Evaluation of saemix extension

Authors: Emmanuelle Comets[1,2], Belhal Karimi[3]

[1] INSERM, IAME, UMR 1137, F-75018 Paris, France; Univ Paris Diderot, Sorbonne Paris Cité, F-75018 Paris, France

[2] INSERM, CIC 1414, 35700 Rennes, France; Univ Rennes-1, 35700 Rennes, France

[3]

1 Introduction

Objectives the main objective is to evaluate the extensions for modelling non continuous data in the **saemix** package. A secondary objective is to evaluate the estimation of the individual parameters for different types of models (continuous and non-continuous responses).

2 Statistical methods

2.1 Statistical models

Let the random variable Y_{ij} denote the observation of the continuous longitudinal data at time t_{ij} for subject $i = 1, ..., N$ and measurement $j = 1, ..., n_i$ and let $\mathbf{Y}_i$ be the n_i-dimensional vector of all repeated measurements for subject i, that is, $\mathbf{Y}_i = (y_{i1}, y_{i2}, ..., y_{in_i})'$. A general formulation of the non-linear mixed-effects model (NLMEM) for the observation y_{ij} can be written as follows:

$$\begin{cases} y_{ij} = f(x_{ij}, \mu, \eta_i) + g(x_{ij}, \mu, \eta_i, \sigma)\, \epsilon_{ij} \\ \theta_i = h(\mu, \eta_i) \\ \eta_i \sim N(0, \Omega) \\ \epsilon_{ij} \sim N(0, 1) \end{cases} \tag{1}$$

where f is the structural model, θ_i denotes the individual parameters, which are related through a function h to μ, the p-dimensional vector containing the fixed effects, and η_i, the q-dimensional vector containing the random effects. ϵ_{ij} is a random variable assumed to be normally distributed. σ denotes variance parameters entering the function g, which expresses the standard deviation of the measurement error and is generally either constant (homoscedastic variance) or a function of f. The random effects η_i and the residual errors ϵ_{ij} are assumed to be independent for different subjects and to be independent of each other for the same subject. We will also denote by ξ_i the sampling times (or design) for subject i.

Without loss of generality, we assume that parameters are estimated by maximum likelihood. In non-linear mixed effect models, the likelihood associated with 1 is intractable as individual likelihoods need to integrate out the unknown parameters θ_i over their distribution $\mathcal{D}_\theta$. In this study, we used the SAEM algorithm, an algorithm of the EM family, to obtain parameter estimates [Kuhn and Lavielle, 2005, Lixoft, 2013]. Alternatively, methods based on model linearisation and using Newton gradient-based algorithms can be used [Boeckmann et al., 1994]. During the stochastic approximation phase, the conditional distribution of the parameters is obtained as it is the distribution in which the unknown parameters η_i are imputed to obtain a complete dataset from which the conditional log-likelihood is derived [Kuhn and Lavielle, 2005]. The estimated individual parameters, called empirical Bayes estimates (EBE), can be defined as the mode or the median of the conditional distribution. With linearisation-based algorithms, EBE are obtained by minimising a Bayesian criterion.

2.2 Standard errors of estimation

In maximum likelihood theory, the MLE $\hat{\theta}$ of θ is asymptotically normally distributed with mean θ and asymptotic covariance matrix given by the inverse of the Fisher information matrix M_F. The asymptotic SE of parameters are then derived from the estimated covariance matrix, which is usually obtained through a first-order approximation of the model for continuous responses [Retout et al., 2002].

For non-continuous responses however, the linearisation based methods have been shown to provide poor estimates of the SE. Exact approaches have been proposed to compute the Fisher information matrix, such as adaptive Gaussian quadrature [**?**] or Markov Chain Monte Carlo integration [**?**], in combination with Monte-Carlo sampling. Alternatively, computationnally intensive approaches like likelihood profiling [], bootstrap [] or Sampling Importance Resampling [] have been suggested but have never been evaluated with non-continuous data.

