Estimation of nonlinear mixed effects model in pharmacokinetics with the SAEM algorithm implemented in MONOLIX

Pr France Mentré, INSERM U738, University Paris Diderot
Pr Marc Lavielle, INRIA, Universities Paris 5 & 11
Paris, France
Outline

1. Introduction
2. Brief history of estimation methods in NLMEM
3. Stochastic EM algorithms
4. MONOLIX software
5. Comparison of Stochastic EM algorithms to NONMEM
6. PKPD example with MONOLIX
7. Conclusion
1. Introduction

- Nonlinear mixed effects models (NLMEM) allow "population" PKPD analyses
  - Global analysis of data in all individuals
  - Rich or sparse design
- Increasingly used in clinical and non clinical drug development
  - Parameter estimation
  - Model selection
  - Covariate testing
  - Predictions & Simulations
  - **Good estimation methods needed**
- Focus here on Maximum Likelihood Estimation (MLE) parametric methods
2. Brief history of estimation methods

**NON linear Mixed Effects Model**

L Sheiner & S Beal, UCSF

- **1972**: The concept and the **FO method**

- **1977**: The first case study

- **1980**: NONMEM - An IBM-specific software
Standard Two-Stage approach (STS)

Stage 1:
- Individual fitting
- Non Linear regression

Subject | Parameters estimate
---|---
#1 | 12.3
#2 | 21.9
#n | 16.1

Stage 2:
Descriptive statistics, linear stepwise regression for covariate effect

From Steimer (1992): « Population models and methods, with emphasis on pharmacokinetics », in M. Rowland and L. Aarons (eds), *New strategies in drug development and clinical evaluation, the population approach*
Population approach

Estimates of individual parameters

Non linear mixed effects model

Single-stage approach (population analysis)

From Steimer (1992) : « Population models and methods, with emphasis on pharmacokinetics », in M. Rowland and L. Aarons (eds), New strategies in drug development and clinical evaluation, the population approach
The population approach

- $N$ individuals ($i = 1, ..., N$)
- Structural model $f$: same shape in all individuals

\[ y_{ij} = f(\theta_i, t_{ij}) + g(\theta_i, t_{ij}) \varepsilon_{ij} \quad (j = 1, ..., n_i) \]

- Assumption on the individual parameters
  \[ \theta_i = \mu + \eta_i \quad \text{or} \quad \theta_i = \mu \exp(\eta_i) \]

  - $\mu = "mean"$ parameters (fixed effects)
  - $\eta_i = \text{individual random effects}$
  - $\eta_i \sim \text{Normal distribution with mean 0 and variance } \Omega$

  $\Omega$: inter-individual variability
The FO method (1)

• Estimation of population parameters by maximum likelihood
  – find parameters that maximise the probability density function of the observations given the model
  – good statistical properties of ML estimator

• Problem: No closed form of the likelihood
  \[
y_{ij} = f(\mu + \eta_i, t_{ij}) + g(\mu + \eta_i, t_{ij}) \varepsilon_{ij}
\]

• First order linearisation of the model around \( \eta = 0 \)
  \[
y_{ij} \approx f(\mu, t_{ij}) + \frac{\partial f}{\partial \eta}(\mu, t_{ij}) \times \eta_i + g(\mu, t_{ij}) \varepsilon_{ij}
\]

→ Extended Least Square criterion
The FO method (2)

Advantages

• Better than Standard Two-Stage approach in many cases
  – STS neglects estimation error
    • Overstimation of inter-individual variability
    • OK for very rich design and small residual error
  – STS cannot be used for rather sparse designs

• Takes into account correlation within individuals
  – better than all naive approaches
  - Naive avering of data (NAD): "population average"
  - Naive pooling of data (NPD): one "giant" individual
More recent statistical developments in estimation methods for NLMEM: three periods

1. **85 – 90**: FOCE + other approaches: nonparametric, Bayesian

2. **The 90’s**: new software, growing interest, new statistical developments, limitations of FOCE

3. **Since 00**: Stochastic methods for parametric ML estimation + ...
## Software for estimation in nonlinear mixed-effects models

<table>
<thead>
<tr>
<th></th>
<th>Maximum likelihood</th>
<th>Bayesian estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parametric</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NONMEM</td>
<td></td>
<td>PK BUGS</td>
</tr>
<tr>
<td>WinNonMix</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nlme (R and Splus)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proc NLMIXED (SAS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPharm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MONOLIX (SAEM)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S-ADAPT (MCPEM)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDX-MCPEM</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Nonparametric</strong></td>
<td></td>
<td>Dirichlet process</td>
</tr>
<tr>
<td>NPML</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NPEM (USC*PACK)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NONMEM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Nonlinear regression in PK and PD

