) nonlinear mixed effects Bayesian analysis of a real data set from population pharmacokinetics. The data set consisted of plasma theophylline concentration-time profiles from 12 subjects for approximately 25 hours following a single oral administration of theophylline (Upton et. al, 1982). The data was modeled using an open 1-compartment pharmacokinetic (PK) model with 1st order absorption as given by:

$$y_{i,j} = f(t_{i,j}, \underline{\theta}_i) + \varepsilon$$

$$f(t_{i,j}, \underline{\theta}_i) = \frac{D_i \cdot ka_i}{\left(\frac{V}{F}\right)_i \cdot (ka_i - ke_i)} \left[\exp(-ke_i \cdot t_{i,j}) - \exp(-ka_i \cdot t_{i,j}) \right]$$

$$ke_i = \frac{\left(\frac{Cl}{F}\right)_i}{\left(\frac{V}{F}\right)_i}$$

where $y_{i,j}$ is the observed plasma theophylline concentration in the i^{th} individual at the j^{th} observation, $f(t_{i,j}, \underline{\theta}_i)$ is predicted plasma concentration, t is the time of the observation, $\underline{\theta}$ is a vector of subject specific random effects, D is the theophylline dose, ka is the first-order absorption rate constant, $\frac{V}{F}$ is the bioavailability (F) normalized volume of distribution, $\frac{Cl}{F}$ is the bioavailability normalized clearance, and ke is the first-order elimination rate constant. As is convention, the subject specific random effects (i.e. ka , $\frac{V}{F}$, and $\frac{Cl}{F}$) were assumed to follow a multivariate lognormal distribution since the parameter values must be positive by definition and for computational convenience. Specifically:

$$\underline{\theta}_i = \left(\log\left(\frac{Cl}{F}\right)_i, \log\left(\frac{V}{F}\right)_i, \log(ka_i) \right)^T$$

$$\underline{\theta}_i | \underline{\mu}, \underline{\Omega} \stackrel{ind}{\sim} N(\underline{\mu}, \underline{\Omega}^{-1})$$

With prior distributions:

$$\underline{\mu} \sim N\left\{\log\left(\left(\frac{Cl}{F}\right)_0\right), \log\left(\left(\frac{V}{F}\right)_0\right), \log(ka_0), \Lambda^{-1}\right\}$$

$$\underline{\Omega}^{-1} \sim W(R, \nu)$$

It should be noted that all normal distributions were parameterized in terms of mean and precision. Finally, residual error ε was modeled with an additive (normally distributed) model. Specifically:

$$y_{i,j} - f(t_{i,j}, \underline{\theta}_i) \sim N\left(0, \frac{1}{\sigma^2}\right)$$

With prior distribution:

$$\frac{1}{\sigma^2} \sim G(a, b)$$

The effect of different priors on the parameters (i.e. mean and variance-covariance parameters) was also explored, comparing the results obtained using informative priors from the literature to using noninformative priors. Finally the Bayesian analysis results were compared to the results obtained using a frequentist maximum likelihood approach.

Materials and Methods:

Data

The data set consisted of plasma theophylline concentration-time profiles from 12 subjects for approximately 25 hours following a single oral administration of theophylline (Upton et. al, 1982). The complete data set is displayed in Appendix A.

Model Fitting

The Bayesian PK model was created using WinBUGS (Version 1.3)/PKBUGS (Version 1.1) with prior distributions of the parameters obtained from other literature sources (Ohnishi et. al 2003, see below) and then the model was exported to WinBUGS (Version 1.4.3)/Pharmaco (PKBUGS Version 2.0) interface to fit the published data set. The fitted model WinBUGS code is displayed in Appendix B. Three MCMC chains of 10,000 iterations were run for each model fitting from different overdispersed initial values for each chain. The first 4,000 iteration in each chain were excluded as the Metropolis algorithm acceptance rate is tuned to a value between 0.2 and 0.4 during the first 4,000 iterations. The convergence and mixing was assed from the history, autocorrelation, and BGR diagnostic plots. To determine model fit to the plasma concentration over time, custom plots were created in R displaying the observed data, the 95% point-wise confidence intervals (CI), and the 95% prediction intervals (PI) for each subject for both the informative and non-informative Bayesian data

analysis.

The frequentist maximum likelihood approach was conducted using the SAS® NLMIXED procedure (SAS® for Windows 9.1.3, Service Pack 4, SAS® Institute, Cary, NC).

The procedure maximizes the marginal likelihood of $y_{i,j}$ using adaptive Gaussian

quadrature integration approximation of the marginal likelihood, integrating out the random effects ($\underline{\theta}$, which enter the model nonlinearly) from the joint distribution of y and $\underline{\theta}$. The fitted PK model was identical to that used for the Bayesian analysis, except obviously there

were no prior distributions specified for $\underline{\mu}$ and Ω^{-1} , and the variance-covariance matrix

(Ω) was simplified to only a diagonal matrix. This matrix had to be simplified to a diagonal matrix due to numerical optimization issues. The fitted model SAS® code is displayed in Appendix C.

Bayesian Prior Distribution Construction:

Relationship between Log-Normal and Normal Distribution

Due to the relationship between the log-normal and normal distribution, it is realized that if $Y \sim N\left(\mu_N, \frac{1}{\sigma_N^2}\right)$ and $X = \exp(Y)$, then:

$$\mu_N = \log(E(X)) - \frac{1}{2} \log\left(1 + \frac{\text{Var}(X)}{(E(X))^2}\right)$$

$$\sigma_N^2 = \log\left(1 + \frac{\text{Var}(X)}{(E(X))^2}\right)$$

Therefore, priors for μ_N and σ_N^2 can be analytically calculated from papers where the sample mean and variance (standard deviation) of the parameters are given on the linear scale (i.e. NOT the log-scale).

