Adaptation of dosing regimen of chemotherapies based on pharmacodynamic models

Inès Paule

N° d'ordre 188-2011 Année 2011

THESE DE L'UNIVERSITE DE LYON

délivrée par

L'UNIVERSITE CLAUDE BERNARD LYON I

ECOLE DOCTORALE : ED 205 Interdisciplinaire Sciences-Santé (EDISS)

DIPLOME DE DOCTORAT

(arrêté du 7 août 2006)

en pharmacométrie

soutenue publiquement le 29 septembre 2011 par

Ines PAULE

Adaptation of dosing regimen of chemotherapies based on pharmacodynamic models

Adaptation de posologie de chimiothérapies basée sur des modèles pharmacodynamiques

Directeurs de thèse
Dr. Pascal GIRARD et Prof. Michel TOD

Jury

Prof. Etienne CHATELUT	rapporteur
Prof. Gilles FREYER	
Dr. Pascal GIRARD	directeur
Dr. Chantal Le GUELLEC	
Dr. Goonaseelan PILLAI	
Prof. Nicolas SIMON	rapporteur
Prof. Michel TOD	directeur

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université	**M. A. Bonmartin**
Vice-président du Conseil d'Administration	M. le Professeur G. Annat
Vice-président du Conseil des Etudes et de la Vie Universitaire	M. le Professeur D. Simon
Vice-président du Conseil Scientifique	M. le Professeur J-F. Mornex
Secrétaire Général	M. G. Gay

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur : M. le Professeur J. Etienne
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directeur : M. le Professeur F-N. Gilly
UFR d'Odontologie	Directeur : M. le Professeur D. Bourgeois
Institut des Sciences Pharmaceutiques et Biologiques	Directeur : M. le Professeur F. Locher
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. le Professeur Y. Matillon
Département de formation et Centre de Recherche en Biologie Humaine	Directeur : M. le Professeur P. Farge

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur : M. le Professeur F. De Marchi
Département Biologie	Directeur : M. le Professeur F. Fleury
Département Chimie Biochimie	Directeur : Mme le Professeur H. Parrot
Département GEP	Directeur : M. N. Siauve
Département Informatique	Directeur : M. le Professeur S. Akkouche
Département Mathématiques	Directeur : M. le Professeur A. Goldman
Département Mécanique	Directeur : M. le Professeur H. Ben Hadid
Département Physique	Directeur : Mme S. Fleck
Département Sciences de la Terre	Directeur : Mme le Professeur I. Daniel
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. C. Collignon
Observatoire de Lyon	Directeur : M. B. Guiderdoni
Ecole Polytechnique Universitaire de Lyon 1	Directeur : M. P. Fournier
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. Pignault
Institut Universitaire de Technologie de Lyon 1	Directeur : M. le Professeur C. Coulet
Institut de Science Financière et d'Assurances	Directeur : M. le Professeur J-C. Augros
Institut Universitaire de Formation des Maîtres	Directeur : M. R. Bernard

REMERCIEMENTS / ACKNOWLEDGEMENTS / PADĖKOS

Je remercie sincèrement / I thank sincerely / Nuoširdžiai dėkoju:

Au Professeur Michel Tod,
codirecteur de cette thèse, d'avoir été un excellent mentor et exemple à suivre, d'une part par vos immenses compétences et connaissances scientifiques et techniques, d'autre part par votre superbe aptitude pédagogique et vos qualités humaines admirables et honorables, notamment votre patience, compréhension, engagement, disponibilité, gentillesse, modestie, soutien et vos encouragements. C'était un honneur et un privilège de travailler avec vous, et une superbe opportunité d'apprendre!

Au Docteur Pascal Girard,
codirecteur de cette thèse, de m'avoir ouvert la porte au monde de la pharmacométrie tout d'abord en m'acceptant en stage de master et puis en trouvant un financement pour continuer en thèse. C'était une réalisation de mon très cher rêve de venir travailler en recherche dans le domaine de santé. Je vous remercie également pour vos conseils, votre confiance en moi et vos encouragements.

Au Professeur Gilles Freyer,
pour votre accueil chaleureux dans votre équipe, votre confiance en moi, votre gentillesse et votre générosité surprenantes, les aides financières que vous m'avez accordées en master et en fin de thèse, mais aussi pour vos conseils en tant que clinicien, votre enthousiasme et énergie très impressionnantes et inspirantes.

Aux Professeurs Etienne Chatelut et Nicolas Simon,
de m'avoir fait l'honneur d'accepter d'être rapporteurs de mes travaux de thèse.

To Novartis and Dr. Goonaseelan (Colin) Pillai in particular,
for the financial support for my PhD studies and the opportunity to experience the world of a leading pharmaceutical company during an internship, and also for accepting to be a part of the jury judging my thesis works.

Au Docteur Chantal Le Guellec,
d'avoir accepté de participer dans mon jury de soutenance.

Au Docteur Brigitte Tranchand,
Pour vos divers conseils, exemple de dynamisme et d'organisation, et surtout pour votre attention chaleureuse, votre soutien moral et vos encouragements.

Au Docteur Emilie Hénin,
Pour beaucoup de choses, qu'il s'agisse de l'encadrement, de dépannages, d'explications et conseils, de la relecture de mes travaux, ou encore de tes nombreuses qualités humaines, ton attitude ouverte et amicale, ton soutien et tes encouragements, tes efforts pour m'intégrer à la vie et au monde français…

Au Docteur Benoit You,
Pour ton apport de point de vue de clinicien à mes travaux « in silico », ta patience pour répondre à mes nombreuses questions, pour ta digne et amicale personnalité, ouverture et sens du travail, qui font un plaisir de travailler avec toi.

A Raymonde Maraval-Gaget et Fabienne Rippe,
pour votre aide dans des tâches administratives et votre attitude amicale.

A tous les autres membres d'équipe, présents et passés,
pour la bonne ambiance dans laquelle nous avons travaillé, et en particulier Alexandre pour la relecture de ma thèse en conditions pressées. Merci à tous pour votre accueil, et pour ne m'avoir jamais fait sentir comme une étrangère.

And finally to my personal support team, in particular:

To Gordon,
for being my home, my rock and my pillow, and all the other countless precious roles you play in my life.

Mamai ir tetei, už jūsų besąlygišką meilę ir paramą, beribį dosnumą ir pasiaukojimą, palaikymą ir tikejimą manimi, suteiktą laisvę ir galimybes siekti mano svajonių…

Seneliams, už jūsų paramą, dosnumą ir begalinį tikejimą mano galimybėm, jūsų ekstremalias svajones apie mano ateitį ir maldas už mane…

Šita dizertacija Jums – mama, tete ir seneliai.

Mano draugėms Raimondai, Rasai, Eugenijai, Anai, Evelinai, Ievai, už tai, kad nepamiršot net man emigravus, už jūsų tikejimą mano sėkme visur, ką pradėjau, už jūsų buvimą įkvepiančiais pavyzdžiais kiekviena savais būdais… Aš visada labai laukiu galimybių susitikti!

ABSTRACT

There is high variability in response to cancer chemotherapies among patients. Its sources are diverse: genetic, physiologic, comorbidities, concomitant medications, environment, compliance, etc. As the therapeutic window of anticancer drugs is usually narrow, such variability may have serious consequences: severe (even life-threatening) toxicities or lack of therapeutic effect. Therefore, various approaches to individually tailor treatments and dosing regimens have been developed: *a priori* (based on genetic information, body size, drug elimination functions, etc.) and *a posteriori* (that is using information of measurements of drug exposure and/or effects). Mixed-effects modelling of pharmacokinetics and pharmacodynamics (PK-PD), combined with Bayesian *maximum a posteriori* probability estimation of individual effects, is the method of choice for *a posteriori* adjustments of dosing regimens.

In this thesis, a novel approach to adjust the doses on the basis of predictions, given by a model for ordered categorical observations of toxicity, was developed and investigated by computer simulations. More technical aspects concerning the estimation of individual parameters were analysed to determine the factors of good performance of the method. These works were based on the example of capecitabine-induced hand-and-foot syndrome in the treatment of colorectal cancer. Moreover, a review of pharmacodynamic models for discrete data (categorical, count, time-to-event) was performed. Finally, PK-PD analyses of hydroxyurea in the treatment of sickle cell anemia were performed and used to compare different dosing regimens and determine the optimal measures for monitoring the treatment.

RESUME

Il existe une grande variabilité dans la réponse aux chimiothérapies anticancéreuses. Ses sources sont diverses: génétiques, physiologiques, comorbidités, médicaments associés, etc. La marge thérapeutique de ces médicaments étant généralement étroite, une telle variabilité peut avoir de graves conséquences: toxicités graves ou absence d'effet thérapeutique. Plusieurs approches pour adapter individuellement les posologies ont été proposées: *a priori* (basées sur l'information génétique, la taille corporelle, les fonctions d'élimination, etc.) et *a posteriori* (sur les informations de mesures d'exposition au médicament et/ou effets). La modélisation à effets-mixtes de la pharmacocinétique et de la pharmacodynamie (PK-PD), combinée avec une estimation bayésienne des effets individuels, est la meilleure méthode pour individualiser des schémas posologiques *a posteriori*.

Dans cette thèse, une nouvelle approche pour ajuster les doses sur la base des prédictions données par un modèle pour les observations catégorielles de toxicité a été développée et explorée par simulation. Les aspects plus techniques concernant l'estimation des paramètres individuels ont été analysés pour déterminer les facteurs de bonne performance de la méthode. Ces travaux étaient basés sur l'exemple du syndrome mains-pieds induit par la capécitabine dans le traitement du cancer colorectal. Une revue des modèles pharmacodynamiques de données discrètes (catégorielles, de comptage, de survie) a été effectuée. Enfin, des analyses PK-PD de l'hydroxyurée dans le traitement de la drépanocytose ont été réalisées pour comparer des différentes posologies et déterminer les modalités optimales de suivi du traitement.

RESUME SUBSTANTIEL EN FRANÇAIS

Il existe une grande variabilité dans la réponse aux chimiothérapies anticancéreuses. Ses sources sont diverses: génétiques, physiologiques, comorbidités, médicaments associés, etc. La marge thérapeutique de ces médicaments étant généralement étroite, une telle variabilité peut avoir de graves conséquences: soit des toxicités graves, soit une absence d'effet thérapeutique. Plusieurs approches pour adapter les posologies au niveau individuel existent: *a priori* (basées sur l'information génétique, la taille corporelle, les fonctions d'élimination, etc.) et *a posteriori* (sur les informations de mesures d'exposition au médicament et/ou effets). La modélisation à effets-mixtes de la pharmacocinétique et de la pharmacodynamie (PK-PD), combinée à une estimation bayésienne des effets individuels, est la meilleure méthode pour individualiser des schémas posologiques *a posteriori*.

Dans cette thèse, une nouvelle approche pour ajuster les doses sur la base des prédictions d'un modèle pour données catégorielles de toxicité a été développée et explorée par simulation (Article 2). Ce travail était basé sur l'exemple du syndrome mains-pieds induit par la capécitabine dans le traitement du cancer colorectal métastatique ou avancé.

Plus de la moitié des patients traités par la capécitabine souffrent de cette toxicité dermatologique qui affecte la peau des paumes des mains et des plantes de pieds et peut aller jusqu'à des cloques avec des douleurs sévères et des difficultés dans les activités quotidiennes comme la marche et la manipulation des objets. Par conséquent, cette toxicité (aux grades supérieures 2 et 3) est dose-limitante et nécessite des interruptions de traitement et des réductions de doses. L'approche standard utilisée dans la pratique clinique est d'interrompre le traitement jusqu'à l'amélioration au moins jusqu'au grade 1, et ensuite de continuer le traitement avec des doses réduites de 25% ou de 50%, et finalement de l'interrompre si plusieurs événements de toxicités sévères sont apparus.

Cette approche « rigide » n'est probablement pas optimale, et il est intéressant d'étudier la faisabilité d'une adaptation individualisée des doses, qui soit basée sur les prédictions données par un modèle de toxicité du patient, et de comparer sa performance à celle de la méthode standard, par simulation d'essais cliniques.

Un modèle de population du syndrome mains-pieds (hand-foot syndrome, HFS) et un modèle de l'inhibition de la croissance tumorale (tumour growth inhibition, TGI) pour la capécitabine dans le traitement du cancer colorectal avancé ou métastatique ont été publiés récemment

[Henin09, Claret09]. Ces modèles ont été utilisés dans la présente étude. Les étapes de ce travail ont consisté à:

- Développer le concept et les techniques numériques afin de déterminer la «meilleure» prochaine dose sur la base de la prédiction du modèle individuel du risque d'HFS intolérable (grade ≥2) au cours du prochain cycle de traitement;
- Définir le protocole d'essai clinique, les modalités d'adaptation de dose cliniquement pertinentes, les critères de comparaison des différentes méthodes d'adaptation de dose, avec la consultation d'un oncologue clinique;
- Implémenter les modèles HFS et TGI, ainsi que les différentes approches d'adaptation de dose, dans un code de simulation d'essais cliniques;
- Effectuer les simulations avec différentes spécifications de la procédure de recherche de dose afin de déterminer par analyses de sensibilité le «calibrage» optimal de l'adaptation individuelle, c'est-à-dire celui aboutissant à la plus grande réduction de la toxicité HFS sans diminuer l'efficacité antitumorale par rapport à l'approche standard.

Le meilleur résultat obtenu a été une réduction moyenne d'environ 10 jours de la durée totale de HFS de grade ≥2, sur la durée d'essai de 30 semaines de traitement maximum + 4 semaines de suivi post-traitement. Il s'agit de la réduction maximale de toxicité obtenue tout en préservant la même efficacité antitumorale.

Cette étude a montré que l'adaptation individualisée de dose sur la base des observations catégorielles de toxicité peut être faisable et bénéfique en termes de résultat clinique global. Dans l'exemple utilisé du syndrome mains-pieds, l'impact a été limité par la faible sensibilité de la toxicité aux changements de dose. L'impact d'une telle méthode serait probablement plus élevé pour les toxicités réversibles avec une dynamique rapide (par exemple, gastro-intestinales).

Au cours du développement de l'approche d'individualisation de dose présentée ci-dessus, la qualité des estimations empiriques bayésiennes (EBEs) des effets aléatoires individuels (permettant d'ajuster le modèle de population pour le patient sur la base de ses observations précédentes) a été examinée. L'exactitude et la précision observées étant plutôt faibles, cela a motivé la recherche des facteurs potentiellement influents. La qualité des EBEs a un impact direct sur l'exactitude des prédictions de risque individuel de toxicité, et est donc une condition importante pour la bonne performance de la méthode basée sur les prédictions.