1980

Linear mixed-effects models
EM – algorithm
NPML
FOCE
Bayesian methods using MCMC

1990

Laplacian
Gaussian Quadrature
ITBS/P-PHARM
NPEM
POPKAN
PKBUGS

2000

Limitations of FOCE
New ML algorithm based on Stochastic EM

The FO and FOCE methods

• First and most popular methods for estimation of population parameters by maximum likelihood in NLMEM

• FO: First order linearisation of the model around random effects = 0

• FOCE: First order linearisation of the model around current estimates of random effects

Implemented in NONMEM, WinNonMix, nlme (R and Splus), Proc NL MIXED (SAS)
Limitations of FO and FOCE

• FO
  – assume that mean response = response for mean parameters
  – not true for nonlinear models!!
    ➢ Bias if "not very small" inter-individual variability

• FOCE
  – not consistent for sparse designs
  – very sensitive to initial estimates:
    ➢ Lot’s of run failed to converge, waste of time for modellers

• Both: Not real Maximum Likelihood Estimates (MLE)
  – good properties of MLE not demonstrated (LRT, standard errors from Fisher Information matrix, ...)

France Mentré, NCS, September 2008
Other approaches for computation of likelihood

• With approximation: linearisation using Laplace (NONMEM)
  ➢ Similar problems of initial values than FOCE

• Integration of the likelihood by Adaptative Gaussian Quadrature (Proc NLMIXED in SAS)
  ➢ Limited to models with small number of random effects
3. Stochastic EM algorithms

EM algorithm

• Developed for MLE in problems with missing data
• Two steps algorithm
  – E-step: expectation of the log-likelihood of the complete data
  – M-step: maximisation of the log-likelihood of the complete data
• Mixed-effects models
  – individual random-effects = missing data

EM in NLMEM

- Problem in EM for NLMEM
  - no analytical solution for integral in E-step

1. Linearisation around current estimates of random effects (PPharm, ITS)
   - Similar problems for sparse design than FOCE


2. Full stochastic E-step
   - Can be very time consuming, not in available software

3. **MCPEM** (in S-ADAPT and PDX-MCPEM): Monte Carlo integration during the E step using importance sampling around current individual estimates


4. **SAEM** (in MONOLIX): Decomposition of E-step in 2 steps
   - S-step: simulation of individual parameters using MCMC
   - SA-step: stochastic approximation of expected likelihood

SAEM in NLMEM

• Do not compute integral of E-step at each iteration
  – less time consuming than MCPEM
• Good statistical properties clearly demonstrated
• Addition of a Simulated Annealing algorithm to converge more quickly around the MLE
  – robust with respect to choice of initial estimates
  – fast
Recent extensions of the SAEM algorithm

• Correct handling of BQL data

• Models defined by ODE or SDE

• REML Estimation

• Inter-occasion variability

• Binary data
  Meza, Jaffrezic, Foulley (2008). Estimation in the probit normal model for binary outcomes using the SAEM algorithm. (in revision)
MONOLIX (MOdèles NOn LInéaires à effets miXtes)

The MONOLIX Group

The MONOLIX Software
The MONOLIX group

- **Multi-disciplinary group** (Pr M Lavielle & F Mentré)
  - created in October 2003
  - meets every month to exchange and develop activities in the field of mixed effect models
  - interest in both in the study and applications of these models

- Involves scientists with varied backgrounds
  - academic statisticians from several universities of Paris (theoretical developments),
  - researchers from INSERM (applications in pharmacology)
  - researchers from INRA (applications in agronomy, animal genetics)
  - Scientists from from the medical faculty of Lyon-Sud University (applications in oncology)
The MONOLIX software (1)

- SAEM algorithm for estimation in NLMEM
- Open-source software
  - MATLAB
  - Stand alone version for models in library
  - Rich PKPD library
  - ODE models in MATLAB or C++
  - MLXTRAN for complex models definition
  - Correct handling of LOQ data
  - Estimation of inter-occasion variability
  - Friendly Graphical User Interface
  - Graphical outputs (GOF, VPC, ...)

