

Evaluation of the Fisher information matrix in nonlinear mixed effect models using adaptive Gaussian quadrature

Thu Thuy Nguyen, France Mentré

► **To cite this version:**

Thu Thuy Nguyen, France Mentré. Evaluation of the Fisher information matrix in nonlinear mixed effect models using adaptive Gaussian quadrature. *Computational Statistics and Data Analysis*, Elsevier, 2014, 80, pp.57 - 69. <10.1016/j.csda.2014.06.011>. <inserm-01077176>

HAL Id: inserm-01077176

<http://www.hal.inserm.fr/inserm-01077176>

Submitted on 24 Oct 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Evaluation of the Fisher information matrix in nonlinear mixed effect models using adaptive Gaussian quadrature

Thu Thuy Nguyen*, France Mentré

*IAME, UMR 1137, INSERM, F-75018 Paris, France;
University Paris Diderot, Sorbonne Paris Cité, F-75018 Paris, France*

Abstract

Nonlinear mixed effect models (NLMEM) are used in model-based drug development to analyse longitudinal data. To design these studies, the use of the expected Fisher information matrix (M_F) is a good alternative to clinical trial simulation. Presently, M_F in NLMEM is mostly evaluated with first-order linearisation. The adequacy of this approximation is, however, influenced by model nonlinearity. Alternatives for the evaluation of M_F without linearisation are proposed, based on Gaussian quadratures. The M_F , expressed as the expectation of the derivatives of the log-likelihood, can be obtained by stochastic integration. The likelihood for each simulated vector of observations is approximated by Gaussian quadrature centred at 0 (standard quadrature) or at the simulated random effects (adaptive quadrature). These approaches have been implemented in R. Their relevance was compared with clinical trial simulation and linearisation, using dose-response models, with various nonlinearity levels and different number of doses per patient. When the nonlinearity was mild, three approaches based on M_F gave correct predictions of standard errors, when compared with the simulation. When the nonlinearity increased, linearisation correctly predicted standard errors of fixed effects, but over-predicted, with sparse designs, standard errors of some variability terms. Meanwhile, quadrature approaches gave correct predictions of standard errors overall, but standard Gaussian quadrature was very time-consuming when there were more than two random effects. To conclude, adaptive Gaussian quadrature is a relevant alternative for the evaluation of M_F for models with stronger nonlinearity, while being more computationally efficient than standard quadrature.

Keywords: Design, Dose-response studies, Fisher information matrix, adaptive Gaussian quadrature, Linearisation, Nonlinear mixed effect model

1. Introduction

Nonlinear mixed effect models (NLMEM) are frequently used in model-based drug development to analyse longitudinal data obtained during clinical trials (Lalonde et al., 2007; Smith and Vincent, 2010). They were introduced about 40 years ago (Sheiner et al., 1972, 1977) and were initially used in pharmacokinetic analyses as an alternative to the non-compartmental approach (NCA) (Gabrielson and Weiner, 2006). This modelling approach is more complex than NCA, but

*Correspondance to: Thu Thuy Nguyen, IAME, UMR 1137, INSERM - University Paris Diderot, 16 Henri Huchard, 75018, Paris, France (Email: thu-thuy.nguyen@inserm.fr, Tel: +33157277540)

allows for the analysis of few samples per subject. It accounts for within and between subject variability, and is appropriate for exploiting the richness of repeated measurements. Consequently, this approach is increasingly used in the biomedical field, not only for pharmacokinetic analyses (Sheiner et al., 1972, 1977), but also for analyses of viral loads (Perelson and Ribeiro, 2008), of bacterial resistance to antibiotics (Nielsen et al., 2007), and of the dose-response relationship. This approach has become the main statistical tool in pharmacometrics, the science of quantitative pharmacology (Van der Graaf, 2012). Parameters of these models are commonly estimated by likelihood maximisation (Dartois et al., 2007). However, the nonlinearity of the structural model prevents a closed form solution for the integration over the random effects in the expression of the likelihood function. Many approaches have been proposed over the years to overcome this difficulty, and implemented in several estimation software packages. These are first-order marginal quasi-likelihood or first-order linearisation (Lindstrom and Bates, 1990) in NONMEM, R, Splus, Laplace approximation (Wolfinger, 1993) in NONMEM and SAS, adaptive Gaussian quadrature (Pinheiro and Bates, 1995) in SAS and in the R package *lme4*, Stochastic Approximation Expectation Maximisation (SAEM) (Kuhn and Lavielle, 2005) in MONOLIX and NONMEM. Pillai et al. (2005) described these estimation methods in a review paper and recently, Plan et al. (2012) compared their performance, showing that adaptive Gaussian quadrature, although the slowest, was generally the best method.

Before the modelling step to estimate parameters, it is important to define an appropriate design, which consists in determining a balance between the number of subjects and the number of samples per subject, as well as the allocation of times and doses, according to experimental conditions. The choice of design is crucial for an efficient estimation of model parameters (Al-Banna et al., 1990; Hashimoto and Sheiner, 1991; Jonsson et al., 1996), especially when the studies are conducted in children or in patients where only a few samples can be taken per subject. The main approach for design evaluation has long been based on clinical trial simulation (CTS), but it is a cumbersome method and so the number of designs that can be evaluated is limited. An alternative approach has been described in the general theory of optimum experimental design used for classical nonlinear models (Atkinson et al., 2007; Walter and Pronzato, 2007; Atkinson et al., 2014), relying on the Rao-Cramer inequality which states that the inverse of the Fisher information matrix (M_F) is the lower bound of the variance-covariance matrix of any unbiased estimate of the parameters and its

diagonal elements are the expected variances of the parameters. Several criteria based on M_F have been developed to evaluate designs. One of the criteria widely used is the criterion of D-optimality, which consists in maximising the determinant of M_F . The computation of this criterion requires a priori knowledge of the model and its parameters, which can usually be obtained from previous experiments. This leads to the concept of "locally optimal designs", which has been studied in several publications (Chernoff, 1953; Box and Lucas, 1959; D'Argenio, 1981). Since there is no closed form of the likelihood in NLMEM, there is no analytical expression of M_F . An approximation of the expected M_F has been proposed for NLMEM, using first order linearisation of the model around the random effect expectation (Mentré et al., 1997; Retout et al., 2002; Bazzoli et al., 2009). This approach has been implemented in several software programs (Bazzoli et al., 2010; Leonov and Aliev, 2012; Gueorguieva et al., 2007; Nyberg et al., 2012) such as PFIM (INSERM, University Paris Diderot), POPED (University of Uppsala), POPDES (University of Manchester), and POPT (University of Otago), frequently used to design new studies in academia as well as in pharmaceutical companies (Mentré et al., 2013).

However, it has been shown that the use of the linearisation (LIN) approach is only appropriate if the variances of the random effects are small, or the nonlinearity is mild (Jones and Wang, 1999; Jones et al., 1999). Consequently, as pointed out by Han and Chaloner (2005), when an optimal design is found using this approximation but the estimation is carried out using a true NLMEM, the performance of the design needs further investigation. The nonlinearity of a model with respect to its parameters is defined from the behaviour of the first order derivatives of the model function. Its consequences on the structural identifiability of a model have been studied by Walter and Pronzato (1995). The notion of level of nonlinearity ("mild" or "strong" as mentioned in this paper) is derived from the term "close to linear" introduced when evaluating a nonlinear model's behaviour (Fletcher and Powell, 1963). Measures of nonlinearity have been studied in several publications (Bates and Watts, 1980; Cook and Goldberg, 1986; Smyth, 2002) in order to evaluate whether the "close to linear" condition is satisfied and to indicate if the linear approximation is reasonable or questionable. When first-order linearisation is to be avoided, alternative approaches are necessary. Various new approaches have been proposed, but these are not always better than the usual first-order linearisation. For instance, linearisation of the model around the individual values of the random effects (Retout and Mentré, 2003) or around the expected mode of the marginal likelihood (Nyberg et al.,

2012) has been proposed but is quite time consuming because Monte Carlo simulations are needed. Other approaches based on the Laplace approximation (Vong et al., 2012) or Monte Carlo integration (Mielke, 2012) give correct predictions for the precision of parameter estimation, but they are also very time consuming. Another possible alternative for computing the Fisher information matrix in designing studies is the use of Gaussian quadrature rules. This consists in approximating integrals of functions with respect to a given probability density by a weighted sum of function values at abscissas chosen within the integration domain. It has been shown that adaptive Gaussian quadrature (AGQ) performs better than standard Gaussian quadrature (GQ) in estimation (Pinheiro and Bates, 1995); the difference between the two approaches is that the grid of nodes is centred at the expectation of the random effects in GQ while it is centred at the conditional modes of the random effects in AGQ. Neither approach has ever been proposed for designing studies with different types of models, except in an example of an HIV dynamic model written with ordinary differential equations (Guedj et al., 2007).

In this context, we aim to propose alternatives to linearisation for evaluating the predicted Fisher information matrix, based on GQ and AGQ. In order to challenge and investigate the performance of both new approaches as well as linearisation compared with CTS, we use examples of dose-response trials inspired by the article of Plan et al. (2012) comparing different estimation methods. Dose-response studies are of critical importance in drug development and need to be planned carefully (Bretz et al., 2010; Pronzato, 2010; McGree et al., 2012) but little has been done to study their design in the context of NLMEM. We consider the sigmoid E_{\max} model, with various degrees of nonlinearity (i.e. different sigmoidicity coefficients), and in addition a linear model where M_F can be calculated exactly.