Ainsi, ce travail (Article 3) a consisté à étudier l'influence de:

- L'estimateur (mode vs. moyenne de la distribution *a posteriori* des effets aléatoires);
- L'algorithme d'optimisation, où des algorithmes de recherche optimale locaux (simplex, quasi-Newton) et global (recherche aléatoire adaptative) ont été comparés;
- La quantité de données par patient;
- Les distributions individuelles des catégories;
- L'amplitude de la variabilité inter-individuelle;
- Les valeurs des paramètres du modèle d'effet.

Dans cette étude, on a constaté que les principaux facteurs affectant la qualité des EBEs étaient les valeurs des paramètres qui régissent la relation dose-réponse et la distribution intra-sujet des catégories.

Dans l'exemple du modèle HFS, les valeurs de la variable qui produit l'effet étaient trop faibles pour permettre de bien identifier les paramètres de la fonction Emax, et les distributions intra-sujet des catégories était fortement non-uniforme. Les EBEs avaient un biais et une faible précision, l'impact sur les décisions cliniques concernant la prochaine dose à prendre était variable, mais considérable. Dans les cas extrêmes, l'impact de l'imprécision des EBEs sur le risque prédit de grade ≥ 2 pouvait être de l'ordre de 4; les doses correspondantes calculées selon les risques vrai et estimé pouvaient varier 10 fois. Cependant, cela n'a pas eu d'impact sur la performance globale de la méthode d'adaptation de dose en raison de la faible sensibilité du HFS pour des changements de dose: les résultats étaient très semblables si les valeurs des effets aléatoires vraies (simulées) ou leurs EBEs étaient utilisées ou si elles n'étaient pas utilisées du tout (c'est-à-dire, avec uniquement le modèle de population pour le patient moyen).

Les résultats cliniques sont souvent décrits en tant que données discrètes : événements (mort, accident vasculaire cérébral, crise d'épilepsie, lésions de sclérose en plaques, progression de la maladie, de graves effets secondaires), le temps jusqu'à leur apparition, leur fréquence et degré de sévérité. Ces données discrètes ont besoin de structures et méthodes de modélisation spécifiques. Une revue des approches de modélisation des données discrètes et des exemples de leurs applications dans l'analyse des données pharmacodynamiques fait partie de cette thèse (Article 1).

Le quatrième article de cette thèse concerne l'hydroxyurée (HU), anticancéreux utilisé également dans le traitement de la drépanocytose, l'une des plus fréquentes maladies génétiques. Elle est due à une mutation dans le gène de l'hémoglobine et se caractérise par des

globules rouges rigides et en forme de faucille, ce qui peut conduire à des crises vaso-occlusives avec diverses complications, telles que les crises douloureuses aiguës, l'ischémie et des lésions de différents organes, syndrome thoracique aigu ou accident vasculaire cérébral. Les objectifs de ce travail ont été:

- De développer des modèles de population de PK-PD pour l'hydroxyurée dans le but de caractériser les relations dose-exposition-réponse et leur variabilités, de rechercher des covariables potentielles;
- D'utiliser ces modèles pour comparer deux schémas posologiques (une quotidienne en continu et l'autre avec des interruptions de 2 jours après tous les 5 jours) par simulation;
- De développer des recommandations pour le suivi du traitement.

Comme aucune toxicité dose-limitante n'est survenue dans cette étude, aucun modèle de toxicité et aucun outil d'ajustement *a posteriori* des doses n'ont été développés. L'efficacité a été mesurée sur deux biomarqueurs: le pourcentage d'hémoglobine fœtale (HbF%) et le volume globulaire moyen (VGM).

La dynamique de VGM a été plus rapide que celle de HbF%, mais la variabilité interindividuelle à l'état d'équilibre (EE) des valeurs de HbF% a été beaucoup plus grande que celle de VGM. Par conséquent, la valeur de VGM à EE à 3 mois n'était pas prédictive de la valeur de HbF% à l'EE au mois 26. Par conséquent, le niveau de HbF%, directement lié à l'allégement des symptômes de la drépanocytose, parait être un meilleur biomarqueur que VGM pour le suivi du traitement par HU.

Concernant la comparaison des deux schémas posologiques, la différence était très petite pour les profils de VGM, mais plus grande pour les profils de HbF%, surtout pour les patients atteignant les plus hauts niveaux de HbF%. Pour ces patients répondant le mieux au médicament, la prise quotidienne de HU peut être plus avantageuse en termes d'augmentation de HbF% qu'avec des interruptions 2 jours par semaine.

L'estimation de l'intensité de l'effet a suggéré une perspective intéressant: le niveau de HbF pourrait être encore accru par des médicaments plus puissants ou par des combinaisons de médicaments.

Malheureusement, aucune covariable expliquant une partie significative de la variabilité n'a pu être identifiée à partir de ce jeu de données, et des informations concernant les polymorphismes génétiques associés à la réponse à HU n'étaient pas disponibles. Des informations sur les polymorphismes dans les gènes de régulation du métabolisme de HU, de ses transporteurs, de l'expression de HbF et de la prolifération des progéniteurs érythroïdes permettraient potentiellement d'ajuster la posologie *a priori*.

KEYWORDS

Pharmacometrics, mixed-effects pharmacokinetic and pharmacodynamic models, discrete data models, cancer chemotherapy, dosing individualization, empirical Bayes estimates, optimization algorithms, capecitabine, hand-foot syndrome, hydroxyurea, sickle cell anemia.

MOTS-CLES

Pharmacométrie, modèles à effets mixtes de la pharmacocinétique et de la pharmacodynamie, modèles pour données discrètes, chimiothérapie anticancéreuse, individualisation de posologie, estimation empirique bayésienne, algorithmes d'optimisation, capécitabine, syndrome mains-pieds, hydroxyurée, drépanocytose.

LABORATOIRE

EMR 3738 Ciblage Thérapeutique en Oncologie
BP12, Faculté de Médecine Lyon-Sud
69921 OULLINS Cedex

TABLE OF CONTENTS

LIST OF ARTICLES

1. Pharmacodynamic models for discrete data

Ines Paule, Pascal Girard, Gilles Freyer, Michel Tod

(Manuscript)

2. Dose adaptation of capecitabine based on individual prediction of limiting toxicity grade: evaluation by clinical trial simulation

Ines Paule, Michel Tod, Emilie Hénin, Benoit You, Gilles Freyer, Pascal Girard

Cancer Chemotherapy and Pharmacology (published online 02-08-2011)

3. Empirical Bayes estimation of random effects of a mixed-effects proportional odds Markov model for ordinal data

Ines Paule, Pascal Girard, Michel Tod

Computer Methods and Programs in Biomedicine (published online 26-05-2011)

4. Population pharmacokinetics and pharmacodynamics of hydroxyurea in sickle cell anemia patients, a basis for optimizing the dosing regimen

Ines Paule, Hind Sassi, Anoosha Habibi, Kim P.D. Pham, Dora Bachir, Frédéric Galactéros, Pascal Girard, Anne Hulin and Michel Tod

Orphanet Journal of Rare Diseases 6(1):30 (published 28-05-2011)

ABBREVIATIONS

5FU	5-fluorouracil
ALP	alkaline phosphatase
ALT	alanine transaminase
ARS	adaptive random search
AST	aspartate transaminase
AUC	area under the curve of concentration-time plot
BLQ	below the limit of quantification
BSA	body surface area
cGMP	cyclic guanidine monophosphate
CI	confidence intervals
CL (CL/F)	(apparent) clearance
CLcr	creatinine clearance
CML	chronic myeloid leukemia
CR	complete response
CV	coefficient of interindividual variability
CYP	cytochrome P450
DPD	dihydropyrimidine dehydrogenase
EBE	empirical Bayes estimate
EC50	concentration providing 50% of maximal effect
ED50	dose providing 50% of maximal effect
EGFR	epidermal growth factor receptor
EE	état d'équilibre
FOCE	first order conditional estimation
G-CSF	granulocyte colony stimulating factor
GFR	glomerular filtration rate
GQ	Gaussian Quadrature
Hb	hemoglobin
HbF	fetal hemoglobin
HbF%	percentage of fetal hemoglobin (of total hemoglobin)
HFS	hand-foot syndrome
HU	hydroxyurea

IIV	interindividual variability
IPBDA	individual prediction-based dose adaptation
IPPSE	individual PK parameter with imprecision (standard error)
ka	the rate constant for absorption
kcp	rate constant for transfer from central to peripheral compartment
kpc	rate constant for transfer from peripheral to central compartment
KPD	kinetic-pharmacodynamic
LDH	lactate dehydrogenase
LOQ	limit of quantification
LV	leucovorin
MAE	mean absolute error
MAP	maximum a posteriori probability
MCH	mean corpuscular haemoglobin
MCMC	Markov chain Monte Carlo
MCV	mean corpuscular volume
NPDE	normalised predictive discrepancy errors
OFV	objective function value
OS	overall survival
PC-VPC	prediction-corrected visual predictive check
PD	pharmacodynamics
PD	progressive disease
PFS	progression-free survival
PK	pharmacokinetics
PMN	polymorphonuclear neutrophils
PR	partial response
RE	random effects
RECIST	Response Evaluation Criteria In Solid Tumors
RSE	relative standard errors
SAEM	Stochatistic Approximation Expectation Maximisation
SCA	sickle cell anemia
SD	stable disease
SD	standard deviation
SE	standard error

SS	steady-state
TDM	therapeutic drug monitoring
TGI	tumour growth inhibition
TK	tyrosine kinase
TPMT	thiopurine methyltransferase
TR	target risks
TTE	time-to-event
Vc (Vc/F)	(apparent) central volume of distribution
VGM	volume globulaire moyen
VPC	visual predictive check
WGT	body weight

INTRODUCTION

1. Variability in response to chemotherapy

Clinical practice of drug therapy of many diseases is challenged by a large variability in the response and tolerance to treatment among (and within) individuals. It can lead to therapeutic failures or intolerable adverse effects, if not handled properly. Such risks are especially high for medicines with a narrow therapeutic index, that is a small difference between efficacious and toxic doses.

This characteristic is particularly true for cancer chemotherapies, where the drug amounts that are big enough to affect cancer cells are harmful for normal cells too. And the "right" dose of cytotoxic drugs is considered to be the highest one that can be tolerated in terms of adverse effects. As the relationship between drug doses and anticancer effect is not well quantified, it is expected that the higher the drug exposure, the stronger the effect. The systemic exposure after fixed doses of cytotoxic drugs may vary between patients 4-10 times or even much more [Evans 1989, Masson 1997, Gurney 1996]. The average proportion of patients for whom anticancer drugs are ineffective is about 75% [Spear 2001].

Variations in many processes contribute to the overall variability in the individual response to the drug: at the pharmacokinetic level (absorption, distribution, metabolism, excretion), as well as at the pharmacodynamic level (mechanisms of the therapeutic and adverse effects) [Ratain 1990, Gibaldi 1992].

The sources of heterogeneity are very diverse [Undevia 2005]:

- genetic (germline or somatic mutations (in cancer tissues)),
- demographic/physiologic (age, body size and composition, sex, ethnicity, performance status),
- diseases (especially hepatic and renal) and special conditions (allergies and intolerancies, dehydration, pregnancy, age, etc.),
- previous or concomitant medications,
- environmental (smoking, alcohol consumption, diet, etc.),
- circadian variation,
- intrinsic or developed tolerance/resistance to the drug,
- behavioural (compliance to the prescribed dosing regimen), etc.

2. “Personalized medicine” tools

As the same dose of the drug may be too high (lead to unacceptable toxicity) for one patient and be too low (have no therapeutic effect) for another, there is a need to individually adjust doses for certain drugs. Advances in science and technology bring new (sophisticated) decision support tools for the practice/implementation of the “personalized medicine” to tailor the medical treatment to the individual characteristics to help achieve the therapeutic objectives without unacceptable toxic effects for each patient [The case for personalized medicine].

2.1. Pharmacogenomics

One of its branches is the new rapidly developing translational science of pharmacogenetics or pharmacogenomics, which consists of analyzing molecular determinants of response to medications at the gene, protein and metabolite levels. Such diagnostic tools are aimed to identify individuals likely to respond to treatment as well as those likely to have strong adverse effects and select treatments and tailor their doses accordingly.

2.1.1. PK level

A major area of pharmacogenetics application is germline mutations in genes coding the drug transporters or metabolizing enzymes, such as cytochrome P450 (CYP) family of liver enzymes, which are responsible for breaking down more than 30 different classes of drugs [Phillips 2001]. Individuals with absent or reduced activity of such enzymes risk severe toxicities due to overexposure to the drug, if doses are not reduced as compared to the population with normal activity of metabolizing enzymes. Examples of such pharmacogenetics applications to anticancer drugs include genetic testing for:

- the UGT1A1*28 polymorphism to identify individuals with reduced UGT1A1 gene expression and decreased glucuronidation of the active metabolite SN38 and resulting increased risk of neutropenia and diarrhoea on treatment with irinotecan for colorectal cancers [Hoskins 2007, Iyer 2002]. Their doses are thus reduced to reduce the risk of toxicity [FDA Camptosar label];
- the drug metabolizing enzyme thiopurine methyltransferase (TPMT) in the treatment by thiopurines (6-mercaptopurine, azathioprine) to avoid the risk of severe hematotoxicity [Black 1998, Evans 2004, Bosch 2006]. The genetic test classifies the patients according to normal, intermediate and deficient TPMT activity [Relling 1999, McLeod 2002]. Doses are reduced for intermediate or deficient patients (about 10% of the population).

- genetic and/or phenotypic approaches to determine the dihydropyrimidine dehydrogenase (DPD) activity for candidates to treatment of solid tumours by 5-Fluorouracil (5FU) or its oral prodrugs (such as capecitabine) [Milano 1996, Milano 1999, Mercier 2006, Bocci 2008, Kaestner 2007b].

A dose individualisation method taking into account the CYP2C genotype has been developed for an investigational anticancer agent indisulam to avoid unacceptable neutropenia [Zandvliet 2009].

2.1.2. PD level

Targeted therapy

Cancer treatment is furthermore challenged by high genetic and phenotypic heterogeneity at the pharmacodynamic level, including somatic mutations in the tumour tissue. Genetic and epigenetic perturbations in signal pathways drive cancer growth, survival, invasion and metastatic spread [Senzer 2005]. They may make malignant cells resistant to one treatment but sensitive to another specifically targeted agent.

