ESTIMATION OF MIXED HIDDEN MARKOV MODELS WITH SAEM
APPLICATION TO DAILY SEIZURES DATA

Maud Delattre (1), Radojka Savic (2), Raymond Miller (3), Mats O. Karlsson (4), Marc Lavielle (5)

(1) University of Paris Sud, (2) INSERM U738, (3) Pfizer Global R & D, (4) Uppsala University, (5) INRIA Saclay Île-de-France

Lewis Sheiner Student Session - PAGE 2010 Berlin

The data

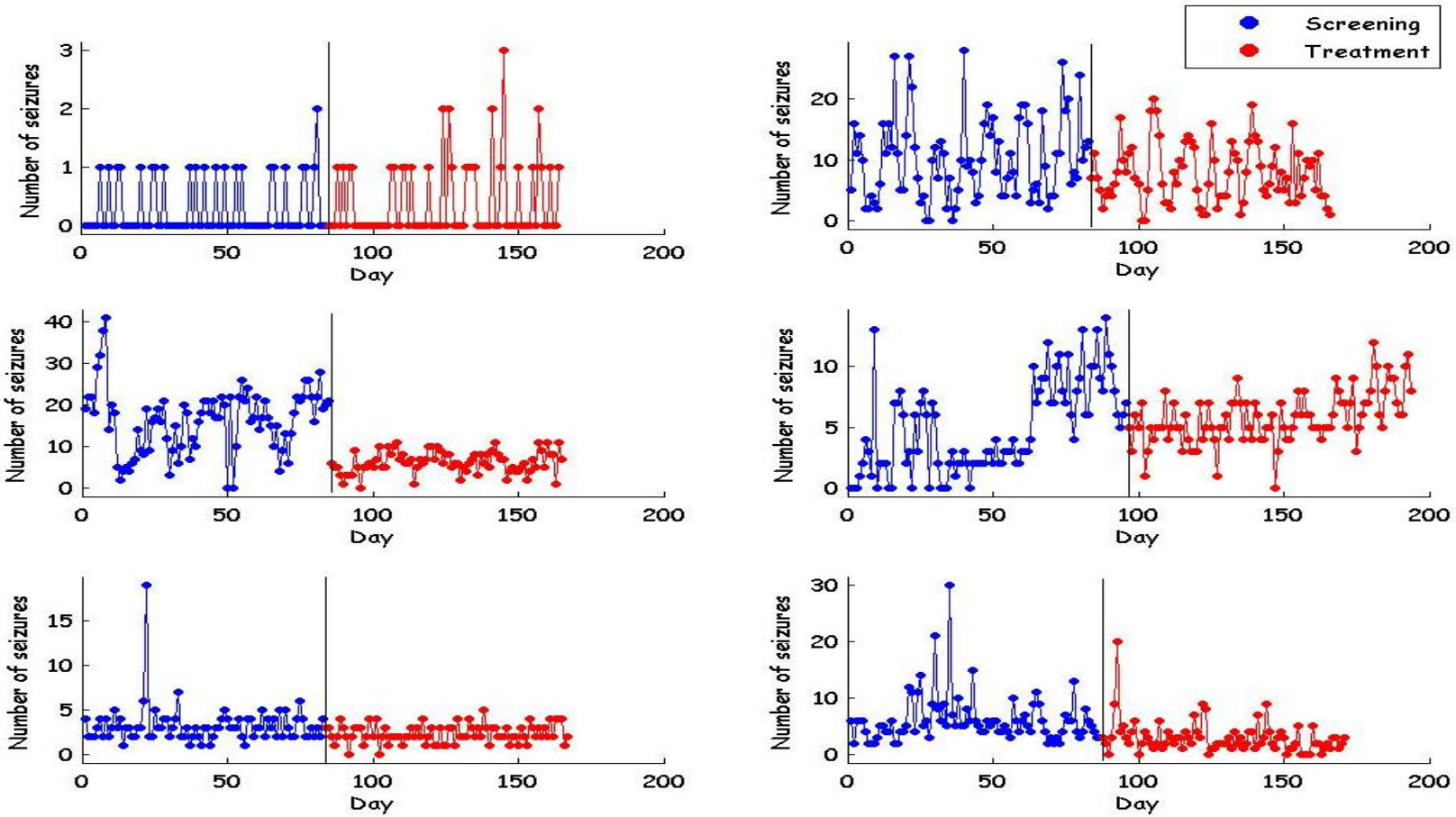
Clinical design:

- double-blind, placebo-controlled, parallel-group and multicenter study
- 788 epileptic patients
- 12 weeks screening phase
  $\longrightarrow$ standard antiepileptic therapy
  12 weeks active treatment phase
  $\longrightarrow$ standard antiepileptic therapy + placebo/pregabalin (0.6, 0.9, 1.2, 1.8g TID)