We introduce the necessary notation for the design and model and present the new GQ and AGQ approaches developed for computing M_F in Section 2. The performance of both approaches is evaluated and compared with CTS and LIN for different scenarios in Section 3 and is discussed in Section 4.

2. Computing Fisher information matrix in NLMEM with Gaussian quadratures

2.1. Design

The elementary design ξ_i of individual i ($i = 1, \dots, N$) is defined by the number n_i of observations and the design variables $(x_{i1}, \dots, x_{in_i})$. Consequently, the population design for N individuals can be defined as $\Xi = \{\xi_1, \dots, \xi_N\}$. Usually, population designs are composed of a limited number Q of groups of individuals with identical designs in each group. Each of these groups is composed of a global elementary design ξ_q and is performed in a number N_q of subjects. The population design can thus be written as $\Xi = \{[\xi_1, N_1]; \dots; [\xi_Q, N_Q]\}$.

2.2. Nonlinear mixed effect model

We denote by y_i the n_i -vector of observations for the individual i obtained with the design ξ_i and by f the known function describing the nonlinear structural model. The NLMEM linking the response y_i to the samples $\xi_i = (x_{i1}, \dots, x_{in_i})$ can be written as

$$y_i = f(\phi_i, \xi_i) + \varepsilon_i, \quad (1)$$

where ε_i is the vector of random errors which follows a standard normal distribution $\mathcal{N}(0, \sigma^2 I_{n_i})$ and I_{n_i} is an $n_i \times n_i$ identity matrix. The vector of individual parameters ϕ_i can be expressed as function g of μ , the P -vector of fixed effects and of b_i , the P -vector of random effects for individual i . g can be additive (for normal parameters), so that the p^{th} component of ϕ_i is written as

$$\phi_{ip} = g(\mu_p, b_{ip}) = \mu_p + b_{ip}, \quad (2)$$

or exponential (for lognormal parameters), so

$$\phi_{ip} = g(\mu_p, b_{ip}) = \mu_p \exp(b_{ip}). \quad (3)$$

It is assumed that $b_i \sim \mathcal{N}(0, \Omega)$, with Ω defined here as a diagonal variance-covariance matrix of size $P \times P$. Each element ω_p^2 of Ω represents the variance of the p^{th} component of b_i . As usual, the following assumptions are made: $\varepsilon_i | b_i$ are independent between subjects, and ε_i and b_i are

independent for each subject. The model can also be written as

$$y_i = f(g(\boldsymbol{\mu}, b_i), \xi_i) + \varepsilon_i. \quad (4)$$

Let $\boldsymbol{\lambda} = (\omega_1, \dots, \omega_p, \sigma)'$ be the vector of standard deviations of random effects and standard deviation of random error. $\boldsymbol{\Psi}$, the vector of all population parameters to be estimated, is $\boldsymbol{\Psi} = (\boldsymbol{\mu}', \boldsymbol{\lambda}')$.

2.3. Fisher information matrix

The individual Fisher information matrix is defined from the likelihood of the observations y_i of individual i , which is expressed by the following integral

$$L(y_i; \boldsymbol{\Psi}) = \int_{R^p} p(y_i|b_i; \boldsymbol{\Psi}) p(b_i) db_i, \quad (5)$$

where $p(y_i|b_i; \boldsymbol{\Psi})$ is the conditional density of the observations y_i given the random effects b_i , and $p(b_i)$ is the density of $b_i \sim \mathcal{N}(0, \Omega)$. Then, the Fisher information matrix for the elementary design ξ_i is the following expectation taken with respect to the distribution of the observations $p(y; \boldsymbol{\Psi})$

$$M_F(\boldsymbol{\Psi}; \xi_i) = E \left(\frac{\partial \log L(y_i; \boldsymbol{\Psi})}{\partial \boldsymbol{\Psi}} \frac{\partial \log L(y_i; \boldsymbol{\Psi})'}{\partial \boldsymbol{\Psi}} \right). \quad (6)$$

As the individuals are independent, the Fisher information matrix $M_F(\boldsymbol{\Psi}, \Xi)$ for a population design Ξ is defined by the sum of the N elementary matrices $M_F(\boldsymbol{\Psi}, \xi_i)$, so that

$$M_F(\boldsymbol{\Psi}, \Xi) = \sum_{i=1}^N M_F(\boldsymbol{\Psi}, \xi_i). \quad (7)$$

In the case of a limited number Q of elementary designs ξ_q , we have

$$M_F(\boldsymbol{\Psi}, \Xi) = \sum_{q=1}^Q N_q M_F(\boldsymbol{\Psi}, \xi_q). \quad (8)$$

However, there is no analytical expression for the likelihood $L(y_i; \boldsymbol{\Psi})$ in NLMEM. Approximations such as linearisation can be used to approximate the Fisher information matrix (Mentré et al., 1997; Retout et al., 2002; Bazzoli et al., 2009). The predicted M_F after linearisation is a block matrix with a block corresponding to derivatives of the log-likelihood with respect to the fixed effects, a

block for derivatives with respect to the standard derivation terms and a block containing mixed derivatives with respect to all parameters. In our work, the block of mixed derivatives was set to 0 for linearisation, based on publications showing the better performance of the block diagonal expression compared with the full one (Mielke and Schwabe, 2010; Fedorov and Leonov, 2014; Nyberg et al., 2014).

2.4. Gaussian quadrature and adaptive Gaussian quadrature

When linearisation is to be avoided, alternatives such as Gaussian quadratures can be used to express the likelihood analytically. For simplicity, we omit the index i for the individual in this section. The elementary Fisher information matrix $M_F(\Psi, \xi)$ in (6) for an individual with elementary design ξ of n observations can also be written as

$$M_F(\Psi, \xi) = \int_{R^n} h(y; \Psi) p(y; \Psi) dy, \quad (9)$$

where $p(y; \Psi)$ is the marginal density of y and

$$h(y; \Psi) = \frac{\partial \log L(y; \Psi)}{\partial \Psi} \frac{\partial \log L(y; \Psi)}{\partial \Psi'} = \frac{\frac{\partial L(y; \Psi)}{\partial \Psi} \frac{\partial L(y; \Psi)}{\partial \Psi'}}{L(y; \Psi)^2}. \quad (10)$$

First, we propose to evaluate the integral of $h(y; \Psi)$ with respect to $p(y; \Psi)$ in (9) by stochastic integration. We generate M vectors of observations y_m ($m = 1, \dots, M$), each vector with the same design ξ of n samples, under the same probability distribution $p(y; \Psi)$. Thus, an approximation of (9) can be obtained by

$$M_F(\Psi, \xi) = \frac{1}{M} \sum_{m=1}^M h(y_m; \Psi). \quad (11)$$

Then, in order to obtain an analytical expression of $h(y_m; \Psi)$ as expressed in (10) for each simulated y_m , we need to compute the likelihood $L(y_m; \Psi)$. We define $\eta = \Omega^{-1/2}b$, then $\eta \sim \mathcal{N}(0, I_P)$, where I_P is the identity matrix of dimension $P \times P$. Consequently, the likelihood in (5) can also be written as

$$L(y_m; \Psi) = \int_{R^P} p(y_m | \eta; \Psi) p(\eta) d\eta, \quad (12)$$

where $p(\eta)$ is the density of $\mathcal{N}(0, I_P)$. This integral can be evaluated numerically by the Gaussian quadrature approach, as a weighted sum of $p(y_m | \eta; \Psi)$ evaluated at pre-determined nodes chosen

within the distribution of η (Pinheiro and Bates, 1995). Quadrature weights and nodes for approximating one-dimensional integrals are those of the standard Gauss-Hermite rule (Abramowitz and Stegun, 1964; Golub and Welsh, 1969; Golub, 1973). The nodes are the roots of the Hermite polynomial of degree κ , obtained from successive derivatives of $\exp(\eta^2)$ (Press et al., 1992). As in Pinheiro and Bates (1995), the P -dimension integral here was transformed into P successive one-dimensional integrals. The distribution of nodes and weights used in standard GQ does not take into account the nature of the integrand, which is why standard GQ requires a high number of nodes in order to provide a correct estimation of parameters. AGQ, by centring and scaling the standard quadrature nodes, places the nodes according to the areas of high density and improves the approximation. Consequently it requires fewer nodes compared with standard GQ, while giving better performance in estimation (Pinheiro and Bates, 1995).