This leads to “personalized medicine” approach to use genetic and phenotypic biomarker tests to identify tumours having specific mutations making them susceptible to specially designed targeted treatment. The first examples include:

- testing for the BCR/ABL oncogene, caused by the so-called Philadelphia chromosome translocation, in chronic myeloid leukemia (CML) patients (positive in more than 90% of cases) to decide about treatment with tyrosine kinase (TK) inhibitors, that specifically target this TK protein. The first such drug created by "rational drug design" was imatinib, named “the magic bullet” to cure cancer by TIME magazine [Saglio 2004]. Later second generation drugs, dasatinib and nilotinib, were developed to overcome imatinib resistance and to increase responsiveness to TK inhibitors; now they are also approved for first-line treatment of CML.
- testing for HER2 oncogene overexpression in breast tumours (about 25-30% of cases) to identify those likely to respond to trastuzumab, a monoclonal antibody blocking HER2 receptors [Piccart-Gebhart 2005, Romond 2005]. Overexpression of this receptor of epidermal growth factor is associated with a poor prognosis, increased tumour formation and metastasis, as well as resistance to chemotherapy [Nanda 2007].
- measurements of epidermal growth factor receptor (EGFR) expression in lung and colorectal cancer patients to identify those susceptible to treatment by tyrosine kinase inhibitors (gefitinib for lung cancer [Paez 2004], erlotinib for non-small-cell lung cancer [Rosell 2009], cetuximab and panitumumab for colorectal cancer [Sartore-Bianchi 2007]). Mutations that

result in increased EGFR expression or activity are associated with uncontrolled cell division and therefore predisposition to cancer [Zhang 2007].

- testing for somatic KRAS gene mutations in tumour tissue before starting treatment with anti-EGFR monoclonal antibody drugs cetuximab and panitumumab indicated for metastatic colorectal cancer (activated in about 40% of cases) [Lievre 2006]. KRAS is a protein involved in many signal transduction pathways, its mutations are thought to be an essential step in the development of many cancers [Kranenburg 2005]. KRAS mutations may lead to increased activation of the Ras/Raf/MAPK pathway and result in a lack of activity of EGFR inhibitors [Lievre 2006, Amado 2008]. It is associated with shorter progression-free survival and overall survival [DiFiore 2007, Lievre 2008].

Screening for individuals at cancer development or recurrence risk

The most common example is concerning BRCA1 and BRCA2 gene germline mutations that lead to inheritable high risks to develop breast and ovarian cancers [Nelson 2005]. Mutation carriers may choose to have a prophylactic surgery, preventive chemotherapy or at least close surveillance with the help of genetic counselling [Genetics].

A test has been developed that measures the expression of 21 genes in women with breast cancer to determine the risk of tumour return within 10 years [Silver 2009]. Such information may help decide on therapy strategy: hormone therapy alone or more aggressive chemotherapy.

Biomarkers alerting for possible recurrence of cancer after treatment

For some tumours, associated tumour markers circulating in blood have been identified that inform of possible cancer existence/regrowth before physical signs and symptoms. Elevated levels of tumour biomarkers would lead to more thorough examinations for cancer and allow to make clinical decisions sooner.

2.2. Non-genetic bases to individualise dose *a priori*

However, the sources of between and within-subject variability of drug response are very diverse. Genetic determinants can explain only a part of it, and they were identified or are relevant for only a part of the drugs. Other individual patient characteristics may be used for adjusting doses if a clinically significant link could be identified between them and the drug exposure and/or response and if they explain a large part of the variability.

2.2.1. BSA, weight

Traditionally, the standard dosing of cytotoxic drugs was based on body surface area (BSA) or body weight. The idea of BSA-based dosing was introduced for cancer chemotherapy about 50 years ago [Kaestner 2007a] and has been very widely applied in clinical practice, despite that it lacks scientific grounds and the numerous clinical studies showing that it fails to reduce the variability in the systemic exposure for most of cytotoxic drugs as compared to one-size-fits-all dosing [Gurney 1996, Mathijssen 2007, Kaestner 2007a, Gao 2008]. Initially it was based on inter-species studies in the beginning of 20th century, where it was established that the basal metabolic rate varied among various species as a function of weight (power function with the exponent of about 0.75), which approximately corresponds to the variation of BSA as a function of weight [Gao 2008]. BSA-based dosing does not take into account the possible physiologic changes in individuals with extreme weight (cachetic or obese). Moreover, many formulas exist for BSA estimation but they may all lack accuracy and precision [Kaestner 2007a]. BSA or weight scaling may have some relevance only for the prediction of the first-in-human dose from preclinical animal data by allometric scaling or the pediatric doses.

2.2.2. Alternatives to BSA

Considering the number of different factors of variability in drug response, it is unlikely that the dose can be well adjusted by body size measure only. This is especially true for various specific subgroups whose pharmacokinetic processes are significantly different from the "average normal" individual, such as children, elderly, patients with impaired renal or hepatic functions, low metabolising enzyme expression/activity, significantly decreased blood protein levels, concomitant interacting medications, poor disease and/or performance status. If a clinically relevant and significant relationship between covariates and PK parameters is identified and showed to explain a large part of variability in drug exposure, this knowledge may be used to adjust the dose before any drug is taken for individuals identified to be at risk (of severe toxicity or lack of response).

If the drug elimination is mainly renal and a large part of the variability of its exposure is explained by glomerular filtration rate (GFR), it may be reasonable to adjust doses by GFR, once the target AUC is determined which would be optimal in terms of clinical response and toxicity risk. Such a method has first been developed and validated by prospective clinical trial for carboplatin [Calvert 1989]; it is now widely used in clinical practice [Mathijssen 2007]. A simpler way to predict the carboplatin clearance without the need of isotopic

determination of GFR (as in Calvert's formula) using body weight, age, sex, and blood creatinine level has been developed by Chatelut *et al.* [Chatelut 1995] and is used in practice [Tranchand 2003]. Similar carboplatin dosing equations have been developed and validated for pediatric patients also [Chatelut 1996, Thomas 2000]. However, even using these approaches, there may still remain a large variability in PK among patients [Kaestner 2007b, Ekhart 2006, Gao 2008].

Since many physiologic and pathologic changes, as well as common multiple medications, in elderly patients may make their drug exposure and response significantly different from the younger adults, special consideration of their conditions is needed to ensure safe dosing of cytotoxic drugs for them [Lichtman 2000, Tranchand 2003, Bach 2010].

Many population PK-PD analyses have been performed and identified covariates for PK and/or PD of anticancer agents, mostly descriptors/indicators of body size, protein binding, renal and hepatic functions, but few covariate-based *a priori* dosing recommendations have been proposed or prospectively validated and accepted in clinical practice [Tranchand 2003, de Jonge 2005, Zandvliet 2008, Bach 2010].

2.2.3. Phenotyping studies

Attempts were made to use probe drugs or substrates to estimate *a priori* the individual exposure of some drugs that are predominantly eliminated by specific enzymes (such as CYP450 family members, DPD) or are concerned by particular transporters (ABCB1) [Gao 2008, Mathijssen 2007]. The probes are ideally cheap, safe, and easily available and determined agents, whose metabolism, distribution and elimination are related to the pharmacokinetic behaviour of the anticancer drug in question [Mathijssen 2007]. These exploratory methods still need to be validated before use in clinical practice.

A randomized prospective clinical trial comparing CYP3A phenotyping-based dose adaptation to standard BSA-based dosing of irinotecan demonstrated that individualization resulted in a substantially reduced interindividual variability of irinotecan and its active metabolite SN-38 AUCs, as well as of absolute neutrophil count nadir; the incidence of severe neutropenia was lower as well [Van der Bol 2008].

2.3. *a posteriori* dose adaptation

Drug disposition and effects involve too many processes that may be affected by numerous heterogeneous factors to allow an exact prediction of them for each individual by a few easily measurable covariates before the drug is taken. Even though *a priori* dose adjustment may be

useful to reduce the risk of severe toxicity of the first doses, subsequent dose modifications based on individual experience with the drug are likely to be needed to optimise the clinical outcome for drugs with a narrow therapeutic index and high inter-individual variability.

2.3.1. Test dose

A part of dose individualization strategy for drugs with linear pharmacokinetics may be to determine the patient's PK after administering a test dose, generally smaller than therapeutic [Cano 1985, Plunkett 1985, Pignon 1994, Bolinger 2001, Bleyzac 2001]. The future doses can then be adapted according to the patient's individual PK characteristics.

2.3.2. Therapeutic drug monitoring

Therapeutic drug monitoring (TDM) usually consists of measuring drug concentrations in biological fluids (mostly blood, but may be urine, saliva) in the course of the treatment in order to determine how the next doses should be adjusted for the patient [Moore 1987].

In order to have benefit from pharmacokinetically guided dosing, the drug should meet several criteria [Rousseau 2002]:

- a narrow therapeutic index;
- a substantial interpatient PK variability;
- relatively low interoccasion variability (i.e. among different occasions in the same individual) [Holford 1999];
- a clinically significant exposure-response relationship, so that the response would be sensitive to the changes in drug concentration.

Moreover, rapid, accurate, precise and reproducible assays for drug concentration quantification must be developed and widely available and feasible in clinical care units by their personnel [Hon 1998, Bates 1998]. Finally, in order to be accepted in clinical practice, such individualized dosing strategies need to be first validated by large prospective randomized clinical studies, comparing them to standard dosing. However this step is rarely carried out.

Due to their high risks and high interpatient variability, many anticancer agents have been identified as likely candidates for TDM [Van Den Bongard 2000]. However, a major limitation for many chemotherapy drugs is that a clinically important relationship between plasma drug concentrations and therapeutic effect or toxicity is hard to identify and thus establish a "therapeutic range". Tumours are highly heterogeneous and drug concentrations in the target tissues are unknown and possibly poorly correlated to the plasma drug

concentrations. Moreover, there is a delay (days or weeks) between drug administration and clinical response. Exposure-toxicity relationships are usually more successfully established. Furthermore, chemotherapies frequently consist of combinations of drugs with overlapping therapeutic and toxic effects, making it difficult to identify the contribution of each drug. TDM is even harder to apply for drugs with nonlinear PK, as well as prodrugs and drugs with active metabolites [Oellerich 2001].

PK-based TDM

Pharmacokinetically-guided TDM methods target a particular predefined drug exposure (AUC or steady-state concentration), which is associated with acceptable level of toxicities, more rarely with an optimum clinical outcome in terms of efficacy and toxicity [Gamelin 1998]. The exposure target can be reached by modifying the dose or the inter-dose interval or both [Plunkett 1985, Ando 1996, Jodrell 1994].

Historical methods to determine the dose

Historically, the needed dose correction to reach the target was determined by taking one or several blood samples during or after administration of the drug and using simple formulas, nomograms or multilinear regressions [de Jonge 2005, Bach 2010]. Estimation of PK parameters by traditional noncompartmental methods (such as trapezoid for AUC) usually requires a large number of samples. To make it more feasible in clinical practice, limited sampling strategies have been developed to minimize the needed number of samples to one or a few at specific optimized time points [Loh 2007]. However, such strategies are sensitive to deviations from those optimum points and can only be used for the same dosing regimen [Kaestner 2007b]. They have been extensively reviewed by, for example, de Jonge [de Jonge 2005].

Bayesian TDM

An alternative to these traditional analyses, the so-called population PK-PD modelling approach, has been introduced by Lewis Sheiner and Stuart Beal [Sheiner 1982, Sheiner 1984]. This statistical method combines data from all study patients to estimate a population model where the main (“fixed”) parameters describe the typical population effect and the individual (“random”) parameters adjusts for the main parameters for each individual and account for correlation among observations in the same subject. The population approach allows to split the overall variability to the between-subject variability and the within-subject

(and residual) variability. Some of between-subject variability may be explained by known individual factors (covariates), but the population models also include individual ("random") parameters which represent unobservable individual sources of variability. The latter are estimated by so-called Bayesian maximum *a posteriori* probability (MAP) method, which combines both the knowledge of effects in other individuals (via the population model) and the individual measured data. This approach is more flexible for the times of measurements and generally requires less of them than the noncompartmental analysis. However, some time points are more informative than others [D'Argenio 1981, Panetta 2002], and the number of measurements per individual also has an impact on how good the individual estimates are [Savic 2009]. The quality of the population model is also of prime importance for the precision of individual predictions used to determine the dose adjustment.

Many clinical studies have tested Bayesian PK-guided TDM for various chemotherapy drugs, only a few of them were prospective [Veal 2003, Rousseau 2002, de Jonge 2005, Kaestner 2007b, Bach 2010, Zandvliet 2008]. Most of them aimed at increasing the part of patients having drug exposure closer to the target exposure. And often the results were quite satisfactory in terms of this objective [de Jonge 2005]. However, few studies have evaluated/showed the gain in clinical outcome in terms of anticancer efficacy and/or toxicity profiles [Evans 1998]. One of reasons is that such studies would require a large number of patients.

PD-guided TDM

The ultimate objective of dose individualization is a gain in terms of therapeutic outcome. PK-guided TDM is not likely to be optimal/sufficient to attain this, in cases where there is high between-patient variability at the pharmacodynamic level, such as disease and performance status, prior therapies, concomitant diseases and medications, different sensitivity of bone marrow and other healthy tissues, as well as high heterogeneity among tumours (acquired resistances etc.). There may be no universal target systemic drug exposure level.

Therefore, PD-guided dose adaptation strategies are developed. Since for cytotoxic chemotherapies the "best" dose is traditionally defined as the highest that is tolerated by the patient, it is natural to base the dose on the most important (dose-limiting) toxicities. Reducing the occurrence of severe dose-limiting toxicity also allows to avoid the consequent dose delays and/or reductions, which may negatively affect the antitumour efficacy.

Another reason why PD-based dose adaptations are more often toxicity-based is that it is easier to identify the relationship between systemic drug exposure and adverse endpoint rather than the anticancer outcome.

The explanations of the difficulty to identify a relationship between administered drug doses and anticancer effect include:

- absence of data of effect site concentrations, their irrelevance to plasma concentrations;
- delays between drug administration and clinical response;
- high heterogeneity between and within the tumours;
- many confounding factors (prior and combined anticancer treatments, concomitant diseases and their medications);
- sparseness of data.

Therapeutic effect models

The primary clinical endpoints in cancer treatment are progression-free survival (PFS) and overall survival (OS). However, earlier (intermediate/surrogate) measures are needed to make early decisions both in drug development [Bruno 2009] and in the treatment of an individual patient. These are lesion size and count, as well as levels of various tumour markers (PSA, hCG, CEA, sKIT, sVEGFR-3, etc.). There are a growing number of studies having identified pharmacodynamic effect relationships between:

- drug doses and tumour response [Claret 2006, Claret 2009, Tham 2008, Stein 2011, Frances 2011],
- systemic drug exposure and tumour response [Houk 2009, Lu 2010, Joerger 2010],
- systemic drug exposure and survival [Houk 2009, Joerger 2010, You 2008],
- change in tumour size (as well as tumour size baseline and performance status) and survival [Claret 2006, Claret 2009, Wang 2009, Lindbom 2009],
- systemic drug exposure and efficacy biomarkers [Jadhav 2006, Tornoe 2007],
- biomarker dynamics and response measures [You 2009, You 2010, Hansson 2011].

It is expected that early assessment of treatment antitumour efficacy and prediction of expected survival, using tumour growth inhibition and survival models, will be beneficial in drug development (screening of candidate drugs, simulation of trials, optimization of trial designs) [Wang 2009, Bruno 2009, Claret 2010].