2.3 Estimation of individual parameters

3 Simulation studies

3.1 Objective and evaluation criteria

Evaluate individual parameters, shrinkage
Evaluate individual predictions (ie maybe individual parameters are off but predictions are similar)

3.2 Continuous data

Simulations have shown good performances for the SAEM algorithm implemented in the **saemix** package [**?**], in the same simulation examples as used by Plan et al. to evaluate **nlme**, **NONMEM** and **Monolix** in 2012 [**?**]. Few papers have evaluated the quality/reliability of the individual parameter estimates: **XXcheckBiblioXXX**

Evaluation of individual parameters:

- Illiadis 1985: Bayesian estimate of clearance both in subjects used in the estimation of the population parameters and in routine subjects [**?**]
  - infusion of high dose methotrexate, preceded by an IV test dose to estimate PK parameters; PK modelled as a double exponential (eg 2 cpt model with fixed duration infusion)
  - procedure integrated in a somewhat different perspective which tries to use data from routine infusion to access the individual parameters, with a reduced sampling approach, so assessment of LSS itself and choice of best strategy
  - evaluation of the estimated clearance, compared to the simulated values, as bias and RMSE
- Nguyen 2013: estimation of individual parameters during HCV treatment
  - viral kinetic model with an ODE system of 3 equations
  - evaluate mean relative error and RRMSE of all parameters, as well as shrinkage
  - impact of design (sparse or informative) and of the number of BQL data
  - good estimation of individual parameters as long as parameters remain identifiable
  - setting may be too complex for simulations with **saemix** (ODE model)

- Studies in therapeutic drug monitoring
  - Preijers 2017: evaluation of LSS strategies in the predicion of selected individual parameters (CL, time to 1% and Vss); uses an initial population model fit followed by Bayesian analysis; simulations to assess the rMPE and the rRMSE of CL, t12 and Vss (note: 3 cpt PK model) for each of 11 LSS schemes
  - several papers evaluating the prediction of an outcome (eg INR response, AUCs, ...) in therapeutic drug monitoring
  -

3.2.1 PK model with covariate

Possibilities:

- one cpt PK model used in Combes 2014
  - diagonal variability 50 or 20%
  - proportional error model 30 or 40%
  - different designs (5, 3 or 2 points, optimised with PFIM) and number of subjects
  - bias or RMSE on individual parameters not estimated/reported, paper focused on power of LRT and Wald tests
- one cpt PK model used in Lavielle 2016
  - impact of Weight on all parameters
  - full covariance matrix
  - but odd error model: exponential error...
  - bias or RMSE on individual parameters not estimated/reported, paper focused on diagnostics
- two-compartment PK model used in Thai 2014 [Thai et al., 2014]
  - variability 30%, correlation between CL and Q ($\rho = 0.5$)
  - proportional error model 25%
  - sparse (N=70/n=4) and rich design (N=30/n=9), and a mixed design for the MM simulation
  - bias or RMSE on individual parameters not estimated/reported

Proportional or combined error model
In this scenario we will vary:

- the number of subjects: $N \in \{.., , ...\}$
- the number of observations per subject: $n_i \in \{, ...\}$, as well as a mixed design with rich and sparse sampling (respectively $n_i = XXX$ and XXX).
- the interindividual variability $\omega \in \{30, 50, 70\}$ (expressed in %)
- the residual variability: $\epsilon \in \{20\%, 50\%\}$

3.2.2 PD Hill model

In this example, we use the same simulation settings as in [**?**]. This example has the following interesting features: first, the non-linearity of the model can be adjusted through the sigmoidicity parameter, to explore its impact on model estimates; second, a covariance between ED50 and Emax is assumed; third, we can compare some results to previously published findings [**?**, **?**].

In this scenario we will vary:

- model non-linearity: $\gamma \in \{1, 2, 3\}$
- the number of subjects: $N \in \{.., 100, ...\}$
- the number of observations per subject: $n_i \in \{2, 4, ...\}$
- the interindividual variability $\omega \in \{30, 50, 70\}$ (expressed in %)
- the residual variability: $\epsilon \in \{.., 2, ..\}$

In this simulation we use a similar framework as in [**?**] to simulate new datasets where we keep the individual parameters for comparison. The model is a sigmoid E_{max} model, a standard model in dose-response studies where the effect of a drug in response to a dose d, E(d), is a sigmoid function corresponding to the following equation:

$$E(d) = E_0 + E_{\max} \frac{d_i^\gamma}{d_i^\gamma + ED_{50}^\gamma} \tag{2}$$

This model involves 4 parameters, the initial effect E_0, the maximum effect E_{max}, the concentration at which half the maximum effect is achieved EC_{50} and the sigmoidicity factor γ which controls the nonlinearity of the model through the curvature. Interindividual variability was modelled through a log-normal distribution for all parameters, except for γ which was assumed to be the same for all subjects (no IIV). A correlation was simulated between E_{max} and EC_{50}. In Plan et al., 3 values of γ were tested, 1, corresponding to the E_{max} model, and 2 and 3, involving increasing amounts of nonlinearity ([**?**]), along with two residual error models, additive or proportional error. In the present work, we focus on the scenarios with proportional variance σ^2. The parameters used in the simulation are given in Table 1.

