Comparison of Models for Baseline

UPPSALA UNIVERSITET

C. Dansirikul, H.E. Silber, M.O. Karlsson

Division of Pharmacokinetics and Drug Therapy, Uppsala University, Uppsala, Sweden

Introduction

Baseline response can be modeled in different ways. While a few approaches have been used for handling baseline responses, an exploration of their performance has not yet been reported. Estimation of the typical value and interindividual variability (IIV) of baseline in the population (Model 1, M1) is considered the gold standard. Inclusion of the observed baseline response as a covariate, acknowledging the residual variability (M2) has been suggested as an alternative.[1] A more general version of M2, also taking into account IIV in baseline is M3. A fourth method is subtraction of baseline from observed responses (M4). In this study, these four methods were compared using Monte Carlo simulations.

Materials and Methods

Baseline models:

$$M1: BL_i = \hat{BL} \cdot e^{\eta}$$

$$M2: BL_i = BL_{i,o} \cdot e^{\eta \cdot \sigma}$$

$$M3: BL_i = \left(\hat{BL} \cdot \frac{\sigma^2}{\omega^2 + \sigma^2} + BL_{i,o} \cdot \frac{\omega^2}{\omega^2 + \sigma^2} \right) \cdot e^{\eta \cdot \sigma \cdot \frac{\omega^2}{\omega^2 + \sigma^2}}$$

$$M4: BL_i = BL_{i,o}$$

where:

BL_i = individual baseline
$\hat{BL}$ = the typical baseline parameter
$BL_{i,o}$ = individual observed baseline
η = interindividual random variability with mean of 0 and variance of ω^2
σ = standard deviation of residual variability

Note: In M2 & M3, variance of η in '$e^{\eta \cdot \sigma}$' was fixed to 1. This allows the residual variability in baseline to be varied among individuals, but to be the same within each individual.

Simulations: PD responses over a single dosing interval were simulated under 22 designs (Table 1), each containing 100 datasets. All analyses were conditioned on individual PK profiles ('sequential' PKPD modelling). Indirect effect models with a drug acting through stimulation or inhibition of the production of response (R), according to an Emax model, were employed (Table 1). Baseline response was simulated using M1 (see Figure 1 for example).

Table 1: Designs, model and model parameters used for simulations

No. Subjects	No. Samples per subject	Stimulation K_{in} Dose (mg)	Stimulation K_{in} RV (%)	Inhibition K_{in} Dose (mg)	Inhibition K_{in} RV (%)
25	4	2	30	4	10
	7				
	13				
50	4				
	7				
	13				
100	4				
	7				
	13				
50	13	2	10	-	-
			50	-	-
		1	30	-	-
		4	30	-	-

RV = residual variability

PK model: $Cp = \frac{D}{V} \cdot e^{-\frac{CL}{V} t}$

PD models:

$$R_i(0) = \hat{BL} \cdot e^{\eta}$$

$$\frac{dR}{dt} = k_{in} \cdot \left(1 + \frac{E_{max} \cdot C_p}{EC_{50} + C_p}\right) - k_{out} \cdot R$$

$$\frac{dR}{dt} = k_{in} \cdot \left(1 - \frac{C_p}{EC_{50} + C_p}\right) - k_{out} \cdot R$$

$$Y = \log(R) + \varepsilon$$

where the typical parameter values (% IIV) are:

CL (L/hr)	6.93 (30%)
V (L)	100 (30%)
Kout (h^{-1})	0.4 (30%)
BL	5 (30%)
E_{max}	12 (30%)
EC_{50} (µg/L)	10 (30%)

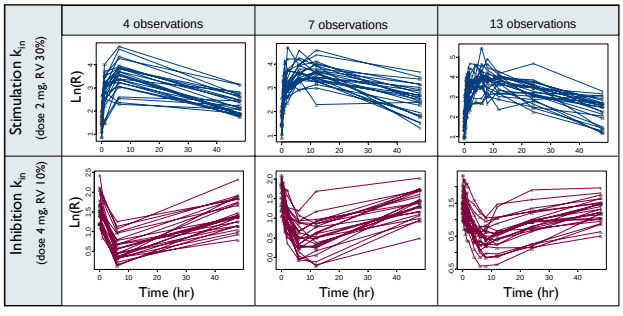

Figure 1: Simulated PD response data (25 subjects, on log scale)

Reference:

[1] Lewis Sheiner. NONMEM Tip#16 – April 2, 2003 – Modeling a "baseline" component and an additive "drug" component.

Materials and Methods (cont.)

Estimations: Four approaches to handling baseline response (M1, M2, M3, and M4) were examined using the FO and the FOCE estimation methods with NONMEM VIβ. Response data were log transformed.

Analyses: The mean error (ME, %) and root mean squared error (RMSE, %) in relation to the values used in the simulation were computed. ME and RMSE were used as an indicator of bias and imprecision, respectively. |ME| and RMSE from 8 methods (4 baseline approaches x 2 estimation methods) were ranked within the same design, the lower the value the better. Average rank of each method from all designs was reported.

