HAL
open science

La pharmacométrie

Caroline Bazzoli, Julie Bertrand, Emmanuelle Comets

To cite this version:

Caroline Bazzoli, Julie Bertrand, Emmanuelle Comets. La pharmacométrie. Techniques de l'Ingénieur, 2014, base documentaire : TIB386DUO (ref. article : pha1010), 56 p. hal-00932097

HAL Id: hal-00932097

https://hal.science/hal-00932097

Submitted on 27 Feb 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.

La pharmacométrie

Caroline Bazzoli [*], Julie Bertrand [†], Emmanuelle Comets [‡]

26 février 2014

[*]Laboratoire Jean Kuntzmann, Département Statistique, University of Grenoble, France
[†]UCL Genetics Institute, University College of London, London, United Kingdom
[‡]INSERM INSERM IAME, UMR 1137, Univ Paris Diderot, Sorbonne Paris Cité, Paris, France ; CIC 0203, Univ Rennes 1, Rennes, France

1 Introduction

La pharmacométrie a été récemment définie comme la science de la pharmacologie clinique quantitative [1]. La pharmacologie étudie l'interaction entre notre organisme et le médicament, ce terme désignant « toute substance ou composition [...] exerçant une action pharmacologique, immunologique ou métabolique »(article L5111-1 du code de la santé publique). Cette interaction bidirectionnelle recouvre la pharmacocinétique (PK) (ce que notre organisme fait subir au médicament) et la pharmacodynamie (PD) (ce que le médicament fait subir à notre organisme).

La PK est principalement étudiée au travers de l'évolution de la concentration en fonction du temps, souvent résumée par des paramètres comme l'aire sous la courbe (AUC), mesurant l'exposition totale, et la demi-vie, définie comme le temps requis pour diminuer de moitié la quantité de médicament dans le système. L'activité PD est, quant à elle, plus diverse dans sa nature. Les marqueurs de cette activité dépendent de la réponse considérée, biologique ou clinique, continue ou discrète. L'étude conjointe de la PK et de la PD d'un médicament permet de définir la marge thérapeutique comprise entre les doses ou les concentrations minimale et maximale requises pour observer un niveau de réponse respectivement efficace et toxique.

Les marqueurs PK/PD sont recueillis au cours du développement du médicament chez des volontaires sains en phase I, puis chez des patients en phase II avec respectivement un accent sur la toxicité puis l'efficacité. Dès la phase II et jusqu'en routine clinique, la variabilité présente entre les sujets (inter-sujet) est quantifiée et explorée afin de pouvoir en définir les sources et évaluer la nécessité d'une personnalisation du traitement à travers la dose et/ou la posologie.

La pharmacométrie regroupe les techniques permettant de caractériser l'activité PK et PD et dans quelle mesure cette activité varie d'un sujet et/ou d'une occasion à l'autre, ainsi que de prédire et de simuler ces activités afin de fournir des critères rationnels à la prise de décision. Cette discipline a pris son essor lors des dernières décades, avec le développement de méthodes dites "de population", reposant sur l'utilisation des modèles non-linéaires à effets mixtes. Ces méthodes contrastent avec les analyses en deux étapes utilisées auparavant : les marqueurs de l'activité PK et PD étaient estimés individuellement puis leur valeur moyenne et leur variabilité calculées. L'estimation individuelle nécessite un nombre élevé de prélèvements par sujet, et n'est utilisable que dans les études précoces lors du développement clinique. Dans les études ultérieures, où le nombre de patients est plus grand mais le nombre de prélèvements par patient réduit, ou dans les populations fragiles (patients diminués, enfants,...) où le nombre de prélèvements doit rester minimal, les approches de population permettent d'analyser toutes les observations simultanément et s'aide des sujets

avec plus de prélèvements pour inférer les sujets moins informatifs, en s'appuyant sur des hypothèses statistiques et mécanistiques.

Cet article présentera les données rencontrées en pharmacologie clinique en section 2, les modèles dynamiques et statistiques qui sous-tendent l'approche de population en section 3, les méthodes d'estimation des paramètres des modèles en section 4, et de construction et de validation de modèles en section 5. Pour finir, les méthodes d'évaluation et d'optimisation des protocoles qui s'appuient sur ces modèles seront présentées en section 6, et l'ensemble de ces outils ainsi que leur extension sera discuté en section 7. Tout au long de cet article, nous illustrerons nos propos par une application à la PK/PD de la warfarine.

2 Données

2.1 Données longitudinales

On appelle couramment "données longitudinales" des données mesurées de façon répétée chez un même sujet, induisant ainsi une corrélation entre les observations. En pharmacométrie, ces données sont obtenues lors d'essais cliniques sur des médicaments. Elles sont recueillies chez l'ensemble des sujets plusieurs fois au cours du temps, au cours de différentes périodes de traitement (occasions) ; leur nombre peut différer d'un sujet et/ou d'une occasion à l'autre. Les données collectées répondent à une structure hiérarchique, en effet les observations peuvent être regroupées comme appartenant à une même entité, le sujet ou individu. Chaque sujet est différent, avec des paramètres PK et/ou PD spécifiques et donc différents de ceux de la population *i.e* du sujet "moyen". Pour chaque sujet, des facteurs démographiques, génétiques, physiologiques et environnementaux sont recueillis. Appelés covariables, ils peuvent expliquer une partie de la diversité des profils individuels au sein d'une population.

Tout au long de cet article, nous noterons y_{ij} la $j^{ième}$ observation du sujet i avec $i = 1,...,N$ et $j = 1,...,n_i$ où n_i est le nombre d'observations pour le sujet i. Nous noterons x_{ij} les variables du protocole expérimental (temps de prélèvements, dose,...) et c_i les covariables propres au sujet i.

2.2 Exemple de travail

Cette partie présente le jeu de données que nous utiliserons comme fil conducteur dans cet article. Il s'agit d'un exemple classique en pharmacologie [2], étudiant la PK/PD de la warfarine, un anticoagulant de la famille des antivitamines K. Des données de PK (concentrations au cours du temps) et de PD (activité du complexe pro-

thrombinique (PCA), un facteur de coagulation) ont été recueillies chez 32 volontaires sains, ayant reçus une dose unique de 1,5 mg/kg de warfarine. N. Holford a étudié la relation dose-concentration-réponse de la warfarine par une approche de population en 1986 [3]. Les concentrations plasmatiques de la warfarine (mg/L) ont été mesurées sur 120 h après la prise du médicament, aux temps : 0.5, 1, 2, 3, 6, 9, 12, 24, 36, 48, 72, 96, 120 h (n=13 prélèvements PK). Pour la mesure de la PCA (%), les prélèvements ont été effectués aux temps : 0, 24, 36, 48, 72, 96, 120, 144 h, après la prise (n=8 prélèvements PD). La figure 1 représente l'évolution des deux réponses en fonction du temps chez les 32 sujets. Une variabilité importante apparait entre les sujets pour les deux réponses. Chaque sujet possède sa propre courbe de concentration de warfarine et de la PCA, illustrant la diversité des profils PK et PD.

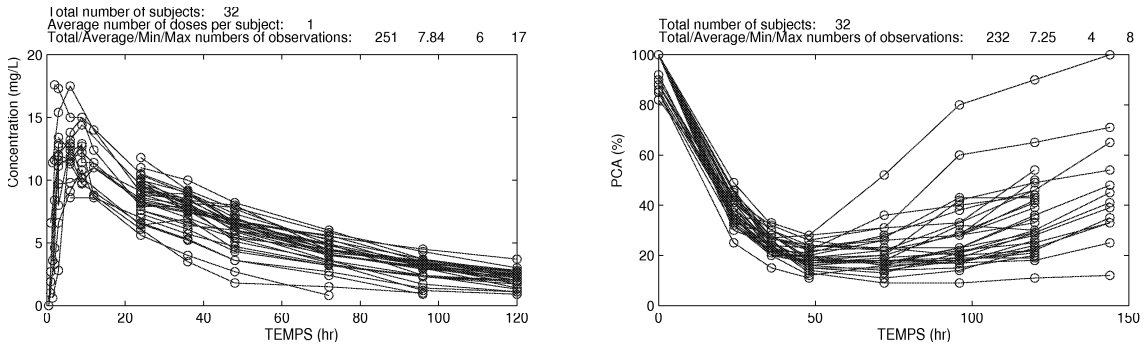

Figure 1 – Courbes individuelles des concentrations plasmatiques de la warfarine (à gauche) et de la PCA (%) (à droite) pour les 32 volontaires sains de l'étude.

La table 1 montre les premières données correspondant à cette étude. Chaque ligne correspond à une observation, soit PK soit PD (colonne Type), avec le temps de prélèvement correspondant et les covariables de ce sujet.

Liste des sigles	
Sigle	Signification
AGQ	Adaptative Gaussian Quadrature
AIC	Akaike Information Criterion
AUC	Aire sous la courbe
BIC	Bayes Information Criterion
DIC	Deviance information Criterion
EAP	Esperance A Posteriori
EBE	Empirical Bayes Estimates
ED	Expectation D-optimal design
EID	Expectation optimal design of Determinant of inverse Fisher information matrix
EM	Expectation-Maximisation
FO	First Order
FOCE	First Order Conditional Estimates
GAM	Generalized Additive Models
hr	Heure
IWRES	Individual Weighted Residuals
LL	Log-Likelihood
LRT	Likeliwood Ratio Test
MAP	Mode A Posteriori
MCMC	Markov Chain Monte Carlo
ODE	Ordinary differential equation
PBPK	Physiologically-Based PharmacoKinetic
PCA	Activité du complexe prothrombinique
PD	Pharmacodynamique
PK	Pharmacocinétique
PPC	Posterior predictive check
SAEM	Stochastic Approximation Expectation Maximisation
SE	Standard Error
VPC	Visual Predictive Check
WRES	Weighted Residuals

Sujet	Temps	Obs	Type	Poids	Age	Sexe
1	0	100	PD	66.7	31	1
1	0.5	0	PK	66.7	31	1
1	2	8.4	PK	66.7	31	1
...						
1	120	1.4	PK	66.7	31	1
1	120	49	PD	66.7	31	1
1	144	54	PD	66.7	31	1
2	0	100	PD	80	40	1
2	3	12	PK	80	40	1
2	6	13.2	PK	80	40	1
...						
2	120	3	PK	80	40	1
2	120	11	PD	80	40	1
2	144	12	PD	80	40	1
...						

Table 1 – Premières données de PK/PD de la warfarine.

3 Modèles non-linéaires à effets mixtes

3.1 Modèles PK/PD

Ces modèles visent à décrire l'évolution des systèmes dynamiques, et ont le temps pour variable de régression.

3.1.1 Modèles pharmacocinétiques

L'évolution des concentrations en fonction du temps peut être modélisée mécanistiquement en décomposant l'organisme en un système de compartiments à l'intérieur desquels la concentration est supposée homogène (hypothèse du "well-stirred system"). Chaque compartiment est caractérisé par un volume, dit volume de distribution, et la concentration s'exprime comme la quantité de la molécule présente dans le compartiment divisée par ce volume. La quantité présente à un instant t est la résultante de processus d'entrée et de sortie dans le compartiment. Dans l'ensemble de l'organisme, ces processus sont regroupés en quatre grandes étapes : **A**bsorption, **D**istribution, **M**étabolisme et **E**xcrétion (d'où l'acronyme ADME).

Absorption : le médicament peut être administré de différentes façons (voie intraveineuse, perfusion, prise orale, injections sous-cutanées et intramusculaires), et passe ensuite dans la circulation systémique. De nombreux mécanismes rentrent en jeu comme des transports actifs ou passifs, et lorsque la voie d'administration n'est pas directement le sang, la totalité de la dose ne parvient pas toujours jusqu'au compartiment observé du fait des nombreuses barrières protégeant notre organisme des éléments extérieurs ; le plus important est l'effet de premier passage pour les administrations orales, par lequel le foie re-excrète une partie plus ou moins importante du médicament absorbé avant qu'il n'atteigne la circulation systémique. Un paramètre important d'un médicament est donc sa biodisponibilité, correspondant à la fraction de la dose qui est effectivement absorbée et donc utilisée dans le système modélisé. La modélisation de l'absorption décrit ces processus par des compartiments où le médicament est administré et par la façon dont il passe dans le compartiment central.

Distribution : une fois la circulation sanguine atteinte, les médicaments se distribuent dans le sang et les tissus. Un arrangement des compartiments du système en ensemble caténaire (un compartiment d'entrée et un compartiment de sortie, avec un ou plusieurs compartiments intermédiaires, chacun d'eux étant simultanément le successeur unique du précédent et le précurseur unique du suivant) correspond à une distribution progressive du médicament à travers l'organisme alors qu'un arrangement

mamillaire (un compartiment central précurseur de tous les compartiments périphériques, eux-mêmes indépendants entre eux) traduit une distribution du médicament dans des organes plus ou moins bien perfusés. La distribution est aussi fonction des propriétés physico-chimiques du médicament qui conditionnent sa diffusion membranaire et son degré de fixation aux protéines plasmatiques et tissulaires.

Métabolisme et Excrétion : l'élimination du médicament par notre organisme s'effectue essentiellement à travers l'action des enzymes métaboliques de l'appareil digestif et par les voies biliaires et rénales. Souvent, aux doses administrées, l'élimination est linéaire *i.e.* la vitesse d'élimination est proportionnelle à la quantité de médicament dans le compartiment, et le volume de médicament éliminé par unité de temps, ou "clairance", est égal au produit de la constante d'élimination par le volume. Lorsque le processus d'élimination est limité par la quantité d'enzyme ou de transporteur disponible, l'élimination est saturable et devient non-linéaire.

Modélisation : la figure 3.1.1 montre deux types de modèles qu'il est possible de construire pour une molécule comme la warfarine. A gauche, un modèle où l'ensemble de l'organisme est décomposé en compartiments reliés entre eux par la circulation sys-

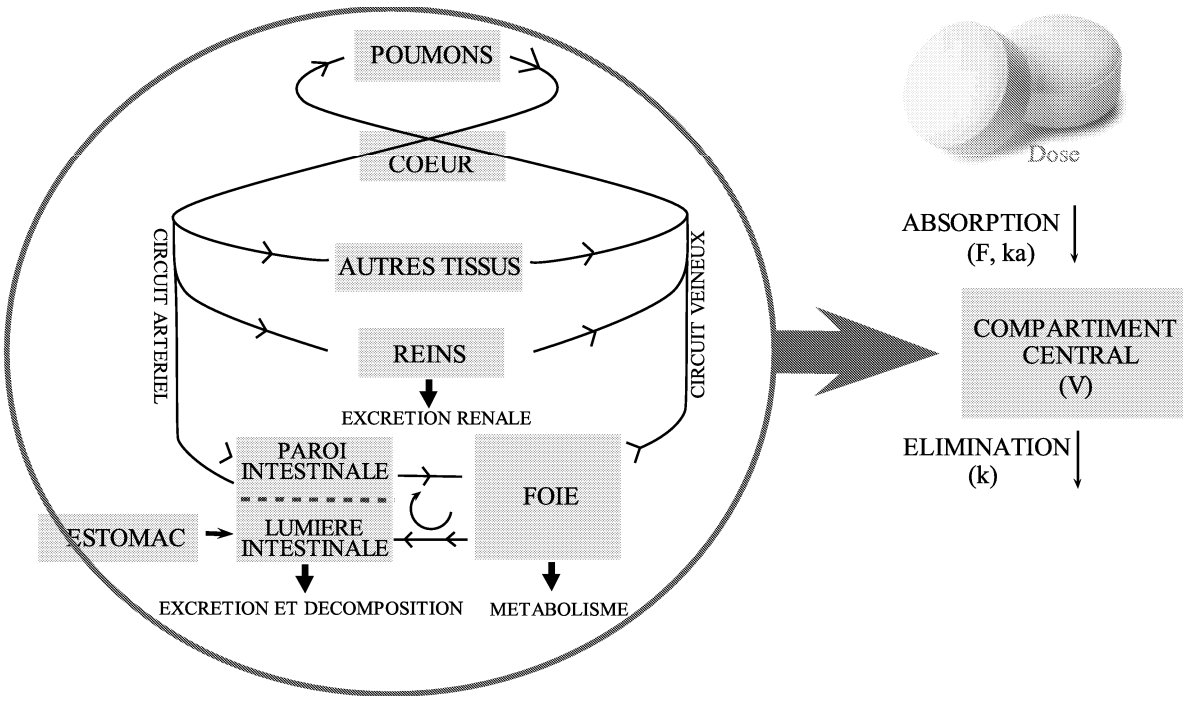

Figure 2 – Modèles de la PK de la warfarine.

témique. A droite, cet ensemble est simplifié et assimilé à un unique compartiment :

la cinétique d'une molécule comme la warfarine qui i) est administrée oralement, ii) se lie facilement aux protéines plasmatiques et va donc difficilement traverser les membranes tissulaires et iii) est métabolisée par le cytochrome P450 sera modélisée par un seul compartiment de distribution avec absorption et élimination d'ordre 1. Ce système peut s'écrire sous forme d'équations différentielles :

$$\begin{aligned} \frac{dG}{dt} &= -k_a\, G \qquad & G(t=0) &= F\ Dose \\ \frac{dQ}{dt} &= k_a\, G - k\, Q \qquad & Q(t=0) &= 0 \end{aligned} \tag{1}$$

G et Q représentent respectivement la quantité de médicament dans le compartiment de dépôt et le compartiment de mesure. Quatre paramètres caractérisent ce modèle ; F, la biodisponibilité, k_a la constante d'absorption, V le volume de distribution et k la constante d'élimination, et la concentration observée est égale à $C(t) = Q(t)/V$. Les constantes de vitesse k et k_a sont exprimées en hr^{-1}, et on peut calculer la demie-vie d'élimination en hr, $t_{1/2} = ln(2)/k$. Ce système peut être résolu grâce notamment à l'utilisation de transformées de Laplace, menant à une fonction dite analytique :

$$C(t) = \frac{F\ Dose}{V}\ \frac{k_a}{k_a - k}\left(e^{-k\ t} - e^{-k_a\ t}\right) \tag{2}$$

Ce système n'est cependant pas structurellement globalement identifiable, ce qui signifie qu'il peut y avoir plusieurs jeux de paramètres équivalents donnant exactement les mêmes prédictions à tous les temps. Pour pouvoir en estimer tous les paramètres, il faudrait disposer de la quantité totale absorbée dans le système, par exemple en recueillant les concentrations et les volumes urinaires, ou disposer également de données par voie intraveineuse (où la biodisponibilité vaut 1). Lorsqu'on ne dispose que de données orales, on estime donc le ratio V/F, ou volume apparent, qui lui est identifiable.