INFORMATIVE PRIORS

Population mean priors ($\underline{\mu}$)

Consider the prior for the population mean parameters to be distributed multivariate

normal with mean vector $\left[\log\left(\left(\frac{Cl}{F}\right)_0\right), \log\left(\left(\frac{V}{F}\right)_0\right), \log(ka_0)\right]^T$ and diagonal precision

matrix Λ^{-1} , for simplicity. The bioavailability of theophylline is near 100%. Therefore, for

$\log\left(\left(\frac{Cl}{F}\right)_0\right)$ and $\log\left(\left(\frac{V}{F}\right)_0\right)$ the F value was fixed to 1.0. The clearance ($\frac{Cl}{F}$) and

volume of distribution (V/F) of theophylline has been well investigated in previous studies and reviewed (Ohnishi et. al 2003). The sample mean and variances were reported for these parameters in the Ohnishi et. al review (Appendix D) and therefore the mean value of the logarithm of clearance and the logarithm of volume can be obtained analytically as described above. The R-code used to conduct these calculations is displayed in Appendix E. The mean (μ_N) of the logarithm of clearances in these different studies had a mean of 3.76 with standard deviation 0.397. Thus, a $N(3.76, 6.35)$ prior was chosen for $\log\left(\left(\frac{Cl}{F}\right)_0\right)$. The mean (μ_N) of the logarithmic distribution volume in these same studies had a mean of 6.13 with standard deviation 0.162. Therefore a $N(6.13, 38.1)$ prior was chosen for $\log\left(\left(\frac{V}{F}\right)_0\right)$.

Theophylline absorption is fast and complete with the time to peak plasma concentrations (t_{max}) ranging from 1 to 2 hours (Hendeles et. al, 1984). Based on previous mean estimates of ke from the elimination half-life ($t_{1/2}$), i.e. $ke = \frac{\log(2)}{t_{1/2}}$, of 0.0941, the absorption rate constant can be determined iteratively by realizing that:

$$t_{max} = \frac{\log\left(\frac{ka}{ke}\right)}{ka - ke}$$

Thus ka_0 ranges from approximately 1.47 to 3.78, and the prior of $N(0.857, 18.0)$ was chosen for $\log(ka_0)$. In summary:

$$\underline{\mu} \sim N\left\{\begin{pmatrix} \log\left(\left(\frac{Cl}{F}\right)_0\right) \\ \log\left(\left(\frac{V}{F}\right)_0\right) \\ \log(ka_0) \end{pmatrix}, \Lambda^{-1}\right\} = N\left\{\begin{pmatrix} 3.76 \\ 6.13 \\ 0.857 \end{pmatrix}, \begin{pmatrix} 6.35 & 0 & 0 \\ 0 & 38.1 & 0 \\ 0 & 0 & 18.0 \end{pmatrix}\right\}$$

Population variance-covariance matrix priors (Ω)

The precision of the random effect Ω^{-1} was assumed to be distributed $W(R, \nu)$, and therefore Ω is distributed $IW(R, \nu)$. Then the expectation of $\Omega_{i,j}$ would be $\frac{R_{i,j}}{\nu - p - 1}$, where p is the dimension of the matrix (in this case $p = 3$). For simplicity, we also assumed Ω to be a diagonal matrix. As with the population mean priors, the variances of the logarithm of the clearance and the logarithm of the volumes can be analytically obtained from the Ohnishi et. al review. The sample mean and variances of clearance are reported in the above referenced studies and therefore the variance of the logarithmic clearances can be obtained analytically as described above. The variance (σ_N^2) of the logarithmic clearances

$(\Omega_{1,1})$ in this review had a mean of 0.187 from 14 studies, therefore we set the ν equal to 14 since it represents the equivalent prior sample size. It follows from above that $R_{1,1} = 1.87$.

The variance (σ_N^2) of the logarithm of volume ($\Omega_{2,2}$) in these different researches had a mean of 0.0523. With ν equal to 14 it follows that $R_{2,2} = 0.523$. Due to lack of data, the coefficient variability (C.V.) of absorption rate constant is assumed to be 25%, since this is a typical variability in the absorption rate constant ($\Omega_{3,3}$) for many oral products. From a lognormal distributed value, the C.V. is given by:

$$C.V. = \sqrt{\exp(\sigma^2) - 1}$$

Hence the mean variance (σ^2) of $\log(ka_i)$ would be 0.0606. With $\nu = 14$ it follows that $R_{3,3} = 0.606$. In summary,

$$\Omega^{-1} \sim W(R, \nu) = W\left(\begin{pmatrix} 1.87 & 0 & 0 \\ 0 & 0.523 & 0 \\ 0 & 0 & 0.606 \end{pmatrix}, 14\right)$$

Residual error prior

The residual error is from many sources, including measurement error and model misspecification, among others. The C.V. value for the theophylline bioanalytical assay (i.e. measurement error) among 100 laboratories was reported to be 11.1% at 10 $\mu\text{g/mL}$ (Amadeo et. al, 1998). To account for other sources of variability, the variance based on the above estimated measurement error was empirically multiplied by two. The C.V. of 11.1% at 10 $\mu\text{g/mL}$ implies the mean value of measurement error variance is about 1.23 mg^2/ml^2 . For an additive error model, after multiplication by two to account for other sources variability, the precision will be ~ 0.4 . If we also assume that the precision is ranged from 0 to 0.813 then the variance of the precision will be ~ 0.04 . Therefore, a $G(4,10)$ was chosen as a reasonable prior for σ^{-2} using an additive error model.

NON-INFORMATIVE PRIORS

The non-informative priors were constructed with the same mean estimates as used for the informative priors, however, the precisions was set to 0.0001 for the population mean

priors ($\underline{\mu}$) and ν was set reduced to 5 (to maintain positive variance expectations) for the precision matrix prior (Ω^{-1}). Finally the variance of the residual error precision (σ^{-2}), was increased 1000 fold as well, while maintaining the same mean estimates. Specifically:

$$\underline{\mu} \sim N \left\{ \begin{pmatrix} 3.76 \\ 6.13 \\ 0.857 \end{pmatrix}, \begin{pmatrix} 0.0001 & 0 & 0 \\ 0 & 0.0001 & 0 \\ 0 & 0 & 0.0001 \end{pmatrix} \right\}$$

$$\Omega^{-1} \sim W(R, \nu) = W \left(\begin{pmatrix} 0.187 & 0 & 0 \\ 0 & 0.0523 & 0 \\ 0 & 0 & 0.0606 \end{pmatrix}, 5 \right)$$

$$\sigma^{-2} \sim G(0.0004, 0.001)$$

Initial Values:

Three sets of initial values were generated for both the informative and non-informative Bayesian analysis (Appendix F). One set was centered at the expectation of the parameters based on the prior distributions (Appendix F.1) and the other two sets were overdispersed initial values, either high (Appendix F.2) or low (Appendix F.3) (NOTE: high and low initial values are in terms of the population means ($\underline{\mu}$) and variances(Ω, σ^2)).

The mean parameter estimates from the non-informative Bayesian analysis were used as initial values for the maximum likelihood analysis.