The model of Wang *et al.* has shown its good predictivity of survival in the investigation of carboplatin + paclitaxel regimen and erlotinib in the treatment of non-small cell lung cancer [Bruno 2009].

Toxicity models

Hematotoxicity models

Neutropenia is the most threatening of the common adverse effects of chemotherapies, and therefore it is closely monitored and is the main dose-limiting toxicity for numerous cytotoxic agents. A relationship between drug exposure and neutropenia has been shown for many anticancer drugs. Simple empirical models were used to relate a summary variable of drug exposure and a summary variable of neutropenia (nadir, grade of neutropenia, percentage decrease in leukocytes or their count under a threshold) [Testart 2007].

However, myelosuppression is a complex transitional process (from the drug effect on bone marrow till reduction of the circulating blood cell count) and a static summary variable cannot describe its full dynamics. More mechanistic models reflecting the physiological processes have been developed lately, with good consistency among drugs [Friberg 2002, Sandstrom 2005, van Kesteren 2005, Kloft 2006, Latz 2006]. Such an analysis of pemetrexed data lead to the suggestion that using folic acid, vitamin B_6 and vitamin B_{12} supplementation may reduce the risk of severe neutropenia induced by pemetrexed [Latz 2006a, Latz 2006b]. The model has been extended to take into account the administration of granulocyte colony stimulating factor (G-CSF) [Sandstrom 2006], effective for the reduction and prevention of severe neutropenia [Aapro 2011]. Another extension of the model for a simultaneous analysis of leukocytes and neutrophils has been recently developed [Quartino 2010]. Improvements for these semi-physiologic myelosuppression models were proposed to better account for the sensitivity of the response to dose changes [Meille 2008].

Applications of hematoxicity models to individualize dosing regimens

The early dose adaptation strategies targeted a chosen interval of a summary variable of hematotoxicity (usually nadir), using a regression of a summary drug exposure variable and other predictive factors, such as baseline neutrophil value, performance status, prior treatment, bone marrow function, etc. [de Jonge 2005, Kaestner 2007b].

Recently, Wallin *et al* have developed a neutrophil-guided dose adaptation tool in Excel based on a semimechanistic myelosuppression model and the Bayesian MAP technique to be used in clinical practice [Wallin 2009]. This tool uses the information of the neutrophil count after the first treatment cycle to select the subsequent doses. The authors did a simulation study to compare the neutrophil nadir after the second treatment course with either no dose adjustment after the first course, standard dose reduction by 25% if grade 4 neutropenia occurred at the

first cycle, or model prediction-based dose adjustment targeting a nadir of 10^9 cells/l with different types and amounts of information used to estimate the individual random effects of the neutropenia model [Wallin 2010].

Barbolosi and Iliadis used a model incorporating drug pharmacokinetics, antitumour and hematotoxic effects to develop an optimized dosing regimen by simulation studies [Iliadis 2000, Barbolosi 2001, Barbolosi 2003]. This model was associated with constraints on the plasma concentrations, drug exposure, and leukopenia. The resulting recommended regimen for etoposide consisted of an initial high-dose chemotherapy up to saturation of constraints associated with hematotoxicity and a maintenance continuous infusion at a rate that keeps the leukocyte count at its lower limit while decreasing the tumour cell population.

A similar framework was developed for docetaxel and epirubicin plus G-CSF in the treatment of metastatic breast cancer; a simulation study was done to investigate optimal time and duration of G-CSF administration [Freyer 2005]. It was concluded that the regimen of 5 consecutive injections starting at day 5 was not inferior to usual 10-15 injections schedule.

The hematotoxicity model proposed by Meille *et al.* [Meille 2008] extended to take into account the effect of G-CSF was prospectively applied in a phase I trial to individualize densified docetaxel + epirubicin administration (dose rate and duration) in the treatment of metastatic breast cancer [You 2007]. This approach aimed to optimize the drug dosing schedule at different dose levels by respecting *a priori* fixed constraints on severity and duration of neutropenia and thrombocytopenia. Such a control was successful and allowed to safely increase the dosing density.

Non-hematological toxicity models

In contrast to numerous models for chemotherapy-induced hematotoxicity, there are very few published models of other toxicities [Xie 2002, Henin 2009, Keizer 2010]. Although not as dangerous as severe hematoxicity, other adverse effects (gastrotoxicites, dermatological, cardiovascular, neurotoxicities) may reach intolerable levels and lead to treatment delays, dose reductions and discontinuations. Standard empirical dose reductions by 25% or 50% after several events with severe toxicity may be not the best strategy to an optimal clinical outcome. Individual model prediction-based dose adaptation is expected to be more advantageous. Since the effect is observed without invasive sampling, and if a model relates it to drug doses (instead of systemic exposure), such methods do not impose the inconvenience of additional blood samples for patients and clinical teams.

CONTRIBUTIONS OF THIS THESIS

1. Review of pharmacodynamic models for discrete data

Clinical outcomes are often described as discrete data: events (death, stroke, epileptic seizure, multiple sclerosis lesions, disease progression, severe toxic side-effect, etc.), their time-to-event, count (rate), and severity grade. Each type of discrete data requires specific modelling structures and methods. A part of this thesis was to review the most common modelling approaches for categorical, count and time-to-event data, together with published applications of such models to analyse pharmacodynamic data. Useful applications of pharmacodynamic modelling include identification of influential factors related to the clinical outcome, characterization and quantification of their impact, making better informed predictions and clinical decisions, assessments of efficacy of therapeutic interventions, optimising the individual treatments and drug development studies.

2. *A posteriori* dose individualization of capecitabine based on ordered categorical toxicity model predictions

One of such mixed-effects pharmacodynamic models for ordinal data has been applied in this thesis to individually adjust the doses of capecitabine in the treatment of metastatic or advanced colorectal cancer to reduce the frequency and duration of hand-and-foot syndrome (HFS). More than half of patients treated with capecitabine develop this dermatologic toxicity, which affects the skin of palm and soles and may develop to blisters, severe pain and difficulties in normal everyday activities like walking and handling objects (Table 1).

Grade		**0**	**1**	**2**	**3**
Symptoms	**Pain**	None	Tingling or burning	Pain	Severe pain
	Skin damage	None	Minimal skin changes or dermatitis (e.g., erythema)	Skin changes (e.g., peeling, blisters, bleeding, edema)	Ulcerative dermatitis, interfering with function

Table 1: Hand-and-foot syndrome severity grades.

Therefore, even if not life-threatening, this toxicity (at the higher grades 2 and 3) is dose-limiting and requires dose interruptions and reductions. The standard approach used in clinical

practice is to interrupt the treatment until recovery to at least grade 1 and then continue the treatment with reduced doses (Table 2).

Grade	Number of events			
	1	**2**	**3**	**4**
2	100%	75%	50%	0
3	75%	50%	0	0

Table 2: Dose reductions after events with intolerable grade of HFS.

This "rigid" approach is probably not optimal and it was interesting to investigate the feasibility of a subject-specific dose adaptation based on predictions given by a toxicity model, and to compare its performance to that of the standard method by simulations of virtual clinical trials.

Recently, population models for both the adverse effect HFS and the tumour growth inhibition (TGI) effect of capecitabine in the treatment of advanced or metastatic colorectal cancer were published [Henin 2009, Claret 2009]. These models were used in this simulation study concerning individual Bayesian *a posteriori* dose adaptations. The stages of this work were:

- to develop the concept and numerical techniques to determine the "best" next dose on the basis of the individual model prediction of the risk of intolerable (grade ≥2) HFS over the next treatment cycle (including the estimation of individual random effects of the HFS model, estimation of the average weekly toxicity risk over the next 3 weeks taking into account that each week's risk is conditional on the previous week's HFS grade, algorithms for the search of the dose whose risk is closest (without exceeding) to the targeted risk);
- to define the clinical trial protocol, the clinically relevant modalities of dose adaptations, the criteria for comparison of the different dose adaptation methods, with the consultation of a clinical oncologist;
- to implement the HFS and TGI models together with the different dose adaptation approaches into a clinical trial simulation code;
- to perform the simulations with various specifications of the dose search procedure in order to determine by sensitivity analysis the optimal "calibration" of the individual prediction-based dose adaptation, that is one that would result in the biggest reduction of HFS toxicity without affecting the antitumour efficacy as compared to the standard dose reduction approach.

In order to individually adjust the capecitabine doses for a particular patient so that the risk of hand-and-foot syndrome can be reduced and/or the chances of anticancer effect increased,

predictions of his/her risks are required. Data gathered from other patients with the same disease taking the same drug were used to construct a so-called population model of the adverse effect HFS. A particular feature of such "population" models is that they contain "population" parameters, representing the common effects, and the parameters representing individual differences, which can be estimated from the population model and the patient's data collected in the course of treatment. Those so-called empirical Bayes estimates (EBEs) of individual effects are used to adjust the population model for each particular patient. The individualized model can then be used to predict the patient's risk of severe HFS during the next treatment cycle. The next dose to take can be calculated according to the predicted risk and predefined criteria of what level of risk defines the "optimal" dose.

The model used to describe the HFS is a *mixed-effects transitional proportional odds* model for longitudinal ordinal data, i.e. it:

- does not model the observed grades themselves but their probabilities; to be more exact it is the *logit* transformation of the probability of being in a cumulative set of categories which is modeled. The *logit* transformation assures that the predicted probability lies in the interval from 0 to 1; it is the *log* of *odds,* which is a ratio of the probability of the event (grade) and of the probability of the opposite event. Because of this required duality of the events, it is the *cumulative* probabilities that are modeled, that is an event is defined as $\leq$ G (where G is a particular grade, one of 0, 1, 2,…), its opposite then is $>$ G. The probabilities of particular grades are then calculated by subtracting the probabilities of being in adjacent cumulative categories: $P(Y = G) = P(Y \leq G) - P(Y \leq G\text{-}1)$, where G is a particular grade;
- uses the assumption of *proportional* effects of predictive variables for all grades (therefore, *proportional odds* model);
- assumes the probabilities of the next week's observation to be dependent on the current week's grade, therefore is called a *transitional* (or *Markov*) model;
- contains typical (population) parameters and individual-specific effects (therefore, *mixed-effects* model).

The best achieved result was an average reduction of about 10 days in the total duration of grade $\geq$2 HFS over the trial duration (max 30 weeks of treatment + 4 weeks post-treatment monitoring). It was a result of less frequent and shorter reoccurring events with intolerable HFS. Consequently, there were fewer dropouts due to toxicity. This was the best found result in terms of HFS reduction while preserving the same antitumour efficacy.

It would be worthwhile and interesting to investigate the possible correlation between the pharmacodynamics of the HFS and the antitumour effects; however, the efficacy data was not available for the HFS model authors and therefore such a relationship could not be investigated.

In conclusion, it was shown that individualized dose adaptation on the basis of ordinal observations of toxicity can be feasible and beneficial in terms of overall clinical outcome. In the used example of HFS, the impact was limited by the low sensitivity of the syndrome to dose changes. It is expected that the impact of such a method would be higher for reversible toxicities with faster dynamics (e.g. gastrointestinal).

A new pharmacogenetics-based perspective to optimize capecitabine dosing regimens arose recently, as polymorphisms in genes of enzymes involved in the transformation of capecitabine into 5-fluorouracil have been found to be significantly associated with efficacy and toxicity. A polymorphism in cytidine deaminase gene was associated to increased risk of severe HFS [Caronia 2011] and a carboxylesterase 2 gene polymorphism was identified as a predictor of capecitabine response and time to progression [Ribelles 2008]. Future work on dose adaptation should take into account these polymorphisms.

The only covariate identified in the HFS model by their authors from the data available to them was the creatinine clearance as calculated by Cockcroft-Gault formula [Henin 2009]. It has been tested on both of the interindividual variability parameters (K and baseline logit). As creatinine clearance reflects the renal function, it was expected that it would explain the variability in the kinetics parameter K. However, it was much more significant as a covariate of baseline logit (baseline risk of HFS), as discussed in [Henin 2009].

In the KPD model, the parameter K governs both the drug accumulation and the rate of effect manifestation, so it is not a purely PK parameter. The relation of the creatinine clearance and the risk of HFS may be linked to the hypothesis about the endogenous substances implicated in the HFS; the renal function may determine the elimination of those endogenous substances and thus influence the development of HFS. The pathophysiology of HFS is not well understood yet.

The HFS model predicts the risk of grade ≥2 to be lower for a patient in grade 1 than for a patient in grade 0. Based on the observed data from Phase III studies used to build the model (cf. Figure 1), the G0→G2 transition rates were higher than G1→G2 transition rates (0.032 vs. 0.014). (G0→G1 rate: 0.021)

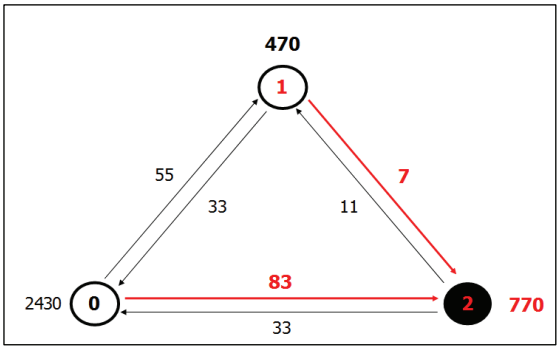

Figure 1: Transitions between grades, as observed in 600 patients in Phase III studies. The circles represent the HFS grades 0, 1, ≥2. The numbers beside arrows and circles represent the counts of observed transitions (once per week).

In clinical practice, it is also observed that some patients develop the grade 2 toxicity apparently straight from grade 0 (no toxicity). One hypothesis is that there may be a subpopulation of patients more sensitive to capecitabine (5FU) due to deficiency of metabolising enzymes (DPD). In contrast, patients who remain in grade 1 for a longer (observable) time seem to be less likely to develop more severe HFS. This hypothesis was tested during model development (mixture of 2 populations: sensitive and non-sensitive), but the available data did not support it.

Another issue may be concerning the grading of the HFS. Attribution of grade 1 or 2 to the intensities of redness, swelling, pain and interference with daily activities is highly subjective. Though physiologically the transition from grade 0 to 2 or higher must pass via grade 1, the duration of grade 1 may be very short and therefore not reported in the dataset. The HFS model assumed once weekly observation consisting of the highest grade during the week, therefore omission of grade 1 was likely.

In case of such missing observations in the dataset, latent variable models might be useful. There a continuous latent (unobservable) response variable is assumed, which is then translated to observable categories via thresholds of its values [Hutmacher 2008, Skrondal 2007]. This methodology should be investigated in the future research involving categorical data.

3. Empirical Bayes estimates of random effects of mixed-effects models for ordinal data

As a part of development of the above introduced dose individualization approach, the quality of empirical Bayes estimates (EBEs) of the individual random effects (that adjust the population model to the individual patient on the basis of their previous observations) was checked. As the observed accuracy and precision of EBEs were rather low, it motivated to investigate their potentially influential factors. The quality of EBEs has a direct impact on the correctness of the predictions of individual risk of severe HFS toxicity, and thus is an important condition for the performance of the prediction-based dose adaptation methods.