Dans le cas d'une élimination non-linéaire (ou saturable), le modèle de Michaelis-Menten [4] est généralement utilisé :

$$\begin{aligned} \frac{dG}{dt} &= -k_a\, G \qquad & G(t=0) &= F\ Dose \\ \frac{dQ}{dt} &= k_a\, G - \frac{V_m\, Q/V}{K_m + Q/V} \qquad & Q(t=0) &= 0 \end{aligned} \tag{3}$$

où V_m est la constante maximale d'élimination et K_m la concentration à laquelle la constante d'élimination vaut $V_m/2$. Ce système n'a pas de solution analytique et doit donc être résolu numériquement à l'aide d'algorithmes dérivés des méthodes de Runge-Kutta [5]. Une absorption retardée à travers une chaîne de compartiments de transit [6] est un autre système sans solution analytique mais qui peut être approximé par une

fonction discontinue du temps ; la fonction $C(t)$ prenant une forme particulière pendant un temps de "latence" T_{lag} donné, puis une autre ensuite.

Lorsque le médicament est administré à plusieurs reprises, le principe de superposition permet de sommer simplement les fonctions $C(t)$ de chaque dose. Si ces doses sont régulièrement espacées d'un temps τ, le système finit par atteindre un état "d'équilibre" (*steady-state*), où les concentrations à un temps t après la dose sont équivalentes entre doses successives. Dans le cas d'une élimination linéaire, cet état d'équilibre est atteint après approximativement 5 demi-vies et le système (1) se résout en :

$$C(t) = \frac{F\ Dose}{V}\ \frac{k_a}{k_a - k}\left(\frac{e^{-k\ t}}{1 - e^{-k\ \tau}} - \frac{e^{-k_a\ t}}{1 - e^{-k_a\ \tau}}\right) \tag{4}$$

Les modèles PK peuvent devenir bien plus complexes à mesure que l'on cherche à se rapprocher de la physiologie en augmentant le nombre de compartiments et d'intéractions. Une présentation plus complète de ces modèles, ainsi que leur formalisation mathématique, peut être trouvée dans [7].

3.1.2 Modèles pharmacodynamiques

En PD, la réponse observée, ou effet, peut être continue, comme la PCA pour la warfarine, ou discrète, comme par exemple la survenue d'une thrombose ou d'une hémorragie.

Réponse continue : une fois sur son site d'action le médicament déclenche une cascade de réactions entraînant directement ou indirectement l'augmentation ou la diminution du facteur responsable ou d'un marqueur de la pathologie ciblée.

Dans les modèles dits de réponse directe, la cinétique de l'effet se déduit directement de l'évolution des concentrations. Les principaux modèles sont des fonctions de type sigmoïde :

$$E(t) = E_0 + E_{max}\ \frac{C(t)^\gamma}{C_{50}^\gamma + C(t)^\gamma} \tag{5}$$

$$E(t) = E_0\left(1 - I_{max}\ \frac{C(t)^\gamma}{C_{50}^\gamma + C(t)^\gamma}\right) \tag{6}$$

Dans ces modèles, E_0 représente le niveau de la réponse en l'absence de médicament, qui peut être augmenté (effet inducteur, équation (5)) ou diminué (effet inhibiteur, équation (6)), mais atteint un plateau ($E_0 + E_{max}$ dans l'équation (5)) reflétant la saturation du processus physiologique aux fortes concentrations. C_{50} est la concentration nécessaire pour obtenir la moitié de l'effet au plateau et γ, aussi appelé facteur de "Hill" contrôle la sensibilité de la réaction. Le niveau basal de la réponse (E_0) peut évoluer

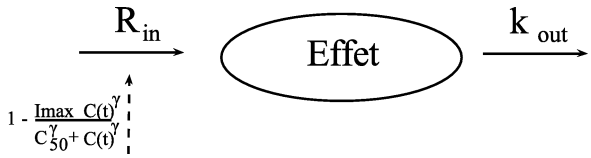

FIGURE 3 – Modèle PD indirect d'inhibition de la synthèse du facteur responsable (ou du marqueur) de la pathologie ciblée.

dans le temps pour tenir compte de changements dans l'état de santé du patient ou de l'évolution de la maladie [8].

Dans les modèles dit de *turnover*, ou encore de réponse indirecte, le médicament n'agit pas directement sur le facteur responsable (ou marqueur) mais plutôt en amont ou en aval, sur les mécanismes générant l'effet observé. C'est le cas pour la warfarine : la réponse observée résulte d'une inhibition du recyclage de la vitamine K qui empêche la formation de facteurs de la coagulation fonctionnels et réduit l'activité anticoagulante, ce qui se reflète dans la diminution de la PCA. Dans ce cas de figure, la dynamique propre du système lui-même est d'abord modélisée, en décrivant la réponse comme la résultante de processus de synthèse et d'élimination, puis l'inhibition ou la stimulation d'un de ces processus par le médicament. Pour la warfarine, ce schéma est illustré par la figure 3.1.2, et peut être décrit par l'équation différentielle ci-dessous :

$$\frac{dE}{dt} = R_{in}\left(1 - \frac{Imax\ C(t)^{\gamma}}{C_{50}^{\gamma} + C(t)^{\gamma}}\right) - k_{out}\ E \qquad \text{et } E(t=0) = R_{in}/k_{out} \tag{7}$$

où R_{in} représente la synthèse de PCA et k_{out} caractérise la dynamique de renouvellement. Il n'existe pas de solution analytique au système et la dépendance à la cinétique du médicament est plus complexe.

L'ensemble des équations des modèles PK et PD (pour une réponse continue) communément utilisés peut être consulté dans la documentation de la librairie de modèles mis en œuvre dans le logiciel Monolix
(http ://www.lixoft.eu/wp-content/resources/docs/PKPDlibrary.pdf).

Réponse discrète : l'action d'un médicament peut aussi être mesurée plus en aval du marqueur physiologique, en critères "cliniques", tels que le niveau de douleur, ou encore le nombre de crises épileptiques sur une durée donnée. Ces marqueurs de l'activité du médicament peuvent être modélisés à l'aide de distributions discrètes caractérisant leur distribution, telle que la loi de Poisson pour les données de comptages ou la loi multinomiale pour les données ordonnées.

Dans le cas de la warfarine, le nombre n de saignements pourrait ainsi suivre une loi de Poisson :

$$P(Y = n) = \frac{\lambda^n \, e^{-\lambda}}{n!} \tag{8}$$

où λ caractérise le nombre moyen et la variance des saignements. La PK du médicament (à travers sa concentration moyenne ou l'AUC par exemple) peut être inclue dans ce modèle comme une covariable sur λ. La loi de Poisson fait l'hypothèse forte d'une équidispersion des données, d'autres distributions comme la binomiale inverse (ou négative) permettent de relaxer cette contrainte.

3.2 Modèles à effets mixtes

Les paramètres des modèles introduits ci-dessus permettent donc de quantifier les différents processus biologiques impliqués dans la réponse au médicament. Ces modèles sont généralement non-linéaires, à la fois en fonction du temps et en fonction des paramètres dont ils dépendent. Comme tout phénomène biologique, la réponse montre une certaine variabilité. L'approche de population permet de décomposer cette variabilité au travers des effets aléatoires et éventuellement d'en identifier les sources. Nous traiterons seulement le cas des réponses continues dans la suite, mais d'autres méthodes permettent de traiter des réponses discrètes. Nous considèrerons le modèle suivant :

$$y_{ij} = f(x_{ij}, \psi_i) + g(x_{ij}, \psi_i)\epsilon_{ij} \tag{9}$$

où f représente le modèle décrivant l'évolution de la réponse considérée, g décrit l'amplitude de l'erreur de mesure ϵ_{ij}, et $\psi_i = (\psi_{i,\ell}\,; 1 \leq \ell \leq n_\psi) \in \mathbb{R}^{n_\psi}$ est le vecteur de n_ψ paramètres individuels propre au sujet i.

3.2.1 Modélisation des effets aléatoires

Dans ce modèle, nous supposons que la même fonction f caractérise l'ensemble de la population étudiée ; chaque sujet a des valeurs de paramètres qui lui sont propres et diffèrent de la valeur associée au sujet moyen ou "effet fixe" par un effet aléatoire. Dans les méthodes paramétriques qui sont le plus couramment utilisées, la distribution est définie *a priori* de l'analyse, et on a coutume de la décomposer en effets fixes, notés $\mu = \{\mu_1, ..., \mu_{n_\mu}\}$, communs à tous les sujets, et en effets aléatoires, notés $\eta_i = \{\eta_{i1}, ..., \eta_{in_\eta}\}$, portant la variabilité interindividuelle. Nous noterons H la fonction décrivant cette relation et ϕ_i le paramètre avant transformation :

$$\psi_i = H(\mu, c_i, \eta_i) = H(\phi_i) \tag{10}$$

Dans la suite nous supposerons, mais cette hypothèse peut être relâchée, que ϕ_i peut s'exprimer comme une combinaison linéaire des effets fixes et aléatoires, soit $\phi_i = \mu_l + \eta_{il}$. Cette relation peut faire intervenir les covariables (supposées connues) du sujet, $c_i = (c_{im}; 1 \leq m \leq M)$. Nous supposerons également connue $\mathcal{D}$, la distribution des effets aléatoires η_i ; elle est le plus souvent prise gaussienne, de matrice de variance-covariance Ω, et nous noterons ω_l^2 les éléments diagonaux de cette matrice : $\eta_i \sim_{i.i.d.} \mathcal{N}(0,\Omega)$. Des formes usuelles pour H incluent :

- $H(\mu, c_i, \eta_i) = \mu + \eta_i$ suit une distribution normale
- $H(\mu, c_i, \eta_i) = \mu e^{\eta_i}$ suit une distribution lognormale, et implique $\psi_i > 0$
- $H(\mu, c_i, \eta_i) = 1/(1 + e^{-(\mu+\eta_i)})$ implique que ψ_i sera compris entre 0 et 1

Lorsque des observations sont recueillies à plusieurs occasions (visites cliniques différentes, changements de régime de dose ou de bras de traitement), les paramètres individuels peuvent montrer une variabilité intra-individuelle supplémentaire, aussi appelée variabilité interoccasion. Il est possible de tenir compte de ce niveau supplémentaire de variabilité dans la définition des ψ_i, en introduisant un vecteur d'effets aléatoires $\kappa_{i,k}$ où k dénote l'occasion. On suppose généralement $\kappa_{i,k} \sim_{i.i.d.} \mathcal{N}(0,\Gamma)$, et les paramètres individuels à l'occasion k deviennent :

$$\psi_{i,k} = H(\mu, c_i, \eta_i, \kappa_{i,k}) \tag{11}$$

3.2.2 Modélisation de l'erreur résiduelle

L'erreur résiduelle $g(x_{ij}, \psi_i)\varepsilon_{ij}$ est l'écart entre l'observation y_{ij} et la prédiction individuelle $f(x_{ij}, \psi_i)$ et représente la variabilité qui ne peut être capturée par la fonction f. Classiquement, les ϵ_{ij} sont supposés normaux et indépendants des effets aléatoires (η_i et $\kappa_{i,k}$). La variance est modélisée par la fonction g, les principaux modèles utilisés en pharmacométrie étant :

- erreur constante (homoscédastique) où $g = \sigma_{inter}$,
- erreur proportionnelle (hétéroscédastique) où $g = \sigma_{slope} f$,
- erreur combinée (hétéroscédastique) où $g = \sigma_{inter} + \sigma_{slope} f$

3.2.3 Modèles de covariable

La variance des effets aléatoires permet de quantifier la variabilité inter-sujet. L'ajout de covariables dans la fonction H, définissant les paramètres du modèle, permet d'expliquer les sources de cette variabilité. Un modèle usuel pour exprimer la dépendance du paramètre ψ_l de distribution log-normale à une covariable continue c_m est :

$$\phi_{il} = \mu_l + \beta_{c_m l} log\left(\frac{c_{im}}{c_{pop\ m}}\right) + \eta_{il} \qquad \text{équivalent à } \psi_{il} = H(\phi_{il}) = \mu_l \left(\frac{c_{im}}{c_{pop\ m}}\right)^{\beta_{c_m l}} e^{\eta_{il}}$$

en centrant sur $c_{pop\,m}$, la valeur typique de la covariable dans la population. Si c_m est une covariable catégorielle prenant les valeurs 1, 2, ..., G, et en supposant que le groupe 1 est la référence ($\beta_1 = 0$), pour un sujet i du groupe g le paramètre peut s'écrire :

$$\phi_{il} = \mu_l + \beta_{gl} + \eta_{il}$$

4 Méthodes d'estimation

Nous noterons θ l'ensemble des paramètres de population à estimer. θ comprend notamment les effets fixes μ, les paramètres de g comme σ_{inter} et σ_{slope}, ainsi que les paramètres décrivant la distribution $\mathcal{D}$ (Ω) : $\theta = (\mu, \Omega, \sigma_{inter}, \sigma_{slope})$. Nous noterons $\ell(y,\theta) = p(y/\theta)$ la vraisemblance des observations $y = (y_{ij}\,; 1 \leq i \leq n\ , 1 \leq j \leq n_i)$, qui s'écrit comme le produit des vraisemblances de chaque sujet i, $\prod_{i=1}^{N} \ell(y_i,\theta)$. Dans les modèles non-linéaires à effets mixtes (MNLEM), on sait écrire $p(y/\psi_i,\theta)$, la vraisemblance des observations d'un sujet i conditionnellement aux paramètres ψ_i de ce sujet, aussi appelée vraisemblance complète, selon le modèle gaussien défini en (9). Les paramètres ψ_i étant inconnus, $\ell(y_i,\theta)$ s'obtient en intégrant sur la distribution $\mathcal{D}$:

$$\ell(y_i,\theta) = \int_{\mathcal{D}} p(y_i/\psi_i,\theta)\, d\psi_i \tag{12}$$

Il existe deux approches pour l'estimation des paramètres, fréquentiste ou bayésienne. Le maximum de vraisemblance, ou approche fréquentiste, consiste à estimer les paramètres maximisant la vraisemblance des données observées $\ell(y,\theta)$ par rapport aux paramètres (section 4.1), tandis que les méthodes bayésiennes maximisent directement la probabilité des paramètres conditionnellement aux données $p(\theta/y)$ (section 4.2). Ces dernières intègrent une information *a priori* sous la forme d'une distribution des paramètres d'intérêt, pour calculer la distribution *a posteriori* des paramètres par des méthodes d'intégration. Une introduction plus complète aux modèles et méthodes d'estimation dans le domaine des MNLEM peut être trouvée dans [9]. Nous terminerons cette section par un panorama des logiciels disponibles, et nous illustrerons la démarche sur l'exemple de la warfarine.

4.1 Méthodes fréquentistes - maximum de vraisemblance

Nous verrons tout d'abord les méthodes pour calculer des approximations stochastiques ou numériques de la vraisemblance qui, nous l'avons vu en (12), s'exprime sous une forme intégrale qui n'admet généralement pas de solution analytique (sauf quand

$\mathcal{D}$ est discrète). Nous présenterons ensuite les algorithmes d'estimation, qui se classent en deux catégories : des algorithmes omnibus de minimisation, ou des algorithmes EM, pour Expectation-Maximisation.