Results and Discussion:

The Metropolis acceptance rates reach the desired 0.2 to 0.4 value by about 1,000 iterations with the informative prior Bayesian analysis (Figure 1). The history and Gelbin-Rubin diagnostic plots of μ^* (Figure 2 and 3, respectively), as well as other parameters that are not shown, indicate that it is likely the Monte Carlo chains converged and mixed well with the informative priors. Furthermore, the autocorrelation of μ^* was relatively low, and the other parameters (not shown) had even less correlation (Figure 4). The Metropolis acceptance rate, history, Gelbin-Rubin diagnostic, and autocorrelations plots were similar for the non-informative prior Bayesian analysis. The first 4,000 samples of each 10,000 sample chain were thrown out for statistical analysis because the Metropolis algorithm is being tuned to the desired acceptance rate during these iterations. Additional iterations were not discarded as it appears the chains had converged by 4,000 iterations. The statistical summaries using both informative and non-informative prior distributions (Table 1 and 2, respectively) are very similar for μ^* parameter posterior mean values, the only noted exception is the mean value of $\mu[3]$. This is not unexpected since the prior information for the mean value of $\mu[3]$ (i.e. $\log(k\alpha_0)$) was based off of less specific data than the other μ^* parameter values. For all parameter estimates, the mean values are similar between the informative and non-informative prior distribution Bayesian analysis. Using informative priors, the 95% credible interval for all values of the covariance between parameters included zero, which suggests that substantial covariance was not apparent between the parameters. The model fits for each using both the informative and non-informative priors are displayed in Figure 5 and 6, respectively, and are very similar. Based on the 95% prediction intervals, the model appears to fit the data well. However, the 95% prediction intervals are modestly smaller for the non-informative prior Bayesian analysis, indicating that the informative priors used did not “match” the observed data completely.

The results from the frequentist analysis (Table 3) are also very similar to the results from the non-informative prior distribution Bayesian analysis (Table 2). In all cases the frequentist results are within one standard error of the non-informative prior Bayesian analysis mean results. The Bayesian 95% credible intervals for all parameters using the non-informative prior appear to be similar to the frequentist confidence intervals. The 95% confidence interval for the variance of random effects includes negative values. This might limit the applicability of frequentist approach in making inferences about variance components. Due to numerical issues with frequentist analysis, the covariance of the parameters of the model could not be determined and were fixed to zero. Numerical optimization and convergence problems are commonly encountered with non-linear mixed effects models, particularly with approaches that seek to directly maximize the joint marginal likelihood of the data, as the utilized NLMIXED procedure does. However, based on the results from the

Bayesian analysis, zero covariance between the model parameters may not be a poor assumption.

Figure 1. Metropolis acceptance rates for the Bayesian analysis with informative priors.

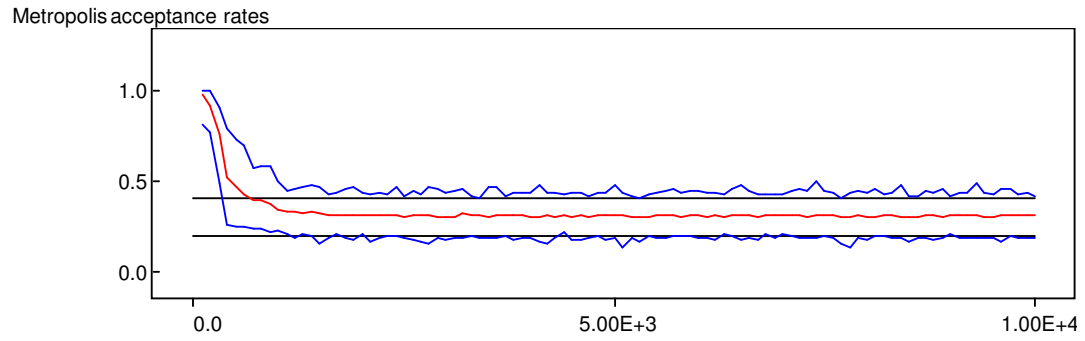


Figure 2. Iteration history plots of three 10000 iteration MCMC chains for μ beginning from overdispersed initial values for the Bayesian analysis with informative priors.