- Supported by several ingeneers from INRIA (Institut National de la Recherche en Informatique et Automatique, France)
The MONOLIX software (2)

- **MONOLIX 1**: v1.1 Feb 2005

- **MONOLIX 2**: v2.1 April 2007
  - v2.3 November 2007 – May 2008
    - C++ package for ODE models
    - Categorical covariates
    - Several distribution for random effects
    - Inter-occasion variability
  - v2.4 September 2008 (beta version in June 2008)
    - MLXTRAN
    - Extension of PKPD library (3 cpt PK, effect compartment)

- **MONOLIX 3**: new MONOLIX project
  - Support from INRIA and several drug companies
Algorithms in MONOLIX

Intensive use of powerful and well-known algorithms in the MONOLIX software:

- Estimation of the population parameters: Maximum likelihood estimation with the SAEM (Stochastic Approximation of EM) algorithm, combined with MCMC (Markov Chain Monte Carlo) and Simulated Annealing.
- Estimation of the individual parameters: Estimation/Maximization of the conditional distributions with MCMC.
- Estimation of the objective (likelihood) function: Monte Carlo and minimum variance Importance Sampling.
- Model selection and assessment: Information criteria (AIC, BIC), Statistical Tests (LRT, Wald test), Goodness of fit plots (Individual fits, Weighted residuals, NPDE, VPC, ...).
Estimation & outputs: without linearisation

- Estimation of all components of variability (even small) and their standard errors
- Estimation of individual random effects
  - From simulated posterior/conditional distribution
  - Mean, Var and Mode without approximation
- Likelihood estimated by importance sampling
- Population and Individual residuals
  - From simulated marginal and posterior distributions
- No real importance of shrinkage
5. Comparison of Stochastic EM algorithms to NONMEM

- Girard & Mentré **PAGE 2005**
  - 100 replicated simulated data sets: one PK and one PD
  - NONMEM V & VI, MCPEM, SAEM (blind evaluation)

- Bauer, Guzy & Chee **AAPS J 2007**
  - Four simulated examples: PK or PKPD
  - NONMEM VI, PDx-MCPEM, S-ADAPT, MONOLIX 1.1

- Bazzoli, Retout & Mentré **PAGE 2007**
  - 1000 replicated simulated data sets: one PKPD model
  - NONMEM V, MONOLIX 2.1

- Laveille, Lavielle, Chatel & Jacqmin **PAGE 2008**
  - 150 simulated examples: PK (linear & nonlinear)
  - NONMEM V & VI, MONOLIX 2.2 & 2.3

- ...
Main conclusions from comparisons

S-ADAPT (MCPEM) and MONOLIX (SAEM)

• Never failed to provide results whatever the models
  – in replicated simulated data sets, often NONMEM results on several data sets are missing
• Much faster than NONMEM FOCE for complex ODE models
• Better results (bias, RMSE) than NONMEM FOCE in sparse designs
• Applied successfully to real complex PKPD data sets (where NONMEM failed to converge)
Simulated sparse design PK example
Classification of methods based on RMSE of the 10 population parameters (%)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MC PEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAS AGQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NM V</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NM VI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nlme</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAS FO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Pts: 89 78 73 70 64 59 44 17
- Run: 100 100 100 100 100* 49 88 100

Girard & Mentré, PAGE 2005
1. A compiled version would probably be considerably faster than actual implementation in Matlab for PEM and SAEM
2. For a successful nls + nlme run
3. For a successful NONMEM run (EST & COV)
### Evaluation of PK library in MONOLIX 2.3

#### CPU times

<table>
<thead>
<tr>
<th>Model Type</th>
<th>Single Dose</th>
<th>Multiple Doses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>120 subjects</td>
<td>120 subjects</td>
</tr>
<tr>
<td></td>
<td>2280 observations</td>
<td>5520 observations</td>
</tr>
<tr>
<td>1 compartment model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IV bolus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>linear elimination</td>
<td>7&quot;</td>
<td>25&quot;</td>
</tr>
<tr>
<td>non linear elimination</td>
<td>34&quot;</td>
<td>1' 36&quot;</td>
</tr>
<tr>
<td>2 compartments model</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st order oral absorption</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lag time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>linear elimination</td>
<td>15&quot;</td>
<td>51&quot;</td>
</tr>
<tr>
<td>non linear elimination</td>
<td>2' 18&quot;</td>
<td>6'52&quot;</td>
</tr>
</tbody>
</table>

Laveille & Lavielle, MONOLIX website, PAGE 2008

France Mentré, NCS, September 2008
6. PKPD example with MONOLIX (v 2.4)

- PKPD analysis of warfarine
  - 32 healthy volunteers
  - 1.5 mg/kg single dose
  - PK: total racemic warfarin plasma concentration
  - PD: prothrombin complex activity (PCA)
  - 250 concentrations, 232 PCA