In order to approximate each one of these integrals in $L(y_m; \Psi)$ using standard GQ, for all simulated y_m , κ quadrature nodes and weights are selected from the standard normal distribution $\mathcal{N}(0, 1)$. Then the integral of dimension P can be approximated by the weighted sum over K nodes η_k , where $K = \kappa^P$ and η_k are vectors in R^P , associated with weights w_k ($k = 1, \dots, K$). So, we obtain the following analytical expression

$$L(y_m; \Psi) = \sum_{k=1}^K w_k p(y_m | \eta_k; \Psi), \quad (13)$$

where $y_m | \eta_k \sim \mathcal{N}(f_k, \sigma^2 I_n)$ with $f_k = f(g(\mu, \Omega^{1/2} \eta_k), \xi)$. Thus, we can write

$$p(y_m | \eta_k; \Psi) = (2\pi\sigma^2)^{-n/2} \exp\left(\frac{-1}{2\sigma^2} (y_m - f_k)' (y_m - f_k)\right), \text{ and} \quad (14)$$

$$\log p(y_m | \eta_k; \Psi) = -\frac{n}{2} \log 2\pi - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} (y_m - f_k)' (y_m - f_k). \quad (15)$$

The likelihood $L(y_m; \Psi)$ can also be approximated by AGQ. Since a vector of standardised random effects η_m is simulated when generating each y_m , we select $K = \kappa^P$ nodes η_{mk} and weights

w_k based on $\tilde{p}(\eta)$ which is the $\mathcal{N}(\eta_m, I_P)$ density. We can write

$$L(y_m; \Psi) = \int_{R^P} \frac{p(y_m | \eta; \Psi) p(\eta)}{\tilde{p}(\eta)} \tilde{p}(\eta) d\eta \quad (16)$$

$$= \sum_{k=1}^K w_k \frac{p(\eta_{mk})}{\tilde{p}(\eta_{mk})} p(y_m | \eta_{mk}; \Psi), \quad (17)$$

where $y_m | \eta_{mk} \sim \mathcal{N}(f_{mk}, \sigma^2 I_n)$ with $f_{mk} = f(g(\mu, \Omega^{1/2} \eta_{mk}), \xi)$, and the weights for AGQ can be written as

$$w_k \frac{p(\eta_{mk})}{\tilde{p}(\eta_{mk})} = w_k \frac{(2\pi)^{-P/2} \exp(-\frac{1}{2} \eta'_{mk} \eta_{mk})}{(2\pi)^{-P/2} \exp(-\frac{1}{2} (\eta_{mk} - \eta_m)' (\eta_{mk} - \eta_m))} \quad (18)$$

$$= w_k \exp\left(-\frac{1}{2} (\eta'_{mk} \eta_{mk} - (\eta_{mk} - \eta_m)' (\eta_{mk} - \eta_m))\right). \quad (19)$$

Finally, using (13) or (17) in (10) and in (11), the individual Fisher information matrix can be approximated as

$$M_F(\Psi, \xi) = \frac{1}{M} \sum_{m=1}^M \frac{\left(\sum_{k=1}^K w_{mk} \frac{\partial p_{mk}(y_m; \Psi)}{\partial \Psi}\right) \left(\sum_{k=1}^K w_{mk} \frac{\partial p_{mk}(y_m; \Psi)}{\partial \Psi'}\right)}{\left(\sum_{k=1}^K w_{mk} p_{mk}(y_m; \Psi)\right)^2} \quad (20)$$

$$= \frac{1}{M} \sum_{m=1}^M \frac{\left(\sum_{k=1}^K w_{mk} p_{mk}(y_m; \Psi) \frac{\partial \log p_{mk}(y_m; \Psi)}{\partial \Psi}\right) \left(\sum_{k=1}^K w_{mk} p_{mk}(y_m; \Psi) \frac{\partial \log p_{mk}(y_m; \Psi)}{\partial \Psi'}\right)}{\left(\sum_{k=1}^K w_{mk} p_{mk}(y_m; \Psi)\right)^2}. \quad (21)$$

When using GQ, $w_{mk} = w_k$ and $p_{mk}(y_m; \Psi) = p(y_m | \eta_k; \Psi)$ as defined in (14). When using AGQ, $w_{mk} = w_k \frac{p(\eta_{mk})}{\tilde{p}(\eta_{mk})}$ as defined in (19) and $p_{mk}(y_m; \Psi) = p(y_m | \eta_{mk}; \Psi)$ as defined in (14) with f_{mk} instead of f_k .

As $\Psi = (\mu', \lambda')'$, the Fisher information matrix $M_F(\Psi, \xi)$ can be written as a block matrix

$$M_F(\Psi; \xi) = \begin{pmatrix} M_F(\mu; \xi) & M_F(\mu, \lambda; \xi) \\ M_F(\mu, \lambda; \xi) & M_F(\lambda; \xi) \end{pmatrix}. \quad (22)$$

To compute $M_F(\mu; \xi)$, we derive the log-likelihood with respect to each element μ_p of the vector

of fixed effects $\mu = (\mu_1, \dots, \mu_p)'$, using

$$\frac{\partial \log p_{mk}(y_m; \Psi)}{\partial \mu_p} = \frac{1}{\sigma^2} \frac{\partial f_k}{\partial \mu'_p} (y_m - f_k), \quad (23)$$

To compute $M_F(\lambda; \xi)$, we derive the log-likelihood with respect to each element ω_p or σ of the vector of standard deviations $\lambda = (\omega_1, \dots, \omega_p, \sigma)'$, using

$$\frac{\partial \log p_{mk}(y_m; \Psi)}{\partial \omega_p} = \frac{1}{\sigma^2} \frac{\partial f_k}{\partial \omega'_p} (y_m - f_k), \text{ and} \quad (24)$$

$$\frac{\partial \log p_{mk}(y_m; \Psi)}{\partial \sigma} = -\frac{n}{\sigma} + \frac{1}{\sigma^3} (y_m - f_k)'(y_m - f_k), \quad (25)$$

with f_{mk} instead of f_k when using AGQ instead of GQ. Finally, using (23, 24, 25) in (21), we obtain each element of $M_F(\mu; \xi)$ in (22).

3. Evaluation by simulation

3.1. Evaluation models and scenarios

Our evaluation examples were inspired by a previous dose-response simulation study (Plan et al., 2012) which mimicked clinical trials including 100 subjects with identical individual design ξ and investigating several dose levels among 0, 100, 300 and 1000 dose units.

Table 1: Different models and designs evaluated

Structural model	Number of parameters P	Number of doses per subject n	Doses	Name	
Linear	2	4	(0, 100, 300, 1000)	M1R	
		2	(100, 300)	M1S	
E_{\max} (E_0 not estimated, fixed at same value in all subjects)	2	$\gamma = 1$	4	(0, 100, 300, 1000)	M2R
			2	(300, 1000)	M2S
		$\gamma = 3$	4	(0, 100, 300, 1000)	M3R
			2	(300, 1000)	M3S
E_{\max} (E_0 estimated, varying between subjects)	3	$\gamma = 1$	4	(0, 100, 300, 1000)	M4R
			3	(0, 300, 1000)	M4S
		$\gamma = 3$	4	(0, 100, 300, 1000)	M5R
			3	(0, 300, 1000)	M5S

First, we considered a linear model of the dose-response relationship describing, for subject i , an effect E_{ij} for the j^{th} dose d_{ij} with two structural parameters: baseline E_0 and slope S .

$$E_{ij} = E_{0_i} + S_i \times d_{ij} + \varepsilon_{ij}. \quad (26)$$

The vector of the fixed effects μ was composed of $\mu_{E_0} = 5$ and $\mu_S = 0.06$. The random effect model was additive with standard deviations $\omega_{E_0} = 1.5$ and $\omega_S = 0.018$. The standard deviation σ of random error ε_{ij} was equal to 2. We studied two designs, a rich one $\xi = (0, 100, 300, 1000)$ and a sparse one $\xi = (100, 300)$. We chose to evaluate first a simple linear mixed effect model (denoted by M1) because in this case the analytical form of the Fisher information matrix is known and so we could easily check whether our calculation and implementation are correct.

Next, we considered a sigmoid E_{\max} model of the dose-response relationship, describing for subject i , an effect E_{ij} for the j^{th} dose d_{ij} with the following structural parameters: baseline E_0 , maximal effect E_{\max} , dose ED_{50} at which 50% of the maximal effect is achieved and sigmoidicity coefficient γ , i.e. the degree of nonlinearity of the function shape.

$$E_{ij} = E_{0_i} + \frac{E_{\max_i} \times d_{ij}^\gamma}{ED_{50_i}^\gamma + d_{ij}^\gamma} + \varepsilon_{ij}. \quad (27)$$

We considered a model with $\gamma = 1$ and a model with $\gamma = 3$. In each model, γ was known with the same value in all subjects (i.e. μ_γ not estimated, $\omega_\gamma = 0$). We first studied models with two structural parameters: E_{\max} and ED_{50} where E_0 was fixed at the same value 5 in all subjects (i.e. μ_{E_0} not estimated, $\omega_{E_0} = 0$). The fixed effects μ were $\mu_{E_{\max}} = 30$, $\mu_{ED_{50}} = 500$ and the random effects were exponential with $\omega_{E_{\max}} = \omega_{ED_{50}} = 0.3$. We denote these models M2 and M3 for $\gamma = 1$ and $\gamma = 3$, respectively. Then, we studied models with three structural parameters: E_{\max} , ED_{50} and E_0 , where E_0 can vary from one subject to another and was estimated. The fixed effects μ were $\mu_{E_{\max}} = 30$, $\mu_{ED_{50}} = 500$, $\mu_{E_0} = 5$ and the random effects were exponential with $\omega_{E_{\max}} = \omega_{ED_{50}} = \omega_{E_0} = 0.3$. We denote these models M4 and M5 for $\gamma = 1$ and $\gamma = 3$ respectively. The standard deviation σ of random error ε_{ij} was equal to 2 in all models. Designs ξ with 2, 3, or 4 doses among (0, 100, 300, 1000) were studied. The list of the 5 models above and the designs studied are detailed in Table 1. The rich design is referred to as R and the sparse design as S. Examples of one dataset simulated