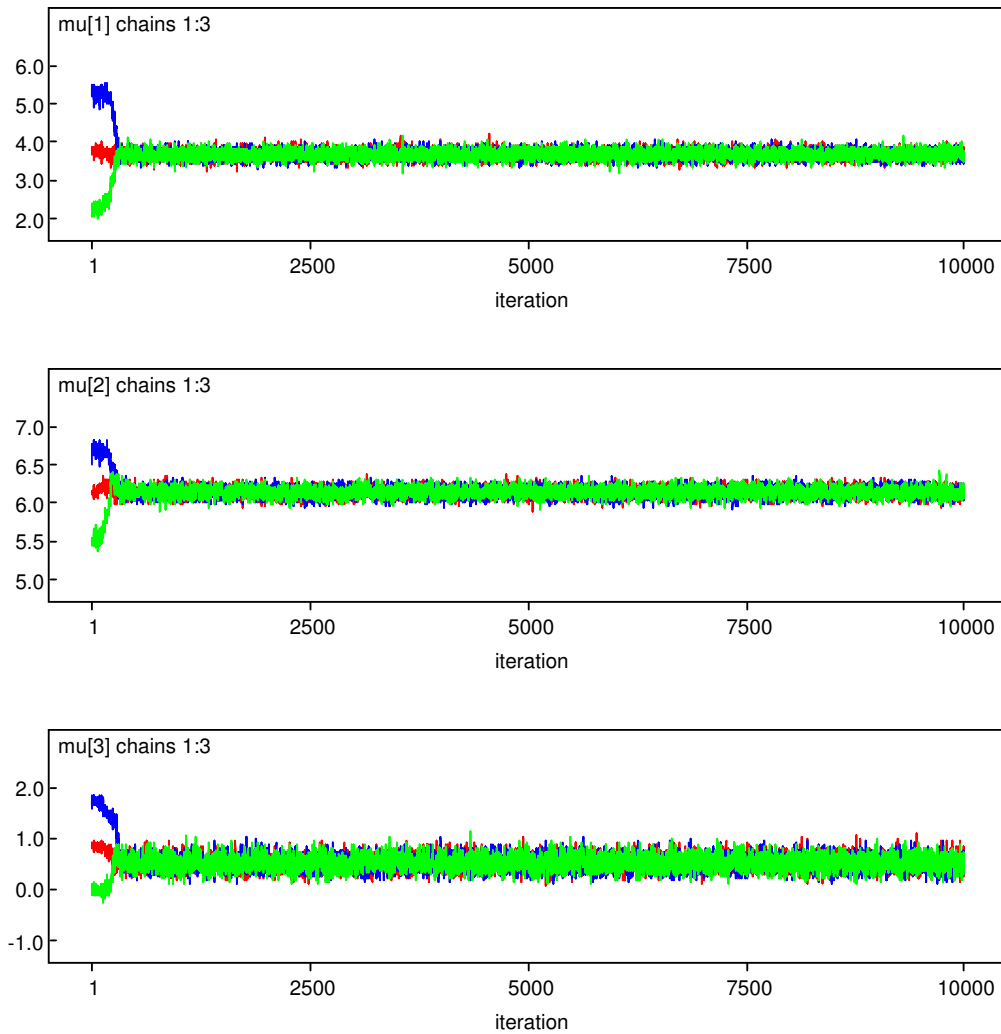


Figure 3. Gelman-Rubin diagnostic plots of three 10000 iteration MCMC chains for $\underline{\mu}$ beginning from overdispersed initial values for the Bayesian analysis with informative priors.

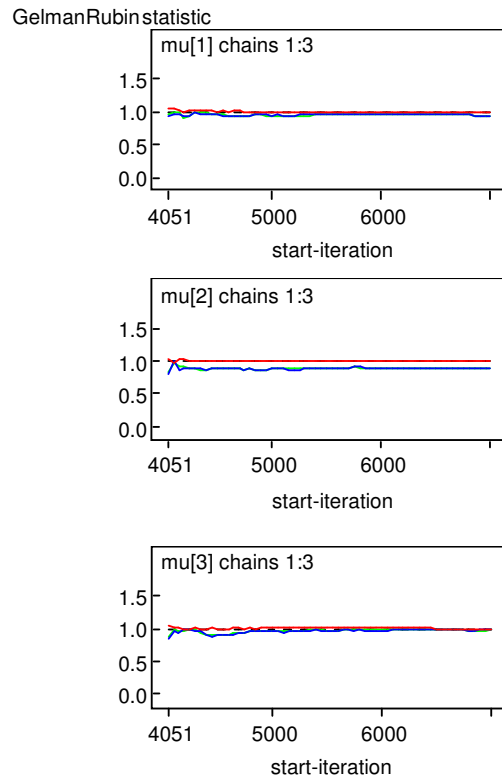


Figure 4. Autocorrelation plots of three 10000 iteration MCMC chains for $\underline{\mu}$ beginning from overdispersed initial values for the Bayesian analysis with informative priors.

Autocorrelation function

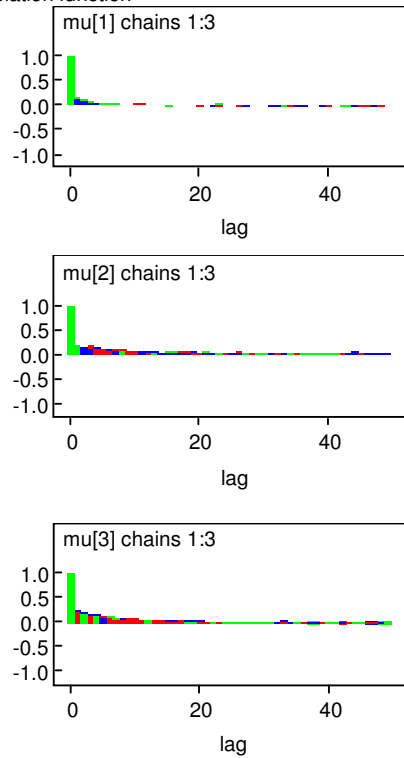


Figure 5. Model fit for the Bayesian analysis with informative priors. Symbols represent the observed plasma concentration, blue lines represent the point-wise 95% CI, and red lines represent the 95% PI.