4.1.1 Calcul de la vraisemblance

Approximation de $\ell(y_i, \theta)$ par linéarisation du modèle. Linéariser f par un développement limité à l'ordre 1 permet de calculer explicitement l'intégrale en (12). Ce développement peut être effectué autour des paramètres fixes μ (méthode FO [10], pour *First-Order*, ou premier ordre) ou autour des paramètres individuels (First-Order Conditional Estimation (FOCE) [11]). La distribution de y_i devient une distribution normale de moyenne $\mathbb{E}(y_i)$ et de variance $\mathrm{Var}(y_i)$, dépendant des matrices des dérivées premières de f et g au voisinage de 0 (méthode FO) ou des estimations des effets aléatoires individuels $\hat{\eta}_i$ (méthode FOCE). Ces méthodes permettent d'obtenir une expression analytique de la vraisemblance linéarisée.

Calcul de $\ell(y_i, \theta)$ par intégration numérique. Les méthodes d'intégration numérique consistent à approcher l'intégrale (12) par une somme discrète de valeurs de la fonction à intégrer, prises en un nombre N_{QG} de points appelés noeuds et pondérées par des poids. N_{QG} représente l'ordre de la quadrature, un grand nombre améliorant la précision mais aussi la complexité et le coût du calcul. Les poids et les noeuds peuvent être déterminés par différents algorithmes : le plus utilisé est la quadrature de Gauss-Hermitte, bien adaptée à l'intégration de fonctions dont la densité est proche de la loi normale [12]. Dans les MNLEM, cette intégration est réalisée sur chacune des dimensions représentant les paramètres aléatoires [13], ce qui rend vite les calculs trop lourds. Une façon d'améliorer le calcul de l'intégrale tout en diminuant le coût calculatoire est d'utiliser des quadratures dites adaptatives (AGQ) dont le support coïncide avec celui de la fonction à intégrer [14]. A l'ordre 1, l'AGQ est équivalente à l'approximation dite de Laplace [15].

Calcul de $\ell(y_i, \theta)$ par intégration stochastique. La même intégrale peut être calculée par une approximation stochastique au lieu de numérique, en utilisant des tirages de Monte-Carlo dans les distributions à intégrer, pour la remplacer à nouveau par une somme discrète. Comme précédemment, la qualité de l'approximation est déterminée par le nombre de points, et le coût calculatoire augmente exponentiellement avec le nombre de paramètres. Pour le limiter, il est possible d'utiliser des méthodes d'échantillonnage préférentiel qui à nouveau cherchent à se rapprocher du support de la distribution pour améliorer la qualité de l'approximation avec moins de tirages (*im-*

portance sampling). Une difficulté supplémentaire est qu'on ne sait généralement pas échantillonner directement dans les distributions d'intérêt pour les MNLEM (distributions des paramètres ou des effets aléatoires individuels), ce qui nécessite de faire appel à des techniques de chaînes de Markov [16], dont une présentation peut être trouvée dans des ouvrages de référence comme [17].

4.1.2 Algorithmes d'estimation

Les algorithmes d'estimation des paramètres de population sont des algorithmes itératifs construisant une suite de valeurs des paramètres de population, θ_0, θ_1, .., θ_k, convergeant vers l'estimateur du maximum de vraisemblance de ces paramètres, $\hat{\theta}$. Il en existe principalement deux types.

Les algorithmes standard de minimisation comme l'algorithme de Newton-Raphson peuvent être utilisés pour trouver le minimum d'une des approximations précédentes de la vraisemblance, par des méthodes de gradient. La première étape consiste à initialiser l'algorithme, avec un choix de $\theta = \theta_0$. Ensuite, à chaque étape k, la vraisemblance de l'ensemble des données conditionnellement à θ_{k-1} est calculée, et les paramètres de population θ_k sont obtenus par maximisation de cette vraisemblance. L'algorithme s'arrête lorsque la différence (relative) entre les composantes de θ_k et θ_{k+1} est moindre qu'un seuil prédéfini (ou tolérance). Des variantes permettent d'éviter le calcul des dérivées (algorithmes de type quasi-Newton) ou au contraire utilisent les dérivées secondes pour améliorer ou accélérer la convergence.

Un algorithme spécifique, l'algorithme EM, a été proposé par Dempster en 1977 pour résoudre les problèmes de données manquantes [18]. Dans le cas des MNLEM, les effets aléatoires individuels η_i peuvent en effet être considérés comme des données manquantes. Egalement itératif, l'algorithme répète à chaque pas deux étapes successives : l'étape E calculant l'espérance conditionnelle de la (log-)vraisemblance complète des observations, en complétant les paramètres manquants par simulation, puis l'étape M maximisant cette quantité en θ. La vraisemblance (12) est calculée à la fin de la minimisation en utilisant les paramètres de population estimés (*plug-in*). Le problème pour les MNLEM réside dans le calcul de l'espérance de la log-vraisemblance complète à l'étape E, et on retrouve les différentes solutions proposées pour le calcul de la vraisemblance : linéarisation autour des effets aléatoires individuels $\hat{\eta}_i$ [19], approximation numérique par quadrature de Gauss [20], approximation stochastique par méthode de Monte-Carlo [21, 22] ou par échantillonnage préférentiel [23]. La plus efficace de ces méthodes aujourd'hui est l'algorithme SAEM (Stochastic Approximation EM), qui utilise des statistiques suffisantes au lieu de calculer la vraisemblance complète, et les actualise progressivement en tirant un seul échantillon à chaque étape, ce

qui accélère l'estimation tout en garantissant la convergence [24].

4.2 Méthodes bayésiennes

4.2.1 Modèle hiérarchique bayésien

Les méthodes bayésiennes cherchent à maximiser la probabilité (densité dans le cas continu) des paramètres θ, conditionnellement aux observations y, $p(\theta/y)$. La relation avec la vraisemblance $p(y/\theta)$ s'exprime d'après le théorème de Bayes :

$$p(\theta/y) = \frac{p(y/\theta)\,\pi(\theta)}{p(y)} \tag{13}$$

en notant $\pi(\theta)$ la distribution *a priori* sur les paramètres, supposée connue avant l'observation des données, et $p(y) = \int p(y|\theta)\pi(\theta)d\theta$ la densité marginale. Cette dernière correspond à l'intégration de $p(y|\theta)$ par rapport à toutes les valeurs possibles de θ, et elle apparaît donc comme une constante par rapport à θ. $\pi(\theta)$ intègre dans la philosophie bayésienne les connaissances accumulées sur le système ; elle permet aussi d'inclure des contraintes de manière naturelle et transparente à travers les *a priori* choisis. $\pi(\theta)$ introduit donc dans le modèle hiérarchique bayésien un niveau supplémentaire par rapport au cas fréquentiste, correspondant aux *a priori* sur tous les paramètres, y compris ceux des modèles décrivant la variabilité inter-sujet et résiduelle :

$$\theta \sim \pi \tag{14}$$

Les méthodes bayésiennes nécessitent de choisir une forme et des paramètres *a priori* pour décrire π, la distribution des effets aléatoires. π peut être choisie peu informative, c'est-à-dire apportant peu d'information par rapport aux données, pour refléter la faiblesse des connaissances acquises à ce stade de l'analyse, ou informative, reflétant par exemple les résultats d'une première étude. Wakefield [25] a proposé un choix de distributions canoniques que l'on retrouve désormais dans beaucoup d'analyses par méthode bayésienne. Ces distributions sont attrayantes pour deux raisons principales ; d'une part, elles sont faciles à échantillonner, d'autre part, les distributions *a posteriori* et *a priori* ont la propriété d'être "conjuguées ». Ceci signifie que la forme de ces deux distributions est la même, avec deux conséquences intéressantes : d'une part, une méthode permettant de simuler dans la distribution *a priori* permettra également de simuler dans la distribution *a posteriori* ; d'autre part, elle permet d'intégrer de manière naturelle de nouvelles informations, et à ce titre, est cohérente avec la philosophie bayésienne selon laquelle l'*a posteriori* d'une analyse sert d'*a priori* dans la suivante.

4.2.2 Algorithmes d'estimation

A partir de $\pi(\theta)$ (distribution *a priori*) et de $p(y/\theta)$ (distribution d'échantillonnage), des méthodes d'intégration sont utilisées pour construire différentes distributions : la distribution jointe $p(y;\theta)$, la distribution marginale $p(y)$ ou distribution prédictive *a priori* (*prior predictive distribution*) des observations en intégrant sur θ, la distribution *a posteriori* $p(\theta/y)$, et la distribution prédictive (*a posteriori*) $p(y*/y)$ de futures observations $y*$. D'un point de vue algorithmique, l'intégration se heurte au même problème qu'en fréquentiste, à savoir l'absence de forme analytique pour l'expression de $p(y/\theta)$, ainsi qu'à un coût calculatoire important notamment pour calculer les constantes de normalisation. Là encore des méthodes d'intégration numérique ou stochastique sont utilisées, avec une préférence pour les méthodes stochastiques car dans un contexte bayésien le fléau de la dimension est particulièrement sensible, l'intégrale portant non seulement sur les effets aléatoires mais aussi sur tous les autres paramètres du modèle de troisième niveau spécifiés dans le modèle (distribution des effets fixes, des variances,...). Ces difficultés ont d'ailleurs conduit au développement des méthodes MCMC avec chaînes de Markov performantes maintenant largement utilisées aussi dans le cas fréquentiste.

Les distributions marginales (distribution d'un paramètre) sont obtenues en intégrant pour éliminer les autres paramètres. Par construction, le résultat "naturel" d'une analyse bayésienne est donc toute une distribution. De même, si on s'intéresse à un paramètre, l'analyse bayésienne fournit toute sa distribution *a posteriori*. Pour résumer cette information, des statistiques descriptives sont calculées :

- le mode de la distribution, c'est-à-dire la valeur de densité maximale pour une loi continue, équivalent à l'estimateur ponctuel en analyse fréquentiste
- la valeur moyenne ou médiane de la distribution, mais ces estimateurs peuvent avoir une faible probabilité (par exemple dans le cas où la distribution du paramètre s'avère bimodale)
- plus généralement, les quantiles de la distribution peuvent être utilisés
- la variabilité de chaque paramètre peut être considérée comme l'incertitude sur ce paramètre, donc une estimation de la précision d'estimation est donnée par l'écart-type de la distribution

La probabilité d'un événement (par exemple, la probabilité que le paramètre estimé soit supérieur à une valeur donnée) est également très facile à calculer puisque toute la distribution est connue.

4.3 Erreurs d'estimation

Les estimations ponctuelles des paramètres doivent s'accompagner de celle de leur incertitude pour permettre de juger de la qualité de l'estimation. Dans les méthodes bayésiennes, l'erreur d'estimation sur chaque paramètre est donnée par l'écart-type de la distribution *a posteriori*. Pour les approches fréquentistes, plusieurs méthodes permettent d'obtenir des erreurs standards d'estimation (SE) ou des intervalles de confiance :

- l'approximation asymptotique, utilisant la matrice d'information de Fisher M_F ;
- le bootstrap, qui rééchantillonne les données de façon répétée pour construire la distribution des estimateurs ;
- le profil de vraisemblance, où un intervalle de confiance est obtenu pour chaque paramètre en fonction de l'évolution de la vraisemblance lorsque ce paramètre est fixé à différentes valeurs

La méthode la plus utilisée est l'approximation asymptotique, qui est généralement celle fournie par les logiciels. Aux termes de l'inégalité de Rao-Cramer, une borne inférieure de l'erreur d'estimation est fournie par l'inverse de la matrice d'information de Fisher. Cette matrice est définie par une des deux formules (asymptotiquement équivalentes) suivantes :

$$M_F(\theta;\Xi) = \mathbb{E}\left(-\frac{\partial^2 LL(y,\theta)}{d\theta d\theta^T}\right) = \mathbb{E}\left(-\frac{\partial LL(y,\theta)}{d\theta}\,\frac{\partial LL(y,\theta)}{d\theta}^T\right) \tag{15}$$

où $LL(y,\theta) = \ln(\ell(y,\theta))$ est la log-vraisemblance et Ξ dénote le protocole expérimental.

4.4 Estimation des paramètres individuels

Dans les MNLEM, l'estimation porte sur les paramètres de population et les paramètres individuels ne sont pas estimés pendant l'analyse, sauf comme intermédiaires de calcul. L'estimation des paramètres individuels se fait donc *a posteriori*, selon une méthode bayésienne utilisant le résultat de l'analyse de population comme distribution *a priori* et les données propres au patient (y_i) :

$$p(\eta_i/y_i,\theta) = \frac{p(y_i/\eta_i,\theta)p(\eta_i/\theta)}{\int p(y_i/\eta,\theta)p(\eta/\theta)d\eta} \propto p(y_i/\eta_i,\theta)\,p(\eta_i/\theta) \tag{16}$$

Dans les méthodes fréquentistes, une estimation *plug-in*, appelée estimation bayésienne empirique (EBE), est obtenue en utilisant l'estimation $\hat{\theta}$. Les paramètres individuels peuvent alors être calculés comme le mode *a posteriori* (MAP) ou l'espérance *a posteriori* (EAP) de la distribution.

Ce critère s'interprète comme la combinaison d'une information *a priori* (ici, la distribution de population) et de l'adéquation entre les données et les prédictions. Plus l'information apportée par un sujet est importante, moins l'information *a priori* va influencer l'estimation des paramètres individuels, mais *a contrario* des sujets avec peu d'observations (information éparse) vont voir leurs paramètres se rapprocher des paramètres de la population. Ce phénomène s'appelle la régression à la moyenne, ou *shrinkage* en anglais. Il peut être quantifié pour un effet aléatoire η par la quantité suivante, souvent appelée *η-shrinkage* :

$$Sh_\eta = 1 - \frac{\mathrm{Var}\,(\hat{\eta}_i)}{\omega_\eta^2} \tag{17}$$

où $\mathrm{Var}\,(\hat{\eta}_i)$ dénote la variance empirique des EBE.

De plus, la régression à la moyenne se traduit par une régression des prédictions individuelles vers les observations (il est "plus facile" d'ajuster peu d'observations en jouant sur la valeur des paramètres), et on peut définir un "*ϵ-shrinkage*" également sur les prédictions : $Sh_\epsilon = 1 - \mathrm{Var}\left(IWRES_{ij}\right)$ où les $IWRES_{ij} = y_{ij} - f(x_{ij}, \hat{\psi}_i)$ sont les résidus individuels.

Le shrinkage a pour effet de diminuer la fiabilité de certains graphes diagnostiques ; en particulier les graphes reposant sur les prédictions ou les résidus individuels sont très sensibles au *shrinkage* et peuvent être non informatifs ou trompeurs lorsque l'information individuelle est insuffisante [26].

4.5 Logiciels d'estimation des paramètres

Il existe aujourd'hui de nombreux logiciels qui permettent de réaliser l'estimation des paramètres de MNLEM. Nous nous contenterons de présenter les plus utilisés, résumés dans le tableau 2.

En maximum de vraisemblance, le premier logiciel et le plus utilisé s'appelle NONMEM [27]. Il a été développé plus spécifiquement dans le cadre des analyses de PK de population [28] et comprend notamment une librairie des modèles PK les plus courants ainsi que des modules pour traiter d'autres types de données (doses, mode d'administration, ...). NONMEM a évolué avec le temps et comprend maintenant, entre autres, les méthodes FO, FOCE, Laplace, SAEM, et une méthode bayésienne. Phoenix est l'un des logiciels commerciaux les plus aboutis, avec une interface graphique et la possibilité d'effectuer des analyses non-compartimentales ou des analyses de population. Monolix [29], utilisant l'algorithme SAEM, connaît actuellement un développement rapide lui permettant d'être à la pointe à la fois des méthodes statistiques et des outils de diagnostics, tout en offrant une interface graphique très complète. Par

ailleurs, des librairies spécialisées ont été développées pour les logiciels statistiques R [30] et SAS [31]. R en particulier, un logiciel gratuit, dispose de plusieurs algorithmes d'estimation dont les méthodes FOCE [32] et SAEM [33].

En analyse bayésienne, le logiciel le plus utilisé est BUGS [34] ou sa version standalone WinBugs [35]. D'autres logiciels ou programmes ont été développés pour mettre en œuvre des algorithmes spécifiques, mais ils sont peu utilisés et nécessitent une bonne connaissance des modèles et des algorithmes pour les adapter aux cas particuliers à traiter.