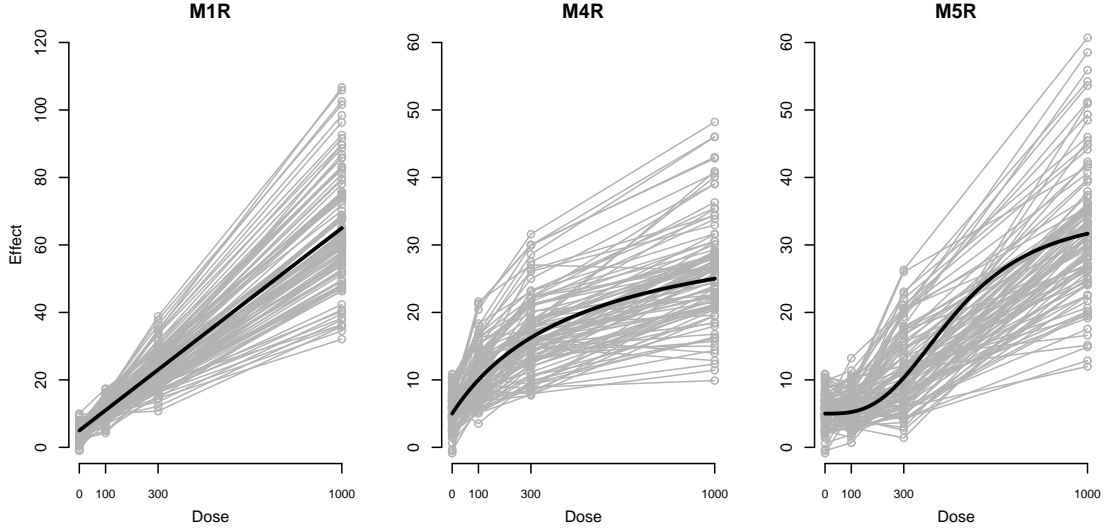


Figure 1: Individual response versus dose profiles (in grey) for one dataset simulated using linear model (M1R), E_{\max} model with sigmoidicity coefficient $\gamma = 1$ (M4R) and E_{\max} model with $\gamma = 3$ (M5R), with 4 doses per subject (0, 100, 300, 1000). The black curves represent the response predicted by the model for the fixed effect parameters.

using linear model, E_{\max} model with 3 structural parameters, $\gamma = 1$ or $\gamma = 3$ and the rich design of 4 doses per subject (M1R, M4R and M5R respectively) are plotted in Figure 1.

We plotted the first and second order derivatives to examine graphically the model nonlinearity (Figure 2). The first derivatives with respect to E_{\max} and ED_{50} vary over design samples and are presented in Figure 2A. We notice that the derivative with respect to E_{\max} depends on ED_{50} and the one with respect to ED_{50} depends on both E_{\max} and ED_{50} . Thus the optimal design depends on values of E_{\max} and ED_{50} but not on E_0 . One class of nonlinearity measures is based on the second order derivatives (Smyth, 2002): if the second derivatives of model f_1 are smaller in absolute values than f_2 , then model f_1 has "closer to linear" behaviour. The second derivatives that are different from 0 are plotted against doses in Figure 2B. The magnitudes of the curves are greater when $\gamma = 3$ than when $\gamma = 1$, indicating that model nonlinearity increases with γ . This observation was confirmed by Plan et al. (2012) who showed decreasing performance of the linearisation approach along the γ -increase when estimating parameters for E_{\max} models.

3.2. Comparison of standard errors between predictions and clinical trial simulation

Our aim was to evaluate and compare the standard errors (SE) predicted using M_F by GQ and AGQ with those obtained by CTS and LIN.

We performed clinical trial simulations in R 2.14.1 with the model, parameters and designs

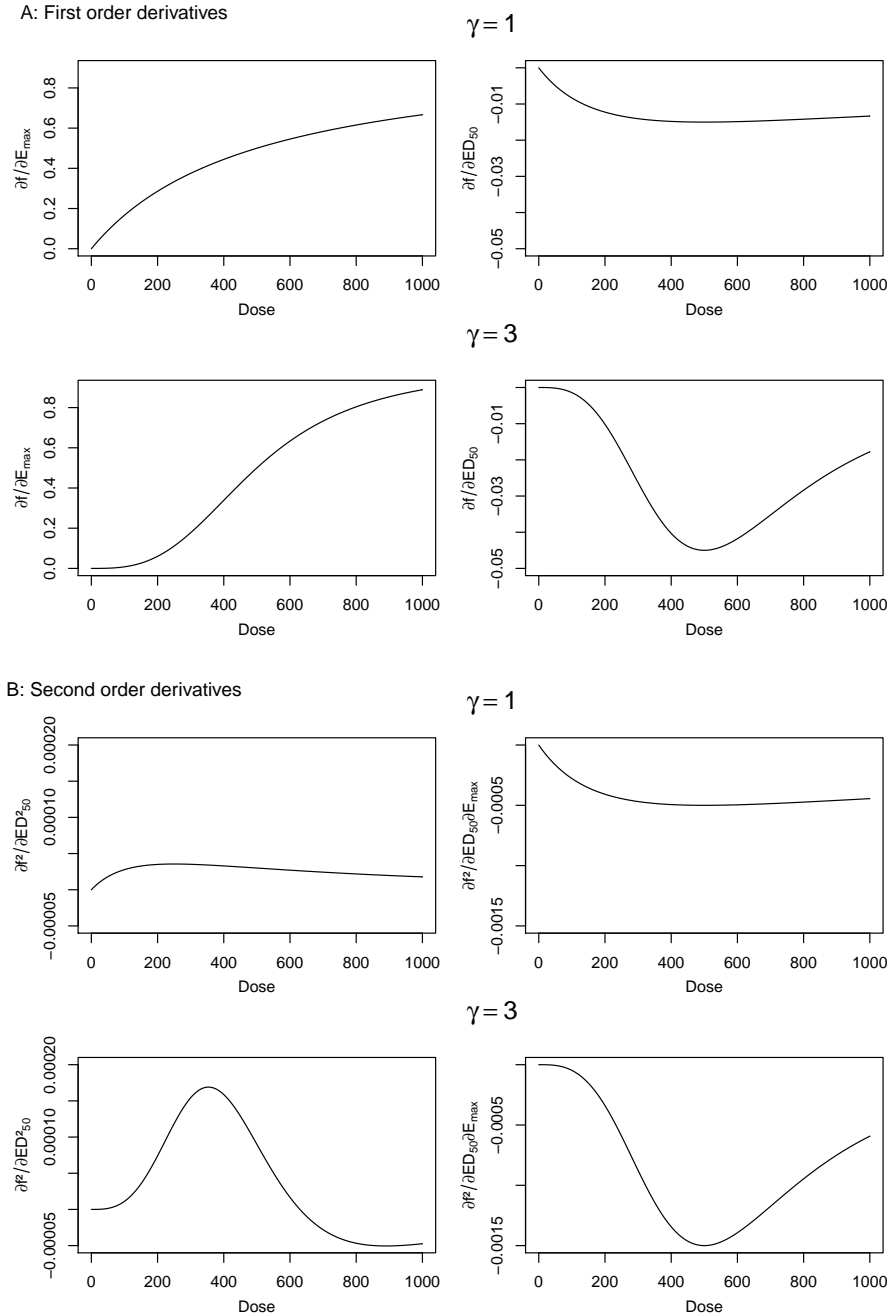


Figure 2: First order (panel A) and second order (panel B) derivatives of the function f of the E_{\max} model with respect to the structural parameters E_{\max} and ED_{50} for two sigmoidicity coefficients $\gamma = 1$ and $\gamma = 3$.

above. For each scenario, 1000 datasets of 100 subjects were simulated. Each simulated dataset was then analysed in MONOLIX 3.2 (www.lixoft.eu). Population parameters were estimated by the SAEM algorithm (Kuhn and Lavielle, 2005). Next, the SE by CTS, defined as the unbiased sample estimate of the standard deviation from the 1000 parameter estimates, were calculated. The relative

standard errors (RSE) by CTS were defined as the ratio of the SE to the true value of parameters. As in the publication by Retout and Mentré (2003), the 95% confidence intervals for each RSE by CTS are given, computed as $\left[\sqrt{\frac{q_1}{999} \text{RSE}^2}; \sqrt{\frac{q_2}{999} \text{RSE}^2} \right]$ where q_1 and q_2 are, respectively, the 2.5% and 97.5% quantiles of the χ^2 distribution with 999 degrees of freedom. We also computed the determinant of the variance-covariance matrix of all parameter estimates, normalised according to the standard definition of D-criterion, i.e. $\det(\mathbf{M}_F)^{-1/(2P+1)}$.

In parallel, we first used PFIM 3.2.2 (www.pfim.biostat.fr) with the model, parameters and designs above to predict M_F by LIN. Second, we implemented the calculations for GQ and AGQ (Section 2.4) in R 2.14.1 in a working version of PFIM, using the function `gauss.quad.prob` of the R package `statmod` and the function `as.weight` of the R package `plink` for multidimensional quadratures, which is an extension of `gauss.quad.prob`. We used $M=1000$ simulations to compute the expectation of h (see equation 11). We found that $\kappa=100$ nodes are needed for GQ and only 30 nodes for AGQ to guarantee stable results with all studied models. The predicted SE of parameters were calculated as the square root of the diagonal terms of M_F^{-1} and the predicted RSE were defined as the ratio of the SE to the true value of parameters. We also calculated the normalised determinant of the predicted variance-covariance matrix of parameter estimates (as defined above).