Results

Bias & Imprecision

- The methods with least bias (ME) in estimates of the typical value and IIV parameters was M3 with the FOCE method (Table 2). When the FO method was used, M2 displayed smallest bias. M4 had the largest bias (Figure 2) for both the FO and the FOCE methods. For all cases, the FOCE method gave smaller bias as compared to the FO method.
- The difference in imprecision (RMSE) of parameter estimates between the FO and the FOCE methods was small. The smallest imprecision was noted with use of M1 and increased, in order, with M3, M2 and M4, respectively.
- Smaller imprecision (RMSE) in estimates of typical parameters and IIV parameters was noticed when the number of subjects, as well as the number of observations, and dose size, increased. This, however, was not seen with bias (ME) (results not shown).

Table 2: Average rank of bias (|ME|) and imprecision (RMSE)

Method	Model	Emax typical value	Emax IIV	EC50 typical value	EC50 IIV	Kout typical value	Kout IIV	BL typical value	BL IIV	RV	Average
Average Rank of |ME| from all designs											
1	FO-M1	5.6	2.4	6.1	2.7	5.6	4.5	6.4	4.5	4.1	4.7
2	FO-M2	4.8	2.6	5.7	3.5	5.7	5.5	3.4[a]	2.6[c]	4.8	4.3
3	FO-M3	6.8	4.2	4.5	5.7	4.1	4.3	5.5	3.9	3.8	4.8
4	FO-M4	7.1	7.6	7.5	6.8	7.3	5.5	3.4[a]	7.1[b]	5.2	6.4
5	FOCE-M1	2.8	4.2	2.7	2.2	3.9	3.8	4.8	4.5	4.5	3.7
6	FOCE-M2	3.3	4.9	2.8	3.1	4.5	4.3	3.4[a]	2.2[c]	4.8	3.7
7	FOCE-M3	2.2	3.5	2.9	4.5	2.0	2.5	5.9	3.9	2.7	3.3
8	FOCE-M4	3.2	6.7	3.7	7.4	2.8	5.6	3.4[a]	7.1[b]	6.1	5.1
Average Rank of RMSE from all designs											
1	FO-M1	3.2	1.2	4.4	2.0	5.4	2.6	4.3	4.6	3.2	3.4
2	FO-M2	4.2	4.2	4.9	3.3	6.3	4.8	5.5[a]	2.1[c]	4.9	4.5
3	FO-M3	4.8	3.8	3.5	3.9	4.5	2.8	4.5	6.2	3.8	4.2
4	FO-M4	3.5	7.1	6.2	7.2	7.9	6.8	5.5[a]	5.1[b]	6.1	6.2
5	FOCE-M1	3.4	2.3	3.5	2.9	3.1	3.5	1.9	4.6	3.7	3.2
6	FOCE-M2	4.7	5.2	4.1	4.0	3.5	5.5	5.5[a]	1.9[c]	4.6	4.3
7	FOCE-M3	5.3	4.8	3.5	4.9	1.7	3.1	3.1	6.3	2.9	4.0
8	FOCE-M4	6.9	7.5	6.0	7.8	3.5	7.0	5.5[a]	5.1[b]	6.7	6.2

[a] = based on the geometric mean of the observed baseline
[b] = based on the standard deviation of the observed baseline
[c] = calculated using $IIV_{M2} = \sqrt{IIV_{M4}^2 - \sigma_{M2}^2}$

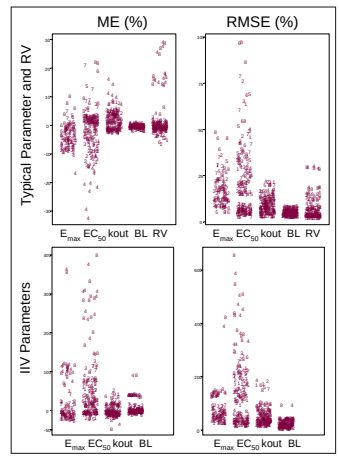

Figure 2: ME (%) and RMSE (%)

Discussion

- M2 estimates individual baseline by correcting the observed value with some error. This error is assumed to have the same magnitude as RV. If RV is large, estimated individual baseline is, then, less reliable. This might also affect the estimates of other parameters.
- M3 offers a more general way of handling baseline as both IIV and RV are accounted during the estimation of individual baseline. Fraction of IIV and RV to their sum is used to weigh the contribution of the observed and the typical value on the estimates of baseline, respectively. If RV is relatively small as compared to IIV, then the observed baseline is given more weight. Hence, M3 is expected to perform as well as M2. On the other hand, if IIV is relatively small as compared to RV, the typical baseline is given more weight. M3 is expected to perform as well as M1.
- Our results showed that the performance of M3 was most similar to M1, and slightly better than M2. The use of subtraction of baseline from other observations (M4), however, led to the largest bias and imprecision in parameter estimates, especially for estimates of IIV.