- Models
  - PK: 1 cp, first order absorption with lag time
    \( T_{lag}, k_a, V, CL \)
  - PD: turnover model with inhibition of input by warfarine
    \( R_{in}, k_{out}, I_{max}, C_{50} \)
The MONOLIX interface is an open-source software for pharmacokinetic and pharmacodynamic modeling. It provides a graphical user interface for modeling, simulation, and analysis of pharmacological data. The interface includes modules for data management, model building, and result visualization. The specific screen shown includes options for data management, model selection, and result display, facilitating the process of building and analyzing pharmacokinetic models.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>tlag</td>
<td>0.683</td>
<td>0.238</td>
</tr>
<tr>
<td>ka</td>
<td>0.772</td>
<td>0.115</td>
</tr>
<tr>
<td>V</td>
<td>7.91</td>
<td>0.211</td>
</tr>
<tr>
<td>Cl</td>
<td>0.132</td>
<td>0.00591</td>
</tr>
<tr>
<td>Imax</td>
<td>1.11</td>
<td>0.0364</td>
</tr>
<tr>
<td>C50</td>
<td>1.72</td>
<td>0.225</td>
</tr>
<tr>
<td>Rin</td>
<td>4.47</td>
<td>0.201</td>
</tr>
<tr>
<td>kout</td>
<td>0.0465</td>
<td>0.00196</td>
</tr>
<tr>
<td>a_1</td>
<td>0.303</td>
<td>0.0479</td>
</tr>
<tr>
<td>b_1</td>
<td>0.0529</td>
<td>0.00934</td>
</tr>
<tr>
<td>c_1</td>
<td>1</td>
<td>-</td>
</tr>
<tr>
<td>a_2</td>
<td>3.84</td>
<td>0.226</td>
</tr>
<tr>
<td>b_2</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>c_2</td>
<td>1</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>omega_tlag</td>
<td>0.707</td>
<td>0.23</td>
</tr>
<tr>
<td>omega_ka</td>
<td>0.885</td>
<td>0.249</td>
</tr>
<tr>
<td>omega_V</td>
<td>0.225</td>
<td>0.031</td>
</tr>
<tr>
<td>omega_Cl</td>
<td>0.285</td>
<td>0.0369</td>
</tr>
<tr>
<td>omega_Imax</td>
<td>0.0374</td>
<td>0.018</td>
</tr>
<tr>
<td>omega_C50</td>
<td>0.392</td>
<td>0.0621</td>
</tr>
<tr>
<td>omega_Rin</td>
<td>0.0487</td>
<td>0.018</td>
</tr>
<tr>
<td>omega_kout</td>
<td>0.0316</td>
<td>0.0204</td>
</tr>
</tbody>
</table>

Estimation by linearization

-2 x log-likelihood: 2143.77
Akaike Information Criteria: 2173.77
Bayesian Information Criteria: 2195.76

Elapsed time is 485.8 seconds.
CPU time is 464 seconds.
Some individual PK fits
Some individual PD fits
VPC provided by MONOLIX
Warfarin PKPD data
MLXTRAN specification

$PROBLEM Turnover model PKPD model
$PSI Tlag ka V Cl Imax C50 Rin kout

$PK
ALAG1 = Tlag
KA1 = ka

$ODE
A_0(1) = 0
A_0(2) = Rin/kout
k = Cl/V
Cc = A(1)/V
DADT(1) = -k*A(1)
DADT(2) = Rin*(1-Imax*Cc/(Cc+C50))-kout*A(2)

$OUTPUT
OUPUT1 = Cc
OUPUT2 = A(2)
7. Conclusion (1)

- NLMEM: good approach for PKPD analysis of preclinical studies with rich or sparse designs
  - accurate estimation of all components of variability
- NLMEM applied for increasingly complex dynamic models
- Drug companies mostly used NONMEM
  - FOCE developed 15 years ago: several drawbacks
- New method based on AGQ (in SAS)
  - limited to problems of small dimension
- New MLE methods based on stochastic EM
  - developed by statisticians
  - fast, consistent, no linearization, ...
  - SAEM cleverly used the iteration process
7. Conclusion (2)

- **New software / algorithms should be used**
- **Extensions of MONOLIX, NONMEM... are ongoing**

- **MONOLIX**
  - User friendly, graphical outputs
  - Open source
  - Several statistical extensions planned in next releases
  - Based on thorough and published statistical methods

- **NONMEM**
  - MCPEM & SAEM available in NONMEM VII
  - Implemented by B. Bauer (S-ADAPT)