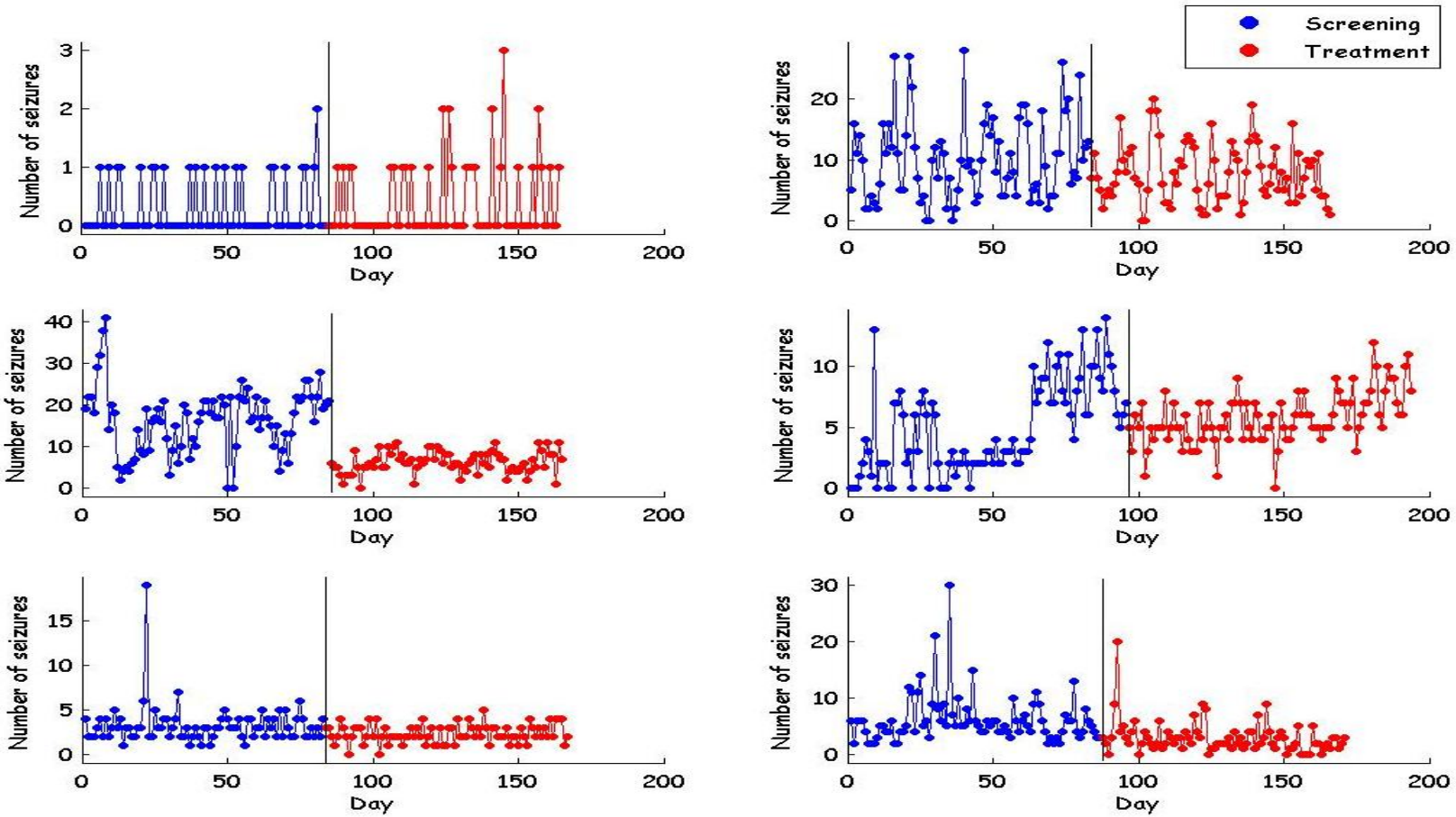
The data:

- the 788 individual sequences of daily seizure counts
  $\longrightarrow$ 134 196 daily seizures counts

