Using Hamiltonian Monte-Carlo to design longitudinal count studies accounting for parameter and model uncertainties

Florence Loingeville[1], Thu Thuy Nguyen[1], Marie-Karelle Riviere[1,2], France Mentré[1]

[1]IAME-UMR 1137, INSERM and University Paris Diderot, Paris, France [2]Biostatistics & Programming Department, Sanofi, Paris, France

CONTEXT

The Fisher Information Matrix (FIM) can be used to design longitudinal studies for nonlinear mixed effect models (NLMEM) [1]. A Monte-Carlo Hamiltonian Monte-Carlo (MC/HMC) method has been developed to evaluate the FIM [2], then the D-optimality can be used to optimize designs. This approach however requires a priori knowledge on models and parameters, leading to design that are only locally optimal.

OBJECTIVES

- To extend the MC/HMC-based method to evaluate the FIM in NLMEM accounting for uncertainty in parameters and/or in models
- To illustrate this approach in robust design optimization for repeated count data

METHODS

Notations

M = population Fisher information matrix

M_R = robust population Fisher information matrix

Ξ ={N,ξ} = population design , N = number of individuals

ξ = elementary design (identical in all individuals)

=> To be optimized

ψ_m = population parameter values for model m

P_m = number of population parameters of model m

$p_m(\psi_m)$ = *a priori* population parameter distribution for model m

b = vector of random effects

y = vector of observations for one individual

α_m = weight quantifying balance between M models ($\sum \alpha_m = 1$)

1) Optimal design for given parameter values ψ_m* of a given model *m*

- Evaluation of **FIM** by **MC/HMC** [2]

$$M(\psi_m, \Xi) = N \, x M(\psi_m, \xi)$$

$$M(\psi_m, \xi) = E_y\left(\frac{\partial \log(L(y,\psi_m))}{\partial \psi_m}\frac{\partial \log(L(y,\psi_m))}{\partial \psi_m}^T\right)$$

with the likelihood $L(y, \psi_m) = \int \underbrace{p(y|b,\psi_m)}_{\text{pdf of observations } y \text{ given random effects } b} \underbrace{p(b|\psi_m)}_{\text{pdf of } b} db$

=>2 integrals to compute: w.r.t y (MC) and w.r.t b (HMC)

- Use of **D-optimality criterion**

$$\Phi_{D,m}(\Xi) = det(M({\psi_m}^*, \Xi))^{1/P_m}$$

2) Robust design accounting for parameter uncertainty for a given model *m*

- Evaluation of **robust FIM** by **MC/HMC**

$$M_R(\Xi) = E_{\psi_m}(M(\psi_m, \Xi))$$

- two integrals **w.r.t. *y*** and **w.r.t. *b*** for evaluation of $M(\psi_m, \Xi)$
- one supplementary integral **w.r.t. ψ_m** for evaluation $M_R(\Xi)$

⇒ Evaluation by **MC-HMC** using Stan (drawing jointly ψ_m and y by MC)

- Use of **DE-optimality criterion**

$$\Phi_{DE,m}(\Xi) = det(M_R(\Xi))^{1/P_m}$$

3) Robust design accounting for model uncertainty for given parameter values

- Proposition of a set of M candidate models
- Evaluation of **FIM** by **MC/HMC** for each model *m*
- Evaluation of **D-optimality criterion** on each model *m*
- Use of the **Compound D-optimality** criterion [3,4]

$$\Phi_{CD}(\Xi) = \prod_{m=1}^{M} \Phi_{D,m}(\Xi)^{\alpha_m} = \prod_{m=1}^{M} (det(M(\psi_m^*, \Xi)))^{\alpha_m/P_m}$$

4) Robust design accounting for parameter and model uncertainties

- Proposition of a set of M candidate models (m=1,...,M)
- Evaluation of **robust FIM** by **MC/HMC** for each model *m*
- Evaluation of **DE-optimality criterion** on each model *m*
- Use of the **Compound DE-optimality criterion**:

$$\Phi_{CDE}(\Xi) = \prod_{m=1}^{M} \Phi_{DE,m}(\Xi)^{\alpha_m} = \prod_{m=1}^{M} (\det(M_R(\Xi)))^{\alpha_m/P_m}$$