Parameter	Value	Parameter	Value
E_0 (-)	5	$\omega^2_{E_0}$	0.09
E_{max} (-)	30	$\omega^2_{E_{max}}$	0.49
EC_{50} (mg)	500	$\omega^2_{EC_{50}}$	0.49
γ (-)	3	$\text{cov}(E_0, E_{max})$	0.245
σ (-)	0.1		

Table 1: Parameters used in the simulation for the E_{max} model (see [**?**]).

.

TODO: change and decide on tuning *The design of the simulation study mimicked that of a clinical trial including 100 individuals and investigating four dose levels (0, 100, 300, and 1000 mg). Two sampling designs were evaluated, a rich design (R) in which all individuals were sampled at the four dose levels, and a sparse design (S) where each individual was randomly allocated to only two of the four dose levels.// We evaluated the influence of the starting values and of the tuning parameters on the performance of* ***saemix****. Initial parameter estimates were either set to the values used in the simulation (*`true`*) or to different values (*`false`*), where the fixed parameters were multiplied by 2 while the variability parameters were set to 0.1. We also used either the default settings of the algorithm, or tuned the settings by increasing the number of chains to 5 and increasing the number of iterations to*

300 and 150 respectively in the burn-in and convergence phases of the algorithm. Combining starting values and tuning parameters, we therefore performed 4 successive parameter estimations for each dataset. In all settings, we used the same random seed for all runs.

3.3 Discrete and time-to-event data

Features of simulation examples:

- fixed and random effects
  - identifiability of IIV for TTE model ?
- design ensuring identifiability
  - optimisation with PFIM for discrete models ?
  - optimisation with PFIM and Bayesian criterion for individual parameter estimation ?

Note: No simulations performed yet to evaluate applications of SIR in discrete data models (purpose of Marilou's MSc course). Two examples in time-to-event data and binary data could be used from the real data examples analysed in [**?**].

3.3.1 Time-to-event data

Simulation settings could be taken from examples in the literature:

- Svensson 2017
  - model with longitudinal evolution of bacterial load, logistic model for probability of finding bacteria in samples, and time-to-event model for time to positivity of bacteria (?) (TTP)
  - somewhat complicated to simulate
- Cerou 2018
  - longitudinal model for PSA informing the hazard function for the risk to dropout
  - the IIV is carried only by the longitudinal model, which is assumed to be perfectly known
- Dosne 2017: three examples of real life data;
  - PD6: RTTE model for epileptic seizures (Abrantes 2014, PAGE abstract)
    - different hazard functions, involving categorical covariates
  - PD7: Cox-proportional hazard model (Karlsson 2014, PAGE abstract)
    - not really applicable in our settings as non-parametric
  - PD11: TTE for conversion to sinus rhythm in acute atrial fibrillation, PK coupled with time-to-event exponential model with covariates (Hennig 2009, PAGE abstract)
    - joint PKPD model with a categorical event
- Schindler 2017
  - complex model involving drug exposure, biomarkers and overall survival modelled as TTE (probably no variability on TTE model, but individual covariate included)

Simple examples used in Monolix:

- Single event

  - Veteran's lung cancer study [Kalbfleisch and Prentice, 1980]
    * 137 with advanced inoperable lung cancer, given either a standard therapy or a test chemotherapy (9 right censored)
    * modelled using an exponential hazard function with 2 covariates (Karnofsky type and histological cell type)
  - North Central Cancer Treatment Group (NCCTG) data [Loprinzi and XXX, 1994]
    * survival of 228 patients with advanced lung cancer, including 63 right censored
    * Gompertz model with variability on scale and prognostic covariates (gender and ECOG score)
  - Oropharynx data set [Kalbfleisch and Prentice, 1980]
    * comparison between radiation therapy alone or radiation therapy together with a chemotherapeutic agent in 195 patients
    * 30% of the survival times are censored owing primarily to patients surviving to the time of analysis
  - Primary Biliary Cirrhosis data set [Fleming and Harrington, 1991]
    * double-blinded randomized trial in primary biliary cirrhosis of the liver (PBC), comparing the drug D-penicillamine (DPCA) with a placebo
    * 312 patients, 125 deaths, 8 lost to follow-up, 19 with liver transplantation

- Repeated TTE
  - not really described, need further check
- Monolix demos
  1. tte1: constant hazard model
  2. tte2: constant hazard model with interval censored data
  3. tte3: RTTE with constant hazard model
  4. tte4: RTTE with constant hazard model for interval censored data
  5. weibullRTTE: Weibull RTTE model (note: count version of this model also in demos weibullCount)

3.3.2 Ordinal data

Question: binary or directly 3 categories ?