The data

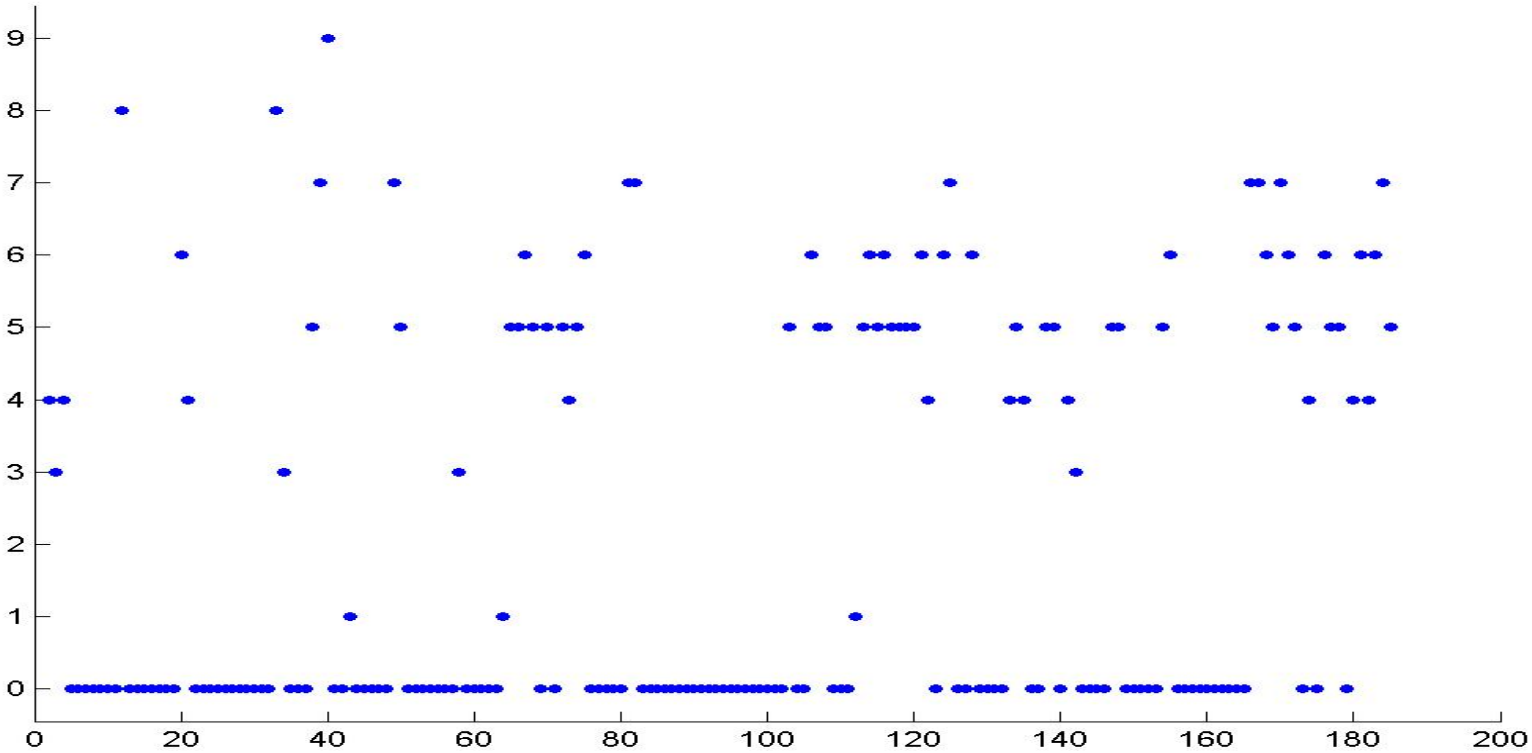


The data



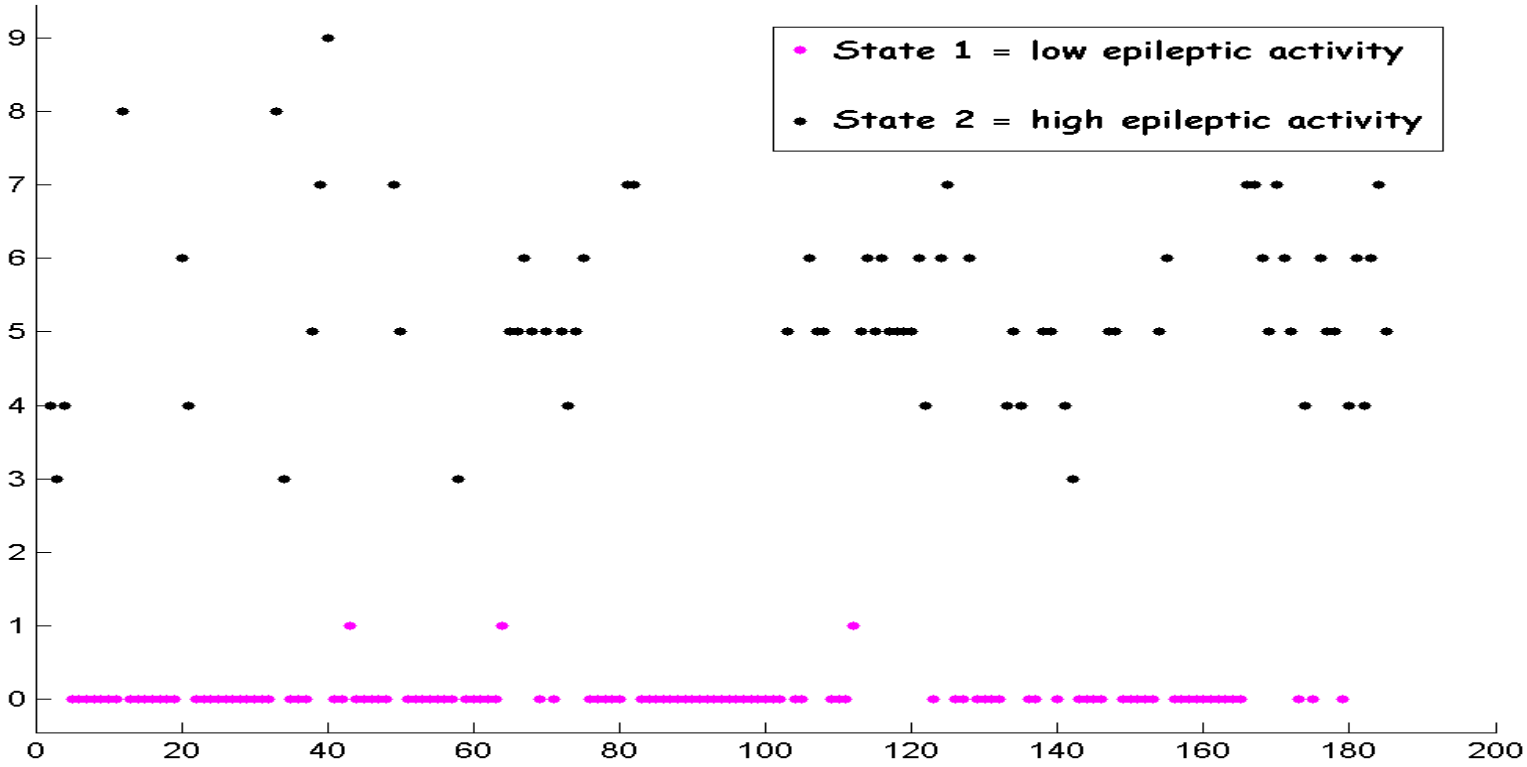
We want to develop a placebo/drug model for this data.

Screening model



Screening model

- The existence of two (hidden) disease stages could be assumed.