4.6 Application à l'exemple

Estimation. Nous avons utilisé Monolix pour estimer les paramètres de population du modèle PK/PD de la warfarine défini précédemment, à partir des données montrées en figure 1. Ce modèle s'inspire d'un projet de démonstration, disponible après installation du logiciel Monolix, le projet `warfarin_PKPD3_project` dans le répertoire `Demos/warfarin`. Dans ce modèle, les concentrations $C(t)$ peuvent être décrites par $C(t) = Q(t)/V$ où $Q(t)$ est donné par l'équation 1, et l'effet inhibiteur sur la synthèse de la PCA est modélisé selon l'équation (7) (*turnover model*). La variabilité inter-sujet sur certains paramètres était faible et a été considérée comme nulle (*i.e*, tous les sujets ont la même valeur du paramètre), et un effet significatif du poids sur la clairance CL et le volume V a été trouvé. Enfin, le modèle d'erreur résiduelle était combiné pour la PK et constant pour la PD. L'ajustement donne les estimations exposées en table 3. Les paramètres moyens sont bien estimés (erreurs d'estimation inférieures à 20%), ainsi que les effets des covariables (SE de 16 et 38%) qui sont tous les deux significatifs selon un test de Wald (p-valeur de 10^{-7} et 0.003 pour respectivement $\beta_{WT,V}$ et $\beta_{WT,CL}$. L'erreur d'estimation sur les variabilités est un peu plus élevée, de 15 à 60%, ce qui est assez classique. Dans ce tableau, les variabilités sont données en écart-type des effets aléatoires η, ce qui correspond à peu près à un coefficient de variation pour les effets fixes. Ainsi, $\omega_{k_a} = 0.63$ correspond environ à 63% de variabilité dans la population pour le paramètre k_a. V et CL dépendent du poids selon une relation de type $\ln(V_i) = \ln(V_0) + \beta_{WT,V}(\ln(WT) - \ln(70)) + \eta_i$. Ainsi, le volume prédit pour un sujet de 75 kg (5 kg de plus que le poids moyen de 70 kg) est de 8.24 L, soit une augmentation de 0.47 L (+6%) par rapport au sujet de poids moyen pour lequel V vaut 7.77 L.

Un exemple d'ajustement est donné en figure 4 pour deux sujets de l'étude, en séparant les deux réponses. On superpose les points correspondant aux données et les prédictions, obtenues en estimant les paramètres individuels pour chaque sujet. Ici, ces graphes montrent que les deux modèles PK et PD reproduisent bien les processus observés alors même que les profils des deux sujets sont assez différents.

Enfin, la table 3 donne également le shrinkage calculé pour chaque paramètre. Ici, l'information est assez riche (6 à 17 points par sujet pour la PK, et 4 à 8 points pour la PD). Le shrinkage est cependant supérieur à 30% pour presque tous les paramètres, et supérieur à 50% pour la moitié d'entre eux.

Utilisation des résultats. Les paramètres ainsi estimés peuvent être utilisés pour évaluer le comportement PK/PD d'un médicament et proposer des régimes de dose adaptés à la population générale. Ces informations sont utilisées pour rédiger les notices d'information données dans les boîtes de médicament. Par ailleurs, les profils individuels montrés en figure 4 permettent de déterminer l'évolution d'un patient spécifique, et peuvent être utilisés pour adapter le traitement à chaque cas particulier ; cette dernière application de la modélisation s'appelle la surveillance thérapeutique.

Logiciel	Algorithmes disponibles	Type	Interface
Monolix	SAEM	Mixte[1]	Oui
NONMEM	FO, FOCE, Laplace, SAEM, Bayes	Payant	Non
Phoenix	FOCE	Payant	Oui
R (librairies)	nlme, lme4 (FOCE, Laplace), saemix (SAEM)	Gratuit	Non[2]
SAS	NLMIXED (FO, FOCE, AGQ)	Payant	Oui
WinBugs	Bayes	Gratuit	Non

[1] *Licence gratuite pour les universitaires et étudiants*

[2] *R n'a pas d'interface graphique spécifique mais des outils comme Rstudio peuvent être utilisés pour le faire tourner*

Table 2 – Logiciels de population les plus utilisés

Paramètres Modèle structural	Effet fixe (μ) (SE en %)	Paramètre de variance (ω) (SE en %)	Shrinkage (%)
k_a (hr^{-1})	0.54 (15)	0.63 (18)	69
V (L)	7.77 (3)	0.13 (22)	38
$\beta_{WT,V}$ (-)	0.82 (19)	-	-
CL ($L.hr^{-1}$)	0.13 (4)	0.22 (14)	7
$\beta_{WT,CL}$ (-)	0.66 (33)	-	-
IC_{50} ($mg.L^{-1}$)	1.20 (9)	0.42 (15)	19
I_{max} (-)	0.93 (2)	-	-
R_{in} (% d'activité.hr^{-1})	5.06 (3)	0.04 (63)	69
k_{out} (hr^{-1})	0.05 (3)	0.04 (41)	54
γ (-)	1.44 (7)	0.13 (38)	79
Modèle d'erreur	σ		
$\sigma_{inter,PK}$ ($mg.L^{-1}$)	0.64 (16)		
$\sigma_{slope,PK}$ (-)	0.07 (24)		
$\sigma_{inter,PD}$ (% d'activité)	3.61 (6)		

Table 3 – Paramètres estimés par Monolix pour le modèle de réponse indirecte de la warfarine, ainsi que leur erreur d'estimation (SE). La SE est donnée en pourcentage, qui est calculé comme l'erreur d'estimation absolue rendue par le logiciel divisée par l'estimation du paramètre de population. La colonne variabilité contient l'estimation de l'écart-type pour l'effet aléatoire correspondant (ω). Les paramètres k_a (vitesse d'absorption), V (volume de distribution) et CL (clairance) caractérisent la PK de la warfarine, tandis que les paramètres R_{in} (synthèse de PCA en l'absence de warfarine), k_{out} (vitesse de renouvellement de la PCA), I_{max} (inhibition maximale du cycle de la vitamine K), et IC_{50} (concentration de warfarine à laquelle l'inhibition est égale à 50% de I_{max}) caractérisent la réponse et l'influence de la warfarine sur son évolution.

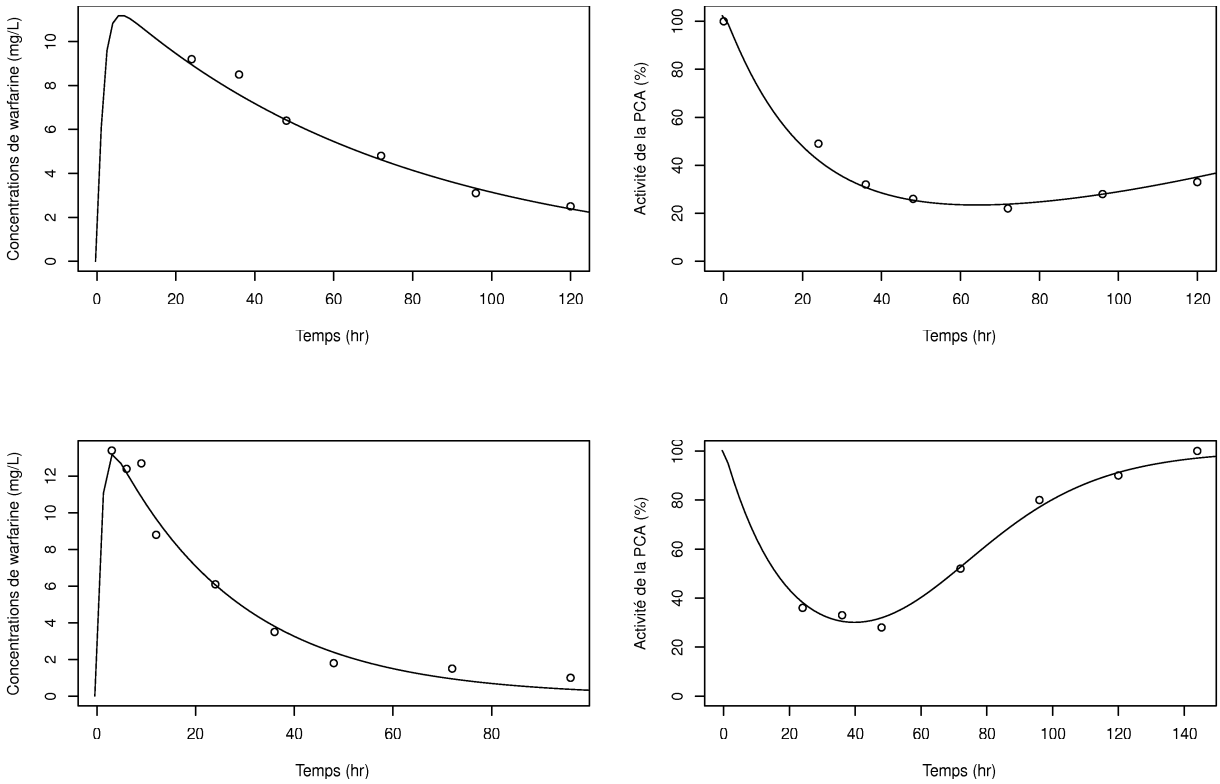

FIGURE 4 – Ajustement des données chez deux sujets (à gauche la PK, à droite la PD.

5 Construction et évaluation de modèles

La construction d'un modèle dans une analyse de pharmacométrie nécessite d'évaluer les modèles proposés, et de les comparer statistiquement. Nous présentons dans cette section les outils d'inférence et de diagnostics permettant de conduire ce processus décisionnel [36].

5.1 Inférence

Inférence en maximum de vraisemblance. En régression, les tests d'inférence classiques sont le test du rapport de vraisemblance (LRT), le test de Wald et le test du score. Dans les MNLEM, le LRT, le test de Wald et les critères statistiques comme le critère d'Akaike sont le plus souvent utilisés.

Le LRT compare les log-vraisemblances LL_c et LL_r de deux modèles M_c (complet) et M_r (réduit), présentant la particularité d'être emboîtés, c'est-à-dire que le modèle M_c, à $q_c = q_r + q_{c-r}$ paramètres, peut être réduit au modèle M_r en supposant que certains paramètres (ici q_{c-r}) sont nuls. L'hypothèse H_0, soit $\{M_c = M_r\}$ ou encore $\{\theta_{q_r+1} = \theta_{q_c} = 0\}$, peut être testée asymptotiquement en comparant la différence entre les vraisemblances à un χ^2 de nombre de degrés de liberté égal à q_{c-r} :

$$\chi = -2\,(LL_r - LL_c) \sim \chi^2_{q_{c-r}} \tag{18}$$

Pour tester si un paramètre θ_l du modèle peut être pris égal à 0, le test de Wald utilise comme statistique :

$$W = \frac{(\hat{\theta}_l)^2}{\hat{SE}(\hat{\theta}_l)^2} \sim \chi^2_1 \tag{19}$$

où $\mathrm{SE}(\hat{\theta}_l)$ est l'erreur d'estimation sur $\hat{\theta}_l$. Ce test s'étend au test simultané de plusieurs paramètres (test de Hotelling).

Lorsque les modèles ne sont pas emboîtés, il est possible de les comparer au travers de critères statistiques tels le critère d'Akaike (AIC) ou le BIC (Bayes Information Criterion). Ces critères pondèrent la (log-)vraisemblance du modèle par une fonction du nombre de paramètres p du modèle :

$$\begin{aligned} AIC &= -2LL(y;\theta) + 2p \\ BIC &= -2LL(y;\theta) + p\,\ln n \end{aligned} \tag{20}$$

D'autres critères ont été proposés pour les MNLEM [37].

Les propriétés statistiques de ces tests sont valables dans des conditions asymptotiques. Dans les MNLEM, cela nécessite que le nombre de sujets et le nombre de prélèvements par sujet tendent simultanément vers l'infini. En pratique, on observe une

augmentation de l'erreur de type I des tests dans les petits échantillons [38, 39]. Cette augmentation dépend également de la méthode d'estimation, et notamment l'approximation FO ne donne pas des estimateurs consistants [40]. Dans les conditions finies qui sont généralement celles des analyses PK, il est possible de corriger le niveau des tests : des seuils par permutations ainsi que des corrections de la statistique du test de Wald ont été ainsi proposés [41].

Application. Dans l'exemple précédent, Monolix nous donnait les p-valeurs du test de Wald pour les deux paramètres du modèle de covariables (section 4.6). Ce test présente l'avantage de pouvoir tester la significativité d'un paramètre en tenant compte des autres covariables présentes dans le modèle. Nous pouvons aussi utiliser le test du rapport de vraisemblance pour tester par exemple si la relation entre le poids et CL est significative. Pour cela, nous refaisons tourner le modèle en enlevant le paramètre $\beta_{WT,CL}$. Les valeurs de -2LL sont respectivement 2276.39 et 2285.16 pour le modèle complet et le modèle sans $\beta_{WT,CL}$, et la p-value du test du LRT est donc de 0.03. Ici les deux tests (Wald et LRT) donnent la même conclusion quant à l'intérêt de cette association entre le poids et CL dans le modèle. De plus, la prise en compte de la covariable a pour effet de diminuer légèrement ω_{CL} qui passe de 0.26 à 0.22, soit une diminution de 15%.

Inférence bayésienne. Si la notion de test dans le sens de test de Neyman-Pearson d'une hypothèse nulle fait débat [42], il est naturel sous la philosophie bayésienne d'évaluer la crédibilité d'une valeur *a priori* de paramètres, ou l'adéquation d'un modèle, en calculant la probabilité *a posteriori* d'une hypothèse à partir des distributions estimées grâce au théorème de Bayes. Formaliser une théorie de la décision intégrant toutes les étapes depuis la construction du modèle jusqu'à son utilisation en passant par l'estimation est toutefois un processus complexe, et nous référons le lecteur à des ouvrages plus complets comme [43, 44].

Deux modèles concurrents, correspondant à deux hypothèses H_0 et H_1, peuvent être comparés au travers du facteur de Bayes [45]. En notant $p(H_0)$ et $p(H_1) = 1 - p(H_0)$ les probabilités *a priori*, nous calculons le rapport des probabilités *a posteriori* des deux hypothèses en tenant compte des données y et nous le transformons à l'aide du théorème de Bayes :

$$\frac{p(H_1|y)}{p(H_0|y)} = \frac{p(y|H_1)}{p(y|H_0)} \frac{p(H_1)}{p(H_0)} \tag{21}$$

où on appelle le rapport $B_{10} = \frac{p(y|H_1)}{p(y|H_0)}$ le facteur de Bayes. Ce terme mesure la force de conviction en faveur de l'hypothèse H_1 par rapport à l'hypothèse H_0 contenue dans les données, et modifiant le rapport de probabilités *a priori*. Les facteurs de Bayes peuvent

être calculés pour toute une série de modèles différents afin d'en combiner les prédictions. Contrairement aux tests fréquentistes, les modèles n'ont pas besoin d'être emboîtés pour être comparés par les facteurs de Bayes. Notons que le BIC est d'ailleurs une approximation de $\log B_{10}$, ce qui explique son intérêt pour la sélection de modèles [46]. Spiegelhalter propose également un critère bayésien spécifique, le DIC pour Deviance Information Criterion [47], calculé en pondérant une mesure d'adéquation du modèle par le nombre de paramètres effectifs.

5.2 Construction de modèles

Construire un modèle consiste à trouver le modèle minimal qui décrive correctement les données, en déterminant ses différentes composantes : modèle structurel, variabilité interindividuelle et résiduelle, influence des covariables.

Il n'y a pas de consensus sur les méthodes de construction de modèles, mais généralement dans les analyses PK/PD, il s'agit d'un processus itératif en deux étapes, où l'on construit d'abord le modèle de base sans covariables, puis on inclut pas à pas les covariables dans le modèle. Pour la première étape, différents modèles structuraux sont testés en utilisant les tests ou critères établis plus haut, puis les modèles de variabilité sont précisés. Pour la deuxième étape, différentes méthodes sont possibles. La première consiste à effectuer une série d'analyses univariées pour évaluer les relations paramètres-covariables en choisissant un seuil de significativité assez élevé, puis à inclure les relations significatives par une méthode pas-à-pas ascendante, descendante, ou combinée dans un modèle complet. Une autre méthode construit le modèle à partir des EBE, soit pas-à-pas, soit en utilisant des modèles additifs généralisés (GAM), avant d'inclure les relations trouvées dans le modèle final [48]. Ces méthodes ont de nombreuses variantes, selon les tests utilisés, la prise en compte de collinéarité entre les covariables, ou l'intégration de relations *a priori* comme les relations allométriques (reliant les paramètres PK au poids du sujet). La pertinence clinique peut également être un critère de sélection, certains auteurs proposant de ne tester dans les modèles que des relations *a priori* plausibles.

D'autres algorithmes ont été proposés notamment pour la construction du modèle de covariables, comme les algorithmes génétiques [49] ou la construction d'un modèle comprenant toutes les covariables cliniquement pertinentes après élimination des collinéarités [50].

Application. Une démarche similaire peut être réalisée pour l'exemple de la warfarine. Ainsi, pour obtenir le modèle structural, nous avons testé différents types de modèles PK et PD, puis nous avons affiné la description des modèles statistiques. Le

but de cet article n'étant toutefois pas de présenter une analyse PK/PD, nous nous sommes arrêtées au modèle proposé en section 4.6, mais le lecteur intéressé peut reprendre l'exemple de démonstration et compléter ce travail en affinant la description de la variabilité inter-sujet et en considérant d'autres covariables dans le modèle.