- Respiratory status data set [Davis, 1991] showcased in Monolix Datxplore
  - 111 patients given a placebo or an active treatment, and followed at baseline and 4 follow-up visits; covariates such as center, sex and age
  - respiratory status categorised as poor or good (binary data)
  - no model associated ?
- Monolix demos
  1. categorical1: cumulative odds ratio model with 3 categories
  2. categorical2: proportional odds ratio model with 2 covariates
  3. discrete Markov model with 2 independent states

4. discrete Markov model with transition matrix and initial state assumed to be balanced (p=0.5)
5. discrete Markov model with transition matrix and estimated initial state
6. discrete Markov model with transition matrix changing with time
7. discrete Markov model with 3 states

3.3.3 Count data

- Epilepsy attacks data set [Leppik and XXX, 1985] showcased in Monolix Datxplore
  - clinical trial of 59 epileptics who were randomized to receive either the anti-epileptic drug progabide or a placebo, as an adjuvant to standard chemotherapy, and followed at baseline and 4 follow-up visits
  - no model associated ?
- Monolix demos
  1. count1a: Poisson model with constant distribution over time
  2. count1b: Poisson model with constant distribution over time, mixture of two Poisson distributions
  3. count2: Poisson distribution with λ changing exponentially over time
- –
- –

4 Results of simulation studies

4.1 Continuous data

4.1.1 PD Hill model

Tables 2 to 4 show the bias on the population parameters in the 3 models in the rich design scenario (4 doses per subject) (table 2), a sparse design scenario with 2 doses per subject (table 3), and a rich design with a large residual variability (table 4). With a rich design, all parameters are well estimated without bias, even when the nonlinearity becomes high ($\gamma = 3$). With a sparse design, some bias appears for the variance components but surprisingly the bias is more important for the Emax model ($\gamma = 1$) than for the more nonlinear Hill model with $\gamma = 3$. In the setting with a large residual variability, the population parameters are well estimated but the variances are extremely biased (with the exception of $\omega^2 E_0$).

The next 3 tables (5 to 7) give the relative biases on the SE reported by **saemix** in the 3 simulation settings. In this table, the estimated SE are compared to the empirical SE derived from the simulated datasets. In general, the SE seem underestimated except in the scenario with high residual variability where the SE on the variance terms tend to be overestimated.

The next tables compare the individual estimated parameters to their simulated values (γ is not shown as it simulated and estimated without variability). Tables 8 to 10 give the relative bias on the different parameters for the MAP estimate, the mean of the conditional distribution at the end of the estimation step (Cmean), and the mean of the conditional distribution after running the conditional distribution algorithm starting from these estimates (Cmean2), in the three simulation settings. In brackets, the average of the standard deviation of the difference between estimated and simulated parameters is shown, representing the average precision of estimation for each parameter. Tables 11 to 13 give the corresponding shrinkage defined for a parameter θ_p as:

$$Shr(\theta_p) = 1 - \frac{var(\hat{\theta_{p,i}})}{\hat{\omega_p^2}} \tag{3}$$

where $\hat{\theta_{p,i}}$ denotes the individual estimate of θ_p in subject i and $\hat{\omega_p^2}$ is the population estimate of the variance of the parameter.

In the rich design, the individual parameters are well estimated with only a few showing some bias (5-12% mostly on E_{max} or ED_{50}). The two conditional mean estimates are very highly correlated, so that running the additional conditional distribution step does not seem to add much information. The MAP estimates seem to have slightly less bias than the conditional estimates. In the sparse design, the bias is predictably much higher, with only 2 samples per subject to estimate 3 or 4 parameters. There is less bias in the design with higher residual error but again ED_{50} is the most difficult parameter to estimate. In all designs, paradoxically the simpler Emax model has more bias than the two Hill models for the ED_{50} parameter. Higher bias seems to be related to higher shrinkage, but in some cases there is no relationship, for instance in the sparse design there is no bias on E_0 but the shrinkage is over 30%. We also note that shrinkage is much larger for the conditional mean estimates compared to the MAP estimates, while the bias is generally comparable.