Finally, we compared the RSE and the normalised determinant of the variance-covariance matrix predicted using LIN, GQ and AGQ with those obtained by CTS.

3.3. Results

For the linear dose-response model (Figure 3), the RSE predicted by LIN, GQ and AGQ were close and in the same range as the ones calculated from CTS. In this case, the linearisation approach provided the exact calculation of M_F and of the true SE. As expected, the RSE were higher in the sparse design with 2 doses (Figure 3, M1S) than in the rich design with 4 doses (Figure 3, M1R), not so much for the fixed effects but more for the standard deviations of the random effects and of the random error.

For the E_{\max} model with 2 structural parameters and 2 random effects, when $\gamma=1$ (Figure 4, M2R-M2S), the three prediction approaches gave RSE close to those obtained by CTS. However, when $\gamma=3$ (Figure 4, M3R-M3S), with the sparse design, the linearisation approach over-predicted the RSE for $\omega_{ED_{50}}$ and especially for σ (42% predicted by LIN versus 15% by GQ, 14% by AGQ

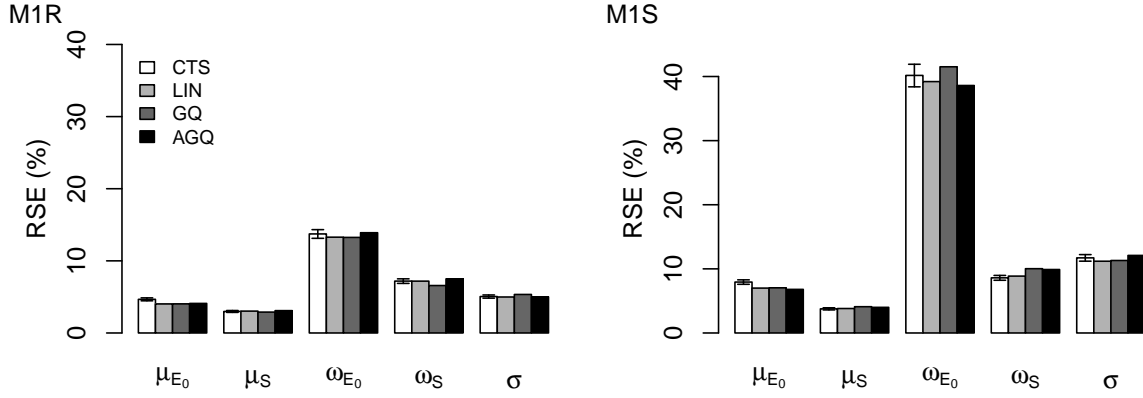


Figure 3: Relative standard error RSE (%) for the linear dose-response model obtained by clinical trial simulation (white bar) versus those predicted by linearisation (light grey bar), Gaussian quadrature (dark grey bar) and adaptive Gaussian quadrature (black bar), for the fixed effects (μ), the standard deviation of the random effects (ω), and of the random error (σ), with rich or sparse designs: M1R (4 doses per subject), M1S (2 doses per subject). The 95% confidence intervals of the RSE obtained by clinical trial simulation are displayed on top of the white bars.

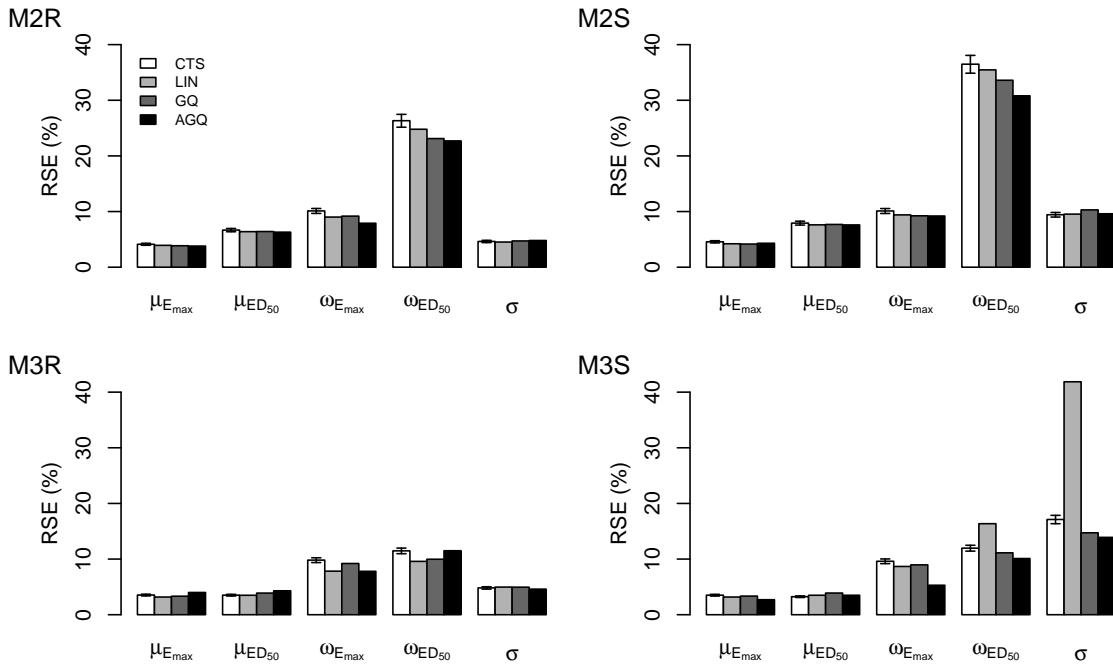


Figure 4: Relative standard error RSE (%) for the E_{\max} dose-response model with fixed baseline, obtained by clinical trial simulation (white bar) versus those predicted by linearisation (light grey bar), Gaussian quadrature (dark grey bar) and adaptive Gaussian quadrature (black bar), for the fixed effects (μ), the standard deviation of the random effects (ω), and of the random error (σ), with rich or sparse designs: M2R ($\gamma = 1$, 4 doses per subject), M2S ($\gamma = 1$, 2 doses per subject), M3R ($\gamma = 3$, 4 doses per subject), M3S ($\gamma = 4$, 2 doses per subject). The 95% confidence intervals of the RSE obtained by clinical trial simulation are displayed on top of the white bars.

and 17% obtained by CTS). With the rich design, the level of nonlinearity seemed to have less impact on the performance of linearisation: like GQ and AGQ, LIN adequately predicted RSE of all model parameters, in the same range as the ones obtained by CTS.

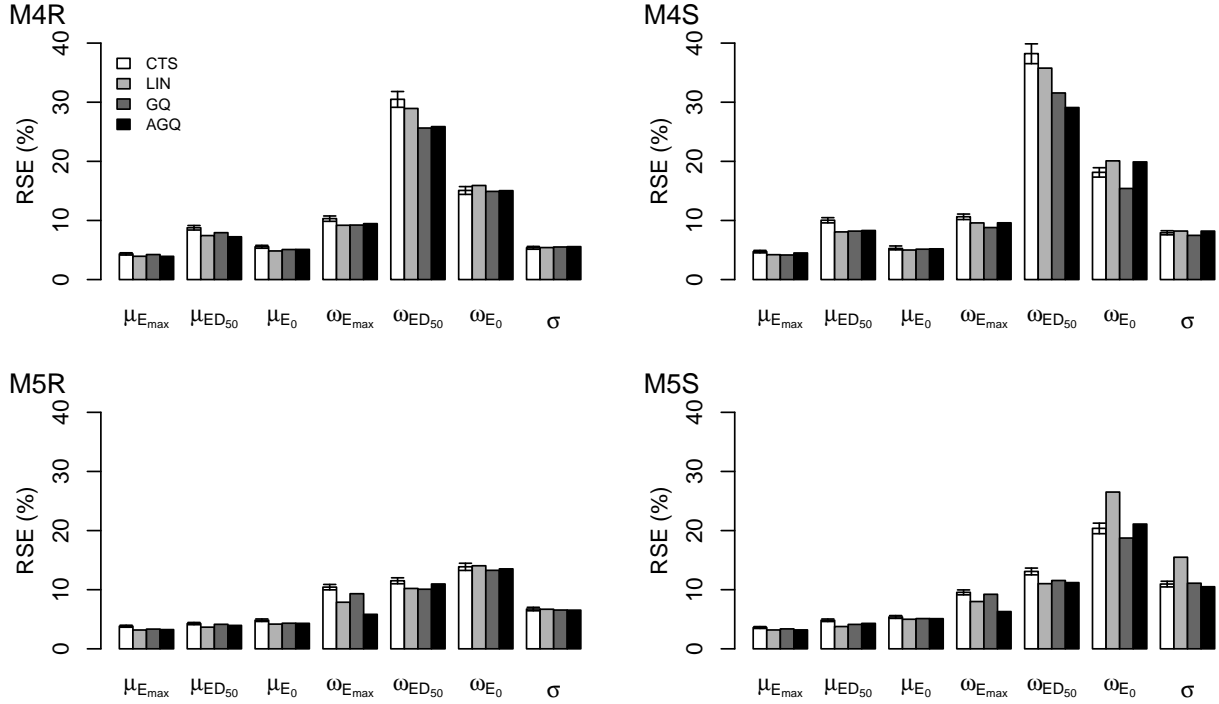


Figure 5: Relative standard error RSE (%) for the E_{max} dose-response model with estimated baseline, obtained by clinical trial simulation (white bar) versus those predicted by linearisation (light grey bar), Gaussian quadrature (dark grey bar) and adaptive Gaussian quadrature (black bar), for the fixed effects (μ), the standard deviation of the random effects (ω), and of the random error (σ), with rich or sparse designs: M4R ($\gamma = 1$, 4 doses per subject), M4S ($\gamma = 1$, 3 doses per subject), M5R ($\gamma = 3$, 4 doses per subject), M5S ($\gamma = 4$, 3 doses per subject). The 95% confidence intervals of the RSE obtained by clinical trial simulation are displayed on top of the white bars.