5.3 Outils diagnostiques

Evaluer un modèle consiste à examiner ses caractéristiques, sa capacité à reproduire les données observées ainsi qu'à en prédire de nouvelles. Bien que son importance soit soulignée notamment par les instances de régulation [51, 52], l'évaluation est rarement bien faite et souvent omise des résultats publiés [53]. Une des difficultés, comme l'indique la citation bien connue de Box [54], « all models are wrong, but some are useful »(tous les modèles sont faux, mais certains sont utiles), est que nous ne pouvons jamais vraiment valider un modèle, mais seulement en évaluer les caractéristiques pour déterminer s'il est adapté à l'usage que nous souhaitons en faire, comme par exemple les utiliser pour le suivi thérapeutique et l'ajustement posologique.

Notons M_B le modèle à évaluer, où B dénote le jeu de données utilisé pour le développer, et supposons que nous voulons évaluer M_B sur un jeu de données V, qui peut être le même que B (évaluation interne) ou différent (évaluation externe). Nous noterons H_0 l'hypothèse nulle que M_B décrit V de façon adéquate.

Graphes. Les graphes sont très utiles pour visualiser la qualité de l'ajustement, et suggérer des améliorations possibles du modèle [55, 56]. Le premier aspect qui nous intéresse est la capacité du modèle à reproduire les données observées, et nous allons donc utiliser les prédictions du modèle. On utilise pour les graphes diagnostiques deux types de prédictions :

- prédictions dites de population, obtenues en utilisant les paramètres de population ; ces paramètres ne tiennent pas compte des données observées chez le sujet, mais uniquement des covariables et du protocole expérimental propre à ce sujet
- prédictions dites individuelles, obtenues en utilisant les paramètres individuels et donc en tenant compte de toute l'information disponible chez le sujet ; ce sont les « meilleures »prédictions qu'il est possible d'obtenir à partir des données

Les prédictions individuelles ont été utilisées dans les graphes en figure 4 pour évaluer l'ajustement chez chaque sujet. Il est possible de superposer également les prédictions de population, représentatives d'un comportement médian du modèle, sur le même graphe.

En remontant au niveau de la population, on peut comparer l'ensemble des prédictions aux observations. Dans la figure 5, nous montrons les prédictions en fonction des

observations pour les deux réponses, afin d'évaluer la qualité du modèle structural et de détecter les biais éventuels. Les deux graphes du haut utilisent les prédictions de population, on s'attend à voir se répartir les points de part et d'autre de la bissectrice, avec un étalement reflétant la variabilité interindividuelle. Dans les deux graphes du bas en revanche, les prédictions individuelles sont utilisées et les points se resserrent autour de la bissectrice, reflétant l'amélioration de la prédiction obtenue en utilisant les observations. Ensemble, ces quatres graphes montrent une bonne adéquation des modèles PK et PD.

Plutôt que de regarder directement les observations, nous pouvons nous intéresser aux résidus. Dans les MNLEM, les résidus bruts (observations-prédictions) n'incluent ni la variabilité interindividuelle ni la variabilité résiduelle, et c'est pourquoi différents types de résidus ont été définis. Linéariser le modèle permet de définir des résidus appelés WRES (*weighted residuals*) ou IWRES (voir section 4.4), faciles à calculer et rendus par la plupart des logiciels, mais dont la distribution théorique n'est pas connue. Des résidus mieux adaptés aux MNLEM sont les npde, pour *normalised prediction distribution errors* [57, 58]. Les npde sont calculées en simulant la distribution prédictive pour chaque observation puis en prenant le quantile de l'observation dans cette distribution (après décorrélation). Des graphes des npde en fonction du temps ou des prédictions du modèle permettent de détecter les écarts au modèle. La figure 6 montre les npde en fonction du temps et des prédictions, à nouveau pour les deux réponses. Dans ces graphes, ce sont les tendances éventuelles du modèle qui peuvent indiquer des zones de sous ou sur-prédiction ; par ailleurs, les npde doivent se distribuer selon une gaussienne et par conséquent leur amplitude est majoritairement contenue entre -1.96 et +1.96. Ici, aucune tendance particulière n'est observée.

Cette notion de simuler sous le modèle pour comparer aux données observées se retrouve dans un graphe diagnostic très utilisé, les VPC (*Visual Predictive Check*). Les VPC sont obtenus en simulant des vecteurs de paramètres sous les distributions estimées dans le modèle, puis en calculant pour des temps donnés l'enveloppe des prédictions correspondantes. On peut également construire des intervalles de prédiction autour de percentiles donnés, comme la médiane et les percentiles à 5 et 95%, et comparer les percentiles correspondants pour les données effectivement observées. Ces graphes représentent le profil moyen ainsi que la variabilité du processus modélisé, et permettent d'évaluer la capacité du modèle à reproduire ces caractéristiques. Les VPC sont représentés en figure 7, et à nouveau illustrent une bonne capacité du modèle à reproduire les profils des deux réponses, avec peut-être toutefois pour la PK des prédictions légèrement inférieures aux concentrations observées car la médiane reste dans la limite haute de la fourchette.

Ces graphes peuvent être complétés avec d'autres graphes comme ceux des résidus individuels absolus tracés en fonction des concentrations prédites, pour évaluer la qualité du modèle d'erreur. Parmi les autres hypothèses effectuées lors de la modélisation, celles sur les variabilités (distribution des effets aléatoires, explication d'une partie de la variabilité par des covariables), peuvent être évaluées par des graphes spécifiques, cependant en raison du shrinkage (qui fait que la distribution des paramètres individuels s'éloigne de la distribution théorique) et du nombre de sujets assez restreint ces hypothèses peuvent être plus difficiles à vérifier. Les distributions des paramètres ou des effets aléatoires individuels estimés peuvent être tracées sous forme d'histogrammes, un écart trop important à la distribution attendue des paramètres pouvant conduire à changer les hypothèses sur la distribution des paramètres aléatoires.

Les corrélations entre les effets aléatoires peuvent être détectées à l'aide de graphes deux à deux, tandis que des graphes systématiques des relations entre paramètres et covariables peuvent permettre de détecter des relations. Ces graphes sont donnés pour deux paramètres du modèle (CL et C_{50}) dans la figure 8, et nous avons utilisé les estimations individuelles des effets aléatoires parce que le modèle contient déjà deux covariables sur la clairance, or nous cherchons à évaluer les autres relations possibles. La ligne noire horizontale est la ligne des ordonnées, tandis que la ligne pointillée rouge représente un lissage par spline. Ici, les points sont bien répartis sur le graphe, ne suggérant pas de tendance particulière, pas plus que les splines (le fait qu'ils s'écartent de la ligne des ordonnées aux bords traduit le fait que le lissage est influencé par les points au bord mais n'a pas de signification particulière).

Tests. En utilisant la distribution théorique de certaines variables, on peut définir des tests de rejet de H_0, toutefois notons tout de suite qu'il n'y a pas vraiment de test d'adéquation dans les MNLEM. Les WRES ont une distribution gaussienne standard lorsque la fonction f est linéaire, mais cette approximation est mauvaise pour les modèles non-linéaires, conduisant à une augmentation sensible des erreurs de type I du test [59]. Les npde montrent de meilleures propriétés statistiques [57], et plus généralement, il est possible de proposer des tests de type Neyman-Pearson pour les diagnostics utilisant des simulations sous le modèle [60]. Ces approches basées sur des simulations sont regroupées sous la dénomination anglaise de *posterior predictive check* (PPC). Le PPC s'appuie d'une part sur l'idée que, si le modèle est correct, des données simulées sous le modèle devraient ressembler aux données observées, et d'autre part, sur la définition de caractéristiques souhaitables du modèle, comme sa capacité à reproduire un trait particulier des données observées. Le test consiste à comparer la valeur d'une statistique donnée, calculée sur le jeu de données à évaluer, avec sa

distribution sous le modèle candidat. Si cette valeur est trop extrême comparée à la distribution attendue, le modèle est rejeté (test de Neyman-Pearson) [61].

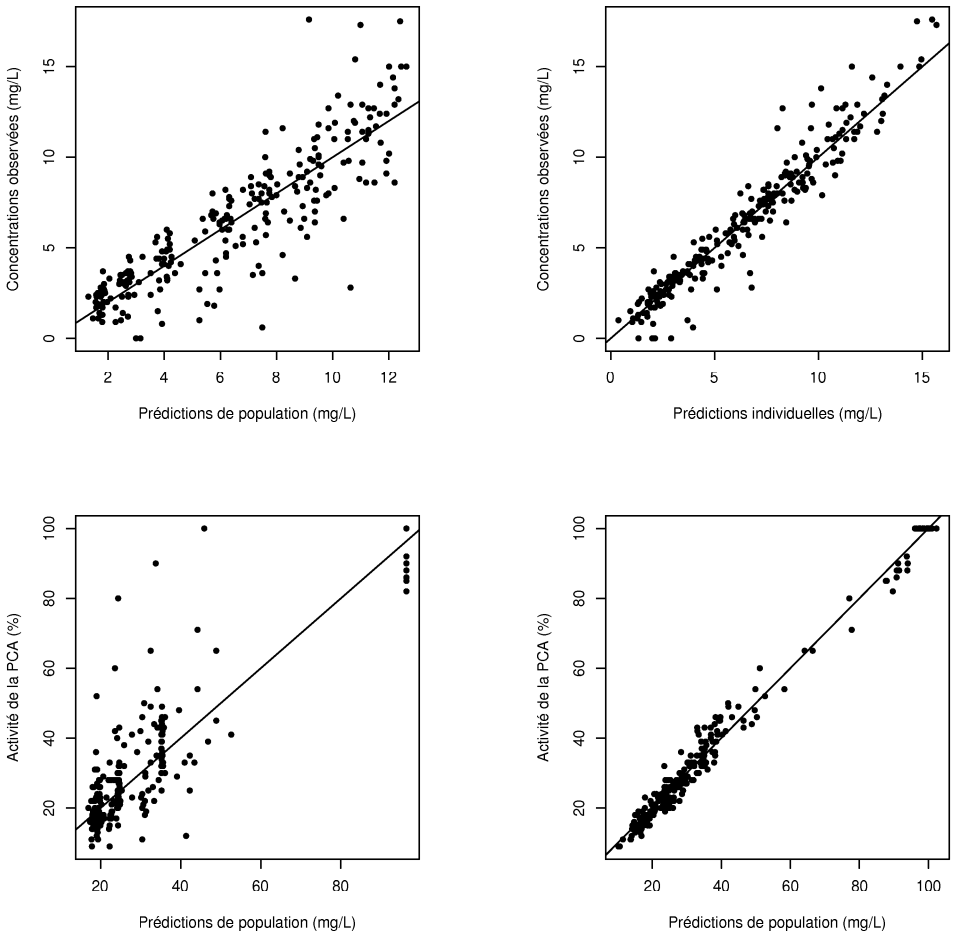

FIGURE 5 – Graphes de diagnostic montrant les observations en fonction des prédictions de population (gauche) et des prédictions individuelles (droite), pour la PK (haut) et la PD (bas).

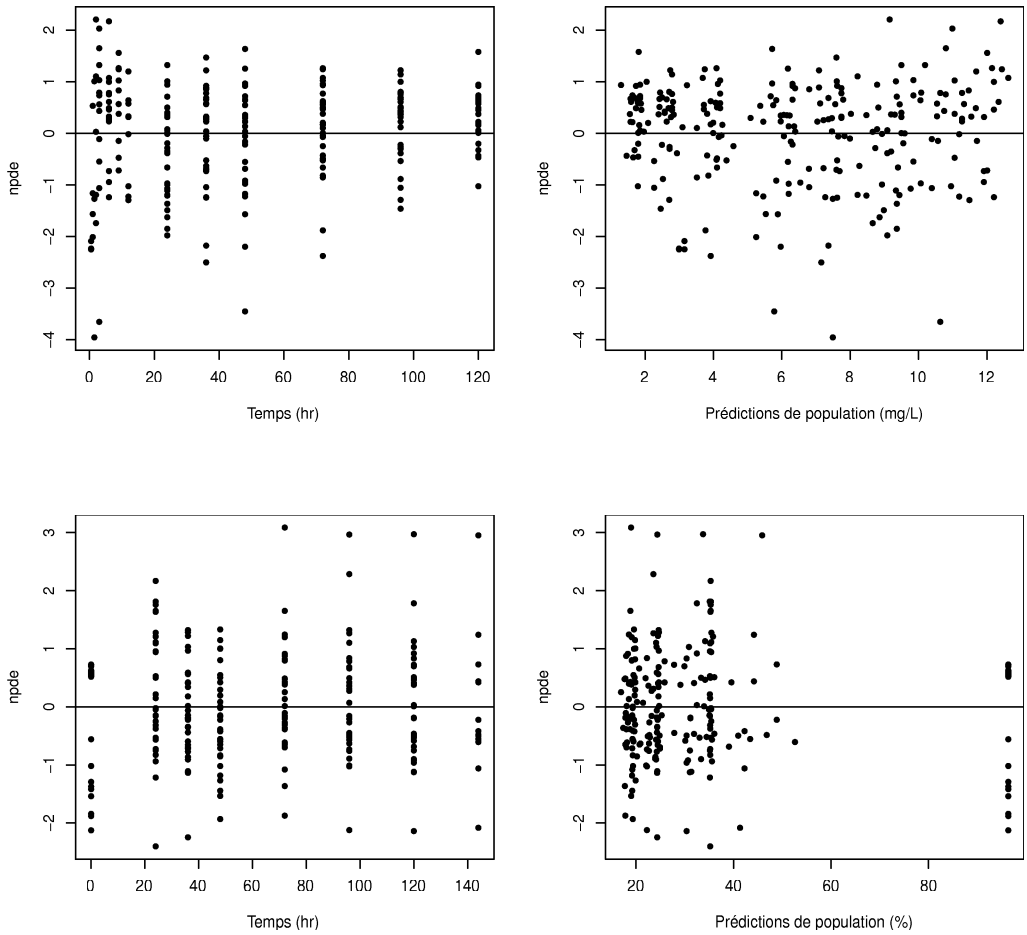

Figure 6 – Résidus (npde) en fonction du temps (gauche) et des prédictions de population (droite) pour le modèle PK (haut) et pour le modèle PD (bas).

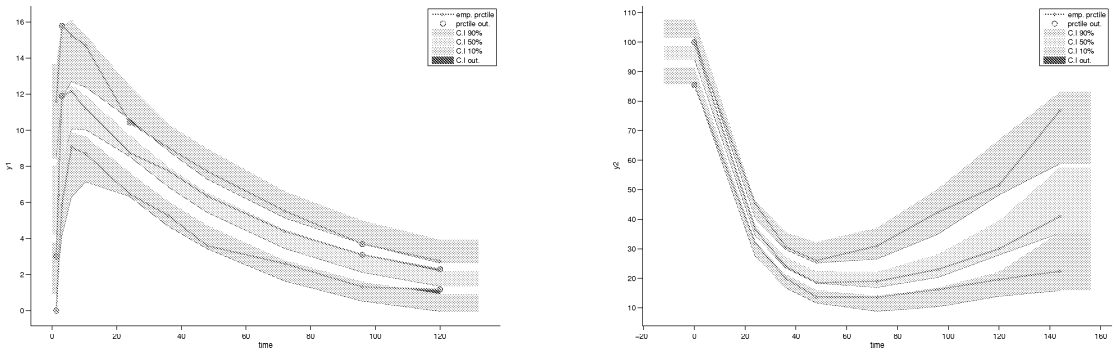

FIGURE 7 – VPC (Visual Predictive Check) pour la PK (gauche) et la PD (droite) de la warfarine.