The precision of estimation for each parameter (numbers in brackets in tables 8 to 10) can be compared to the SE predicted by **PFIM** assuming Bayesian estimation. These predictions could not be obtained in the sparse design, however we estimated them for the rich and IIV designs. For the Emax model in the rich scenario, precisions of 9, 18 and 39% respectively were predicted by PFIM, and precisions of 25 to 29% were expected for the IIV scenario. In both cases the precision is accurate for E_0 but worse than expected for ED_{50}. This was qualitatively similar for the models with more non-linearity.

Action points: these results prompt the following questions.

- Estimation of population parameters
  - no bias in the rich design, even with large IIV and a covariance term, seems ok
  - bias on variances (both residual and interindividual with either less information or more noise): similar results to those in JSS paper [**?**]
- Estimation of SE
  - underestimation could be due to the asymptotic nature of the FIM (lower bound of the Fisher information matrix)

  ⇒ check with previous simulation (Christelle) if same magnitude
- Estimation of the individual parameters
  - little or no bias in the rich design, with the MAP estimate, so that seems ok
  - some bias in the IIV design (to be expected ?)
  - however... why is conditional mean worse than MAP for the sparse design...
  - why no improvement when running the conditional distribution step after convergence ?
  - worse precision than expected according to PFIM except for E_0

  ⇒ **implement diagnostic graphs for the estimation of the conditional distribution; debug (maybe check acceptance algorithm?)**

	γ=1	γ=2	γ=3
Design	Rich	Rich	Rich
E0	0.64	0.18	-0.33
Emax	0.140	1.461	-0.026
ED50	-0.031	1.037	0.808
gamma	0.00	0.24	0.64
b.1	0.22	-1.33	-1.04
omega2.E0	-1.62	-2.38	-1.15
omega2.Emax	-1.0004	-0.0911	0.0088
omega2.ED50	-3.42	2.43	-3.21
cov.Emax.ED50	-2.01	3.13	-3.31

Table 2: Relative bias on population parameters for the 200 simulated datasets (%)

	γ=1	γ=2	γ=3
Design	Sparse	Sparse	Sparse
E0	0.49	0.61	0.42
Emax	6.957	-3.435	-3.121
ED50	9.607	-2.078	-1.680
gamma	0.00	4.24	5.42
b.1	-14.49	-17.73	-9.28
omega2.E0	1.04	2.89	0.77
omega2.Emax	10.3404	-4.2647	1.0712
omega2.ED50	53.46	-11.37	-10.77
cov.Emax.ED50	47.50	-20.60	-14.82

Table 3: Relative bias on population parameters for the 200 simulated datasets (%)

	γ=1	γ=2	γ=3
Design	IIV	IIV	IIV
E0	0.18	0.17	-0.16
Emax	7.972	1.222	4.393
ED50	7.207	0.134	5.536
gamma	0.00	9.07	3.72
b.1	-2.60	-3.05	-2.29
omega2.E0	10.15	3.19	6.45
omega2.Emax	219.0393	124.8689	115.1995
omega2.ED50	906.11	159.72	86.62
cov.Emax.ED50	814.11	227.91	146.53

Table 4: Relative bias on population parameters for the 200 simulated datasets (%)

	γ=1	γ=2	γ=3
Design	Rich	Rich	Rich
SE.E0	0.25	-9.42	-1.08
SE.Emax	1.11	-21.25	-18.04
SE.ED50	-7.42	-14.34	-14.26
SE.gamma	0.00	-24.51	-24.65
SE.b.1	-7.83	-7.50	-11.75
SE.omega2.E0	11.444	-3.902	-7.272
SE.omega2.Emax	-8.73	-20.82	-13.99
SE.omega2.ED50	-27.99	-26.06	-21.70
SE.cov.Emax.ED50	-20.27	-32.47	-30.59

Table 5: Relative bias on estimated SE for the 200 simulated datasets (%)