5) Implementation

- Extension of R package MIXFIM [5] using Stan to draw HMC samples and to calculate partial derivatives of the log-likelihood [6]
- *Robust_fisher_evaluation()* function to evaluate $M_R(\Xi)$
- *Combin_optimization()* function to perform combinatorial optimization of design elements in ξ
- *Compound_optimality()* function to evaluate the CD and CDE-optimality criteria

APPLICATION TO DESIGN OPTIMIZATION FOR COUNT DATA

Count data example

- Daily count of events that we want to prevent
- Poisson model for repeated count response [2]: $P(y = k|b) = \frac{\lambda^k e^{-\lambda}}{k!}$

λ: mean number of events / day

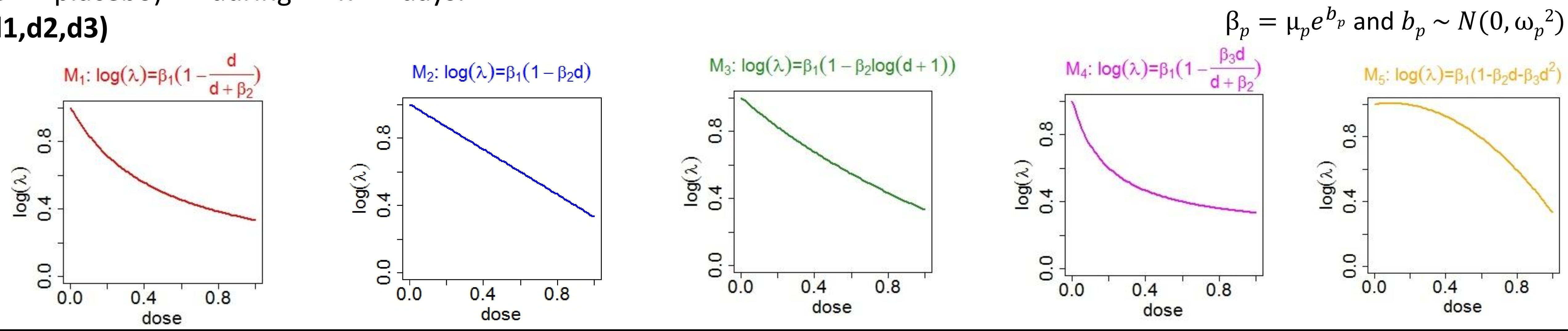
- Each patient observed at 3 dose levels (one placebo) during x days: **ξ=(d1,d2,d3)**

Candidate models

- $M_1: \log(\lambda) = \beta_1(1 - \frac{d}{d+\beta_2})$,
- $M_2: \log(\lambda) = \beta_1(1 - \beta_2 d)$,
- $M_3: \log(\lambda) = \beta_1(1 - \beta_2 \log(d+1))$,
- $M_4: \log(\lambda) = \beta_1\left(1 - \frac{\beta_3 d}{d+\beta_2}\right)$,
- $M_5: \log(\lambda) = \beta_1(-\beta_2 d^2 + \beta_3 d + 1)$.

	ψ_m*					$p_m(\psi_m)$				
	μ_1*	μ_2*	μ_3*	ω_1*	ω_2*	μ_1	μ_2 (CV(μ_2)=70%)	μ_3	ω_1	ω_2 (CV(ω_2)=90%)
M_1	1	0.5		0.3	0.3	1	*LN*(-0.89,0.63)		0.3	*LN*(-1.50,0.77)
M_2	1	0.67		0.3	0.3	1	*LN*(-0.60,0.63)		0.3	*LN*(-1.50,0.77)
M_3	1	0.96		0.3	0.3	1	*LN*(-0.24,0.63)		0.3	*LN*(-1.50,0.77)
M_4	1	0.2	0.8	0.3	0.3	1	*LN*(-1.81,0.63)	0.8	0.3	*LN*(-1.50,0.77)
M_5	1	0.8	0.13	0.3	0.3	1	*LN*(-0.60,0.63)	0.13	0.3	*LN*(-1.50,0.77)