For the E_{max} model with 3 structural parameters and 3 random effects, when $\gamma = 1$ (Figure 5, M4R-M4S), the three prediction approaches gave RSE in the same range as those obtained by CTS. However, as in the previous model, when $\gamma = 3$ (Figure 5, M5R-M5S), the performance of the linearisation approach deteriorated with the sparse design and it over-predicted the RSE of ω_{E_0} (27% predicted by LIN versus 19% by GQ, 21% by AGQ and 20% obtained by CTS). With the rich design, the four approaches gave RSE that were close.

It is of note, as reported in Figures 3-4-5, that the RSE obtained by CTS were representative of the true RSE because rather small 95% confidence intervals were found. Figure 6 represents the normalised determinants of the variance-covariance matrix obtained by the four approaches for

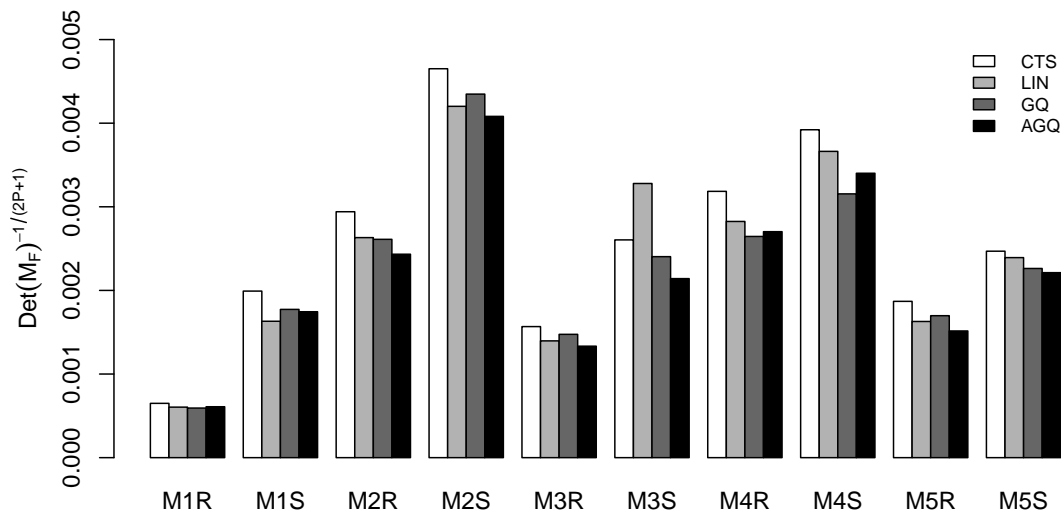


Figure 6: Determinant of the variance-covariance matrix of parameter estimates, normalised by inverse of the total number of parameters $(2P + 1)$, obtained by clinical trial simulation (white bar) versus those predicted by linearisation (light grey bar), Gaussian quadrature (dark grey bar) and adaptive Gaussian quadrature (black bar), for different models as defined in Table 1.

different scenarios. Overall, the predicted variances were lower than the ones observed from CTS. However, for the scenario M3S, LIN predicted a determinant of the variance-covariance matrix that is larger than the one obtained by CTS.

Regarding computing time, the linearisation approach was much faster than the two other prediction approaches. For the models with 2 random effects, the runtimes were less than 1 s by LIN in PFIM 3.2.2, about 6 h by GQ, 20 minutes by AGQ, and 10 h by CTS. For the models with 3 random effects, the runtimes were less than 1 s by LIN, about 110 h by GQ, 11 h by AGQ and 12 h by CTS.

4. Discussion

The present work proposes new approaches to the evaluation of the Fisher information matrix in NLMEM, using Gaussian quadrature and adaptive Gaussian quadrature when designing longitudinal studies. This is an alternative to first-order linearisation, the most commonly used approach in the field of optimal design in NLMEM. We investigated the performance of these approaches in predicting parameter standard errors and determinants of the Fisher matrix for dose-response models, as compared with linearisation and clinical trial simulation. These predictions are important

for design evaluation which is always the first step when planning an upcoming study and correct values of the D-criterion are needed for design optimisation.

When the nonlinearity was mild ($\gamma = 1$), the linearisation approach gave correct predictions of the RSE, close to the empirical ones obtained by simulation. When nonlinearity increased ($\gamma = 3$), linearisation correctly predicted the RSE of fixed effects but over-predicted, in the case of the sparse design, the RSE of $\omega_{ED_{50}}$ and σ . This poor performance could be partly explained by the behaviour of the first derivatives of the model function, which are used in the calculation of the M_F block corresponding to the standard deviation terms (Bazzoli et al., 2009). As shown in Figure 2A, the derivatives of f when $\gamma = 3$ (especially with respect to ED_{50}) are much more sensitive to design samples as compared with $\gamma = 1$. This might be why when $\gamma = 3$, with the sparse design of 2 doses (300, 1000) only, RSE of some standard deviation terms were not well predicted by linearisation. Further evaluations of other examples are needed to understand better the difference in performance of this approach between the sparse and rich designs.

GQ and AGQ gave correct predictions of RSE overall, close to the RSE obtained by CTS, in spite of a slight discrepancy versus CTS for some variability terms. When experimenters design population studies, they are more interested in the magnitude of the RSE than the exact value, so these approaches are relevant to evaluate M_F when linearisation is to be avoided. Here, we used $\kappa = 100$ nodes for GQ and $\kappa = 30$ nodes for AGQ to approximate the integrals correctly, when compared with CTS, in all studied models (so the total number of model evaluations was $K = \kappa^P$, P is the number of random effects). Even fewer nodes were sufficient to obtain stable results by AGQ for models with milder nonlinearity. Clarkson and Zhan (2002) needed $\kappa = 15$ nodes for spherical-radial quadrature when estimating parameters of a logistic mixed model with two random effects. Gotwalt et al. (2009) also used spherical-radical quadrature to compute the Bayesian design criteria (in a nonlinear model without random effects) and showed that the number of function evaluations to be performed increases as the square of the number of parameters P . In order to evaluate each elementary Fisher information matrix, we also needed to perform an integration over the marginal distribution $p(y, \Psi)$ of the observations y , which depends on the design Ψ (equation 9). Theoretically, this could have been evaluated by Gaussian quadrature rules as well, based on the distribution of y . However, this integral has the same dimension as the number of samples in the individual design ξ , and in practice Gaussian quadrature rules can only be applied when

the number of design samples is small. We proposed an alternative method based on stochastic integration, simulating $M = 1000$ vectors of observations under the same probability distribution as y , which was less time consuming than using a second quadrature, especially for rich designs. Guedj et al. (2007) also found in their study, with a similar approach using a simulation step, that $M = 1000$ was sufficient to evaluate M_F in a model of HIV dynamics. Further evaluations are needed to select the best balance between M and κ .

As a consequence, AGQ is much more computationally efficient than GQ for designing trials. For instance, in our study, for the models with 2 random effects, the runtimes were about 6 h by GQ, 20 min by AGQ and 10 h by CTS; and for the models with 3 random effects, about 110 h by GQ, 11 h by AGQ and 12 h by CTS. Here, we considered variability in all parameters; the number of random effects is the number of structural parameters, which is not always the case. Indeed, there are usually only one or two random effects in discrete data mixed models (Savic et al., 2011; Abebe et al., 2014) or survival frailty models (Vigan et al., 2012), so the computing times by these new approaches are still reasonable compared with CTS. Another advantage of GQ and AGQ compared with CTS is that the computing time in GQ or AGQ does not depend on the number of subjects in the population design because of the properties of the population M_F (equations 7 and 8). Consequently, a gain in computing time would be more obvious when designing studies in large populations. For instance, for a model with two random effects, simulation and fitting of 1000 datasets of 1000 subjects or 10 000 subjects will take about 30 hours and 800 hours respectively, while GQ will always take 6 hours and AGQ 20 minutes to evaluate the same design in all subjects. GQ and AGQ would more easily allow evaluation of a larger number of designs because of the property (8), while CTS is certainly more time consuming because one has to simulate and analyse again new datasets whenever adding or removing an elementary design. One perspective of this work is to evaluate the performance of alternative sampling techniques in the stochastic integration part of our AGQ approach, for example we would try using Latin Hypercube sampling, which might be less time consuming than standard random sampling, as suggested Ueckert et al. (2010).

Here we considered a diagonal variance-covariance matrix Ω , but one may want to take into account the correlation between random effects in the model, for example, between those on E_{\max} and ED_{50} as in Plan et al. (2012). Therefore, it would be interesting to include the full Ω in GQ and AGQ approaches as well, based on calculations proposed by Dumont et al. (2014). Moreover, we

should also evaluate the GQ and AGQ with different levels of between-subject variability or with increasing variance as in growth data. It would also be necessary to evaluate these approaches not only for dose-response models but also for discrete data mixed models and survival frailty models where the linearisation approach to design would work poorly.