	γ=1	γ=2	γ=3
Design	Sparse	Sparse	Sparse
SE.E0	-11.50	-17.11	-11.74
SE.Emax	-33.50	-30.06	-39.85
SE.ED50	-30.68	-24.69	-30.53
SE.gamma	0.00	-24.60	-31.87
SE.b.1	-0.90	79.93	21.85
SE.omega2.E0	-8.793	-0.032	0.619
SE.omega2.Emax	-23.55	-23.85	-26.93
SE.omega2.ED50	-30.95	-25.77	-25.51
SE.cov.Emax.ED50	-31.32	-24.80	-22.83

Table 6: Relative bias on estimated SE for the 200 simulated datasets (%)

	γ=1	γ=2	γ=3
Design	IIV	IIV	IIV
SE.E0	-4.45	-10.75	-13.48
SE.Emax	-18.17	-16.54	-12.35
SE.ED50	-22.43	-13.14	-6.24
SE.gamma	0.00	-15.58	-10.95
SE.b.1	26.81	29.72	23.85
SE.omega2.E0	41.505	30.700	19.548
SE.omega2.Emax	82.47	44.81	48.36
SE.omega2.ED50	104.28	30.90	18.29
SE.cov.Emax.ED50	95.29	30.27	25.41

Table 7: Relative bias on estimated SE for the 200 simulated datasets (%)

	MAP	Cmean	Cmean2
$\gamma = 1$ E0	-1.206 (9.14)	0.55 (9.40)	0.53 (9.49)
$\gamma = 1$ Emax	-0.89 (23.66)	3.34 (26.71)	3.18 (27.16)
$\gamma = 1$ ED50	4.68 (48.31)	11.53 (56.46)	11.51 (56.86)
$\gamma = 2$ E0	-0.980 (8.30)	0.36 (8.52)	0.32 (8.58)
$\gamma = 2$ Emax	0.65 (33.91)	6.56 (40.51)	6.19 (41.56)
$\gamma = 2$ ED50	2.06 (36.79)	5.76 (37.87)	5.57 (39.43)
$\gamma = 3$ E0	-0.794 (7.55)	0.33 (7.69)	0.29 (7.76)
$\gamma = 3$ Emax	3.86 (51.87)	7.85 (53.20)	7.36 (53.20)
$\gamma = 3$ ED50	2.70 (39.28)	4.17 (38.21)	3.75 (36.60)

Table 8: Relative bias on individual parameters for the 200 simulated datasets (%) - Design Rich

	MAP	Cmean	Cmean2
$\gamma = 1$ E0	0.410 (20.70)	2.36 (21.84)	2.37 (22.23)
$\gamma = 1$ Emax	8.82 (47.15)	17.10 (55.84)	16.76 (57.41)
$\gamma = 1$ ED50	34.10 (121.65)	45.69 (145.75)	46.89 (151.32)
$\gamma = 2$ E0	0.977 (18.81)	2.01 (19.61)	2.22 (20.09)
$\gamma = 2$ Emax	11.52 (74.34)	14.09 (75.66)	13.83 (77.13)
$\gamma = 2$ ED50	24.39 (124.31)	22.27 (114.34)	22.44 (115.07)
$\gamma = 3$ E0	0.476 (16.76)	1.39 (16.80)	1.51 (17.08)
$\gamma = 3$ Emax	18.17 (91.33)	17.42 (87.33)	17.47 (92.56)
$\gamma = 3$ ED50	24.20 (119.23)	18.04 (96.52)	18.30 (98.71)

Table 9: Relative bias on individual parameters for the 200 simulated datasets (%) - Design Sparse

	MAP	Cmean	Cmean2
$\gamma = 1$ E0	-1.216 (24.33)	3.03 (25.46)	3.02 (25.54)
$\gamma = 1$ Emax	11.64 (36.72)	13.51 (35.82)	13.66 (36.00)
$\gamma = 1$ ED50	25.58 (66.41)	19.28 (56.26)	19.23 (56.61)
$\gamma = 2$ E0	-0.091 (23.31)	3.26 (24.20)	3.36 (24.40)
$\gamma = 2$ Emax	3.56 (30.99)	5.80 (31.76)	6.05 (32.30)
$\gamma = 2$ ED50	7.26 (35.41)	5.56 (34.21)	5.73 (34.73)
$\gamma = 3$ E0	-1.753 (21.81)	2.18 (22.83)	2.20 (23.00)
$\gamma = 3$ Emax	6.70 (31.12)	8.96 (32.32)	9.33 (32.89)
$\gamma = 3$ ED50	11.48 (32.88)	9.90 (32.34)	10.08 (32.74)