Therefore, this work consisted of investigating the influence of:

- the estimator (mode vs. mean of the *a posteriori* distribution of random effects);
- the optimisation algorithm, where local (simplex, quasi-Newton) and global (adaptive random search) optimum search algorithms were compared;
- the amount of data per patient;
- the distribution of categories within patients;
- the magnitude of the inter-individual variability;
- the values of the effect model parameters.

In this study, it was found that the main factors affecting the quality of EBEs were the values of parameters governing the dose-response relationship and the within-subject distribution of categories.

In the HFS model example, the values of the effect-driving variable were too low to well identify the E_{max} function parameters, the within-subject distributions of categories were far from uniform. The EBEs had some bias and low precision, the impact on the clinical decision concerning the next dose to take was variable but considerable. In extreme cases, the impact of incorrect EBEs on the predicted risk of grade ≥2 can be 4-fold; the corresponding doses calculated according to the true and to the estimated risk prediction can differ 10-fold. However, this almost did not have any impact on the overall performance of the dose adaptation method due to low sensitivity of the HFS to dose changes: the results were very similar with the true (simulated) values of random effects, with their EBEs, and without random effects (that is using the population model for the average patient).

In general, the practical interest of knowing how good the EBEs are, is their direct relation to the quality of predictions and thus dosage decision based on that individual prediction. For a real patient, the quality of his/her EBEs cannot be exactly known (because the true values of individual random effects are not known).

An assessment of the quality of EBEs can be based on the information about the main factors known to influence the estimates: the range of individual's predictor variable (dose, exposure, etc.), the number of repeated observations and distribution of observed toxicity grades.

More precise evaluation of general/population EBE quality can be obtained by simulations from the model and comparison of EBEs and simulated random effect values. Awareness of the quality of EBEs can then help decide whether to use them for predictions and dosage decisions.

Potential ways to influence the quality of EBEs and therefore individual predictions:

- make sure that all grades are recorded (→frequent enough observations) and no information is thrown away, e.g. by transforming the actual observations into one (maximal) per week;
- investigate if a simpler model (with less parameters) would give better individual predictions, even if the fit is worse, since a model which best fits the data is not necessarily the one who gives best predictions;
- in case the original dataset contains the times of grade changes (as opposed to one grade per time unit), investigate if a continuous-time transition model would be more predictive.

4. Population PKPD of hydroxyurea in sickle cell anemia patients

The fourth work of this thesis concerned the chemotherapeutic drug hydroxyurea (HU) in the treatment of sickle cell anemia, one of the most common genetic diseases. It is caused by a mutation in the hemoglobin gene and characterized by rigid sickle-shaped red blood cells, which may lead to vaso-occlusion with various complications, such as acute painful crises, ischemia and damage of various organs, acute chest syndrome or stroke.

The objectives of this work were:

- to develop population PK-PD models for hydroxyurea in order to characterize the dose-exposure-response relationships and their variability, investigate for possible covariates;
- to use these models to compare two dosing regimens (one continuous daily and the other with interruptions of 2 days after every 5 days) by simulation;
- to develop some recommendations for monitoring the treatment.

Since no dose-limiting toxicity occurred in this study, no toxicity model and related *a posteriori* dose adjustment tool was developed. Efficacy was measured on two biomarkers: fetal hemoglobin percentage (HbF%) and mean corpuscular volume (MCV).

The change in MCV was more rapid than that in HbF%, the interindividual variability of the steady-state (SS) values of HbF% was much greater than that of MCV. Consequently, the SS value of MCV at month 3 was not predictive of the SS HbF% value at month 26. Therefore, the HbF% level, which is also directly related to the relief of sickle cell disease symptoms, may be the better biomarker for monitoring HU treatment.

Concerning the comparison of the two dosing regimens, the difference was very small for the MCV profile, but larger for the HbF% profile, especially for patients reaching the highest

levels of HbF%. For these patients who have the strongest response to the drug, continuous dosing may be more beneficial in terms of increase in HbF% than the intermittent schedule.
The estimate of effect intensity suggested an exciting perspective that HbF could be further increased by more potent drugs or by drug combinations.
Unfortunately, no covariates explaining a significant part of variability could be identified from this dataset, no data concerning compliance or genetic polymorphisms found to be associated with the response to HU was available. Information about polymorphisms in genes regulating HU metabolism or transporters, HbF expression and erythroid progenitor proliferation would potentially adjust the dosing schedule *a priori*.

DISCUSSION

Cancer chemotherapy is a delicate balance between anticancer efficacy and toxicity on healthy tissues. In order to receive the highest benefit from the available treatments, the potentially high variability in response among patients has to be handled properly. Various "personalized medicine" approaches have been and are developed to individually tailor the treatments and their dosing regimens to patients: based on genetic and phenotypic markers, indicators of drug disposition functions, monitoring of drug exposure or effects, etc.

Neither *a priori*, nor *a posteriori* methods are sufficient and optimal alone in all cases; they are complementary in their roles: *a priori* knowledge of high risk of severe toxicity or of lack of efficacy may help avoid these serious consequences, while *a posteriori* dose adjustments may allow to achieve the best possible clinical benefit from the treatment. For *a priori* decisions, the role of pharmacogenomics is rapidly growing and promising. For *a posteriori* dose adjustments, Bayesian method using population PK-PD analyses is the method of choice. A dosing scheme combining both *a priori* and *a posteriori* approaches has been proposed by [Gao 2008].

Dose-individualization strategies have been developed for many anticancer drugs. However, there is a gap between the research and the clinical practice; only a few examples are regularly used in clinical setting. Large prospective randomized clinical trials comparing the individualized and standard dosing approaches are needed to demonstrate whether individualized dosing results in a clinically significant therapeutic benefit (decreased toxicity and/or improved therapeutic outcome). Most of the reported prospective studies have been rather small, though useful in assessing the feasibility of individualized dosing protocols [de Jonge 2005, Zandvliet 2008].

Moreover, as there may be increased costs in terms of assays, logistics, clinical staff training and work time, measurement related supplemental hospitalization, the clinical benefit has to be substantial enough to outweigh the additional costs and demonstrate cost-effectiveness of the individualized approach.

Wide-spread implementation is likely to be further hindered by practical issues concerning technical feasibility, where required costly analytical equipment may not be available. Moreover, TDM may require special training (for assay analysis, specialised computer software and interpretation of results) and a considerable collaborative effort of the whole clinical service team (oncologists, pharmacists, nurses, laboratory technicians, etc.) [Hon 1998].

In some cases, TDM may lead to some inconvenience for the patients: more visits to the hospital, more blood samples (for PK or hematotoxicity-based), more care for compliance with orally taken drugs when dosing regimens are changing. There are also ethical issues concerning the handling and access to genetic data [Squassina 2010].

Considering all the practical and economical constraints for launching large prospective randomized clinical trials to demonstrate the clinical benefit of individual dose adaptations, virtual studies by computer simulations offer a particularly useful assessment of possible impacts and constitute a valuable aid in planning and making decisions. It allows to compare different dosing regimens, determine the most influential factors by sensitivity analyses, determine the impact on various endpoints on different scales, evaluate and compare different trial designs (sample size, dosing regimens, measurement times, etc.), assess the impacts in special populations (pediatrics, hepatic or renal impairment, etc.), while taking into account the variabilities and uncertainties at various levels. An important benefit of clinical trial simulation is the necessity to explicitly state all the assumptions related to the trial. Therefore, it may help reduce risks, time and monetary costs, as well as ethical concerns.

However, simulations require the specification of mathematical models which are a simplification of reality, representing only what we already know about the drug, disease and factors affecting variability, and thus reproduce only the known outcomes. While protocol violations such as dropout and non-compliance can be reproduced in the simulations, manifestation of unexpected rare events cannot. Therefore, simulations of clinical trials are not meant to replace real clinical studies, but to complement and to help optimize them.

Population PK-PD analyses are strongly encouraged, as they may be highly beneficial in enabling:

- to integrate information and knowledge from various sources (blood samples, tissue samples, imagery, animals, in vitro, genetics, knowledge about similar compounds), at various effect scales (cellular, tissue, organ, disease, survival);
- to identify relevant relationships between drug doses, exposures, effects;
- to quantify different types of variabilities in such relationships, identify factors useful in explaining them, and quantify their relevance for clinical outcome;
- to investigate for drug interactions, optimal regimens for combined therapies common in oncology;
- to make better predictions about future or unstudied situations (extrapolate);
- to develop individual dose optimization strategies.

More mechanistic models, better reflecting the physiology, are needed for *in silico* analyses more sophisticate than mere description of observed data.

In order to be able to extract knowledge from data, they have to be extensive, precise and informative. Growing availability of advanced medical imaging techniques, such as PET or PET-CT scans, to measure malignant tissues, is expected to facilitate the identification of drug dose/exposure-efficacy relationships.

To be able to identify well the dose/exposure-response relationship (therapeutic or toxic), a wide range of doses/exposures has to be tested, from those having weak (subtherapeutic) effect to (quasi)maximal effect (too toxic).

Concerning the identification of relationships between doses/exposures and categorical toxicities, it is very important that all grades (including the "tolerable" grade 1) are rigorously reported to well inform the dynamics of the toxic effect. Hidden Markov models may be an interesting option to investigate, to account for possible imprecisions/heterogeneity in toxicity grade reporting.

For the *a posteriori* dose modifications to be effective, the toxic endpoint has to be sensitive enough to dose changes. In addition, it must not to be much influenced by unknown (uncontrollable) factors (i.e. have low interoccasion variability), so that predictions of future risk are sufficiently accurate. Moreover, the model used for TDM has to be able to reproduce realistically that sensitivity to dose changes. Extensive validations of the models are needed.

The EBE quality is an issue that may be difficult to overcome, especially for toxicities evaluated only by grades, since a wide range of drug exposures, high number of observations, uniform within-patient distributions of grades are required for good estimation but not very likely in clinical reality.

Advances in science, technology, understanding of genetic and molecular mechanisms of cancer and effect of treatments on malignant and healthy tissues bring possibilities to better tailor treatments and their dosing regimens for individual patients, so that best possible therapeutic effect is attained with least of adverse effects. Various complementary "personalized medicine" tools have been and are developed to allow earlier and better informed clinical decisions. Many still are at exploratory stage, encountering diverse technical and practical difficulties on their way to clinical implementation, but promising significant improvements in optimizing healthcare of cancer patients.

REFERENCES

Aapro MS, Bohlius J, Cameron DA et al. 2010 update of EORTC guidelines for the use of granulocyte-colony stimulating factor to reduce the incidence of chemotherapy-induced febrile neutropenia in adult patients with lymphoproliferative disorders and solid tumours. Eur J Cancer. 2011; 47(1):8-32.

Amado RG, Wolf M, Peeters M et al. Wild-type KRAS is required for panitumumab efficacy in patients with metastatic colorectal cancer. J Clin. Oncol 2008; 26(10):1626-34.

Ando Y, Minami H, Saka H, et al. Therapeutic drug monitoring in 21-day oral etoposide treatment for lung cancer. Jpn J Cancer Res. 1996; 87:856-61.

Bach DM, Straseski JA, Clarke W. Therapeutic drug monitoring in cancer chemotherapy. Bioanalysis. 2010; 2(5):863-79.

Barbolosi D, Iliadis A. Optimizing drug regimens in cancer chemotherapy: a simulation study using a PK-PD model. Comput Biol Med. 2001; 31(3):157-72.

Barbolosi D, Freyer G, Ciccolini J, et al. Dosage regimen optimization in cancer chemotherapy using a mathematical model. Bull Cancer 2003; 90:167-75.

Bates DW, Soldin SJ, Rainey PM, et al. Strategies for physician education in therapeutic drug monitoring. Clin Chem 1998; 44:401-7.

Black AJ, McLeod HL, Capell HA, et al. Thiopurine methyltransferase genotype predicts therapy-limiting severe toxicity from azathioprine. Ann Intern Med 1998; 129:716-8.

Bleyzac N, Souillet G, Magron P, et al. Improved clinical outcome of paediatric bone marrow recipients using a test dose and Bayesian pharmacokinetic individualization of busulfan dosage regimens. Bone Marrow Transplant. 2001; 28:743-51.

Bocci G, Di Paolo A, Barbara C, et al. Pharmacokinetics, a main actor in a many-sided approach to severe 5-FU toxicity prediction. Br J Clin Pharmacol. 2008; 67(1):132-4.

Bolinger AM, Zangwill AB, Slattery JT, et al. Target dose adjustment of busulfan in pediatric patients undergoing bone marrow transplantation. Bone Marrow Transplant. 2001; 28(11):1013-8.

Bosch TM, Meijerman I, Beijnen JH, et al. Genetic polymorphisms of drug-metabolising enzymes and drug transporters in the chemotherapeutic treatment of cancer. Clin Pharmacokinet. 2006; 45(3):253-85.

Bruno R, Lu J, Sun Y, et al. Simulation of survival with first and second-line non-small cell lung cancer (NSCLC) therapy using a public domain drug-disease modeling framework. J Clin Oncol. 2009; 27: Abstr 8087.

Bruno R, Claret L. On the use of change in tumor size to predict survival in clinical oncology studies: toward a new paradigm to design and evaluate phase II studies. Clin Pharmacol Ther. 2009; 86(2):136-8.

Bruno R, Lu JF, Sun YN, et al. A modeling and simulation framework to support early clinical drug development decisions in oncology. J Clin Pharmacol. 2011; 51(1):6-8.

Calvert AH, Newell DR, Gumbrell LA, et al. Carboplatin dosage: prospective evaluation of a simple formula based on renal function. J Clin Oncol. 1989; 7(11):1748-56.

Canal P, Chatelut E, Guichard S. Practical treatment guide for dose individualisation in cancer chemotherapy. Drugs. 1998; 56:1019-38.

Cano JP, Bruno R, Lena N, et al. Dosage predictions in high-dose methotrexate infusions. Part 1: Evaluation of the classic test-dose protocol. Cancer Drug Deliv. 1985; 2:271-6.

Caronia D, Martin M, Sastre J, et al. A polymorphism in the cytidine deaminase promoter predicts severe capecitabine-induced hand-foot syndrome. Clin Cancer Res. 2011; 17(7):2006-13.

Chatelut E, Canal P, Brunner V, et al. Prediction of carboplatin clearance from standard morphological and biological patient characteristics. J. Natl Cancer Inst. 1995; 87(8):573-80.

Chatelut E, Boddy AV, Peng B, et al. Population pharmacokinetics of carboplatin in children. Clin Pharmacol Ther. 1996; 59:436-43.