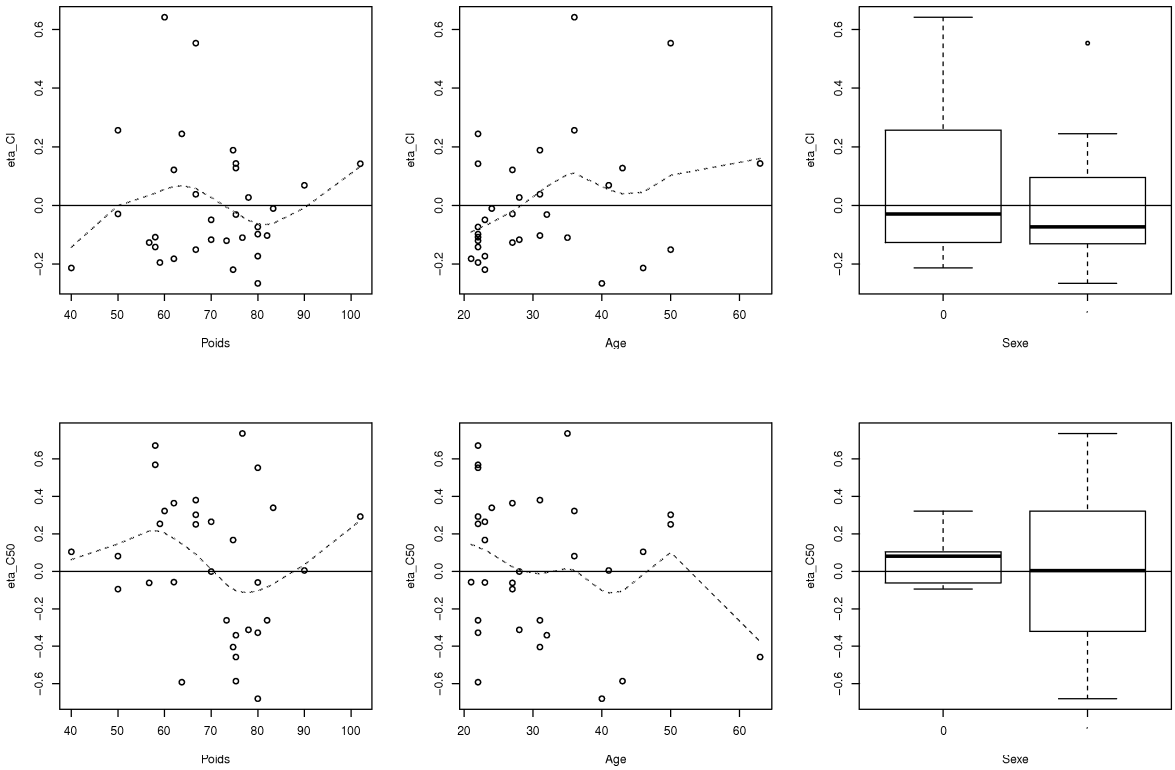

FIGURE 8 – Relation entre les paramètres CL (haut) et C_{50} (bas) et les 3 covariables du jeu de données, poids (kg), âge (an) et sexe (0 : hommes, 1 : femmes).

6 Protocoles

Le développement de nouvelles méthodes d'estimation des MNLEM et leur mise en œuvre dans des logiciels facilite aujourd'hui l'analyse de données longitudinales en pharmacologie. Préalable à toute analyse, le choix du protocole expérimental reste une problématique fondamentale et complexe. Dans cette partie, nous reviendrons sur la définition d'un protocole expérimental, puis nous introduirons la méthodologie actuelle pour l'évaluation et l'optimisation de protocoles. Enfin, nous décrirons les outils informatiques existants et nous déterminerons avec l'un entre eux de nouveaux protocoles sur l'exemple sur les données de warfarine.

6.1 Protocoles de prélèvements

Le protocole de population Ξ est composé des N protocoles dits élémentaires $\xi_i (i = 1...N)$ réalisés chez chacun des sujets de l'étude. Un protocole élémentaire ξ_i est défini par un nombre de prélèvements n_i ainsi que par les temps auxquels ils ont été effectués : $\xi_i = \left\{t_{i1}, t_{i2}, \ldots, t_{i_{N1}}\right\}$. En regroupant les sujets ayant le même protocole élémentaire, on obtient Q groupes et les individus étant indépendants, la matrice d'information de Fisher de la population se décompose comme la somme de N matrices d'information de Fisher élémentaires ou en sommant sur les groupes :

$$M_F(\theta, \Xi) = \sum_{i=1}^{N} M_F(\theta, \xi_i) = \sum_{q=1}^{Q} N_q M_F(\theta, \xi_q) \tag{22}$$

Déterminer un protocole de population implique d'une part d'identifier les temps à allouer pour chaque sujet et d'autre part de définir la structure du protocole de population (nombre de groupes, nombre de prélèvements par groupe et proportion de sujets associée à chacun des groupes), en tenant compte de contraintes (typiquement, un nombre total de prélèvements n fixé). Tous ces éléments sont importants et influencent fortement la précision d'estimation des paramètres, conduisant parfois à des études non-interprétables [62, 63]. Ces choix sont d'autant plus essentiels lorsque les études sont conduites dans des populations particulières telles que les enfants, les personnes âgées ou fragiles chez qui un nombre limité de sujets et de prélèvements par sujet peut être obtenu. Un "bon" protocole peut donc améliorer l'efficacité d'une étude en améliorant l'estimation des paramètres et en la rendant plus éthique et/ou moins coûteuse.

La principale approche pour évaluer les protocoles a longtemps été basée sur des simulations, mais c'est une méthode coûteuse en temps et limitée en nombre de protocoles évaluables. Basée sur la théorie des protocoles utilisée en régression non linéaire

classique [64, 65], une approche générale pour optimiser un protocole consiste à minimiser la matrice de variance covariance d'estimation des paramètres, les éléments diagonaux de la matrice d'information de Fisher étant les erreurs standards attendues des paramètres.

6.2 Calcul de la matrice d'information de Fisher(M_F)

Comme pour la vraisemblance, le calcul de $M_F(\theta, \Xi)$ (équation 15) implique des intégrales, et différentes méthodes ont été proposées pour en obtenir une approximation.

Approche par linéarisation au premier ordre : une première approximation de l'expression de la matrice d'information a été proposée par [66]. Cette expression est basée sur une linéarisation au premier ordre du modèle f par un développement de Taylor autour de l'espérance des effets aléatoires, comme dans la méthode d'estimation FO présentée en section 4.1.1. Cette approximation permet d'obtenir un modèle linéaire par rapport aux effets aléatoires et conduit à une expression analytique de la matrice d'information de Fisher [67, 68]. Dans l'expression la plus couramment utilisée, la dépendance de la variance des observations aux paramètres moyens n'est pas prise en compte, et M_F est exprimée sous une forme bloc diagonale. L'expression de la matrice complète est plus complexe du fait de l'introduction des dérivées secondes du modèle [69]. Il n'y a pas de consensus clair sur le choix de la meilleure approximation. Des travaux de comparaison des deux méthodes ont montré que l'approche utilisant une M_F diagonale par blocs était plus fiable que celle avec une M_F complète dans certains modèles [70]. Basée sur cette approche par linéarisation, des expressions de la matrice d'information de Fisher dans des cas très généraux ont été proposées (variabilité inter-occasion, prédiction de la puissance du test de Wald,...) [69, 71]. Des approximations plus poussées inspirées de l'algorithme FOCE ont été également développées [69, 72] donnant des résulats plutôt proches de ceux obtenus avec l'approche FO mais pour un temps de calcul beaucoup plus élevé.

Alternatives à la linéarisation : plusieurs études ont mis en avant les limites de la linéarisation quand la non-linéarité des modèles augmente et/ou les variances des effets aléatoires sont très élevées [73, 74]. Plusieurs approches ont été proposées, basées sur des simulations [75] ou sur la quadrature de Gauss [76]. Ces approches donnent des prédictions correctes mais sont coûteuses en temps de calcul lorsque le nombre d'effets aléatoires devient important. Des alternatives basées sur une approche stochastique ou

sur des AGQ [14] sont à explorer, laissant un champ de recherche important sur cette thématique complexe.

6.3 Critères et algorithmes d'optimisation

Critères d'évaluation et d'optimisation de protocoles : l'expression de M_F dans les MNLEM dépend du protocole mais également du vecteur des paramètres de population θ. Nous cherchons à estimer le plus précisement possible ce vecteur *i.e.* minimiser les variances d'estimation de ses éléments, ce qui revient à "maximiser" M_F. Comme il n'existe pas de relations d'ordre sur les matrices, différents critères d'optimisation ont été développés pour les MNLEM, étendus de la théorie générale d'optimisation des protocoles en régression non linéaire classique [77, 64]. Les plus classiques sont le critère de D-optimalité, qui maximise le déterminant de M_F (le plus utilisé en pharmacométrie) ; le critère de A-optimalité qui minimise la somme des variances asymptotiques (la trace de l'inverse de la matrice) ; la C-optimalité qui minimise la somme des carrés des coefficients de variation des estimations et le critère de E-optimalité qui maximise la plus petite valeur propre de M_F. Pour calculer tous ces critères, il est nécessaire d'avoir des connaissances *a priori* sur le modèle et le vecteur des paramètres, ce qui peut sembler paradoxal puisque l'expérience envisagée a pour but de les estimer. Lorsque ces connaissances sont basées sur des résultats d'études précédentes, on parle de "planification locale" [77, 64]. L'incertitude sur ces connaissances *a priori* est une limite évidente de la planification locale : si les valeurs des paramètres sont différentes de celles supposées *a priori*, la planification locale peut conduire à un protocole peu optimal.

Algorithme d'optimisation de protocoles : de nombreux algorithmes ont été développés, conduisant soit à l'optimisation d'un protocole dit entier, soit à l'optimisation d'un protocole dit statistique (ou continu). Un protocole dit entier comprend Q groupes de protocoles élémentaires $\Xi = \left\{[\xi_1, N_1], \ldots, [\xi_Q, N_Q]\right\}$ (où $N = \sum_{q=1}^{Q} N_q$ et $n = \sum_{q=1}^{Q} n_q N_q$), et on peut attribuer à chaque groupe un poids rationnel $m_q = \frac{N_q}{N}$, correspondant à la proportions de sujets m_q ayant le même protocole élémentaire ξ_q tel que $\sum_{q=1}^{Q} m_q$. Ξ peut alors s'écrire sous la forme $\Xi = [\xi_1, m_1]; \ldots ; [\xi_Q, m_Q]$. On peut étendre ce protocole statistique à des masses m_q quelconques comprises entre 0 et 1 et vérifiant $\sum_{q=1}^{Q} m_q$, et obtenir un protocole dit statistique. Ce protocole est rarement réalisable en pratique puisque nous n'avons pas un nombre $m_q \times N$ entier. Il faut alors se ramener à un protocole entier en approchant les proportions à l'entier supérieur le plus proche. Ce passage peut engendrer une perte d'efficacité (*i.e* quantité d'information), mais qui en général est très faible.

Un algorithme adapté à l'optimisation de protocoles statistiques est l'algorithme du Fedorov-Wynn, qui converge vers un procotole optimal quelque soit le protocole initial [77, 78]. Il a été developpé initialement dans un contexte de regression non-linéaire classique [79] puis étendu aux MNLEM par [66] pour la contruction de protocoles D-optimaux. C'est un algorithme itératif qui permet de trouver la meilleure combinaison de protocoles élémentaires ξ parmi l'ensemble de tous les protocoles possibles, afin de maximiser le déterminant de M_F, pour un nombre total de prélèvements fixe. Cet algorithme optimise la structure du protocole de population optimal Ξ.

Pour l'optimisation de protocoles entiers, des algorithmes spécifiques ont été développés, qui dans le cas de protocoles D-optimaux, sont des algorithmes d'échanges. Leur principe est de partir d'un protocole initial et de l'améliorer successivement en échangeant un temps de prélèvement par un ou plusieurs meilleurs temps [77, 78].

La détermination de protocoles optimaux entiers ou statistiques peut se faire aussi par l'utilisation d'algorithmes généraux, comme celui du gradient, de Newton ou du Simplex [80], basés sur les dérivées respectivement première et secondes du critère. Ces algorithmes sont des méthodes d'optimisation locale et ne permettent de trouver un optimum que dans le voisinage du point initial, ce qui peut être problématique dans le cas de d'optima locaux multiples. L'algorithme de recherche aléatoire ou le *Simulated Annealing* sont en revanche des algorithmes d'optimisation globale plus adaptés à la recherche de protocoles optimaux, particulièrement pour les protocoles plus complexes [81].

L'introduction d'intervalles ou de fenêtres de temps dans l'étape d'optimisation de protocole a été proposée comme une solution afin de trouver des temps de prélèvements optimaux réalisables en pratique. Le protocole est alors non plus défini par des temps de prélèvements discrets, mais par des intervalles de temps, permettant ainsi une meilleure flexibilité en pratique. Diverses approches ont été proposées dans ce contexte [82].

6.4 Logiciels

Cinq logiciels sont aujourd'hui disponibles pour l'évaluation et l'optimisation de protocoles de population : PFIM , PKStamp , PopDes, PopED and POPT. Quatre d'entre eux ont été developpés par des équipes de recherche du monde académique. PFIM est le seul outil utilisant le logiciel gratuit R [30], les autres logiciels ont été developpés sous l'environnement Matlab [83].

Ces dernières années, tous ces logiciels ont évolué en incluant différentes spécificités en termes d'implémentation du modèle et de calcul de M_F. La table 4 décrit les caractéristiques de chaque logiciel dans le cadre de l'évaluation d'un protocole, notam-

mant les différents critères et algorithmes d'optimisation, et les approximations pour le développement de M_F.

Ces logiciels ont été comparés sur deux exemples, l'un de PK simple et l'autre sur un modèle PK/PD. En utilisant la même approximation de M_F, tous ces logiciels présentaient des performances similaires sur ces deux exemples [84].

	Logiciels				
	PFIM	PkStaMp	PopDes	PopED	POPT
Langage	R	Matlab	Matlab	Matlab	Matlab
Disponible à partir d'un site web	✓		✓	✓	✓
Interface	✓	✓	✓	✓	
Librairies de modèles	✓	✓	✓	✓	✓
Ecriture de son propre modèle	✓	✓	✓	✓	✓
Modèles avec plusieurs réponses	✓	✓	✓	✓	✓
Temps différents entre les réponses	✓	✓	✓	✓	✓
Modèle sous la forme ODE	✓	✓	✓	✓	✓
M_F complète	✓	✓	✓	✓	
Variabilité inter-occasion	✓		✓	✓	
Covariables discrètes / Puissance	✓/✓		✓/	✓/✓	✓/

Table 4 – Caractéristiques des logiciels disponibles pour l'évaluation et l'optimisation de protocoles de population.

6.5 Application à l'exemple

Pour l'application aux données de l'exemple, nous avons choisi le logiciel PFIM [85]. Nous avons utilisé le modèle à deux réponses, PK/PD, présenté dans la section 4.6 et les valeurs de paramètres données dans la Table 3.

Nous avons évalué dans un premier temps le protocole utilisé dans l'étude initiale, le protocole empirique. Noté Ξ_{emp} il est défini par un groupe de 32 sujets, ayant reçus une dose de 100 mg de warfarine, avec 13 prélèvements pour les mesures de concentrations de la warfarine et 8 mesures de l'activité de la prothrombine. Ξ_{emp} est donc composé d'un total de 672 mesures. Nous avons ensuite optimisé un protocole de population grâce à l'algorithme du FedorovWynn disponible dans PFIM. Ce protocole, noté $\Xi_{opt.iden}$, devait répondre aux contraintes suivantes : i) inclure seulement cinq temps de prélèvements par protocole élémentaire pour les mesures de PK et de PD (nombre minimum de prélèvements pour l'estimation des paramètres du modèle) ii) un nombre total de sujets identique à celui du protocole empirique *i.e* 32 sujets iii) des

temps de prélèvements parmi ceux du protocole empirique.

Les temps de prélèvements du protocole optimal ainsi que le critère D-optimal ($det(M_F)^{\frac{1}{n_\theta}}$) associé, où n_θ est le nombre total de paramètres du modèle, sont reportés dans la Table 5. La figure 9 représente la sortie produite par le logiciel PFIM suite à l'optimisation du protocole $\Xi_{opt.iden}$. $\Xi_{opt.iden}$ comprend trois protocoles élémentaires alors que Ξ_{emp} n'en comprend qu'un seul. L'algorithme du Fedorov-Wynn optimisant des proportions de sujets, ces dernières ont été arrondies à l'entier le plus proche.

Certains protocoles élémentaires sont associés à très peu de sujets. Pour faciliter la réalisation pratique, $\Xi_{opt.ident}$ a été simplifié en un procotole $\Xi_{opt.ident.simp}$, composé d'un seul groupe de 32 sujets. $\Xi_{opt.ident.simp}$ a un nombre total de prélèvements inférieur à celui du protocole empirique et donc un critère plus petit soulignant une plus faible efficacité du protocole. En augmentant le nombre de sujets pour nous ramener à une efficacité similaire, nous obtenons le protocole $\Xi_{opt.ident.simp.d}$ (45 sujets pour un critère de 273.1, très proche de la valeur de 276.6 obtenue pour Ξ_{emp}).

Les erreurs standards relatives des paramètres obtenues par ce dernier protocole et par le protocole empirique sont présentées sur la Figure 10. Pour le protocole empirique les erreurs standards relatives prédites sont inférieures ou de l'ordre de 30%. Les précisions d'estimations prédites sont très proches entre les deux protocoles, bien que $\Xi_{opt.ident.simp}$ ne requiert que 450 prélèvements au total contre 672 pour le protocole empirique. Cette réduction du nombre de prélèvements pour une efficacité similaire, donne un protocole éthiquement plus correct et moins coûteux.