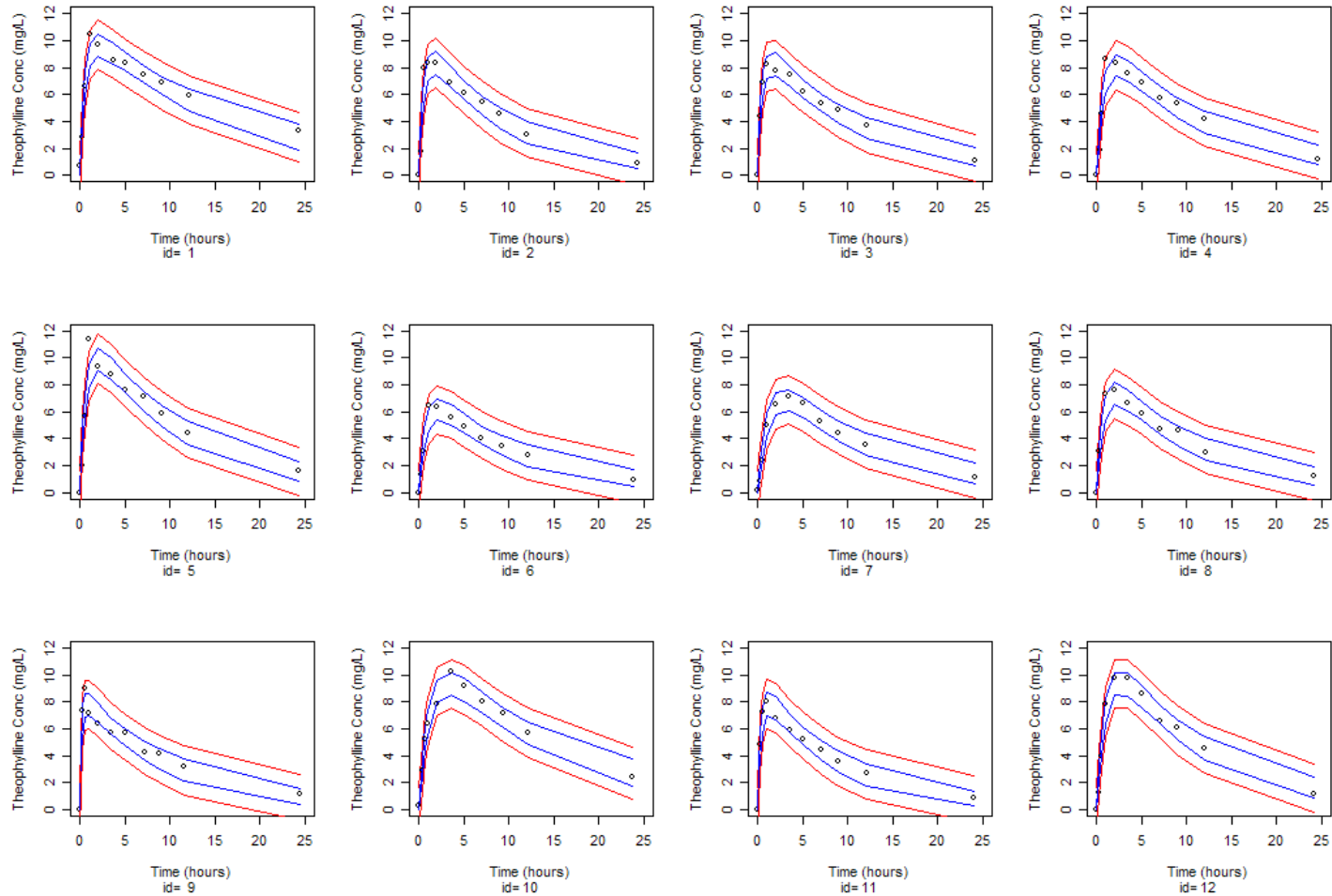


Figure 6. Model fit for the Bayesian analysis with non-informative priors. Symbols represent the observed plasma concentration, blue lines represent the point-wise 95% CI, and red lines represent the 95% PI.

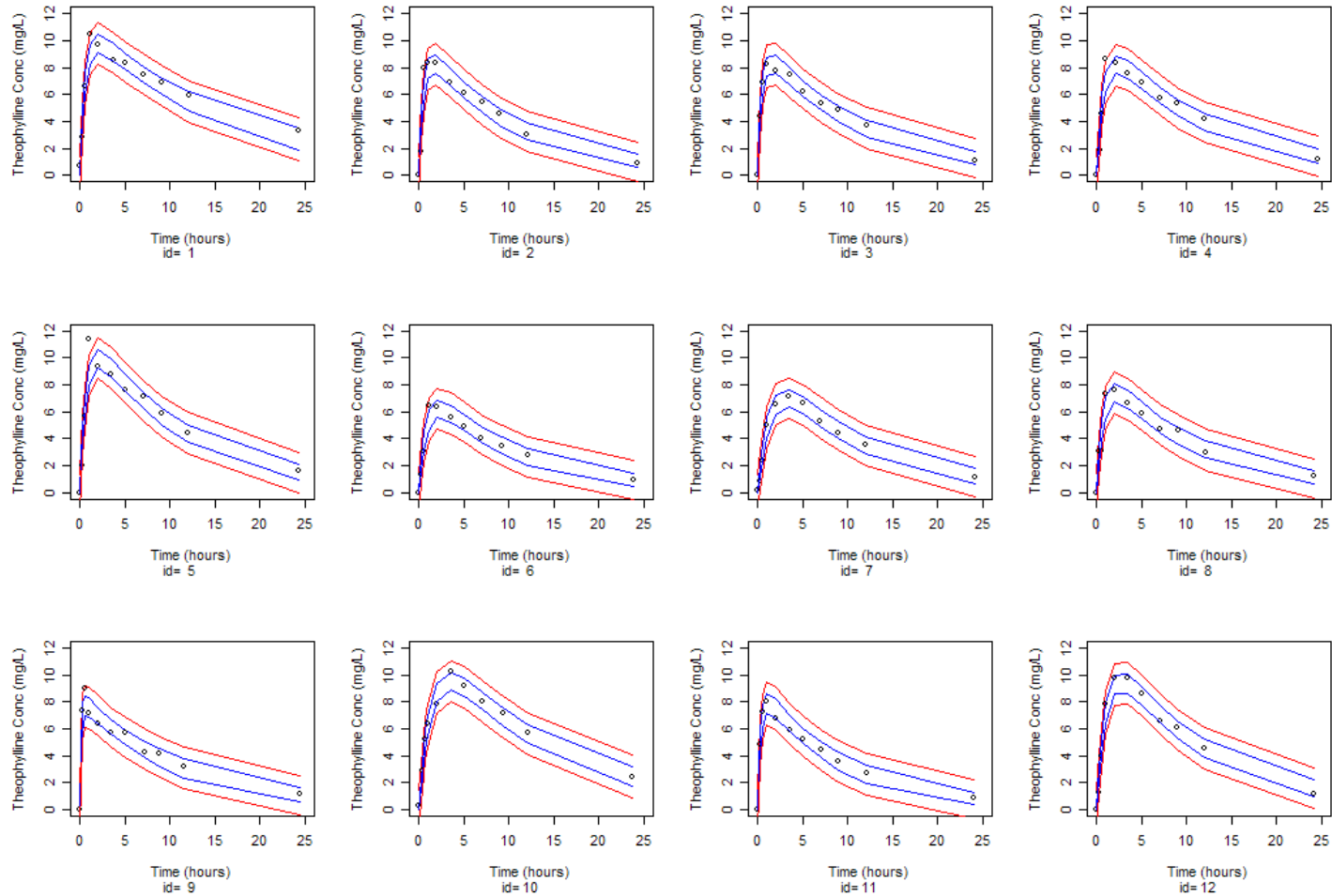


Table 1. Statistical summary of $\underline{\mu}$, $\underline{\Omega}$, and $\underline{\sigma}$ from three 10,000 iteration MCMC chains

for beginning from overdispersed initial value for the Bayesian analysis with informative priors (NOTE: the 1st 4,000 iterations from each chain are excluded due to tuning of the Metropolis algorithm acceptance rates during this period).

node	mean	sd	MC error	2.5%	5.0%	median	95.0%	97.5%	start	sample
mu[1]	3.676	0.1073	0.001272	3.466	3.501	3.675	3.853	3.889	4001	18000
mu[2]	6.14	0.05733	8.777E-4	6.027	6.046	6.14	6.235	6.253	4001	18000
mu[3]	0.5187	0.1294	0.001991	0.2748	0.3135	0.515	0.7357	0.7823	4001	18000
omega[1,1]	0.1314	0.04449	4.26E-4	0.07046	0.07648	0.1233	0.2132	0.238	4001	18000
omega[1,2]	0.01228	0.01628	1.575E-4	-0.01747	-0.01234	0.01127	0.03973	0.04813	4001	18000
omega[1,3]	-0.001796	0.04045	5.189E-4	-0.0866	-0.06795	-8.044E-4	0.06118	0.0775	4001	18000
omega[2,1]	0.01228	0.01628	1.575E-4	-0.01747	-0.01234	0.01127	0.03973	0.04813	4001	18000
omega[2,2]	0.03768	0.01264	1.482E-4	0.02016	0.02199	0.03531	0.06132	0.06844	4001	18000
omega[2,3]	-0.01148	0.02231	3.334E-4	-0.05833	-0.04771	-0.01096	0.02349	0.03222	4001	18000
omega[3,1]	-0.001796	0.04045	5.189E-4	-0.0866	-0.06795	-8.044E-4	0.06118	0.0775	4001	18000
omega[3,2]	-0.01148	0.02231	3.334E-4	-0.05833	-0.04771	-0.01096	0.02349	0.03222	4001	18000
omega[3,3]	0.2095	0.09242	0.00216	0.08769	0.09938	0.1913	0.3805	0.4388	4001	18000
sigma	0.8172	0.05835	9.75E-4	0.7133	0.7271	0.8135	0.9195	0.9404	4001	18000