Other possible alternatives to linearisation for evaluation of M_F when designing studies could be spherical radial integration (Monahan and Genz, 1997) or stochastic approach and importance sampling. Stochastic approach to evaluate M_F can be inspired by the SAEM algorithm (Kuhn and Lavielle, 2005), using Louis's (Louis, 1982) or Oakes's (Oakes, 1999) formulas for M_F , with simulation of individual parameters and stochastic approximation of expectations. Importance sampling is more time consuming than AGQ in estimation, but has the advantage of being versatile in handling distributions other than the normal for both random effects and residual errors (Pinheiro and Bates, 1995). However, it has not yet been assessed for design evaluation.

Once an appropriate approach to evaluating designs and to computing the D-criterion is chosen, the determinant of the Fisher information matrix can be maximised using the iterative Fedorov-Wynn algorithm (Fedorov, 1972; Wynn, 1972) within a finite set of possible designs, based on the Kiefer-Wolfowitz equivalence theorem (Kiefer and Wolfowitz, 1960), as implemented in design software such as PFIM (Retout et al., 2007), PkStaMP and PopDes. As suggested by Fedorov and Leonov (2014), we would compute the normalised predicted variance for all possible combinations of candidate points, which hits its maximum value of $\dim(\Psi)$ at the support points of the D-optimal design. However, a limitation of the current approaches proposed here is that they require a priori knowledge of the model and its parameters, which leads to designs that are only locally optimal. Sensitivity analyses with respect to the model and the parameter values would be necessary to quantify how the results vary. Alternatives, such as adaptive designs (Leonov and Miller, 2009; Foo and Duffull, 2012) or robust designs based on Bayesian criteria (Han and Chaloner, 2004; Dette and Pepelyshev, 2008; Gotwalt et al., 2009; Abebe et al., 2014) would be interesting to explore.

To conclude, when computing the Fisher information matrix in NLMEM for design evaluation and optimisation, the linearisation approach is accurate for models with mild nonlinearity, as has already been demonstrated for several PK (Retout and Mentré, 2003; Nguyen et al., 2012), PK/PD (Bazzoli et al., 2009) and viral dynamic models (Retout et al., 2007; Guedj et al., 2011). This procedure is very fast and relatively simple, is available in several design software programs,

and is a very useful tool for designing longitudinal studies while avoiding extensive simulations. However, when the models are complex and have never been evaluated in design approaches, non-linearity measures (Bates and Watts, 1980; Cook and Goldberg, 1986) of the studied model should be investigated for the evaluated or the optimised design before drawing final conclusions. If the nonlinearity is strong and therefore the linearisation approximation is to be avoided, then GQ and AGQ are relevant alternatives: GQ is very time consuming and is only useful for a small number of random effects, while AGQ is much less time consuming while providing adequate results. An appropriate choice for the number of quadrature nodes is to be defined, depending on the level of nonlinearity. Further evaluations for other types of models are needed before implementing AGQ in PFIM software for efficiently designing trials while reducing the number of samples per subject, which can be very important both ethically and practically when performing studies in patients.

Acknowledgements

The authors thank IFR02 - Paris Diderot University and Hervé Le Nagard for the use of the "Centre de Biomodélisation".

References

- Abebe, H. T., Tan, F. E., Breukelen, G. J., Berger, M. P., 2014. Bayesian D-optimal designs for the two parameter logistic mixed effects model. *Computational Statistics and Data Analysis* 71, 1066–1076.
- Abramowitz, M., Stegun, I., 1964. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York.
- Al-Banna, M. K., Kelman, A. W., Whiting, B., 1990. Experimental design and efficient parameter estimation in population pharmacokinetics. *Journal of Pharmacokinetics and Biopharmaceutics* 18, 347–360.
- Atkinson, A. C., Donev, A. N., Tobias, R. D., 2007. *Optimum Experimental Designs, with SAS*. Oxford University Press, Oxford.

- Atkinson, A. C., Fedorov, V. V., Herzberg, A. M., Zhang, R., 2014. Elemental information matrices and optimal experimental design for generalized regression models. *Journal of Statistical Planning and Inference* 144, 81–91.
- Bates, D. M., Watts, D. G., 1980. Relative curvature measures of nonlinearity. *Journal of the Royal Statistical Society: Series B* 42, 1–25.
- Bazzoli, C., Retout, S., Mentré, F., 2009. Fisher information matrix for nonlinear mixed effects multiple response models: evaluation of the appropriateness of the first order linearization using a pharmacokinetic/pharmacodynamic model. *Statistics in Medicine* 28 (14), 1940–1956.
- Bazzoli, C., Retout, S., Mentré, F., 2010. Design evaluation and optimisation in multiple response nonlinear mixed effect models: PFIM 3.0. *Computer Methods and Programs in Biomedicine* 98 (1), 55–65.
- Box, G. E. P., Lucas, H. L., 1959. Design of experiments in nonlinear situations. *Biometrika* 46, 77–90.
- Bretz, F., Dette, H., Pinheiro, J. C., 2010. Practical considerations for optimal designs in clinical dose finding studies. *Statistics in Medicine* 29 (7-8), 731–742.
- Chernoff, H., 1953. Locally optimal designs for estimating parameters. *Annals of Mathematical Statistics* 24, 586–602.
- Clarkson, D., Zhan, Y., 2002. Using spherical-radial quadrature to fit generalized linear mixed effects models. *Journal of computational and graphical statistics. Journal of Computational and Graphical Statistics* 11, 639–659.
- Cook, R. D., Goldberg, M. L., 1986. Curvatures for parameter subsets in nonlinear regression. *Annals of Statistics* 14, 1399–1418.
- D’Argenio, D. Z., Dec 1981. Optimal sampling times for pharmacokinetic experiments. *Journal of Pharmacokinetics and Biopharmaceutics* 9 (6), 739–756.

- Dartois, C., Brendel, K., Comets, E., Laffont, C. M., Laveille, C., Tranchand, B., Mentré, F., Lemenuel-Diot, A., Girard, P., 2007. Overview of model-building strategies in population PK/PD analyses: 2002-2004 literature survey. *British Journal of Clinical Pharmacology* 64 (5), 603–612.
- Dette, H., Pepelyshev, A., 2008. Efficient experimental designs for sigmoidal growth models. *Journal of Statistical Planning and Inference* 138, 2–17.
- Dumont, C., Chenel, M., Mentré, F., 2014. Influence of covariance between random effects in design for nonlinear mixed effect models with an illustration in paediatric pharmacokinetics. *Journal of Biopharmaceutical Statistics* 24 (3), 471–492.
- Fedorov, V. V., 1972. *Theory of Optimal Experiments*. Academic Press, New York.
- Fedorov, V. V., Leonov, S., L., 2014. *Optimal Design for Nonlinear Response Models*. Chapman and Hall/ CRC Press, Boca Raton.
- Fletcher, R., Powell, M. J. D., 1963. A rapidly convergent descent method for minimization. *The Computer Journal* 6, 163–168.
- Foo, L. K., Duffull, S., 2012. Adaptive optimal design for bridging studies with an application to population pharmacokinetic studies. *Pharmaceutical Research* 29 (6), 1530–1543.
- Gabrielson, J., Weiner, D., 2006. *Pharmacokinetic and Pharmacodynamic Data Analysis: Concepts and Applications*, 4th Edition. Swedish Pharmaceutical Press, Stockholm.
- Golub, G. H., 1973. Some modified matrix eigenvalue problems. *SIAM Review* 15, 318–334.
- Golub, G. H., Welsh, J. H., 1969. Calculation of Gauss quadrature rules. *Mathematics of Computation* 23, 221–230.
- Gotwalt, C. M., Jones, B. A., Steinberg, D. M., 2009. Fast computation of designs robust to parameter uncertainty for nonlinear settings. *Technometrics* 51, 88–95.
- Guedj, J., Bazzoli, C., Neumann, A. U., Mentré, F., 2011. Design evaluation and optimization for models of hepatitis C viral dynamics. *Statistics in Medicine* 30 (10), 1045–1056.