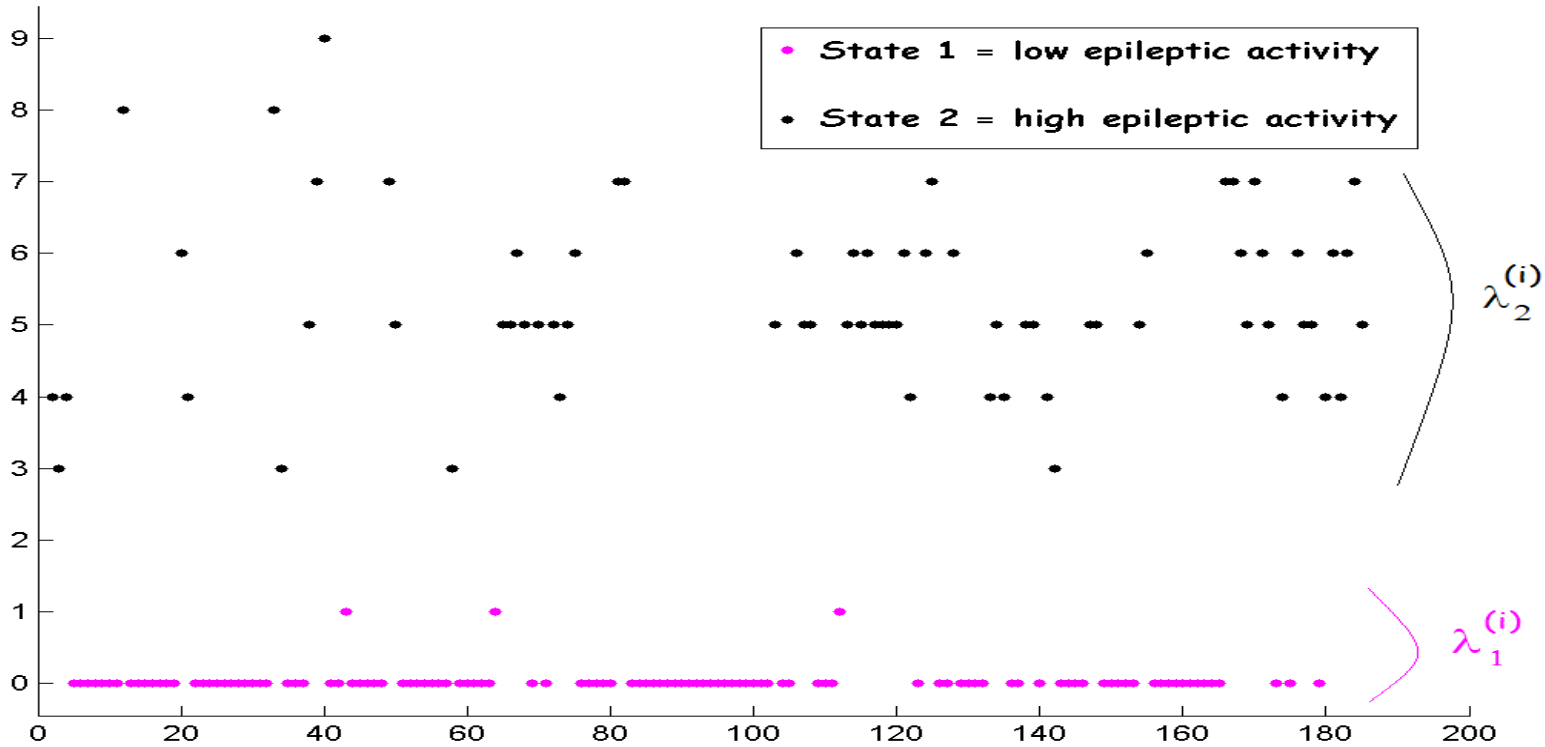


Screening model

- The number of seizures at day j (y_{ij}) is a random variable.
- The distribution of y_{ij} depends on the hidden state z_{ij}

$\longrightarrow$ Poisson distribution with parameter $\lambda_1^{(i)}$ in state 1