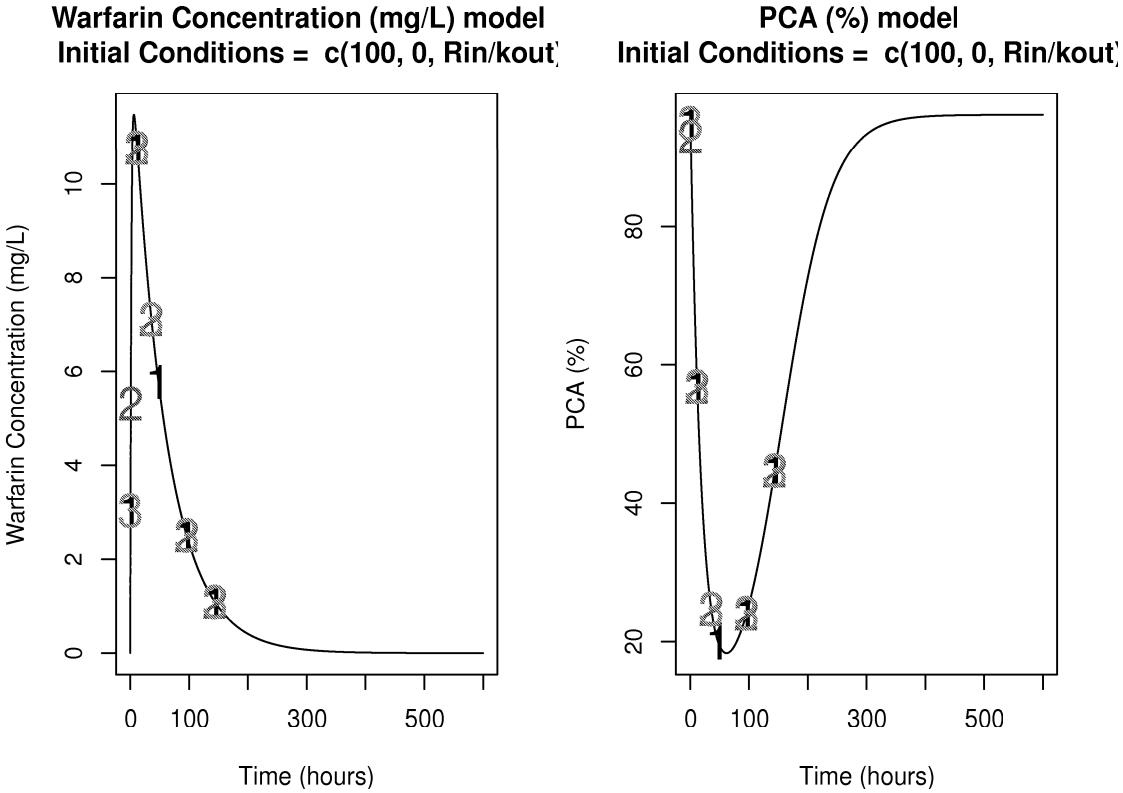

FIGURE 9 – Sortie R du logiciel PFIM 3.2.2 après l'optimisation du protocole $\Xi_{opt.iden}$. Ce graphique représente les profils des concentrations de warfarine et de la PCA en fonction du temps pour les paramètres de population du modèle présenté en section 4.6. Les temps de prélèvements obtenus avec le protocole optimal$\Xi_{opt.iden}$ sont représentés par les numéros 1, 2 et 3 pour les prèlevements du respectivement premier, second et troisième protocole élémentaire.

	N_{prot}	Temps de prélèvements pour chaque réponse (heures)		N_{suj}	N_{tot}	Critère
		PK	PD			
Ξ_{emp}	1	0.5,1,2,3,6,9,12,24,36,48,72,96,120	0,24, 36, 48,72, 96, 120,144	32	672	276.6
$\Xi_{opt.iden}$	3	0.5, 12, 48, 96, 144	0.5, 12, 48, 96, 144	16	320	195.8
		1, 12, 36, 96, 144	1, 12, 36, 96, 144	8		
		0.5, 12, 36, 96, 144	0.5, 12, 36, 96, 144	8		
$\Xi_{opt.iden.simp}$	1	0.5, 12, 48, 96, 144	0.5, 12, 48, 96, 144	32	320	194.6

N_{prot} : nombre de protocoles élémentaires composant le protocole de population
N_{suj} : nombre de sujets (arrondi)
N_{tot} : nombre total de prélèvements

Table 5 – Protocoles optimaux pour le modèle conjoint des concentrations de Warfarine et de la PCA, obtenus sous différentes contraintes.

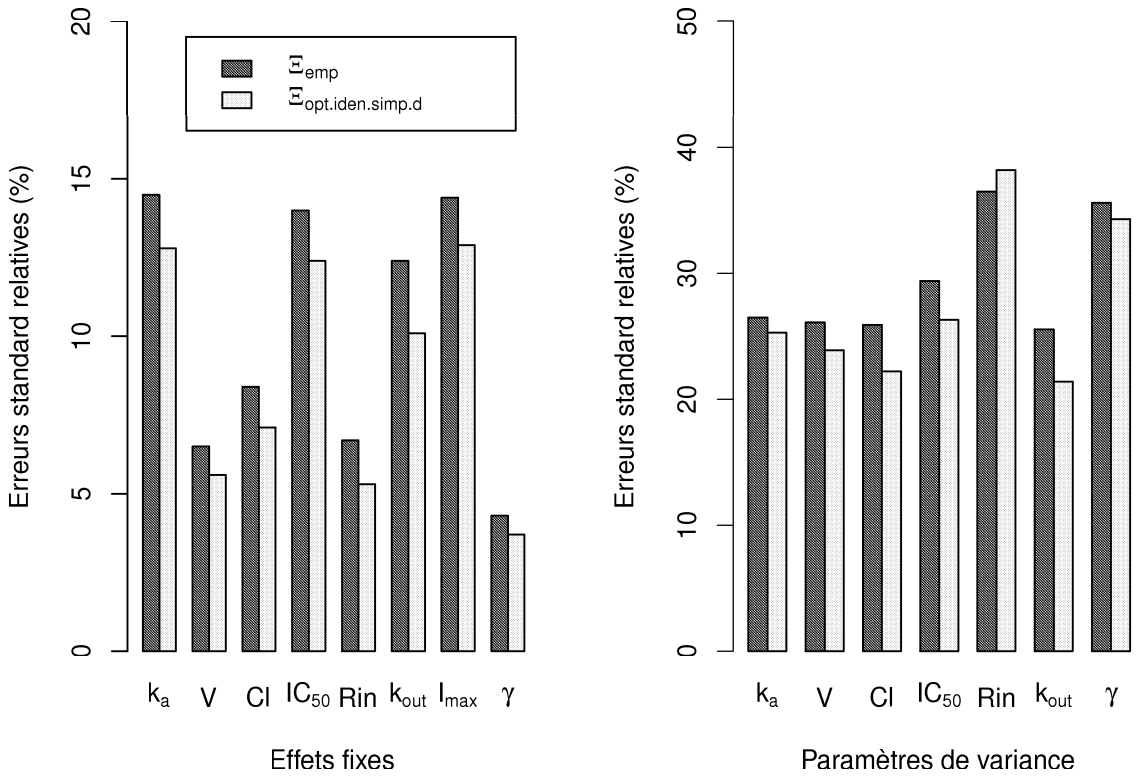

FIGURE 10 – Erreurs standards relatives prédites par le logiciel PFIM pour les protocoles Ξ_{emp} (N=32 sujets) et $\Xi_{opt.ident.simp.d}$ (N=45 sujets).

7 Conclusion et perspectives

Le développement de la pharmacométrie comme discipline scientifique, ainsi que sa diffusion progressive dans la vie du médicament depuis son développement jusqu'au lit du malade, doit beaucoup au travail de Lewis Sheiner (voir notamment [86] pour une liste de travaux et de publications). En tant que statisticien, il a contribué à la fois au développement des méthodes statistiques et à leur mise en œuvre dans le logiciel NONMEM [27], avec, en tant que médecin, la volonté d'améliorer la prescription médicamenteuse et le bon usage des médicaments [10]. Il a également cherché à diffuser la méthodologie au travers de formations et de cours. Lors des deux dernières décades, les méthodes d'estimation se sont diversifiées et raffinées, et la pharmacométrie est maintenant couramment utilisée pour les analyses PK/PD dans les dossiers d'autorisation de mise sur le marché [87]. Elle est également présente pour guider

le développement d'un médicament, notamment à travers l'optimisation de protocole et la simulation d'essais cliniques [88], menant au concept de *Model-Based drug development* [89]. Dans certaines pathologies, comme l'anesthésie où l'administration des médicaments est très précisément calibrée en fonction des concentrations cibles, ou dans l'administration de médicaments anti-suppresseurs après transplantation [90], la modélisation est aussi utilisée au quotidien dans la routine clinique et constitue un élément essentiel d'une future médecine personnalisée.

Les modèles PK et PD discutés dans cet article sont structurellement simples mais puissants dans la diversité des profils qu'ils peuvent capturer, et permettent d'introduire en douceur la modélisation. Ces modèles peuvent être rendus plus complexes pour se rapprocher de la physiologie ; à l'extrême on trouve les modèles dits **PBPK**, qui font correspondre un compartiment à chaque organe avec des types et des constantes de transferts, et tiennent compte de la répartition des enzymes, ligands et transporteurs constituant ces différents organes. Combinant recherches exhaustives de la littérature, expériences *in vitro* et données cliniques [91], ces modèles sont utilisés pour étudier finement des systèmes complexes et prédire la réponse du système à des modifications de voie d'administration, à des changements d'échelle (extrapolation inter-espèces), ou à des modifications physiologiques (changements dans les processus de maturation durant l'enfance). Dans les analyses pédiatriques par des fonctions PK moins sophistiquées, on utilise des fonctions allométriques représentant la relation entre le poids et le volume de distribution et la clairance [92, 93]. CDs corrections sont aussi employées quand l'écart entre les poids dans la population à l'étude est grand, comme chez les sujets obèses. De même, il existe de nombreux modèles PD plus complexes que ceux présentés dans cet article, tels que le modèle de la dynamique des virus de l'immunodéficience humaine et de l'hépatite C [94], de la croissance tumorale [95] ou de l'équilibre du glucose sanguin [96].

En ce qui concerne les modèles statistiques, nous avons évoqué en section 3.1.2 des modèles pour réponses non-continues (comptages, scores) qui peuvent être inclus dans les MNLEM, en considérant que chaque individu est caractérisé par une distribution de probabilité dont les moments diffèrent de la distribution moyenne par autant d'effets aléatoires [97]. Une alternative est de modéliser l'intervalle de temps au bout duquel la réponse est observée, chaque individu ayant alors sa propre fonction de survie [98]. Des modèles à variable latente peuvent être introduits pour modifier le comportement du système en réponse à des variables non observées (état du patient, stade de la maladie) [99], et l'évolution de la maladie est de plus en plus souvent modélisée en propre dans les maladies chroniques ("disease models") [100]. Les méthodes d'estimation et de test peuvent être étendues à ces cas plus complexes, avec des outils

diagnostics principalement basés sur des simulations de type VPC [101].

Sur un plan méthodologique, il existe de nombreuses variantes des méthodes d'estimation que nous n'avons pas citées faute de place. Précisons simplement que les hypothèses effectuées *a priori* sur les distributions des effets aléatoires, tant dans les méthodes bayésiennes que fréquentistes, peuvent être relaxées et donnent lieu aux méthodes dites non-paramétriques (voir par exemple [102, 103] pour des algorithmes fréquentistes). Ces méthodes sont moins utilisées et disponibles seulement dans certains logiciels, mais peuvent permettre d'explorer les distributions des paramètres avec moins de contraintes que les méthodes paramétriques.

Une limite très souvent rencontrée dans les études en pharmacométrie, est la complexité des modèles (ajoutée à la non-linéarité) car elle peut faire naître des problèmes d'identifiabilité [65, 104]. L'identifiabilité structurelle définit l'aspect théorique de l'identifiabilité d'un modèle, permettant de déterminer le nombre de solutions admissibles, en fonction du système d'équations définissant le modèle, et des mesures effectuées. Il est parfois nécessaire de fixer la valeur de certains paramètres du modèle ou de reparamétriser le modèle afin de rendre les autres paramètres identifiables. L'identifiabilité pratique est complémentaire de l'identifiabilité structurelle, et évalue si les paramètres sont identifiables pour un protocole donné. C'est une notion relative qui dépend de l'information disponible, qualitativement et quantitativement. Ainsi le modèle présenté en équation (2) est structurellement identifiable, mais son identifiabilité pratique dépend de la disponibilité de données urinaires ou de données par voie intraveineuse et orale chez les mêmes sujets pour estimer la biodisponibilité F.

Une limite à l'optimisation de protocoles pour les MNLEM est la nécessité d'avoir de l'information *a priori*. La planification globale ou robuste est une alternative possible qui se base sur la distribution *a priori* des paramètres avec des critères différents : ED, EID,... [105, 106]. Une autre alternative est la planification adaptative, largement utilisée dans les essais cliniques ou les études de recherche de dose [107], mais étudiée seulement depuis peu dans les MNLEM [108, 109]. Les protocoles adaptifs permettent une certaine flexibilité au cours de l'étude car ils autorisent des changements dans le protocole, guidés par l'analyse de données recueillies à certains temps intermédiaires de l'essai clinique.

Nous avons présenté dans cet article une vue d'ensemble et très abrégée de la pharmacométrie, une discipline en plein essor tant dans la théorie que dans les applications pratiques et les mises en œuvre logicielles et logistiques. Le lecteur intéressé est invité à se reporter à des ouvrages plus complets, notamment [1, 7, 9, 65, 110, 111].

Références

[1] Ette E. et Williams P., *Pharmacometrics : The Science of Quantitative Pharmacology*. Wiley-Interscience, New-York, USA (2007).

[2] O'Reilly R. et Aggeler P., *Studies on the coumarin anticoagulant drugs. Initiation of warfarin therapy without a loading dose*. Circulation 38, 169–177 (1968).

[3] Holford N.H., *Clinical pharmacokinetics and pharmacodynamics of warfarin. Understanding the dose-effect relationship*. Clin. Pharmacokinet. 11, 483–504 (1963).

[4] Stryer L., *Biochemistry (Third Edition)*. NY : W.H. Freeman and Company, New York, USA (1988).

[5] Kutta W., *Beitrag zur naeherungsweisen integration totaler differentialgleichungen*. Zeit. ang. Math. Physik 46, 435–453 (1901).

[6] Savic R.M., Jonker D.M., Kerbusch T. et Karlsson M.O., *Implementation of a transit compartment model for describing drug absorption in pharmacokinetic studies*. J. Pharmacokinet. Pharmacodyn. 34, 711–726 (2007).

[7] Gabrielsson J. et Weiner D., *Pharmacokinetic and Pharmacodynamic Data Analysis : Concepts and Applications*. Apotekarsocieteten, Stockholm, Sweden (1999).

[8] Vu T.C., Nutt J.G. et Holford N.H., *Progression of motor and nonmotor features of Parkinson's disease and their response to treatment*. Br. J. Clin. Pharmacol. 74, 267–283 (2012).

[9] Davidian M. et Giltinan D., *Nonlinear models for repeated measurement data*. Chapman & Hall, London, UK (1995).

[10] Sheiner L., Rosenberg B. et Melmon K., *Modelling of individual pharmacokinetics for computer aided drug dosage*. Comput. Biomed. Res. 5, 441–59 (1972).

[11] Lindstrom M. et Bates D., *Nonlinear mixed effects models for repeated measures data*. Biometrics 46, 673–87 (1990).

[12] Lemenuel-Diot A., Mallet A., Laveille C. et Bruno R., *Estimating heterogeneity in random effects models for longitudinal data*. Biom. J. 47, 329–45 (2005).

[13] Davidian M. et Gallant A., *The nonlinear mixed effects model with a smooth random effects density*. Biometrika 80, 475–488 (1993).

[14] Pinheiro J. et Bates D., *Approximations to the log-likelihood function in the nonlinear mixed-effect models*. J. Comput. Graph. Stat. 4, 12–35 (1995).

[15] Wolfinger R., *Laplace's approximation for nonlinear mixed models*. Biometrika 80, 791–5 (1993).

[16] Kuhn E. et Lavielle M., *Coupling a stochastic approximation version of EM with a MCMC procedure*. Eur. Series Appl. Indus. Math. : P&S 8, 115–131 (2004).

[17] Robert C. et Casella G., *Monte Carlo statistical methods, 2nd edition*. Springer-Verlag, Springer Texts in Statistics, Paris, France (2004).

[18] Dempster A.P., Laird N.M. et Rubin D.B., *Maximum likelihood from incomplete data via the EM algorithm*. JRSSB 39, 1–38. With discussion (1977).

[19] Mentré F. et Gomeni R., *A two-step algorithm for estimation on non-linear mixed-effects with an evaluation in population pharmacokinetics*. J. Biopharm. Stat. 5, 141–58 (1995).

[20] Wang J., *EM algorithms for nonlinear mixed effects models*. Comput. Stat. Data Anal. 51, 3244–56 (2007).

[21] Wei G. et Tanner M., *Calculating the content and boundary of the highest posterior density region via data augmentation*. Biometrika 77, 649–52 (1990).

[22] Leary B., Jelliffe R., Schumitzky A. et Port R., *Accurate maximum likelihood estimation for parametric population analysis*. 13th Annual Meeting of the PAGE, Uppsala, Sweden Abstr 491 (2004).

[23] Delyon B., Lavielle M. et Moulines E., *Convergence of a stochastic approximation version of the EM algorithm*. Ann. Stat. 27, 94–128 (1999).