Table 2. Statistical summary of $\underline{\mu}$, $\underline{\Omega}$, and $\underline{\sigma}$ from three 10,000 iteration MCMC chains

for beginning from overdispersed initial value for the Bayesian analysis with non-informative priors (NOTE: the 1st 4,000 iterations from each chain are excluded due to tuning of the Metropolis algorithm acceptance rates during this period).

node	mean	sd	MC error	2.5%	5.0%	median	95.0%	97.5%	start	sample
mu[1]	3.685	0.09044	0.001203	3.507	3.539	3.684	3.834	3.868	4001	18000
mu[2]	6.132	0.04902	9.997E-4	6.038	6.053	6.131	6.214	6.232	4001	18000
mu[3]	0.3927	0.2011	0.002687	-0.01102	0.0646	0.3938	0.7213	0.7897	4001	18000
omega[1,1]	0.08316	0.0405	5.833E-4	0.03374	0.03791	0.0736	0.1589	0.1858	4001	18000
omega[1,2]	0.02561	0.01608	1.779E-4	0.003958	0.006639	0.02239	0.05539	0.06576	4001	18000
omega[1,3]	-0.01707	0.06574	0.001057	-0.1636	-0.1273	-0.01235	0.07757	0.1033	4001	18000
omega[2,1]	0.02561	0.01608	1.779E-4	0.003958	0.006639	0.02239	0.05539	0.06576	4001	18000
omega[2,2]	0.02202	0.01145	2.065E-4	0.008554	0.009678	0.01938	0.04321	0.0511	4001	18000
omega[2,3]	-0.01659	0.03439	5.933E-4	-0.091	-0.07259	-0.01552	0.03489	0.04922	4001	18000
omega[3,1]	-0.01707	0.06574	0.001057	-0.1636	-0.1273	-0.01235	0.07757	0.1033	4001	18000
omega[3,2]	-0.01659	0.03439	5.933E-4	-0.091	-0.07259	-0.01552	0.03489	0.04922	4001	18000
omega[3,3]	0.4369	0.2274	0.004133	0.1653	0.1877	0.384	0.8567	1.011	4001	18000
sigma	0.6991	0.05038	7.896E-4	0.6084	0.6214	0.6964	0.786	0.8064	4001	18000

Table 3. Parameter estimates of $\underline{\mu}$, $\underline{\Omega}$, and $\underline{\sigma}$ from the frequentist maximum likelihood analysis.

Parameter	Estimate	Standard		DF	t Value	Pr > t	Alpha	Lower	Upper
		Error							
mu1	3.6895	0.08422		9	43.81	<.0001	0.05	3.4990	3.8800
mu2	6.1259	0.04590		9	133.47	<.0001	0.05	6.0221	6.2297
mu3	0.4582	0.1985		9	2.31	0.0464	0.05	0.009130	0.9073
omega11	0.07056	0.03466		9	2.04	0.0722	0.05	-0.00784	0.1490
omega22	0.01812	0.009551		9	1.90	0.0903	0.05	-0.00349	0.03973
omega33	0.4315	0.2030		9	2.13	0.0624	0.05	-0.02765	0.8907
sigma	0.6916	0.04893		9	14.14	<.0001	0.05	0.5809	0.8022

References:

- Amadeo J. Pesce, Mitchell Rashkin, and Uma Kotagal. (1998) Standards of laboratory practice: theophylline and caffeine monitoring. *Clin. Chem.* 44: 1124-1128.
- Hendeles L, Iafrate RP, Weinberger M. (1984) A clinical and pharmacokinetic basis for the selection and use of slow release theophylline products. *Clin Pharmacokinet.* Mar-Apr;9(2):95-135
- Ohnishi A, Kato M, Kojima J, Ushiyama H, Yoneko M, Kawai H (2003). Differential pharmacokinetics of theophylline in elderly patients. *Drugs Aging.* 20(1):71-84
- Upton RA, Thiercelin JF, Guentert TW, Wallace SM, Powell JR, Sansom L, Riegelman S. (1982). Intraindividual variability in theophylline pharmacokinetics: statistical verification in 39 of 60 healthy young adults. *J Pharmacokinet Biopharm.*;10(2):123-34.

Appendices:**Appendix A.** Theophylline data set.

id	dose = amt	time	dv = conc	abs
1	4020	0	.	1
1	.	0	0.74	1
1	.	0.25	2.84	1
1	.	0.57	6.57	1
1	.	1.12	10.5	1
1	.	2.02	9.66	1
1	.	3.82	8.58	1
1	.	5.1	8.36	1
1	.	7.03	7.47	1
1	.	9.05	6.89	1
1	.	12.12	5.94	1
1	.	24.37	3.28	1
2	4400	0	.	1
2	.	0	0	1
2	.	0.27	1.72	1
2	.	0.52	7.91	1
2	.	1	8.31	1
2	.	1.92	8.33	1
2	.	3.5	6.85	1
2	.	5.02	6.08	1
2	.	7.03	5.4	1
2	.	9	4.55	1
2	.	12	3.01	1
2	.	24.3	0.9	1
3	4530	0	.	1
3	.	0	0	1
3	.	0.27	4.4	1
3	.	0.58	6.9	1
3	.	1.02	8.2	1
3	.	2.02	7.8	1
3	.	3.62	7.5	1
3	.	5.08	6.2	1
3	.	7.07	5.3	1
3	.	9	4.9	1
3	.	12.15	3.7	1