$\beta_p = \mu_p e^{b_p}$ and $b_p \sim N(0, {\omega_p}^2)$

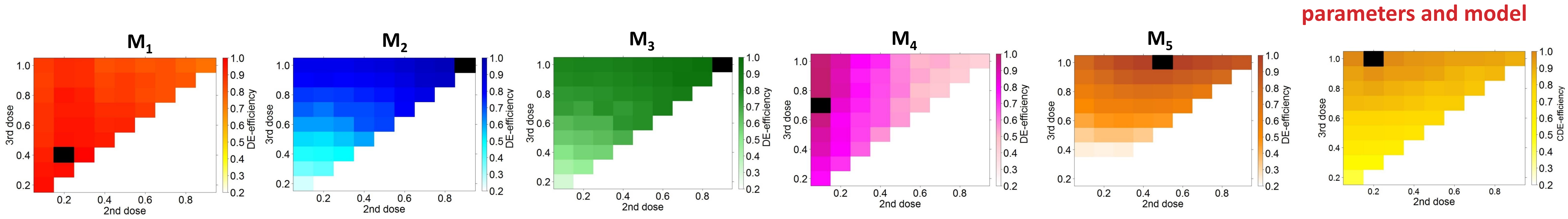


- N=60 subjects, nrep=10 rep/subject/dose
- Combinatorial optimization of 2 dose levels between 0 and 1 with step=0.1, without repetition, with d_1=0 (placebo)
- For computation of M_R: 5000 MC, 200 HMC
- **Robust design optimization w.r.t. parameters and model**

RESULTS

Robust design with respect to parameters for each model

Robust design with respect to parameters and model



$\xi_{DE,1}$=(0,0.2,0.4) $\xi_{DE,2}$=(0,0.9,1) $\xi_{DE,3}$=(0,0.9,1) $\xi_{DE,4}$=(0,0.1,0.7) $\xi_{DE,5}$=(0,0.5,1) ξ_{CDE}=(0,0.2,1)

$$E_{X,m}(\Xi) = \frac{\Phi_{X,m}(\Xi)}{\Phi_{X,m}(\Xi_{X,m})} \quad \text{for X=DE,CDE}$$

	$E_{DE,1}(\Xi)$	$E_{DE,2}(\Xi)$	$E_{DE,3}(\Xi)$	$E_{DE,4}(\Xi)$	$E_{DE,5}(\Xi)$
$\xi_{DE,1}$=(0, 0.2, 0.4)	100%	46.9%	56.7%	77.5%	23.6%
$\xi_{DE,2}$=(0, 0.9, 1)	73.3%	100%	100%	43.5%	87.1%
$\xi_{DE,3}$=(0, 0.9, 1)	73.3%	100%	100%	43.5%	87.1%
$\xi_{DE,4}$=(0, 0.1, 0.7)	89.1%	68.1%	73.9%	100%	51.4%
$\xi_{DE,5}$=(0, 0.5, 1)	83.1%	87.8%	89.6%	58.5%	100%
ξ_{CDE}=(0, 0.2, 1)	90.9%	83.8%	83.9%	84.6%	82.8%

CONCLUSIONS

- Proposition of a new MC/HMC-based method for computation of the FIM avoiding linearization, introducing robustness w.r.t. parameters and/or models
- First applications to design optimization for count data
- Accounting or not for uncertainty in parameters may lead to different allocations of optimal doses but has little impact on efficiencies
- Misspecification of models can lead to low efficiencies
- The CD/CDE-optimal designs provided a good compromise for different candidate models

PERSPECTIVES

- Replacement of MC in MC/HMC by more efficient approach: quasi-random sampling [7]
- Implementation of a more efficient optimization algorithm
- Combination of these methods with adaptive designs [8]

References

[1] Fedorov and Leonov, CRC Press, 2013
[2] Riviere, Ueckert and Mentré, Biostatistics, 2017
[3] Atkinson, Donev and Tobias, Oxford University press, 2009
[4] Nguyen, Benech, Delaforge and Lenuzza, Pharm Stat, 2016
[5] Riviere and Mentré, R package MIXFIM, 2015
[6] Stan developement team, RStan, 2012
[7] Pan and Thompson, Comput Stat Data Anal, 2007
[8] Lestini, Dumont and Mentré, Pharm Res, 2015

Acknowledgments

Integrated DEsign and AnaLysis of small population group trials

This work was supported by the European Union 7th Framework Programme under Grant Agreement n602552. The authors thank Francois Cohen as well as Hervé Le Nagard for the use of the "Centre de Biomodélisation".