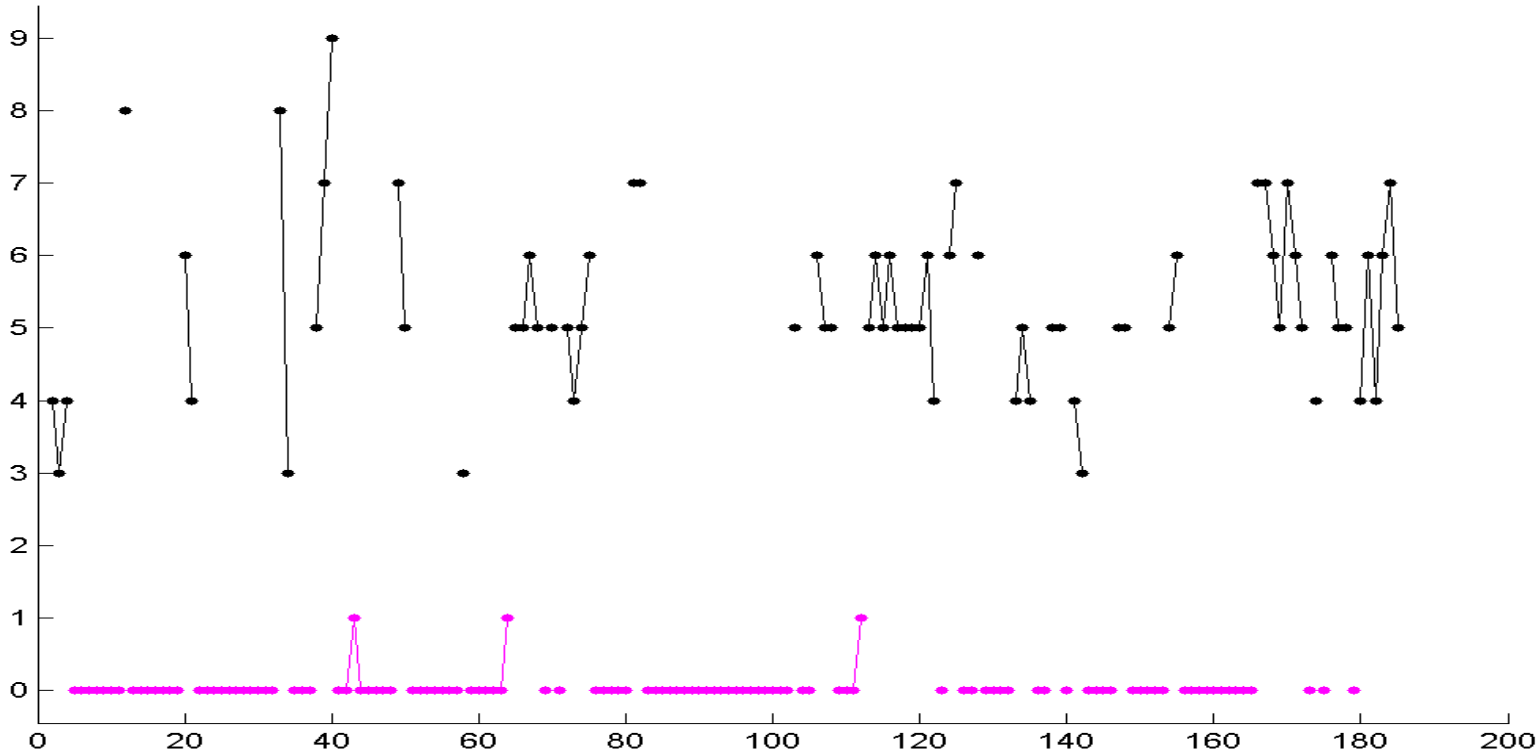
$\longrightarrow$ Poisson distribution with parameter $\lambda_2^{(i)}$ in state 2 $\qquad \left(\lambda_2^{(i)} > \lambda_1^{(i)}\right)$



Screening model

- The data show that epileptic patients are more likely to stay in the same state than to switch to the other state.

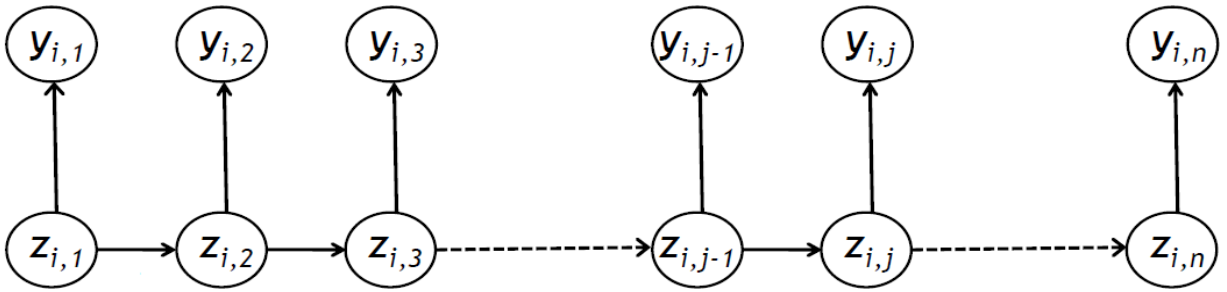
  $\longrightarrow$ The sequences of hidden states have a Markovian dynamics.



Screening model

Consider individual i

- y_{ij}: number of seizures at day j
- z_{ij}: hidden state at day j



- (z_{ij}) is a hidden Markov chain with transition matrix

$$P_i = \begin{pmatrix} p_{11}^{(i)} & p_{12}^{(i)} \\ p_{21}^{(i)} & p_{22}^{(i)} \end{pmatrix}$$

- $y_{ij}|z_{ij} = 1 \backsim \mathcal{P}(\lambda_1^{(i)})$ and $y_{ij}|z_{ij} = 2 \backsim \mathcal{P}(\lambda_2^{(i)})$

Screening model

1. *n* individual hidden Markov models

   The transition matrix of Z_i is defined by $p_{11}^{(i)}$ and $p_{21}^{(i)}$.
   Poisson distributions are chosen for the observations in each state $(\lambda_1^{(i)}, \lambda_2^{(i)})$.