3	.	24.17	1.05	1
4	4400	0	.	1
4	.	0	0	1
4	.	0.35	1.89	1
4	.	0.6	4.6	1
4	.	1.07	8.6	1
4	.	2.13	8.38	1
4	.	3.5	7.54	1
4	.	5.02	6.88	1
4	.	7.02	5.78	1
4	.	9.02	5.33	1
4	.	11.98	4.19	1
4	.	24.65	1.15	1
5	5806	0	.	1
5	.	0	0	1
5	.	0.3	2.02	1
5	.	0.52	5.63	1
5	.	1	11.4	1
5	.	2.02	9.33	1
5	.	3.5	8.74	1
5	.	5.02	7.56	1
5	.	7.02	7.09	1
5	.	9.1	5.9	1
5	.	12	4.37	1
5	.	24.35	1.57	1
6	4000	0	.	1
6	.	0	0	1
6	.	0.27	1.29	1
6	.	0.58	3.08	1
6	.	1.15	6.44	1
6	.	2.03	6.32	1
6	.	3.57	5.53	1
6	.	5	4.94	1
6	.	7	4.02	1
6	.	9.22	3.46	1
6	.	12.1	2.78	1
6	.	23.85	0.92	1
7	4950	0	.	1
7	.	0	0.15	1

7	.	0.25	0.85	1
7	.	0.5	2.35	1
7	.	1.02	5.02	1
7	.	2.02	6.58	1
7	.	3.48	7.09	1
7	.	5	6.66	1
7	.	6.98	5.25	1
7	.	9	4.39	1
7	.	12.05	3.53	1
7	.	24.22	1.15	1
8	4530	0	.	1
8	.	0	0	1
8	.	0.25	3.05	1
8	.	0.52	3.05	1
8	.	0.98	7.31	1
8	.	2.02	7.56	1
8	.	3.53	6.59	1
8	.	5.05	5.88	1
8	.	7.15	4.73	1
8	.	9.07	4.57	1
8	.	12.1	3	1
8	.	24.12	1.25	1
9	3100	0	.	1
9	.	0	0	1
9	.	0.3	7.37	1
9	.	0.63	9.03	1
9	.	1.05	7.14	1
9	.	2.02	6.33	1
9	.	3.53	5.66	1
9	.	5.02	5.67	1
9	.	7.17	4.24	1
9	.	8.8	4.11	1
9	.	11.6	3.16	1
9	.	4.43	1.12	1
10	5500	0	.	1
10	.	0	0.24	1
10	.	0.37	2.89	1
10	.	0.77	5.22	1
10	.	1.02	6.41	1

10	.	2.05	7.83	1
10	.	3.55	10.21	1
10	.	5.05	9.18	1
10	.	7.08	8.02	1
10	.	9.38	7.14	1
10	.	12.1	5.68	1
10	.	23.7	2.42	1
11	4920	0	.	1
11	.	0	0	1
11	.	0.25	4.86	1
11	.	0.5	7.24	1
11	.	0.98	8	1
11	.	1.98	6.81	1
11	.	3.6	5.87	1
11	.	5.02	5.22	1
11	.	7.03	4.45	1
11	.	9.03	3.62	1
11	.	12.12	2.69	1
11	.	24.08	0.86	1
12	5300	0	.	1
12	.	0	0	1
12	.	0.25	1.25	1
12	.	0.5	3.96	1
12	.	1	7.82	1
12	.	2	9.72	1
12	.	3.52	9.75	1
12	.	5.07	8.57	1
12	.	7.07	6.59	1
12	.	9.03	6.11	1
12	.	12.05	4.57	1
12	.	24.15	1.17	1

Appendix B. PKBUGS model code generated from export model command.

```

model {
  for (i in 1:n.ind) {
    for (j in off.data[i]:(off.data[i + 1] - 1)) {
      data[j] ~ dnorm(model[j], tau)
      preddata[j]~dnorm(model[j], tau)
      model[j] <- pk.model(1, theta[i, 1:p], time[j], hist[off.hist[i]:(off.hist[i + 1] -
        1),
          1:n.col], pos[j])
    }
    theta[i, 1:p] ~ dnorm(theta.mean[i, 1:p], omega.inv[1:p, 1:p])
    theta.mean[i, 1] <- mu[1]
    theta.mean[i, 2] <- mu[2]
    theta.mean[i, 3] <- mu[3]
  }
  tau ~ dgamma(tau.a, tau.b)
  sigma <- 1 / sqrt(tau)
  mu[1:q] ~ dnorm(mu.prior.mean[1:q], mu.prior.precision[1:q, 1:q])
  omega.inv[1:p, 1:p] ~ dwish(omega.inv.matrix[1:p, 1:p], omega.inv.dof)
  omega[1:p, 1:p] <- inverse(omega.inv[1:p, 1:p])
}

```

Appendix C. SAS model code.

```

proc nlmixed data=theoph;
  parms beta1=3.7 beta2=6.1 beta3=0.39
    s2b1=0.084 s2b2=0.022 s2b3=0.44 s2=0.49;
  cl = exp(beta1 + b1);
  v = exp(beta2 + b2);
  ka = exp(beta3 + b3);
  ke = cl/v;
  pred = ((dose*1000*ka)/(v*(ka - ke)))*(exp(-ke*time)-exp(-ka*time));
  model conc ~ normal(pred,s2);
  random b1 b2 b3 ~ normal([0,0,0],[s2b1,0,s2b2,0,0,s2b3]) subject=subject;
run;

```

Appendix D. Data tables for prior constructions from Oshini et. al, 2003.