Claret L, Girard P, O'Shaughnessy J, et al. Model-based predictions of expected anti-tumor response and survival in phase III studies based on phase II data of an investigational agent. J Clin Oncol. J Clin Oncol. 2006; 24: Abstr 6025.

Claret L, Girard P, Zuideveld KP, et al. A longitudinal model for tumor size measurements in clinical oncology studies. PAGE 15 (2006) Abstr 1004 [www.page-meeting.org/?abstract=1004]. Last accessed 2011-06-24.

Claret L, Girard P, Hoff PM, et al. Model-based prediction of phase III overall survival in colorectal cancer on the basis of phase II tumor dynamics. J Clin Oncol. 2009; 27:4103-8

Claret L, Lu JF, Sun YN, et al. Development of a modeling framework to simulate efficacy endpoints for motesanib in thyroid cancer patients. Cancer Chemother Pharmacol. 2010; 66(6):1141-9.

D'Argenio DZ. Optimal sampling times for pharmacokinetic experiments. J Pharmacokinet Biopharm. 1981; 9:739-56.

de Jonge ME, Huitema AD, Schellens JH, et al. Individualised cancer chemotherapy: strategies and performance of prospective studies on therapeutic drug monitoring with dose adaptation: a review. Clin Pharmacokinet. 2005; 44:147-73.

Desoize B, Robert J. Individual dose adaptation of anticancer drugs. Eur J Cancer. 1994; 30A (6): 844-51.

Di Fiore F, Blanchard F, Charbonnier F, et al. Clinical relevance of KRAS mutation detection in metastatic colorectal cancer treated by cetuximab plus chemotherapy. Br J Cancer. 2007; 96:1166-9.

Ekhart C, de Jonge ME, Huitema AD, et al. Flat dosing of carboplatin is justified in adult patients with normal renal function. Clin Cancer Res. 2006; 12:6502-8.

Evans WE, Relling MV. Clinical pharmacokinetics-pharmacodynamics of anticancer drugs. Clin Pharmacokinet. 1989; 16(6):327-36.

Evans WE, Relling MV, Rodman JH, et al. Conventional compared with individualized chemotherapy for childhood acute lymphoblastic leukemia. N Engl J Med. 1998; 338:499-505.

Evans WE. Pharmacogenetics of thiopurine S-methyltransferase and thiopurine therapy. Ther Drug Monit. 2004; 26(2):186-91.

Frances N, Claret L, Bruno R, et al. Tumor growth modeling from clinical trials reveals synergistic anticancer effect of the capecitabine and docetaxel combination in metastatic breast cancer. Cancer Chemother Pharmacol. 2011 Apr 8.

Freyer G, Tredan O, Meille C, et al. Optimal duration of G-CSF prophylaxis after docetaxel (D) + epirubicin (E) in metastatic breast cancer (MBC): prediction using a mathematical model. J Clin Oncol. 2005; 23: Abstr 2044.

Friberg LE, Henningsson A, Maas H, et al. Model of chemotherapy-induced myelosuppression with parameter consistency across drugs. J Clin Oncol. 2002; 20:4713-21.

Gamelin E, Boisdron-Celle M, Delva R, et al. Long-term weekly treatment of colorectal metastatic cancer with fluorouracil and leucovorin: results of a multicentric prospective trial of fluorouracil dosage optimization by pharmacokinetic monitoring in 152 patients. J Clin Oncol. 1998; 16:1470-8.

Galpin AJ, Evans WE. Therapeutic drug monitoring in cancer management. Clin Chem. 1993; 39:2419-30.

Gao B, Klumpen HJ, Gurney H. Dose calculation of anticancer drugs. Expert Opin Drug Metab Toxicol. 2008; 4(10):1307-19.

Genetics. [http://www.breastcancer.org/risk/factors/genetics.jsp]. Last accessed 2011-06-24.

Gibaldi M. Revisiting some factors contributing to variability. Ann Pharmacother. 1992; 26: 1002-7.

Gurney H. Dose calculation of anticancer drugs: a review of the current practice and introduction of an alternative. J Clin Oncol. 1996; 14:2590-611.

Hansson EK, Westwood P, Houk B, et al. PKPD modeling of VEGF, sVEGFR-2, sVEGFR-3 and sKIT as biomarkers of tumor response and overall survival following sunitinib treatment in GIST. PAGE 2011 presentation [www.page-meeting.org/?abstract=2183].

Henin E, You B, VanCutsem E, et al. A dynamic model of hand-and-foot syndrome in patients receiving capecitabine. Clin Pharmacol Ther. 2009; 85:418-25.

Holford NH. Target concentration intervention: beyond Y2K. Br J Clin Pharmacol. 1999; 48(1):9-13.

Hon YY, Evans WE. Making TDM work to optimize cancer chemotherapy: a multidisciplinary team approach. Clin Chem. 1998; 44:388-400.

Hoskins JM, Goldberg RM, Qu P, et al. UGT1A1*28 genotype and irinotecan-induced neutropenia: dose matters. J Natl Cancer Inst. 2007; 99:1290-5.

Houk BE, Bello CL, Poland B, et al. Relationship between exposure to sunitinib and efficacy and tolerability endpoints in patients with cancer: results of a pharmacokinetic/ pharmacodynamic meta-analysis. Cancer Chemother Pharmacol. 2010; 66(2):357-71.

Hutmacher MM, Krishnaswami S, Kowalski KG. Exposure-response modeling using latent variables for the efficacy of a JAK3 inhibitor administered to rheumatoid arthritis patients. J Pharmacokinet Pharmacodyn. 2008; 35: 139-57.

Iliadis A, Barbolosi D. Optimizing drug regimens in cancer chemotherapy by an efficacy-toxicity mathematical model. Comput Biomed Res. 2000; 33(3):211-26.

Iyer L, Das S, Janisch L, et al.: UGT1A1*28 polymorphism as a determinant of irinotecan disposition and toxicity. Pharmacogenomics J. 2002; 2:43-7.

Jadhav PR, Agerso H, Tornoe CW, et al. Semi-mechanistic pharmacodynamic modeling for degarelix, a novel gonadotropin releasing hormone (GnRH) blocker. J Pharmacokinet Pharmacodyn. 2006; 33:609-34.

Jodrell DI, Reyno LM, Sridhara R, et al. Suramin: development of a population pharmacokinetic model and its use with intermittent short infusions to control plasma drug concentration in patients with prostate cancer J Clin Oncol. 1994; 12:166-75.

Kaestner SA, Sewell GJ. Chemotherapy dosing part I: scientific basis for current practice and use of body surface area. Clin Oncol. 2007; 19:23-37.

Kaestner SA, Sewell GJ. Chemotherapy dosing part II: alternative approaches and future prospects. Clin Oncol. 2007; 19:99-107.

Keizer RJ, Gupta A, Mac Gillavry MR, et al. A model of hypertension and proteinuria in cancer patients treated with the anti-angiogenic drug E7080. J Pharmacokinet Pharmacodyn. 2010; 37(4):347-63.

Kloft C, Wallin J, Henningsson A, et al. Population pharmacokinetic–pharmacodynamic model for neutropenia with patient subgroup identification: comparison across anticancer drugs. Clin Cancer Res. 2006; 12(18):5481-90.

Kobayashi K, Ratain MJ. Individualizing dosing of cancer chemotherapy. Semin Oncol. 1993; 20(1):30-42.

Kranenburg O. The KRAS oncogene: past, present, and future. Biochim Biophys Acta 2005; 1756(2):81-2.

Latz JE, Karlsson MO, Rusthoven JJ, et al. A semimechanistic–physiologic population pharmacokinetic/ pharmacodynamic model for neutropenia following pemetrexed therapy. Cancer Chemother Pharmacol. 2006; 57(4):412-26.

Latz JE, Rusthoven JJ, Karlsson MO, et al. Clinical application of a semimechanistic-physiologic population PK/PD model for neutropenia following pemetrexed therapy. Cancer Chemother Pharmacol. 2006; 57(4):427-35.

Lennard L. Therapeutic drug monitoring for cytotoxic drugs. Br J Clin Pharmacol. 2001; 52 Suppl 1:75S-87S.

Lichtman SM, Villani G. Chemotherapy in the elderly: pharmacologic considerations. Cancer Control. 2000; 7(6):548-56.

Lievre A, Bachet JB, Le Corre D, et al. KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer. Cancer Res. 2006; 66:3992-5.

Lievre A, Bachet JB, Boige V, et al. KRAS mutations as an independent prognostic factor in patients with advanced colorectal cancer treated with cetuximab. J Clin Oncol. 2008; 26:374-9.

Lindbom L, Claret L, Andre V, et al. A drug independent tumor size reduction-survival model in advanced ovarian cancer to support early clinical development decisions. American Conference of Pharmacometrics 2009 poster. [http://2009.go-acop.org/sites/all/assets/webform/ACoP%202009%20Poster%20-%20Lars%20Lindbom.pdf]. Last accessed 2011-06-24.

Loh GW, Ting LS, Ensom MH. A systematic review of limited sampling strategies for platinum agents used in cancer chemotherapy. Clin Pharmacokinet. 2007; 46(6):471-94.

Lu JF, Claret L, Sutjandra L, et al. Population pharmacokinetic/pharmacodynamic modeling for the time course of tumor shrinkage by motesanib in thyroid cancer patients. Cancer Chemother Pharmacol. 2010; 66(6):1151-8.

Masson E, Zamboni WC. Pharmacokinetic optimization of cancer chemotherapy: effect on outcomes. Clin Pharmacokinet. 1997; 32:324-43.

Mathijssen RH, de Jong FA, Loos WJ, et al. Flat-fixed dosing versus body surface area based dosing of anticancer drugs in adults: does it make a difference? Oncologist 2007; 12(8):913-23.

McLeod HL, Siva C. The thiopurine S-methyltransferase gene locus – implications for clinical pharmacogenomics. Pharmacogenomics 2002; 3(1):89-98.

Meille C, Iliadis A, Barbolosi D, et al. An interface model for dosage adjustment connects hematotoxicity to pharmacokinetics. J Pharmacokinet Pharmacodyn. 2008; 35(6):619-33.

Mercier C, Ciccolini J. Profiling dihydropyrimidine dehydrogenase deficiency in patients with cancer undergoing 5-fluorouracil/capecitabine therapy. Clin Colorectal Cancer. 2006; 6:288-96.

Milano G, Etienne MC. Individualizing therapy with 5-fluorouracil related to dihydropyrimidine dehydrogenase: theory and limits. Ther Drug Monit. 1996; 18:335-40.

Milano G, Etienne MC, Pierrefite V, et al. Dihydropyrimidine dehydrogenase deficiency and fluorouracil-related toxicity. Br J Cancer. 1999; 79:627-30.

Moore MJ, Erlichman C. Therapeutic drug monitoring in oncology. Problems and potential in antineoplastic therapy. Clin Pharmacokinet. 1987; 13:205-27.

Nanda R. Targeting the human epidermal growth factor receptor 2 (HER2) in the treatment of breast cancer: recent advances and future directions. Rev Recent Clin Trials. 2007; 2:111-6.

Nelson HD, Huffman LH, Fu R. Genetic risk assessment and brca mutation testing for breast and ovarian cancer susceptibility: systematic evidence review for the U.S. Preventive Services Task Force. Ann Intern Med. 2005; 143:362-79.

Oellerich M, Armstrong VW. Prodrug metabolites: implications for therapeutic drug monitoring. Clin Chem. 2001; 47:805-6.

Paez JG, Jänne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science. 2004; 304(5676):1497-500.

Panetta JC, Wilkinson M, Pui CH, et al. Limited and optimal sampling strategies for etoposide and etoposide catechol in children with leukemia. J Pharmacokinet Pharmacodyn. 2002; 29:171-88.

Phillips KA, Veenstra DL, Oren E, et al. Potential role of pharmacogenomics in reducing adverse drug reactions: a systematic review. JAMA. 2001; 286:2270-9.

Piccart-Gebhart MJ, Procter M, Leyland-Jones B, et al. Trastuzumab after adjuvant chemotherapy in her2-positive breast cancer. N Eng J Med. 2005; 353:1659-72.

Pignon T, Lacarelle B, Duffaud F, et al. Pharmacokinetics of high-dose methotrexate in adult osteogenic sarcoma. Cancer Chemother Pharmacol. 1994; 33(5):420-4.

Plunkett W, Iacoboni S, Estey E, et al. Pharmacologically directed Ara-C therapy for refractory leukemia. Semin Oncol. 1985; 12 Suppl. 3:20-30.

Quartino AL, Friberg LE, Karlsson MO. A simultaneous analysis of the time-course of leukocytes and neutrophils following docetaxel administration using a semi-mechanistic myelosuppression model. Invest New Drugs. 2010 Dec 14.

Ratain MJ, Schilsky RL, Conley BA, et al. Pharmacodynamics in cancer therapy. J Clin Oncol. 1990; 8:1739-53.

Relling MV, Hancock ML, Rivera GK, et al. Mercaptopurine therapy intolerance and heterozygosity at the thiopurine S-methyltransferase gene locus. J. Natl Cancer Inst. 1999; 91: 2001-8.

Ribelles N, López-Siles J, Sánchez A, et al. A carboxylesterase 2 gene polymorphism as predictor of capecitabine on response and time to progression. Curr Drug Metab. 2008; 9(4):336-43.

Romond EH, Perez EA, Bryant J, et al. Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer. N Eng J Med. 2005; 353:1673-84.

Rosell R, Moran T, Queralt C, et al. Screening for epidermal growth factor receptor mutations in lung cancer. N Engl J Med. 2009; 361: 958-67.

Rousseau A, Marquet P, Debord J, et al. Adaptive control methods for the dose individualisation of anticancer agents. Clin Pharmacokinet. 2000; 38:315-53.

Rousseau A, Marquet P. Application of pharmacokinetic modelling to the routine therapeutic drug monitoring of anticancer drugs. Fund Clin Pharmacol. 2002; 16:253-62.

Saglio G, Morotti A, Mattioli G, et al. Rational approaches to the design of therapeutics targeting molecular markers: the case of chronic myelogenous leukemia. Ann N Y Acad Sci. 2004; 1028(1):423-31.

Sandstrom M, Lindman H, Nygren P, et al. Population analysis of the pharmacokinetics and the haematological toxicity of the fluorouracil–epirubicin–cyclophosphamide regimen in breast cancer patients. Cancer Chemother Pharmacol. 2006; 58(2):143-56.

Sandstrom M, Lindman H, Nygren P, et al. Model describing the relationship between pharmacokinetics and hematologic toxicity of the epirubicin–docetaxel regimen in breast cancer patients. J Clin Oncol. 2005; 23(3):413-21.

Sartore-Bianchi A, Moroni M, Veronese S, et al. Epidermal growth factor receptor gene copy number and clinical outcome of metastatic colorectal cancer treated with panitumumab. J Clin Oncol. 2007; 25:3238-45.