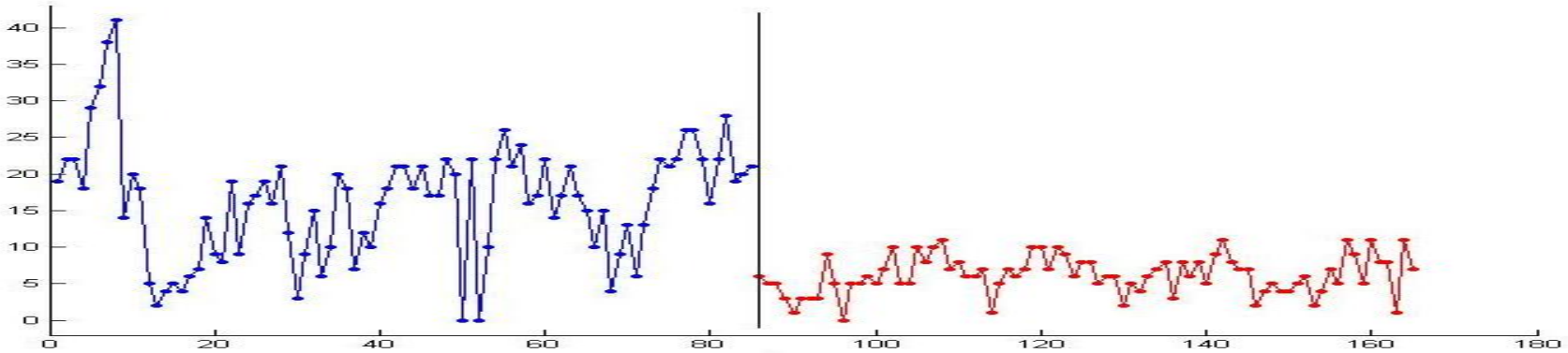
2. population approach

$$\begin{aligned}
\text{logit}(p_{11}^{(i)}) &= \beta_1 + \eta_{1i} \\
\text{logit}(p_{21}^{(i)}) &= \beta_2 + \eta_{2i} \\
\log(\lambda_1^{(i)}) &= \log(\lambda_1) + \eta_{3i} \\
\log(\alpha^{(i)}) &= \log(\alpha) + \eta_{4i} \\
\lambda_2^{(i)} &= \lambda_1^{(i)} + \alpha^{(i)}
\end{aligned}$$

$$\eta_i = (\eta_{1i}, \eta_{2i}, \eta_{3i}, \eta_{4i}) \underset{i.i.d.}{\sim} \mathcal{N}(0, \Omega)$$

3. Here, the population parameters θ are β_1, β_2, λ_1, α and the elements of Ω.

Placebo/drug model



- We want to develop two mixed hidden Markov models simultaneously (screening phase vs treatment phase).
- In the treatment model, the treatment dose could influence both the mean number of seizures in each state and the transition structure of the hidden Markov chain.

Placebo/drug model (1)

Screening phase

$$
\begin{aligned}
\text{logit}(p^S_{11i}) &= \beta^S_1 + \eta_{1i} \\
\text{logit}(p^S_{21i}) &= \beta^S_2 + \eta_{2i} \\
\log(\lambda^S_{1i}) &= \lambda^S_1 + \eta_{3i} \\
\log(\alpha^S_i) &= \alpha^S + \eta_{4i} \\
\lambda^S_{2i} &= \lambda^S_{1i} + \alpha^S_i
\end{aligned}
$$

Treatment phase

$$
\begin{aligned}
\text{logit}(p^T_{11i}) &= \beta^S_{1i} + \delta_{1i} + \gamma_1 D_i \\
\text{logit}(p^T_{21i}) &= \beta^S_{2i} + \delta_{2i} + \gamma_2 D_i \\
\log(\lambda^T_{1i}) &= \log(\lambda^S_{1i}) + \delta_{3i} + \gamma_3 D_i \\
\log(\alpha^T_i) &= \log(\alpha^S_i) + \delta_{4i} + \gamma_4 D_i \\
\lambda^T_{2i} &= \lambda^T_{1i} + \alpha^T_i
\end{aligned}
$$

$$
\begin{aligned}
\delta_{1i} &= \delta_1 + \eta_{5i} \\
\delta_{2i} &= \delta_2 + \eta_{6i} \\
\delta_{3i} &= \delta_3 + \eta_{7i} \\
\delta_{4i} &= \delta_4 + \eta_{8i}
\end{aligned}
$$

Placebo/drug model (2)

Screening phase

$$
\begin{aligned}
\text{logit}(p^S_{11i}) &= \beta^S_1 + \eta_{1i} \\
\text{logit}(p^S_{21i}) &= \beta^S_2 + \eta_{2i} \\
\log(\lambda^S_{1i}) &= \lambda^S_1 + \eta_{3i} \\
\log(\alpha^S_i) &= \alpha^S + \eta_{4i} \\
\lambda^S_{2i} &= \lambda^S_{1i} + \alpha^S_i
\end{aligned}
$$

Treatment phase

$$
\begin{aligned}
\text{logit}(p^T_{11i}) &= \beta^S_{1i} + \delta_{1i} + \gamma_1 D_i \\
\text{logit}(p^T_{21i}) &= \beta^S_{2i} + \delta_{2i} + \gamma_2 D_i \\
\log(\lambda^T_{1i}) &= \log(\lambda^S_{1i}) + (\delta_{3i} + \gamma_3 D_i)(1 - e^{-K_3 t}) \\
\log(\alpha^T_i) &= \log(\alpha^S_i) + (\delta_{4i} + \gamma_4 D_i)(1 - e^{-K_4 t}) \\
\lambda^T_{2i} &= \lambda^T_{1i} + \alpha^T_i
\end{aligned}
$$

$$
\begin{aligned}
\delta_{1i} &= \delta_1 + \eta_{5i} \\
\delta_{2i} &= \delta_2 + \eta_{6i} \\
\delta_{3i} &= \delta_3 + \eta_{7i} \\
\delta_{4i} &= \delta_4 + \eta_{8i}
\end{aligned}
$$

A complete inference methodology

1) **Estimation of the population parameters (M.L.E.)**

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}}\, p\left(Y; \theta\right)$$

$\longrightarrow$ SAEM algorithm

$\longrightarrow$ The Baum Welch algorithm is used to compute

$$p(Y_i, \Psi_i; \theta) = \sum_{Z_i} p(Y_i, Z_i, \Psi_i; \theta)$$

at each iteration of SAEM.