Age (y)	Sex	No. of subjects	Dose (mg/kg)	CL (mL/h/kg)	$t_{1/2}$ (h)	Vd (L / kg)	CL_m (L/h/kg) $\times 10^4$			CL_R (L /h/kg) $\times 10^4$ of theophylline	Reference
							3-MX	1-MU	1,3-DMU		
23-40	M	8	5	50.8 \pm 3.7	7.1 \pm 0.5	0.48 \pm 0.03	108.8 \pm 14.7	131.4 \pm 13.5	186.5 \pm 12.8	81.4 \pm 9.8	57
65-76	M	8	5	40.5 \pm 4.1	9.5 \pm 0.6	0.51 \pm 0.03	78.3 \pm 9.5	97.1 \pm 13.1	160.7 \pm 12.5	69.1 \pm 9.0	57
22-33	F	8	5	50.4 \pm 4.5	6.5 \pm 0.5	0.43 \pm 0.02	93.4 \pm 11.0	129.4 \pm 22.7	187.1 \pm 13.1	94.2 \pm 8.6	57
65-78	F	8	5	38.5 \pm 2.6	7.8 \pm 0.6	0.40 \pm 0.02	58.8 \pm 9.4	99.1 \pm 11.9	147.7 \pm 10.2	79.1 \pm 5.6	57
20-27	M	6	6	40.7 \pm 12.4	8.0 \pm 2.0	0.43 \pm 0.02	56.4 \pm 14.2	88.6 \pm 30.3	174.7 \pm 54.1	87.4 \pm 27.8	58
65-77	M	6	5	34.7 \pm 4.9	8.3 \pm 1.0	0.40 \pm 0.03	48.7 \pm 16.6	72.3 \pm 15.0	157.1 \pm 17.7	68.9 \pm 9.1	58
20-33	M	10	4.5	48.8 \pm 2.6	7.8 \pm 0.4	0.54 \pm 0.02	73.3 \pm 6.6	159.9 \pm 11.8	205.6 \pm 13.9	48.7 \pm 5.2	59
68-88	M	10	4.5	34.0 \pm 2.5	10.7 \pm 0.9	0.50 \pm 0.02	41.8 \pm 3.8	103.5 \pm 11.8	158.6 \pm 11.1	36.0 \pm 3.4	59
20-30	M	6	4.8	37.3 \pm 5.8	8.8 \pm 1.4	0.46 \pm 0.03	50.7 \pm 10.7	101.5 \pm 22.2	150.6 \pm 30.9	70.4 \pm 6.0	^a
65-77	M & F	10 (7 & 3)	2.6-6.2	32.0 \pm 16.1	11.4 \pm 6.8	0.41 \pm 0.09	35.4 \pm 22.6	77.9 \pm 39.5	144.9 \pm 74.9	62.2 \pm 44.7	^a

No. of subjects	CL (mL/h/kg)	$t_{1/2}$ (h)	Vd (L / kg)	Reference
10	86.0 \pm 8.4	5.7 \pm 0.6	0.67 \pm 0.05	59
10	72.4 \pm 8.0	5.6 \pm 0.5	0.55 \pm 0.04	59
10	67.8 \pm 6.3	5.4 \pm 0.4	0.48 \pm 0.02	72
10	54.7 \pm 4.7	6.1 \pm 0.4	0.44 \pm 0.02	72

Appendix E. R-code used for prior calculations.

```
n <- c(8,8,8,8,6,6,10,10,6,10,10,10,10)
```

```
cl <- c(50.8,40.5,50.4,38.5,40.7,34.7,48.8,34.0,37.3,32.0,86.0,72.4,67.8,54.7)
```

```
cl.se <- c(3.7,4.1,4.5,2.6,12.4,4.9,2.6,2.5,5.8,16.1,8.4,8.0,6.3,4.7)
```

```
cl.var <- (cl.se*sqrt(n))^2
```

```
mu.cl <- log(cl) - 0.5*log(1 + ((cl.var)/(cl^2)))
```

```
sigma.cl <- log(1 + ((cl.var)/(cl^2)))
```

```
mean(mu.cl)
```

```
[1] 3.75719
```

```
sd(mu.cl)
```

```
[1] 0.3965297
```

```
mean(sigma.cl)
```

```
[1] 0.1865652
```

```
sd(sigma.cl)
```

```
[1] 0.3260218
```

```
v <- 1000*(c(0.48, 0.51, 0.43, 0.4, 0.43, 0.4, 0.54, 0.5, 0.46, 0.41, 0.67, 0.55, 0.48,  
0.44))
```

```
v.se <- 1000*(c(0.03, 0.03, 0.02, 0.02, 0.02, 0.03, 0.02, 0.02, 0.03, 0.09, 0.05, 0.04,  
0.02, 0.02))
```

```
v.var <- (v.se*sqrt(n))^2
```

```
mu.v <- log(v) - 0.5*log(1 + ((v.var)/(v^2)) )
```

```
sigma.v <- log(1 + ((v.var)/(v^2)))
```

```
mean(mu.v)
```

```
[1] 6.134554
```

```
sd(mu.v)
```

```
[1] 0.1623317
```

```
mean(sigma.v)
```

```
[1] 0.05232289
```

```
sd(sigma.v)
```

```
[1] 0.0989878
```

```
t12 <- c(7.1, 9.5, 6.5, 7.8, 8.0, 8.3, 7.8, 10.7, 8.8, 11.4, 5.7, 5.6, 5.4, 6.1)
```

```
ke <- log(2)/t12
```

```
mean(ke)
```

```
[1] 0.09405957
```

Appendix F. Initial values.**Appendix F.1.** Centered initial values*.

$$\underline{\mu} = \begin{pmatrix} 3.76 \\ 6.13 \\ 0.857 \end{pmatrix}$$

$$\sigma^{-2} = 0.4$$

INFORMATIVE PRIORS

$$\Omega^{-1} = \begin{pmatrix} 7.49 & 0 & 0 \\ 0 & 26.8 & 0 \\ 0 & 0 & 23.1 \end{pmatrix}$$

NON-INFORMATIVE PRIORS

$$\Omega^{-1} = \begin{pmatrix} 26.7 & 0 & 0 \\ 0 & 95.6 & 0 \\ 0 & 0 & 82.5 \end{pmatrix}$$

Appendix F.2. Over-dispersed (high) initial values*.

$$\underline{\mu} = \begin{pmatrix} 5.35 \\ 6.78 \\ 1.80 \end{pmatrix}$$

$$\sigma^{-2} = 0.1$$

INFORMATIVE PRIORS

$$\Omega^{-1} = \begin{pmatrix} 1.87 & 0 & 0 \\ 0 & 6.69 & 0 \\ 0 & 0 & 5.78 \end{pmatrix}$$

NON-INFORMATIVE PRIORS

$$\Omega^{-1} = \begin{pmatrix} 6.68 & 0 & 0 \\ 0 & 23.9 & 0 \\ 0 & 0 & 20.6 \end{pmatrix}$$

Appendix F.3. Over-dispersed (low) initial values*.

$$\underline{\mu} = \begin{pmatrix} 2.17 \\ 5.48 \\ -0.0875 \end{pmatrix}$$

$$\sigma^{-2} = 1.6$$

INFORMATIVE PRIORS

$$\Omega^{-1} = \begin{pmatrix} 29.9 & 0 & 0 \\ 0 & 76.5 & 0 \\ 0 & 0 & 92.4 \end{pmatrix}$$

NON-INFORMATIVE PRIORS

$$\Omega^{-1} = \begin{pmatrix} 107 & 0 & 0 \\ 0 & 382 & 0 \\ 0 & 0 & 330 \end{pmatrix}$$

*In all cases $\underline{\theta}_i$ was set equal to $\underline{\mu}$ for all subjects.