Savic RM, Karlsson MO. Importance of shrinkage in empirical Bayes estimates for diagnostics: problems and solutions. The AAPS Journal 2009; 11(3):558-69.

Senzer N, Shen Y, Hill C, et al. Individualised cancer therapeutics: dream or reality? Expert Opin Ther Targets. 2005; 9(6):1189-201.

Sheiner LB, Beal SL. Bayesian individualization of pharmacokinetics: simple implementation and comparison with non-Bayesian methods. J Pharm Sci. 1982; 71(12):1344-8.

Sheiner LB. The population approach to pharmacokinetic data analysis: rationale and standard data analysis methods. Drug Metab Rev. 1984; 15(1-2):153-71.

Silver M. The case for personalized medicine. Personalized medicine coalition. 2009. [http://www.personalizedmedicinecoalition.org/sites/default/files/TheCaseforPersonalizedMedicine_5_5_09.pdf]. Last accessed 2011-06-24.

Skrondal A, Rabe-Hesketh S. Latent variable modelling: a survey. Scand J Statistics. 2007; 34: 712-45.

Spear BB, Heath-Chiozzi M, Huff J. Clinical application of pharmacogenetics. Trends Mol Med. 2001; 7(5):201-4.

Stein AM, Carter A, Hollaender N, et al. Quantifying the effect of everolimus on both tumor growth and new metastases in metastatic renal cell carcinoma (RCC): A dynamic tumor model of the RECORD-1 phase III trial. J Clin Oncol. 2011; 29: Abstr 4602^.

Thomas H, Boddy AV, English MW, et al. Prospective validation of renal function-based carboplatin dosing in children with cancer: a United Kingdom Children's Cancer Study Group trial. J Clin Oncol. 2000; 18(21):3614-21.

Tornoe CW, Agerso H, Senderovitz T, et al. Population pharmacokinetic/pharmacodynamic (PK/PD) modelling of the hypothalamic-pituitary-gonadal axis following treatment with GnRH analogues. Br J Clin Pharmacol. 2007; 63:648-64.

Tranchand B, Laporte S, Glehen O, et al. Pharmacology of cytotoxic agents: a helpful tool for building dose adjustment guidelines in the elderly. Crit Rev Oncol Hematol. 2003; 48(2):199-214.

Van Den Bongard HJGD, Mathôt RAA, Beijnen JH, et al. Pharmacokinetically guided administration of chemotherapeutic agents. Clin Pharmacokinet. 2000; 39:345-67.

Undevia SD, Gomez-Abuin G, Ratain MJ. Pharmacokinetic variability of anticancer agents. Nat Rev Cancer. 2005; 5(6): 447-58.

United States Food and Drug Administration. Camptosar label. [http://www.accessdata.fda.gov/drugsatfda_docs/label/2005/020571s026lbl.pdf]. Last accessed 2011-06-24.

Van der Bol JM, Mathijssen RH, Verweij J, et al. CYP3A pheynotype-based individualized dosing of irinotecan to reduce interindividual variability in pharmacokinetics and toxicity: results from a randomized trial. J Clin Oncol 2008; 26: Abstr 2506.

Van Kesteren C, Zandvliet AS, Karlsson MO, et al. Semi-physiological model describing the haematological toxicity of the anti-cancer agent indisulam. Invest New Drugs. 2005; 23(3):225-34.

Wallin JE, Friberg LE, Karlsson MO. A tool for neutrophil guided dose adaptation in chemotherapy. Comput Methods Programs Biomed. 2009; 93(3):283-91.

Wallin JE, Friberg LE, Karlsson MO. Model-based neutrophil-guided dose adaptation in chemotherapy: evaluation of predicted outcome with different types and amounts of information. Basic Clin Pharmacol Toxicol. 2010; 106(3):234-42.

Wang Y, Sung C, Dartois C, et al. Elucidation of relationship between tumor size and survival in non-small-cell lung cancer patients can aid early decision making in clinical drug development. Clin Pharmacol Ther. 2009; 86:167-74.

Xie R, Mathijssen RH, Sparreboom A, et al. Clinical pharmacokinetics of irinotecan and its metabolites in relation with diarrhea. Clin Pharmacol Ther. 2002; 72:265-75.

You B, Meille C, Barbolosi D, et al. Mechanistic model predicting hematopoiesis and tumor growth to optimize docetaxel + epirubicin (ET) administration in metastatic breast cancer (MBC): Phase I trial. J Clin Oncol. 2007; 25: Abstr 13013.

You B, Tranchand B, Girard P, et al. Lung Cancer. Etoposide pharmacokinetics and survival in patients with small cell lung cancer: a multicentre study. Lung Cancer. 2008; 62(2):261-72.

You B, Girard P, Paparel P, et al. Prognostic value of modeled PSA clearance on biochemical relapse free survival after radical prostatectomy. Prostate. 2009; 69(12):1325-33.

You B, Pollet-Villard M, Fronton L, et al. Predictive values of hCG clearance for risk of methotrexate resistance in low-risk gestational trophoblastic neoplasias. Ann Oncol. 2010; 21(8):1643-50.

Zandvliet AS, Schellens JH, Beijnen JH, et al. Population pharmacokinetics and pharmacodynamics for treatment optimization in clinical oncology. Clin Pharmacokinet. 2008; 47(8):487-513.

Zhang H, Berezov A, Wang Q, et al. ErbB receptors: from oncogenes to targeted cancer therapies. J. Clin. Invest. 2007; 117(8):2051-8.

ARTICLES

1. Pharmacodynamic models for discrete data

Ines Paule, Pascal Girard, Gilles Freyer, Michel Tod

(manuscript)

ABSTRACT

Clinical outcomes are often described as events: death, stroke, epileptic seizure, multiple sclerosis lesions, recurrence of cancer, disease progression, pain, infection and bacterial/viral eradication, severe toxic side-effect, resistance to treatment, etc. They may be quantified as time-to-event, counts of events per time interval (rates), their severity grade, or a combination of them. Such data are discrete and require specific modelling structures and methods. This article references the most common modelling approaches for categorical, count and time-to-event data, and reviews examples of such models applied in the analysis of pharmacodynamic data. Modelling is useful for identification of influential factors related to the clinical outcome, characterization and quantification of their impact, for making better informed predictions and clinical decisions, assessments of efficacy of therapeutic interventions, optimising the individual treatments and drug development studies.

INTRODUCTION

Clinical outcomes are often described as events: death, stroke, epileptic seizure, multiple sclerosis lesions, recurrence of cancer, disease progression, pain, infection and bacterial/viral eradication, severe toxic side-effect, resistance to treatment, etc. They may be quantified as time-to-event, counts of events per time interval (rates), their severity grade, or a combination of them. Specific modelling approaches required for each type of discrete data, as well as examples of their applications to study pharmacodynamic data, are reviewed in this paper.

First, it reviews the potential explanatory variables and functions for quantification of their effect. Then, specific model structures for binary, categorical (ordered and nominal), count, time-to-event data are briefly described and referenced. Next, illustrations of those models with literature examples of applications in medical research are given. Finally, a brief discussion concludes the review.

MODELS

Explanatory variables

Any model consists of three main components: the response (dependent variable), the explanatory variables, and the link function that relates them.

The usual explanatory variables for the pharmacodynamic responses are drug dose, steady-state concentration or AUC (area under the curve of concentration-time plot) of the drug, individual patient characteristics (age, gender, body size (weight, body surface area), ethnicity, renal and hepatic functions, genotypes, performance status, concomitant diseases and medications) and other external factors, such as time (diurnal fluctuations, seasonality).

Explanatory variable functions

The effect of explanatory variables can be described by various functions: linear, nonlinear (Emax, sigmoid Emax, exponential, power), they may assume interactions of predictors (particularly time). The main "typical" parameters describe the average population effects, the individual "random" effects adjust them to each individual and account for the correlation among observations of the same subject.

Autocorrelation

In some cases, special terms may be used to account for autocorrelation between subsequent observations. For categorical data modelled by discrete time models, probabilities of the next observation may be assumed to be dependent on the current category.

Hidden states

Sometimes, the observed response (symptom) may be related to a latent (unobservable) categorical variable, which more directly represents the disease state. The time course of the disease is described by a hidden Markov process (containing an assumed number of states), and the observed variable is related to the hidden disease state by some probabilistic model (depending on the type of the observed variable). The probabilities of transitions between the states of the hidden layer are described as a function of explanatory variables, including drug exposure.

Categorical data models

A common example of categorical data is grades of severity of pain or adverse effects ("none", "mild", "moderate" and "severe"). The simplest case consists of only two categories: "presence" and "absence" of event (relapse, death, stroke, etc.). They are modelled

by generalized (non)linear models using a link function between the probability of event and the function of explanatory variables:

$$g(p) = f(B,X),$$

where p – probability of event/presence, X – explanatory variables, B – coefficients of the explanatory variables function.

The most common link function is logit:

$$f(B,X) = \log (p/(1-p))$$

Alternative functions may be used:

- Probit: $f(B,X) = \Phi^{-1}(p)$, where Φ is the standard normal cumulative distribution function,
- Complementary log-log: $f(B,X) = \log(-\log(1-p))$.

The choices of link function correspond to different assumptions about the distribution of the underlying continuous variable [Raman 2005, Liu 2005, Lee 2003]:

- Logistic for logit link,
- Normal for probit link,
- Extreme value (log-Weibull) for complementary log-log link.

The binary data model can be naturally extended to categorical data: instead of two categories, there are K, and so K-1 probabilities are modelled instead of one.

In case of ordered categorical (ordinal) data, the most common approach is to specify cumulative probabilities of categories not higher (or lower) than a certain value, so that the order of categories is taken into account:

$$p(Y_t \leq k),\ k = 1, \ldots, K-1.$$

If an underlying continuous response variable is assumed, the values of k represent the threshold (cut-off) values for the separation into category intervals [McCullagh 1980, Anderson 1981].

For interpretation of the results, single category probabilities may be later obtained by the following transitions:

$$P(Y = 1) = P(Y \leq 1)$$

$$P(Y = k) = P(Y \leq k) - P(Y \leq k - 1)$$

The explanatory variable function *f(B,X)* includes constants α_k, which represent baseline probabilities of categories. For ordinal data, they are restricted to be ordered $\alpha_1 < \ldots < \alpha_{K-1}$ (or in decreasing order, depending on the specification of the cumulative probabilities).

Proportional odds models

When the logit model assumes the same effects of explanatory variables for all cumulative probabilities, it is commonly referred to as "proportional odds" model [McCullagh, 1980]. This "parallel slopes" assumption states that the effect of changes in the predictor values X is homogenous for all the $K - 1$ cumulative probabilities.

This assumption may not always be appropriate, especially when the ordinal data does not represent a categorization of a continuous scale, but composite ranks [Zingmark 2005 thesis]. Many different ways to relax the proportional odds assumption have been proposed [Peterson 1990, McCullagh 1980, Zingmark 2005 thesis]. Some of the simplest approaches are, for example, to consider a separate vector of parameters β_k for each response level, or to allow a heterogeneous effect for only some of the explanatory variables (that is assume partial proportional odds [Peterson 1990]).

Continuation ratio models

The probability of the ordinal response may be specified in other ways than the cumulative probability. Continuation ratio models specify the conditional probability of a particular grade given that the higher (or lower) grades are not possible:

$$\frac{P(Y_t = k)}{P(Y_t \leq k)} \quad \text{or} \quad \frac{P(Y_t = k)}{P(Y_t \geq k)}.$$

These models are therefore particularly suited for sequential irreversible processes (e.g. deterioration or growth). In the case of logit link, they are also referred to as proportional hazards models [Agresti 1999, Liu 2005].

The continuation ratio models have the advantage over the cumulative link models that the intercept parameters in the latter have to be ordered ($\alpha_1 < \ldots < \alpha_{K-1}$), while this restriction is not used in the continuation ratio models. Moreover, these models may be estimated by software dealing with binary mixed-effects regression models [Lindsey 1997].

Adjacent category models

Adjacent category models are intended for cases where the response variable is perceived as truly categorical [Vermunt 2004] and one wants to describe the effects that apply

to pairs of adjacent categories rather than to cumulative probabilities [Hartzel 2001]. Here, the odds between two adjacent categories are modelled [Agresti 1999]:

$$\frac{P(Y_t = k)}{P(Y_t = k+1)}.$$

This model is also invariant to changes in the definition of categories, in case of underlying continuous response variable and assumed proportional effect of explanatory variables [McCullagh 1980].

Nominal data models

In case of nominal (unordered) data, the most common approach is to specify one category as a reference/baseline, and model the ratio of probabilities of each other category to the probability of the "reference" category:

$p(Y_t = k)/p(Y_t = K), k = 1, \ldots, K-1.$

Count data models

Responses described as number of events per time interval (for example, number of epileptic seizures per week or month) constitute count data. This type of data is different from the ordinal data in that the number of categories is not definite.

The simplest model for such data (n = 0, 1, 2, …) is the Poisson model:

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

where $\lambda>0$ is a parameter (function of explanatory variables) corresponding to the event rate (a number of expected events per time interval) and to its variance:

$E(Y) = var(Y) = \lambda.$

Since λ has to be a positive number, the effect of explanatory variables is modelled on its logarithm:

$log(\lambda) = f(B,X).$

However, its assumption of equidispersion may often be not valid. Alternative distributions are proposed to relax this assumption:

- Generalized Poisson,
- Zero-inflated Poisson (allows for high counts of zeros),
- Zero-truncated Poisson (excludes zero values),
- Negative Binomial (allows for overdispersion),
- Zero-inflated Negative Binomial (allows for high counts of zeros and overdispersion),

- Zero-truncated Negative Binomial (excludes zero values and allows for overdispersion),
- Hurdle model (Poisson-logit or Negative Binomial-logit),
- Weibull (allows for both over- and underdispersion, nests both Poisson and negative Binomial as special cases).

When there is more than one observation per subject, their correlation can be accounted for by including random effects (having Gaussian, Gamma or Beta distributions).

Time-to-event data models

TTE functions

Rare or single events (relapse, stroke, death, etc.) are usually described using a different approach from frequent events: instead of modelling the number of events per unit of time, the time to event is modelled. Time-to-event (TTE, or survival) data are usually described and modelled by three related functions: the survivor, the hazard, and the cumulative hazard (risk) functions.

The survivor function (or survival probability) S(t) describes the proportion of individuals in the population surviving beyond time t,

$$S(t) = \frac{number\ of\ individuals\ surviving\ longer\ than\ time\ t}{total\ number\ of\ individuals}.$$

At the individual level, it can be interpreted as the probability for the individual to survive beyond time t (or a cumulative probability of having no event up to time t). It is equal to 1 at t=0 and is monotonously decreasing to 0 as time goes to infinity.