A complete inference methodology

2) **Estimation of the individual parameters (M.A.P.)**

$$\hat{\psi}_i = \underset{\psi_i}{\operatorname{argmax}}\ p\left(\psi_i | y_i; \hat{\theta}\right)$$

3) **Estimation of the sequences of hidden states (MAP)**

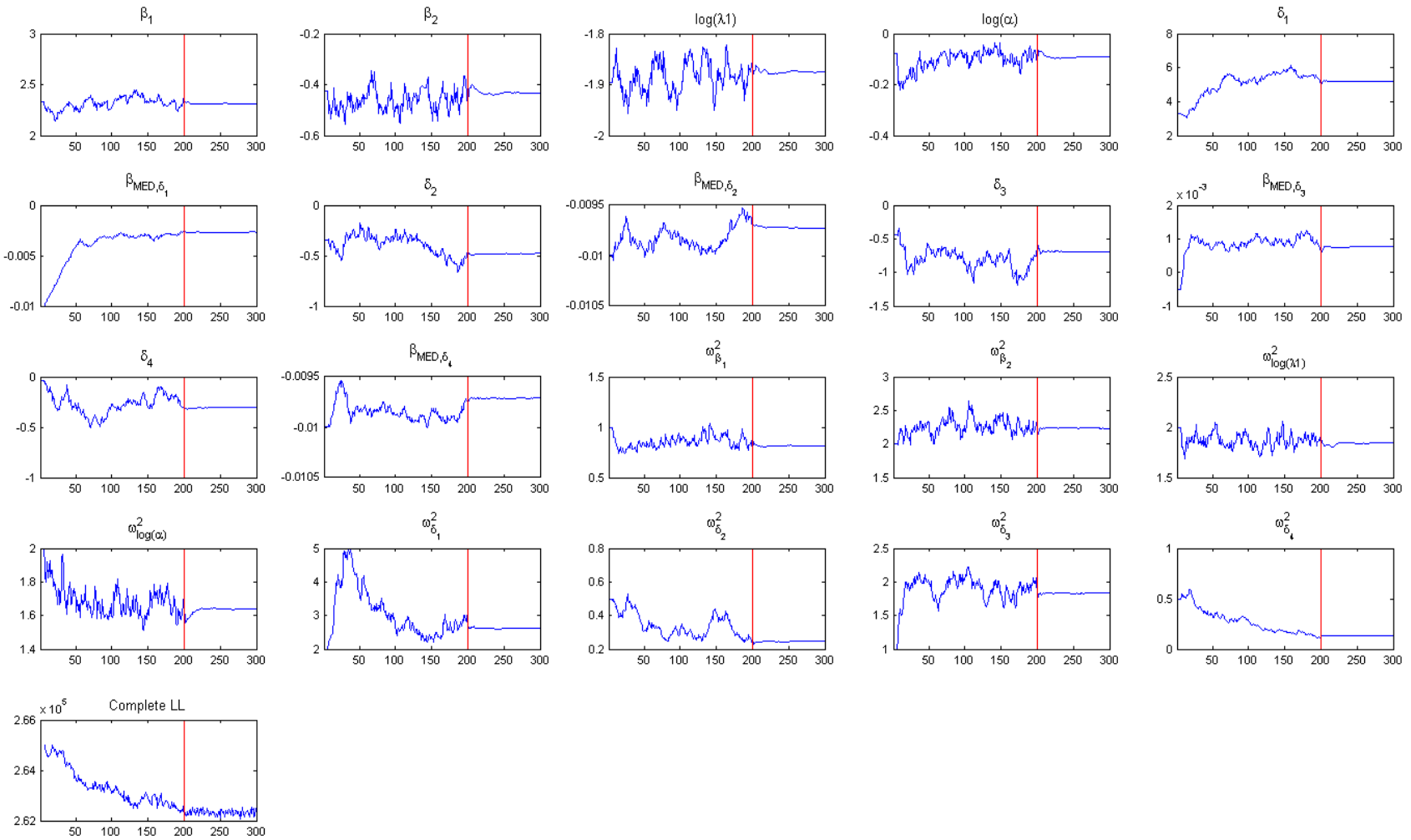
$$\hat{z}_i = \underset{z_i}{\operatorname{argmax}}\ p\left(z_i | y_i, \hat{\psi}_i; \hat{\theta}\right)$$

$\longrightarrow$ Viterbi algorithm

Our methodology has been implemented in MONOLIX 3.1.

Convergence of SAEM

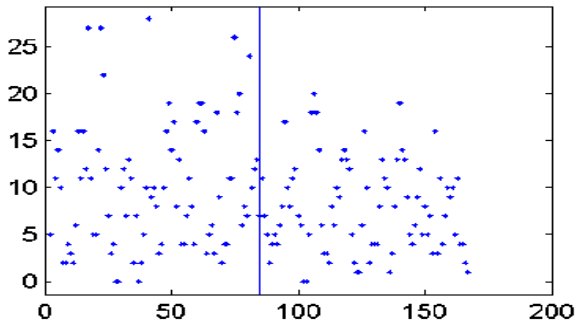
SAEM converges in few iterations (25' for the complete data).