[24] Kuhn E. et Lavielle M., *Maximum likelihood estimation in nonlinear mixed effects models*. Comput. Stat. Data Anal. 49, 1020–1038 (2005).

[25] Wakefield J., *An expected loss approach to the design of dosage regimens via sampling–based methods*. The Statistician 43, 13–29 (1994).

[26] Karlsson M. et Savic R., *Diagnosing model diagnostics*. Clin. Pharmacol. Ther. 82, 17–20 (2007).

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

[28] Sheiner L., *Computer-aided long-term anticoagulation therapy*. Comput. Biomed. Res. 2, 507–18 (1969).

[29] Lavielle M., *MONOLIX (MOdèles NOn LInéaires à effets miXtes)*. MONOLIX group, Orsay, France (2005).

[30] R Development Core Team, *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0 (2006).

[31] SAS Institute Inc, Cary, NC, USA, *SAS/STAT User's Guide, Version 9, Edition* (2000).

[32] Pinheiro J., Bates D., DebRoy S., Sarkar D. et the R Core team, *nlme : Linear and Nonlinear Mixed Effects Models*. R package version 3.1-96 (2009).

[33] Comets E., Lavenu A. et Lavielle M., *SAEMIX, an R version of the SAEM algorithm*. 20th annual meeting of the PAGE, Athens, Greece Abstr 2173 (2011).

[34] Best N., Tan K., Gilks W. et Spiegelhalter D., *Estimation of population pharmacokinetics using the Gibbs sampler*. J. Pharmacokinet. Biopharm. 23, 407–24 (1995).

[35] Spiegelhalter D.J., Thomas A. et Best N.G., *WinBUGS version 1.3 User Manual*. Imperial College, London, United Kingdom (2000).

[36] Pinheiro J. et Bates D., *Model building for nonlinear mixed-effect models*. Technical Report 91, Department of Biostatistics, University of Wisconsin-Madison (1995).

[37] Bertrand J., Comets E. et Mentré F., *Comparison of model-based tests and selection strategies to detect genetic polymorphisms influencing pharmacokinetic parameters*. J. Biopharm. Stat. 18, 1084–102 (2008).

[38] Wählby U., Jonsson N. et Karlsson M., *Assessment of actual significance levels for covariate effects in NONMEM*. J. Pharmacokin. Pharmacodyn. 28, 231–52 (2001).

[39] Comets E. et Mentré F., *Evaluation of tests based on individual versus population modelling to compare dissolution curves*. J. Biopharm. Stat. 11, 107–23 (2001).

[40] Vonesh E., *A note on the use of Laplace's approximation for nonlinear mixed-effects models*. Biometrika 83, 447–52 (1996).

[41] Bertrand J., Comets E., Chenel M. et Mentré F., *Some alternatives to asymptotic tests for the analysis of pharmacogenetic data using nonlinear mixed effects models*. Biometrics 68, 146–55 (2012).

[42] Gelfand A., Det D. et Chang H., *Bayesian statistics*. University Press (1992).

[43] Robert C., *Le choix bayésien : principes et pratique*. Springer-Verlag, Collection Statistiques et Probabilités, Paris, France (2006).

[44] Gelman A., Carlin J., Stern H. et Rubin D., *Bayesian Data Analysis, 2nd edition*. Chapman & Hall, London, UK (2004).

[45] Jeffreys H., *Some tests of significance, treated by the theory of probability*. Proc. Cambridge Phil. Soc. 31, 203–22 (1935).

[46] Kass R. et Raftery A., *Bayes factors*. J. Am. Stat. Assoc. 90, 773–795 (1995).

[47] Spiegelhalter D., Best N., Carlin B. et van der Linde A., *Bayesian measures of model complexity and fit*. JRSS 64, 583–639 (2002).

[48] Mandema J., Verotta D. et Sheiner L., *Building population pharmacokinetic-pharmacodynamic models. I. Models for covariate effects*. J. Pharmacokinet. Biopharm. 20, 511–528 (1992).

[49] Bies R., Muldoon M., Pollock B., Manuck S., Smith G. et Sale M., *A genetic algorithm-based, hybrid machine learning approach to model selection*. J. Pharmacokin. Pharmacodyn. 33, 195–221 (2006).

[50] Gastonguay M., *Full covariate models as an alternative to methods relying on statistical significance for inferences about covariate effects : a review of methodology and 42 case studies*. 20th Annual Meeting of the PAGE, Athens, Greece Abstr 2229 (2011).

[51] Food and Drug Administration, *Guidance for Industry : Population Pharmacokinetics*. FDA, Rockville, USA (1999).

[52] EMEA, *Draft guideline on reporting the results of population pharmacokinetic analyses*. Committee for Medicinal Products for Human Use, European Medicines Agency (2006).

[53] Brendel K., Dartois C., Comets E., Lemenuel-Diot A., Laveille C., Tranchand B., Girard P., Laffont C. et Mentré F., *Are population pharmacokinetic and/or pharmacodynamic models adequately evaluated ? A survey of the literature from 2002 to 2004*. Clin. Pharmacokinet. 46, 221–234 (2007).

[54] Box G., *Science and statistics*. J. Am. Stat. Assoc. 71, 791–9 (1976).

[55] Ette E. et Ludden T., *Population pharmacokinetic modeling : the importance of informative graphics*. Pharm. Res. 12, 1845–55 (1995).

[56] Karlsson M., Jonsson N., Wiltse C. et Wade J., *Assumption testing in population pharmacokinetic models : illustrated with an analysis of moxonidine data from congestive heart failure patients*. J. Pharmacokinet. Biopharm. 26, 207–46 (1998).

[57] Brendel K., Comets E., Laffont C., Laveille C. et Mentré F., *Metrics for external model evaluation with an application to the population pharmacokinetics of gliclazide*. Pharm. Res. 23, 2036–49 (2006).

[58] Comets E., Brendel K. et Mentré F., *Model evaluation in nonlinear mixed effect models, with applications to pharmacokinetics*. JSFdS 151, 106–28 (2010).

[59] Mentré F. et Escolano S., *Prediction discrepancies for the evaluation of nonlinear mixed-effects models*. J. Pharmacokinet. Biopharm. 33, 345–67 (2006).

[60] Brendel K., Comets E., Laffont C. et Mentré F., *Evaluation of different tests based on observations for external model evaluation of population analyses*. J. Pharmacokin. Pharmacodyn. 37, 49–65 (2010).

[61] Girard P., Blaschke T., Kastrissiosr H. et Sheiner L., *A Markov mixed effect regression model for drug compliance*. Stat. Med. 17, 2313–33 (1998).

[62] Al-Banna M., Kelman A. et Whiting B., *Experimental design and efficient parameter estimation in population pharmacokinetics*. J. Pharmacokinet. Biopharm. 18, 347–60 (1990).

[63] Hashimoto Y. et Sheiner L.B., *Designs for population pharmacodynamics : Value of pharmacokinetic data and population analysis*. J. Pharmacokinet. Biopharm. 19, 333–53 (1991).

[64] Atkinson A. et Donev A., *Optimum experimental designs*. Oxford University Press, USA (1992).

[65] Walter E. et Pronzato L., *Identification de modèles paramétriques à partir de données expérimentales*. Modélisation, Analyse, Simulation, Commande, Masson, Paris, France (1997).

[66] Mentré F., Mallet A. et Baccar, *Optimal design in random-effects regression models*. Biometrika 84, 429–42 (1997).

[67] Retout S., Mentré F. et Bruno R., *Fisher information matrix for non-linear mixed-effects models : evaluation and application for optimal design of enoxaparin population pharmacokinetics*. Stat. Med. 21, 2623–2639 (2002).

[68] Bazzoli C., Retout S. et Mentré F., *Fisher information matrix for nonlinear mixed effects multiple response models : evaluation of the appropriateness of the first order linearization using a pharmacokinetic/pharmacodynamic model*. Stat. Med. 28, 1940–56 (2009).

[69] Retout S. et Mentré F., *Further developments of the Fisher information matrix in nonlinear mixed-effects models with evaluation in population pharmacokinetics*. J. Biopharm. Stat. 13, 209–27 (2003).

[70] Mielke T. et Schwabe S., *mODa 9 Advances in Model-Oriented Design and Analysis*. Physica-Verlag HD (2010).

[71] Nguyen T., Bazzoli C. et Mentré F., *Design evaluation and optimisation in crossover pharmacokinetics studies analysed by nonlinear mixed effects models*. Stat. Med. 31, 1043–58 (2011).

[72] Nyberg J., Ueckert S., Stromberg E., Hennig S., Karlsson M. et Hooker A., *Poped : an extended, parallelized, nonlinear mixed effects models optimal design tool*. Comput. Meth. Prog. Biomed. 108, 789–805 (2012).

[73] Jones B. et Wang J., *Constructing optimal designs for fitting pharmacokinetic models*. Stat. Comput. 209–18 (1999).

[74] Jones B., Wang J., Jarvis P. et Byrom W., *Design of cross-over trials for pharmacokinetic studies*. J. Stat. Plan. Infer. 78, 307–16 (1999).

[75] Mielke T., *Impact of information approximations on experimental design in nonlinear mixed effects models*. PODE, Paris, France (2012).

[76] Nguyen T. et Mentré F., *Evaluation of Fisher information matrix using gaussian quadrature in nonlinear mixed effects models : application to dose-response trials*. 20th Annual Meeting of the PAGE, Athens, Greece Abstr 2114 (2011).

[77] Fedorov V., *Theory of optimal experiments*. Academic Press Inc, USA (1972).

[78] Wynn H., *Results in the theory and construction of d-optimum experimental designs*. JRSS B 80, 791–5 (1972).

[79] Kiefer J. et Wolfowitz J., *Optimum designs in regression problems*. Ann. Math. Stat. 30, 271–94 (1995).

[80] Nelder J.A. et Mead R., *A simplex method for function minimization*. Comput. J. 7, 308–13 (1965).

[81] Dufful S., Retout S. et Mentré F., *The use of simulated annealing for finding optimal population designs*. Comput. Meth. Prog. Biomed. 69, 25–65 (2002).

[82] Jonsson E.N., Wade J. et Karlsson M.O., *Comparison of some practical sampling strategies for population pharmacokinetic studies*. J. Pharmacokinet. Biopharm. 245–63 (1996).

[83] Little J. et Moler C., *MATLAB and Statistics Toolbox Release 2012b*. The MathWorks Inc, Natick, United States (1994).

[84] Mentré F., Nyberg J., Ogungbenro K., Leonov S., Aliev A., Dufful S., Bazzoli C. et Hooker A.H., *Comparison of results of the different software for design evaluation in population pharmacokinetics and pharmacodynamics*. 20th Annual Meeting of the PAGE, Athens, Greece Abstr 2066 (2011).

[85] Bazzoli C., Retout S. et Mentré F., *Design evaluation and optimisation in multiple response nonlinear mixed effect models : PFIM 3.0*. Comput. Meth. Prog. Biomed. 98, 55–65 (2010).

[86] Pillai G., Mentré F. et Steimer J., *Non-linear mixed effects modeling - from methodology and software development to driving implementation in drug development science*. J. Pharmacokin. Pharmacodyn. 32, 161–83 (2005).

[87] Lee J.Y., Garnett C.E., Gobburu J.V., Bhattaram V.A., Brar S., Earp J.C., Jadhav P.R., Krudys K., Lesko L.J., Li F., Liu J., Madabushi R., Anshu Marathe N.M., Tornoe C., Wang Y. et Zhu H., *Impact of pharmacometric analyses on new drug approval and labelling decisions : a review of 198 submissions between 2000 and 2008*. Clin. Pharmacokinet. 50, 627–635 (2011).

[88] Jönsson S., Henningsson A., Edholm M. et Salmonson T., *Role of modelling and simulation : a European regulatory perspective*. Clin. Pharmacokinet. 51, 69–76 (2012).

[89] Lalonde R., Kowalski K., Hutmacher M., Ewy W., Nichols D., Milligan P., Corrigan B., Lockwood P., Marshall S., Benincosa L., Tensfeldt T., Parivar K., Amantea M., Glue P., Koide H. et Miller R., *Model-based drug development*. Clin. Pharmacol. Ther. 82, 21–32 (2007).

[90] Saint-Marcoux F., Vandierdonck S., Prémaud A., Debord J., Rousseau A. et Marquet P., *Large scale analysis of routine dose adjustments of mycophenolate mofetil based on global exposure in renal transplant patients*. Ther. Drug Monit. 33, 285–94 (2011).

[91] Jamei M., Marciniak S., Feng K., Barnett A., Tucker G. et Rostami-Hodjegan A., *The Simcyp population-based ADME simulator*. Expert Opin. Drug Metab. Toxicol. 5, 211–223 (2009).

[92] Duffull S.B., Dooley M.J., Green B., Poole S.G. et Kirkpatrick C.M., *A standard weight descriptor for dose adjustment in the obese patient*. Clin. Pharmacokinet. 43, 1167–1178 (2004).

[93] West G.B., Brown J.H. et Enquist B.J., *The fourth dimension of life : fractal geometry and allometric scaling of organisms*. Science 284, 1677–1679 (1999).

[94] Shudo E., Ribeiro R.M. et Perelson A.S., *Modeling HCV kinetics under therapy using PK and PD information*. Expert Opin. Drug. Metab. Toxicol. 5, 321–332 (2009).

[95] Simeoni M., Magni P., Cammia C., De Nicolao G., Croci V., Pesenti E., Germani M., Poggesi I. et Rocchetti M., *Predictive pharmacokinetic-pharmacodynamic modeling of tumor growth kinetics in xenograft models after administration of anticancer agents*. Cancer Res. 64, 1094–1101 (2004).

[96] Silber H.E., Jauslin P.M., Frey N. et Karlsson M.O., *An integrated model for the glucose-insulin system*. Basic Clin. Pharmacol. Toxicol. 106, 189–194 (2010).

[97] Troconiz I.F., Plan E.L., Miller R. et Karlsson M.O., *Modelling overdispersion and Markovian features in count data*. J. Pharmacokinet. Pharmacodyn. 36, 461–477 (2009).

[98] Karlsson K.E., Plan E.L. et Karlsson M.O., *Performance of three estimation methods in repeated time-to-event modeling*. AAPS J. 13, 83–91 (2011).

[99] Ueckert S., Plan E., Ito K., Karlsson M., Corrigan B. et Hooker A., *Application of item response theory to ADAS-cog scores modelling in Alzheimer's disease*. 21st Annual Meeting of the PAGE, Venice, Italy Abstr 2318 (2012).

[100] Delor I., Charoin J.E., Gieschke R., Retout S. et Jacqmin P., *Modeling Alzheimer's disease progression using disease onset time and disease trajectory concepts applied to CDR-SOB scores from ADNI*. CPT : PSP 2, e78 (2013).

[101] Hénin E., You B., VanCutsem E., Hoff P.M., Cassidy J., Twelves C., Zuideveld K.P., Sirzen F., Dartois C., Freyer G., Tod M. et Girard P., *A dynamic model of hand-and-foot syndrome in patients receiving capecitabine*. Clin. Pharmacol. Ther. 85, 418–25 (2009).

[102] Mallet A., *A makimum likelihood estimation method for random coefficient regression models*. Biometrika 73, 645–56 (1986).

[103] Leary R., Jelliffe R., Schumitzky A. et Guilder M.V., *Improved computational methods for statistically consistent and efficient PK/PD population analysis*. 12th Annual Meeting of the PAGE, Verona, Italy Abstr 421 (2003).

[104] Miao H., Xia X., Perelson A. et Wu H., *On identifiability of nonlinear ode models with applications in viral dynamics*. Soc. Indust. Appl. Math. 53 (2011).

[105] Walter E. et Pronzato L., *Optimal experiment design for nonlinear models subject to large prior uncertainty*. J. Pharmacokin. Pharmacodyn. 253, R530–34 (1987).

[106] D'Argenio D.Z., *Incorporating prior parameter uncertainty in the design of sampling schedules for pharmacokinetic parameter estimation experiments*. Math. Biosci. 99, 105–118 (1990).

[107] Chow S. et Chang M., *Adaptative design methods in clinical trials - a review*. Orphanet J. Rare Dis. 3 (2008).

[108] Zamuner S., Lorio V.D., Nyberg J., Gunn R., Cunningham V., Gomeni R. et Hooker A., *Adaptative optimal design in pet occupancy studies*. Clin. Pharmacol. Ther. 87, 563–71 (2012).

[109] Dumont C., Chenel M. et Mentré F., *Planification adaptative en deux étapes pour des modèles non linéaires à effets mixtes : application en pharmacocinétique*. 45e Journées de Statistique, Toulouse, France (2013).

[110] Rowland M. et Tozer T.N., *Clinical Pharmacokinetics and Pharmacodynamics : Concepts And Applications ; 4th edition*. Lippincott Williams and Wilkins, Baltimore, USA (2007).

[111] Simon N., *Pharmacocinétique de population*. Solal Editions, Marseille, France (2006).