



UPPSALA
UNIVERSITET

New Estimation Methods in NONMEM 7: Evaluation of Bias and Precision

Åsa M. Johansson, Sebastian Ueckert, Elodie L. Plan, Andrew C. Hooker, Mats O. Karlsson
Dept of Pharmaceutical Biosciences, Uppsala University, Sweden

Background

NONMEM is the most widely used software for population PKPD analyses. The latest version, NONMEM 7 (NM7), includes several new sampling-based estimation algorithms in addition to the classical methods. In this study the bias and precision of the new methods have been evaluated whereas the runtimes and the methods' sensitivity to initial estimates have been investigated elsewhere [See PAGE poster 1915].

Objective

To investigate the performance with respect to bias and precision of the estimation methods available in NM7 for a diverse set of PKPD models.

Methods

Models: Five PKPD models with different types of PD data were included in the study: Continuous (1), Binary (2), Ordered Categorical (3), Repeated Time-To-Event (4) and Count (5). The models were investigated with both standard parameterization and MU parameterization.

Algorithms: The estimation methods investigated were: FOCE (for model 1) LAPLACE (for models 2-5), ITS, SAEM, IMP, IMPMAP and BAYES. All estimation settings were kept as default except that a convergence criteria (CTYPE=3) was applied to all new methods.

Stochastic Simulations and Estimations (SSE): Each of the models were used to simulate 500 datasets, except model 4 for which only 100 datasets were simulated. The datasets were then analyzed with the different estimation algorithms and the different parameterizations.

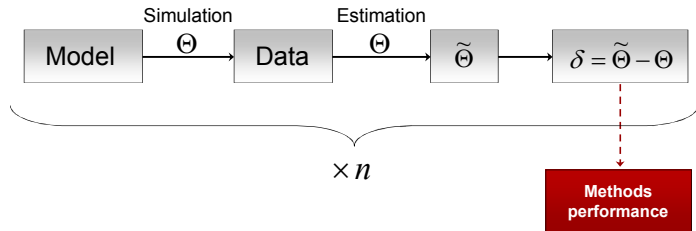


Figure 1: Schematic representation of the simulation and estimation study performed where $n=500$ for models 1-3 and 5 and $n=100$ for model 4.

Bias and precision: If $\Delta_{pa} = \{\delta_{pa1}, \delta_{pa2}, \dots, \delta_{pan}\}$ for parameter p and algorithm a then the subset $\Delta_{pa,90\%}$ contains all δ_{pai} between the 5th and the 95th percentiles of Δ_{pa} and if $\Delta_p = \{\delta_{p1}, \delta_{p2}, \delta_{p3}, \delta_{p4}, \delta_{p5}, \delta_{p6}\}$ (estimation error for all methods) then the subset $\Delta_{p,out}$ contains all values δ_{pi} below the 5th or above the 95th percentile of Δ_p . The bias and precision of the estimation methods were investigated in two ways for each model: (i) calculation and comparison of mean RMSE_{100%} and RMSE_{90%} relative to the FOCE or LAPLACE RMSE_{100%} and RMSE_{90%} (eq 1-3) for each algorithm using Δ_{pa} and $\Delta_{pa,90\%}$ respectively and (ii) investigation of how much each estimation method contributed to the fraction of estimates below the 5th or above the 95th percentile using mean $\Delta_{p,out}$ over all parameters and all methods of each model.

$$RMSE_{pa} = \sqrt{Mean(\delta^2)} \quad (\text{Eq.1})$$

$$RRMSE_{pa} = \frac{RMSE_{pa}}{RMSE_{pFOCE/LAPLACE}} \quad (\text{Eq.2})$$

$$RRMSE_a = Mean(RRMSE_a) \quad (\text{Eq.3})$$

p : parameters, a : algorithms

Conclusions

- ❖ Standard performance measures heavily driven by outliers
- ❖ Outlier robust measures show little difference between methods, apart from BAYES, which is not intended for point estimation
- ❖ MU parameterization did not impact estimates markedly

Results

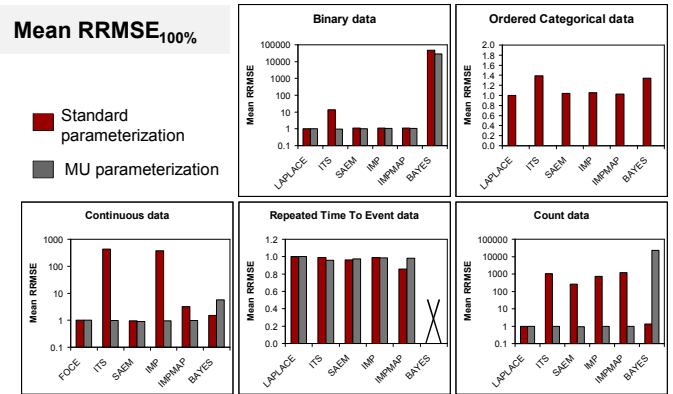


Figure 2: Mean RRMSE_{100%} relative FOCE/LAPLACE for the different estimation methods with standard parameterization and MU parameterization. MU parameterization was not possible for model 3 (ordered categorical data) and there were not enough successful runs for the BAYES method for model 4 (repeated time to event data).

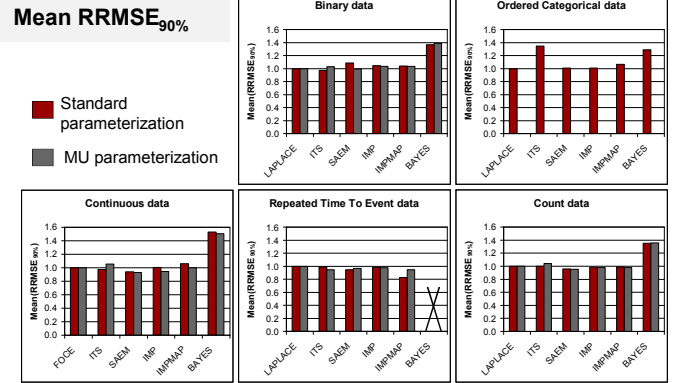


Figure 3: Mean RRMSE_{90%} relative FOCE/LAPLACE for the different estimation methods with standard parameterization and MU parameterization.

Extreme percentiles

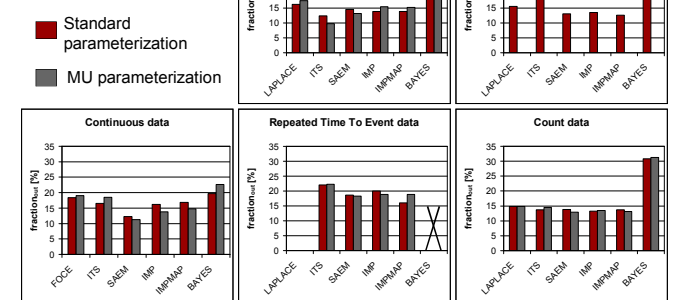


Figure 4: The fraction of each estimation method that contributes to the total fraction of estimates below the 5th or above the 95th percentile of all estimates.