- Guedj, J., Thiébaud, R., Commenges, D., 2007. Practical identifiability of HIV dynamics models. *Bulletin of Mathematical Biology* 69 (8), 2493–2513.
- Gueorguieva, I., Ogungbenro, K., Graham, G., Glatt, S., Aarons, L., 2007. A program for individual and population optimal design for univariate and multivariate response pharmacokinetic-pharmacodynamic models. *Computer Methods and Programs in Biomedicine* 86 (1), 51–61.
- Han, C., Chaloner, K., 2004. Bayesian experimental design for nonlinear mixed-effects models with application to HIV dynamics. *Biometrics* 60 (1), 25–33.
- Han, C., Chaloner, K., 2005. Design of population studies of HIV dynamics. In: Tan, W-Y and Wu, H. (Eds), *Deterministic and Stochastic Models of AIDS Epidemics and HIV Infections with Intervention*. World Scientific Publishing Co., Fitchburg, MA, pp. 525–548. Chapter 21.
- Hashimoto, Y., Sheiner, L. B., 1991. Designs for population pharmacodynamics: value of pharmacokinetic data and population analysis. *Journal of Pharmacokinetics and Biopharmaceutics* 19, 333–353.
- Jones, B., Wang, J., 1999. Constructing optimal designs for fitting pharmacokinetic models. *Statistics and Computing* 9, 209–218.
- Jones, B., Wang, J., Jarvis, P., Byrom, W., 1999. Design of cross-over trials for pharmacokinetic studies. *Journal of Statistical Planning and Inference* 78, 307–316.
- Jonsson, E. N., Wade, J. R., Karlsson, M. O., 1996. Comparison of some practical sampling strategies for population pharmacokinetic studies. *Journal of Pharmacokinetics and Biopharmaceutics* 24, 245–263.
- Kiefer, J., Wolfowitz, J., 1960. The equivalence of two extremum problems. *Canadian Journal of Mathematics* 12, 363–366.
- Kuhn, E., Lavielle, M., 2005. Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics and Data Analysis* 49, 1020–1038.
- Lalonde, R. L., Kowalski, K. G., Hutmacher, M. M., Ewy, W., Nichols, D. J., Milligan, P. A., Corrigan, B. W., Lockwood, P. A., Marshall, S. A., Benincosa, L. J., Tensfeldt, T. G., Parivar,

- K., Amantea, M., Glue, P., Koide, H., Miller, R., 2007. Model-based drug development. *Clinical Pharmacology and Therapeutics* 82 (1), 21–32.
- Leonov, S., Aliev, A., 2012. Optimal design for population PK/PD models. *Tatra Mountains Mathematical Publications* 51, 115–130.
- Leonov, S., Miller, S., 2009. An adaptive optimal design for the Emax model and its application in clinical trials. *Journal of Biopharmaceutical Statistics* 19 (2), 360–385.
- Lindstrom, M. L., Bates, D. M., 1990. Nonlinear mixed effects models for repeated measures data. *Biometrics* 46 (3), 673–687.
- Louis, T., 1982. Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society Series B* 44, 226–233.
- McGree, J., Drovandi, C., Thompson, M., Eccleston, J., Duffull, S., Mengersen, K., Pettitt, A., Goggin, T., 2012. Adaptive Bayesian compound designs for dose finding studies. *Journal of Statistical Planning and Inference* 142, 1480–1492.
- Mentré, F., Chenel, M., Comets, E., Grevel, J., Hooker, A. C., Karlsson, M. O., Lavielle, M., Gueorguieva, I., 2013. Current use and developments needed for optimal design in pharmacometrics: A study performed among DDMoRe’s European Federation of Pharmaceutical Industries and Associations members. *CPT: Pharmacometrics and Systems Pharmacology* 2, e46.
- Mentré, F., Mallet, A., Baccar, D., 1997. Optimal design in random effect regression models. *Biometrika* 84, 429–442.
- Mielke, T., 2012. Approximations of the Fisher information for the construction of efficient experimental designs in nonlinear mixed effects models. Ph.D. thesis, Otto-von-Guericke-Universität Magdeburg, Germany.
- Mielke, T., Schwabe, R., 2010. Some considerations on the Fisher information in nonlinear mixed effects models. In: Giovagnoli, A., Atkinson, A. C., Torsney, B., May, C. (Eds.), *Proceedings of the 9th International Workshop in Model-Oriented Design and Analysis*. Physica Verlag, Bertinoro, Italy, pp. 129–136.

- Monahan, J., Genz, A., 1997. Spherical radial integration rules for Bayesian computations. *Journal of the American Statistical Association* 92, 664–674.
- Nguyen, T. T., Bazzoli, C., Mentré, F., 2012. Design evaluation and optimisation in crossover pharmacokinetic studies analysed by nonlinear mixed effects models. *Statistics in Medicine* 31 (11-12), 1043–1058.
- Nielsen, E. I., Viberg, A., Lowdin, E., Cars, O., Karlsson, M. O., Sandstrom, M., 2007. Semimechanistic pharmacokinetic/pharmacodynamic model for assessment of activity of antibacterial agents from time-kill curve experiments. *Antimicrobial Agents and Chemotherapy* 51 (1), 128–136.
- Nyberg, J., Bazzoli, C., Ogungbenro, K., Aliev, A., Leonov, S., Duffull, S., Hooker, A. C., Mentré, F., 2014. Methods and software tools for design evaluation for population pharmacokinetics-pharmacodynamics studies. *British Journal of Clinical Pharmacology* (in press).
- Nyberg, J., Ueckert, S., Strömberg, E. A., Hennig, S., Karlsson, M. O., Hooker, A. C., 2012. PopED: an extended, parallelized, nonlinear mixed effects models optimal design tool. *Computer Methods and Programs in Biomedicine* 108 (2), 789–805.
- Oakes, D., 1999. Direct calculation of the information matrix via the EM algorithm. *Journal of the Royal Statistical Society Series B* 61, 479–482.
- Perelson, A. S., Ribeiro, R. M., 2008. Estimating drug efficacy and viral dynamic parameters: HIV and HCV. *Statistics in Medicine* 27 (23), 4647–4657.
- Pillai, G., Mentré, F., Steimer, J. L., 2005. Non-linear mixed effects modeling - from methodology and software development to driving implementation in drug development science. *Journal of Pharmacokinetics and Pharmacodynamics* 32, 161–183.
- Pinheiro, J. C., Bates, D. M., 1995. Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics* 4, 12–35.
- Plan, E. L., Maloney, A., Mentré, F., Karlsson, M. O., Bertrand, J., 2012. Performance comparison of various maximum likelihood nonlinear mixed-effects estimation methods for dose-response models. *The AAPS Journal* 14 (3), 420–432.

- Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., 1992. Numerical Recipes in C, The Art of Scientific Computing. Cambridge University Press, New York.
- Pronzato, L., 2010. Penalized optimal designs for dose-finding. *Journal of Statistical Planning and Inference* 140, 283–296.
- Retout, S., Comets, E., Samson, A., Mentré, F., 2007. Design in nonlinear mixed effects models: Optimisation using the Fedorov-Wynn algorithm and power of the wald test for binary covariates. *Statistics in Medicine* 26, 5162–5179.
- Retout, S., Mentré, F., 2003. Further developments of the Fisher information matrix in nonlinear mixed effects models with evaluation in population pharmacokinetics. *Journal of Biopharmaceutical Statistics* 13, 209–227.
- Retout, S., Mentré, F., Bruno, R., 2002. Fisher information matrix for non-linear mixed-effects models: evaluation and application for optimal design of enoxaparin population pharmacokinetics. *Statistics in Medicine* 21, 2623–2639.
- Savic, R. M., Mentré, F., Lavielle, M., 2011. Implementation and evaluation of the SAEM algorithm for longitudinal ordered categorical data with an illustration in pharmacokinetics-pharmacodynamics. *The AAPS Journal* 13 (1), 44–53.
- Sheiner, L. B., Rosenberg, B., Marathe, V. V., 1977. Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. *Journal of Pharmacokinetics and Biopharmaceutics* 5 (5), 445–479.
- Sheiner, L. B., Rosenberg, B., Melmon, K. L., 1972. Modelling of individual pharmacokinetics for computer-aided drug dosage. *Computers and Biomedical Research* 5, 411–459.
- Smith, B. P., Vincent, J., 2010. Biostatistics and pharmacometrics: quantitative sciences to propel drug development forward. *Clinical Pharmacology and Therapeutics* 88 (2), 141–144.
- Smyth, G., 2002. Nonlinear regression. In: El-Shaarawi, A.H., Piegorsch, W.W. (Eds), *Encyclopedia of Environmetrics* 3. John Wiley and Sons Ltd, Chichester, pp. 1405–1411.

- Ueckert, S., Nyberg, J., Hooker, A. C., 2010. Application of Quasi-Newton algorithms in optimal design. Workshop of Population Optimum Design of Experiments, Berlin, Germany.
URL http://www.maths.qmul.ac.uk/bb/PODE/PODE2010_Slides/
- Van der Graaf, P. H., 2012. CPT: Pharmacometrics and Systems Pharmacology. CPT: Pharmacometrics and Systems Pharmacology 1 (9), e8.
- Vigan, M., Stirnemann, J., Mentré, F., 2012. Evaluation of estimation methods for repeated time to event models: application to analysis of bone events during treatment of Gaucher Disease. Abstract 156. 44th Journées de Statistiques, Bruxelles, Belgium.
URL <http://jds2012.ulb.ac.be/showabstract.php?id=156>
- Vong, C., Ueckert, S., Nyberg, J., Hooker, A. C., 2012. Handling below limit of quantification data in optimal trial design. Abstract 2578. 21st Population Approach Group in Europe, Venise, Italy.
URL www.page-meeting.org/?abstract=2578
- Walter, E., Pronzato, L., 1995. Identifiabilities and nonlinearities. In: Fossard, A.J. and Normand-Cyrot, D. (Eds), Nonlinear Systems 1. Chapman & Hall, London, pp. 111–143. Chapter 3.
- Walter, E., Pronzato, L., 2007. Identification of Parametric Models from Experimental Data. Springer, New York.
- Wolfinger, R., 1993. Laplace's approximation for nonlinear mixed models. *Biometrika* 80, 791–795.
- Wynn, H. P., 1972. Results in the theory and construction of D-optimum designs. *Journal of the Royal Statistical Society Series B* 34, 133–147.