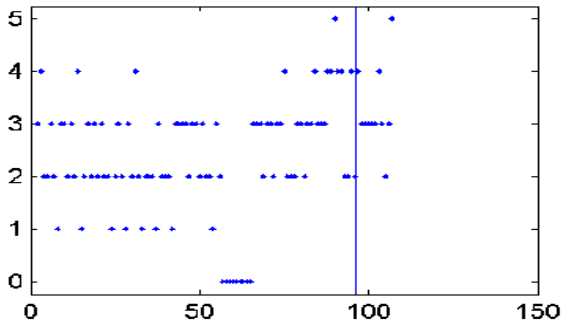


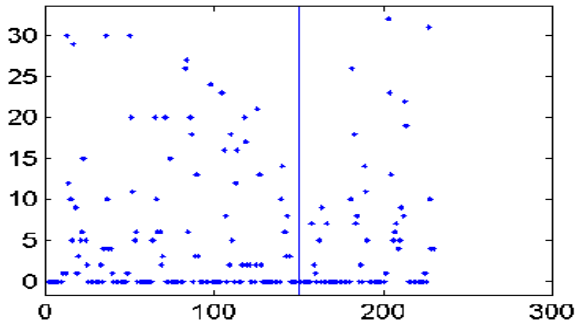
Population parameters estimates

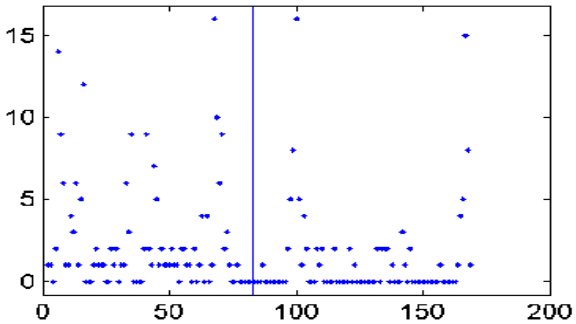
	Fixed Effect				Variance Term (ω^2)		
parameter	estimate	s.e.	r.s.e. (%)	p-value	estimate	s.e.	r.s.e. (%)
β_1^S	2.31	0.057	2		0.814	0.09	11
β_2^S	-0.435	0.074	17		2.23	0.18	8
$\log(\lambda_1^S)$	-1.87	0.056	3		1.85	0.12	6
$\log(\alpha^S)$	-0.09	0.055	61		1.64	0.12	7
δ_1	5.19	0.67	13		2.63	1.5	49
γ_1	-0.00266	0.00065	24	4.10^{-5}	0	-	-
δ_2	-0.478	0.21	43		0.246	0.14	58
γ_2	-0.00972	0.00046	5	0	0	-	-
δ_3	-0.69	0.17	25		1.83	0.25	14
γ_3	0.000769	0.00016	20	10^{-6}	0	-	-
δ_4	-0.307	0.13	42		0.129	0.076	59
γ_4	-0.00971	0.00024	2	0	0	-	-

MAP


0
5
10
15
20
25
0
50
100
150
200

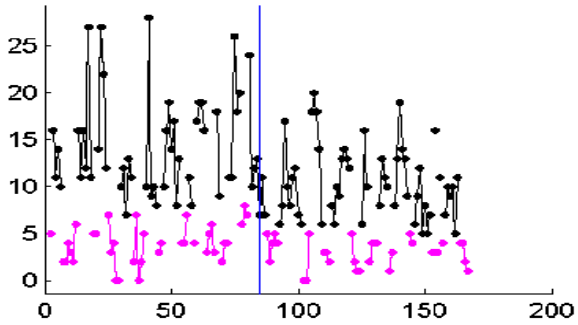

0
1
2
3
4
5
0
50
100
150

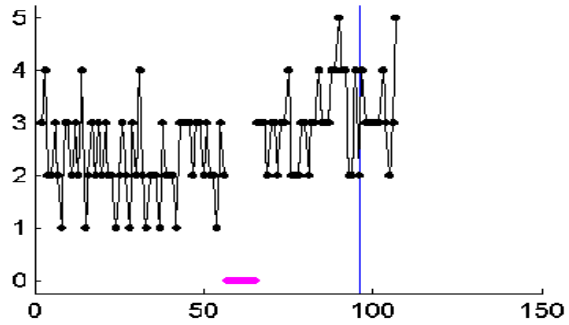

0
5
10
15
20
25
30
0
100
200
300

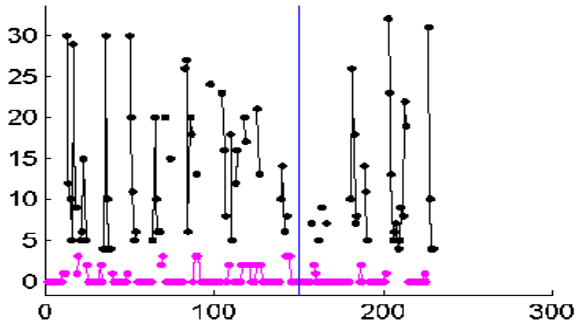

0
5
10
15
0
50
100
150
200

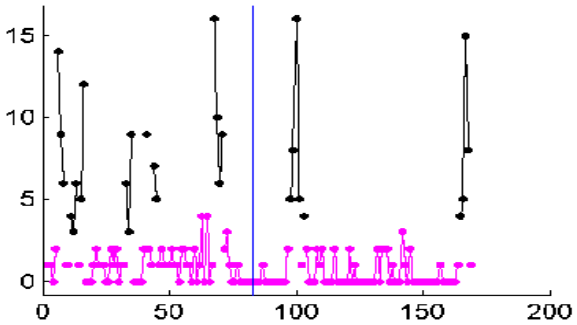
MAP

The individual parameters and the sequences of hidden states have then been estimated.









Conclusion

(A) **Application of MHMM to seizure count data**

- Mixed hidden Markov models are easy to interpret and provide a possible description of the seizure dynamics.
- Our models can also handle a dose effect.

Conclusion

(B) **Our new methodology**

→ Monte Carlo studies showed that our methodology has **good practical properties**:

- The population parameters are accurately estimated with SAEM (small bias and RMSE).
- The estimated s.e. give a good evaluation of the estimates' uncertainty.
- SAEM is fast.

→ Our algorithms are implemented in the **Monolix** software.

→ Our methodology for discrete state space models can be extended for continuous state space models (ex: SDE, see poster PAGE 2010).

Perspectives

- Dose **and** time-dependent drug effect.
- Extension to Generalized Poisson distributions.
- Selection of the number of hidden states.
- ...