Table 10: Relative bias on individual parameters for the 200 simulated datasets (%) - Design IIV

	MAP	Cmean	Cmean2
$\gamma = 1$ E0	-1.27	8.72	8.89
$\gamma = 1$ Emax	-2.28	12.21	10.35
$\gamma = 1$ ED50	-10.40	41.53	35.61
$\gamma = 2$ E0	-1.64	6.97	6.92
$\gamma = 2$ Emax	-1.74	16.32	13.20
$\gamma = 2$ ED50	-2.37	16.90	11.94
$\gamma = 3$ E0	-1.35	5.26	5.43
$\gamma = 3$ Emax	-2.47	20.52	15.94
$\gamma = 3$ ED50	-4.91	9.13	7.69

Table 11: Shrinkage on individual parameters for the 200 simulated datasets (%) - Design Rich

	MAP	Cmean	Cmean2
$\gamma = 1$ E0	-0.52	41.14	36.41
$\gamma = 1$ Emax	4.16	32.44	28.72
$\gamma = 1$ ED50	20.76	71.37	67.12
$\gamma = 2$ E0	1.13	26.44	22.14
$\gamma = 2$ Emax	-11.69	50.11	42.27
$\gamma = 2$ ED50	-30.40	52.41	54.61
$\gamma = 3$ E0	0.13	18.62	17.27
$\gamma = 3$ Emax	-6.03	56.80	50.58
$\gamma = 3$ ED50	-25.77	46.33	48.24

Table 12: Shrinkage on individual parameters for the 200 simulated datasets (%) - Design Sparse

	MAP	Cmean	Cmean2
$\gamma = 1$ E0	0.84	66.88	64.52
$\gamma = 1$ Emax	59.53	83.10	83.01
$\gamma = 1$ ED50	87.56	81.53	84.90
$\gamma = 2$ E0	-2.76	57.98	57.92
$\gamma = 2$ Emax	44.24	80.03	78.10
$\gamma = 2$ ED50	49.29	77.02	78.12
$\gamma = 3$ E0	-1.54	51.71	52.72
$\gamma = 3$ Emax	43.95	79.96	76.69
$\gamma = 3$ ED50	32.34	73.04	71.85

Table 13: Shrinkage on individual parameters for the 200 simulated datasets (%) - Design IIV

5 Discussion

References

[Boeckmann et al., 1994] Boeckmann, A., Sheiner, L., and Beal, S. (1994). *NONMEM Users Guide: Part V*. University of California, NONMEM Project Group, San Francisco.

[Davis, 1991] Davis, C. (1991). Semi-parametric and non-parametric methods for the analysis of repeated measurements with applications to clinical trials. *Statistics in Medicine*, 10:1959–80.

[Fleming and Harrington, 1991] Fleming, T. and Harrington, D. (1991). *Counting Processes and Survival Analysis*. John Wiley & Sons, New York.

[Kalbfleisch and Prentice, 1980] Kalbfleisch, D. and Prentice, R. (1980). *The Statistical Analysis of Failure Time Data*. John Wiley & Sons, New York.

[Kuhn and Lavielle, 2005] Kuhn, E. and Lavielle, M. (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics and Data Analysis*, 49:1020–1038.

[Leppik and XXX, 1985] Leppik, I. and XXX (1985). A double-blind crossover evaluation of progabide in partial seizures. *Neurology*, 35:285–.

[Lixoft, 2013] Lixoft (2013). *MONOLIX (MOdèles NOn LInéaires à effets miXtes)*. INRIA, Orsay, France.

[Loprinzi and XXX, 1994] Loprinzi and XXX (1994). Prospective evaluation of prognostic variables from patient-completed questionnaires. north central cancer treatment group. *Journal of Clinical Oncology : Official Journal of the American Society of Clinical Oncology*, 12:601–7.

[Retout et al., 2002] Retout, S., Mentré, F., and Bruno, R. (2002). Fisher information matrix for nonlinear mixed-effects models: evaluation and application for optimal design of enoxaparin population pharmacokinetics. *Statistics in Medicine*, 21(18):2623–2639.

[Thai et al., 2014] Thai, H., Mentré, F., Holford, N. H., Veyrat-Follet, C., and Comets, E. (2014). Evaluation of bootstrap methods for estimating uncertainty of parameters in nonlinear mixed-effects models: a simulation study in population pharmacokinetics. *Journal of Pharmacokinetics and Pharmacodynamics*, 41:15–33.