The hazard h(t) is the (instantaneous) death (event) rate at time t, conditional on surviving until time t:

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t | T \ge t)}{\Delta t}.$$

It is also described as the instantaneous probability for the event to occur at time t, knowing that there was no event up to that time, or as the force of mortality, or the conditional failure rate, or as a function representing the change in the immediate future of an individual's probability of death, given the survival up to the current time. From a life table, it is calculated as the fraction of those who died at time t out of the number alive at that time. The values of the hazard function are non-negative, but can vary in all directions over time. As the descriptor of event rate, it has units of 1/time.

The cumulative hazard function represents the risk of event over the time interval (from 0 to t), it is the integral of the hazard (or the area under the hazard function) over the interval:

$$R(t) = \int_0^t h(u)du .$$

It can also be seen as the total number of subject's events over the time interval.

Mathematically, the three functions are simply different expressions of the same function:

$$S(t) = P(T > t) = e^{-R(t)} = e^{-\int_0^t h(u)du} .$$

The probability of having an event at a particular time t is described by the probability density function, given by a product of survival and hazard functions:

$$pdf(t) = S(t) \cdot h(t) .$$

TTE analysis methods

The most common approach of analysing survival data is the semiparametric (Cox) model, where no parametric form of the hazard function is specified. Parametric relation assumptions are only used to estimate the effect of explanatory variables on the baseline hazard function (i.e., the hazard for which all covariates are equal to zero).

The survivor functions are estimated by nonparametric Kaplan-Meier method, as ratio of the number of deaths at a given death time to the number of individuals at risk at that time.

However, efficiency (parameter precision) is gained when using parametric approaches, that is assuming a certain distribution for the survivor function:

- exponential: $h_0(t) = \beta_0$
- Gompertz: $h_0(t) = \beta_0 \cdot e^{\beta_1 \cdot t}$
- Weibull: $h_0(t) = \beta_0 \cdot \gamma \cdot t^{\gamma-1}$ or reparameterized $h_0(t) = \beta_0 \cdot e^{\beta_1 \cdot \ln(t)}$

TTE: effect of explanatory variables

Explanatory variables are commonly assumed to have a multiplicative/proportional effect on the hazard (this assures that hazard is always non-negative):

$$h(t) = h_0(t) \cdot e^{f(B,X)} .$$

In the simplest analyses, the explanatory variables are independent of time (constant), their effects are assumed to be summed by a linear function (may include interactions). Such

models are called “proportional odds” models, and the coefficients B are used to calculate the hazard ratio between groups.

Various extensions of this basic model are proposed to accommodate different more complex situations:

- different baseline hazards,
- time-varying explanatory variables,
- time-dependent effects,
- repeated correlated events,
- several possible causes of the event, etc.

Multistate (or continuous-time transitional/Markov) models

Multinomial (multistate) data can be modelled by an extension of time-to-event models to model rates of transitions between various states (for example, stages of illness), instead of just from “no event” to “event” [Meira-Machado 2009].

EXAMPLES

Models for ordered categorical data

Representative examples of drug evaluations based on ordered categorical data analysis are summarized in table I. Frequently, the categorical response studied was a toxic or an adverse effect (in 6 studies out of 17). This is because adverse effects, which are usually collected for qualitative or semi-quantitative assessment only, are frequently measured on a categorical scale. The number of grades in the scale was typically 4 and never greater than 6. A higher number of grades would not yield much more information but would result in large uncertainty on parameter estimates. In most cases, the measurements were repeated several times in each subject.

Model structure

In all cases, the link function was the logit. In almost all studies, at least one random effect (η_Y) was introduced in the model to account for correlation between measurements of a same individual. In case of frequent measurements, serial correlation between successive measurements must be accounted for to avoid biased estimation of model parameters [Lunn 2001]. One way to overcome this difficulty has been to condition each observation on the previous measurement via a Markov model [Lunn 2001, Zingmark 2005].

In several studies, the effect of placebo was accounted for, using an empirical time function. The metrics for drug exposure was either dose, AUC, drug concentration profile in plasma or in the effect compartment. In a few studies, drug exposure was not measured but calculated using a population PK model. In this case, the population PK model was validated by VPC or external validation. The drug effect model was usually an Emax model unless the metrics for drug exposure was the dose and only one dose level was studied. Sometimes, the response or the dose-effect relationship appeared only in a subset of patients [Kowalski 2003]. In this case, the probability of the graded response had to be conditioned on the observation of a response, i.e. the logistic model was applied only to the subset of responders [Kowalski 2003].

Model building and validation

Model building was rarely described in details. In all cases, the pharmacokinetic model, if any, and the logistic model were built separately. The building tools for the logistic model, when reported, were the residual plots (in the probability domain), the likelihood ratio test or Bayes factors, and the confidence or credibility intervals for parameter estimates. Validation of the model was more frequently reported. The main approach consisted in a visual comparison between observed frequencies and predicted probabilities of each grade, as a function of time or drug exposure. The predictions were based either on the typical values (all ηs set to zero) or on repeated simulations (by drawing in ηs distributions). The latter approach is a form of predictive check. This approach should be preferred for all models involving random effects, because it is general, flexible, easy to interpret, and takes into account all parts of the model.

Main results

The main result of the analysis was, in most cases, to identify the exposure measure best related to the response, and to describe the probability of each grade of response as a function of drug exposure and time. Some covariates influencing these relationships were sometimes identified, such as the effect of age for sildenafil [Claret 2006], or the number of courses of previous treatment with platinum-based regimens for topotecan [Mould 2002]. The effect of seasonality on the response to an anti-IgE antibody used for rhinitis could even be characterized [Lunn 2001]. These models ultimately allowed to reach conclusions of broader interest, such as the optimal dose for topotecan topotecan [Mould 2002] or the therapeutic

index for oxybutynin [Gupta 1999]. In the latter case, however, two separate models had to be defined, one for the therapeutic effect, the other for the main side effect.

Models for count data

Model structure

In the three studies (table II), a Poisson model was used to describe the frequency (count per unit time) of the event (incontinence episode or seizure episode). In these studies, the design included a placebo group and each patient received several doses. This design allowed the placebo effect to be accounted for, and the dose-response relationship to be identified. In all cases, the statistical analysis, based on the combination of several studies, was akin to meta-analysis. However, no statistical test for heterogeneity between studies was reported, i.e. comparability between studies was assumed.

In the study on pregabalin (an antiepileptic drug) as add-on therapy, 25% of patients did not respond to the treatment, while the remaining 75% exhibited a dose-response relationship, reaching a maximal percentage of seizure reduction from baseline of 100% for women and 80% for men. The occurrence of these two populations (responders / non-responders) was accommodated through a mixture model on η. The same team assessed the statistical performances of mixture modelling in the context of a Poisson model for count data, i.e. (1) what is the probability of concluding that a mixed population exists when there truly is a mixture, (2) what is the probability of concluding that two subpopulations exist when there is truly a mixed population, and (3) how well can the mixture be estimated, both in terms of the population parameters and the individual subject classification [Frame 2003]. When no mixture was present, the false positive probability was less than 0.078. When mixtures were present, they were characterized with varying degrees of success, depending on the nature of the mixture. As expected, when the difference between subpopulations was greater the mixtures became easier to characterize. The median proportion of subjects classified correctly ranged from 0.59 to 0.96.

Model building and validation

Model building was based on the likelihood ratio test and confidence intervals of estimated parameters. Validation was based on posterior predictive checks, using the percentage of responders or the decrease in event frequency at each dose level.

Main results

Quoting Miller [2003], "The traditional binary approach of analyzing responder rate (> 50% reduction in seizures) ignores the information implicit in a graded exposure-response relationship. Using seizure frequency as the basic unit of response permits the exploration of the full exposure-response relationship". In turn, the influence of covariates on response is expected to be characterized with greater power.

Time-to-event models

Model structure

The probability of time-to-event measures can be expressed as functions of the hazard. The log hazard may be modelled directly in terms of covariates, so that the time-varying nature of the hazard can be explained in part by time-varying covariates such as drug concentration. This approach was used to model the effect of ondansetron concentration on time to emesis after ipecac administration [Cox 1999], or the effect of docetaxel cumulative AUC on time to disease progression. In the absence of time-varying covariates, the expectation or the median of the time-to-event density may be modelled as a function of covariates.

Model building and validation

The tools used for building the model were rarely described. Cox [1999] described a number of "residuals". These residuals involve comparison of observations and predictions for the typical individual (i.e. calculated by setting the fixed effect parameters to their maximum likelihood estimate and η to zero). The first two residuals are residuals for observed event times, the first on the time scale, and the second on the log probability scale. These residuals are defined only for subjects with at least one event. A third type of residual, based on the probability of being event-free for the entire experiment, is defined for all subjects.

Validation was based on visual or posterior predictive checks in almost all studies. When only one event is recorded, the percentage of patients with or without event versus time may be used as a criterion for validation. This was the case in the study on time to anemia in patients treated by interferon and ribavirin for hepatitis C [Tod 2005]. When the event is measured repeatedly, other criteria may be used, such as minimal and maximal time to event, or maximal and median number of events.

Main results

For a single event, time-to-event analysis is complementary with logistic regression. The latter deals with the probability of an event at a given time point, while the former looks at the time required to observe the event, if it happens. For repeated events, time-to-event analysis is complementary with count data analysis, which deals with the frequency of the event. Hence both types of methods are used, sometimes jointly, in the analysis of categorical data. This was the case for the study of sildenafil effect on female sexual activity, where both the satisfaction score and the time between sexual events were of primary interest to characterize drug effect. Time-to-event analysis is amenable to a number of evaluations that are relevant for pharmacological studies, such as ED_{50} or EC_{50} estimation, dose- or concentration-effect relationship and covariate analysis. For example, higher cumulative AUC of docetaxel, lower alpha-1 glycoprotein and fewer organs involved in cancer were associated with longer time to progression and time to death [Veyrat-Follet 2000].

Like any other statistical approach, time-to-event analysis may lead to inconsistent results which deserve an in depth exploration. For example, a Cox survival analysis was carried out to examine the relationship between capecitabine metabolites exposure and time to progression and death in cancer patients [Gieschke 2002]. By a univariate analysis, a negative association was found between the AUC of 5-FU (the active metabolite) in plasma and time to disease progression (hazard ratio = 1.626, $P = 0.0056$). The authors explained that (1) 5-FU concentration in plasma may not be predictive of 5-FU concentration in tumor, due to variable expression, in tumor tissues, of thymidine phosphorylase required to form 5-FU, and (2) patients with high 5-FU AUC in plasma had also a high level of serum alkaline phosphatase (ALP), which was associated with a shorter time to disease progression. In a bivariate analysis of time to progression versus 5-FU AUC and ALP, the relationship with 5-FU became non significant.

Finally, survival data are less informative than continuous data, because intrasubject variability cannot be reduced beyond a certain level. As a result, there remains a large difference between observed time-to-events and their predictions. Simulation of the model allows to predict the typical survival curve, but not the time-to-event in a single individual.

Transitional models

Model structure

The four examples in table IV describe a number of possible settings and usage. The first example deals with non-ordered categorical data. The hypnotic effect of temazepam was evaluated by recording every 30 s during sleep in which of 6 sleep stages were the patients

[Karlsson 2000]. The aim of the model was to estimate the probability of moving from one sleep stage to another, given the fact that the patient was in a different stage during the previous interval, and to determine what factors influence these probabilities. The transition probabilities were assumed to be the sum, in logit domain, of changes attributable to nighttime, stage time, and drug effect.

The second example deals with the comparison of the effect of two drugs (naratriptan and sumatriptan) in acute migraine in a hidden Markov model with two layers [Maas 2006]. The model considers the time course of migraine as transitions between three disease states. The disease states are the first layer, while the headache scores are ascribed to the second layer. The first layer is non-observable (the hidden layer), while the second is observable. The hidden layer states are, by order of apparition, full migraine attack, pain relief and pain free status. Headache scores ranges from 0 to 3. The model focuses on the rate of transition between the disease states, given the measurements of headache scores at several occasions after drug administration for a migraine attack. The time spent in each state before moving in the previous or the next state is a random variable, which is assumed to follow an exponential distribution. The latter assumption is equivalent to assume a constant hazard for transition. Hence the rate of each transition is characterized by a single rate constant. In this formulation, the Markov model is akin to a time-to-event model. The log transition rates are assumed to be the sum of changes attributable to drug exposure and covariates.

Model building and validation

Model building may be carried out by using residual plots comparing the observed rates or frequencies of transitions with their predictions as goodness-of-fit plots, the likelihood ratio test to compare alternative models, and the confidence interval of parameter estimates. The way to obtain confidence intervals on response predictions has been described by Anisimov [2007] and these confidence intervals are important to assess the predictive performances of the model. Validation per se is most conveniently carried out by posterior predictive check based on outcomes such as the proportion of patients or the time spent in each state.

Main results

The primary result of this kind of analysis is to describe the influence of drug exposure on the probability or the rate of transition between states, with respect to time. This may convey much information about drug action. For example, temazepam was shown to reduce

the time spent awake by four mechanisms: (1) facilitation of transition to “deeper” sleep, (2) inhibition of transition to “lighter” sleep, (3) regardless of sleep stage, inhibition of the transition to wake state, and (4) facilitation of return to sleep. This detailed assessment allows in turn to compare drugs. For example, naratriptan was shown to have greater potency than sumatriptan (EC_{50} ratio = 3.3) but lower maximal effect (Emax ratio = 0.74) for the first transition, leading to pain relief in headache. Ultimately, this kind of model may be used to find the optimal dosing schedule and/or the target exposure.

Table I : models for ordered categorical data.

Drug	Response	Nb of grades	Pop size	Nb occa-sions	Rando m effect	Exposure measure	Validation	Software	Main result	Comment	Author year
Investiga-tional	Toxicity	5	3 groups of 6 rats	1	Subject-specific	Dose and AUC	none	NONMEM and BUGS	Probability of score vs dose or AUC	Placebo effect accounted for	Aarons 2001
Sildénafil	Orgasm satisfaction score	5	614 women several doses	Every day, 12 weeks	Subject specific	Dose	PPC	NONMEM	Dose effect relationship vs age	Placebo effect accounted for	Claret 2006
Inolimomab	Acute GVH	4	21 patients	9 to 28	Subject specific	Cumulated AUC	VPC and PC	NONMEM	Identification of efficacy vs exposure relationship		Dartois 2007
Hydrogen sulfide	Toxicity	4	6860 subjects several concentration and duration	1	No	Concen-tration and duration	none	MCsim	LC50	Bayesian framework	Diack 2005
Tramadol	Pain relief	5 items, 3 grades each	104 children	9	Subject specific	Ce(t) of tramadol and metabolite	VPC	NONMEM	Probability of score vs time. EC95 for T and M1		Garrido 2006
Fluvoxamine	Inhibition of PCA effect	3 items, 4 grades each	66 rats 3 dose levels	8	No	C(t)	Visual comparison of observed and predicted prob. vs conc. curves	NONMEM	EC50 for each item and grade		Geldof 2007
Capecitabine	Safety parameters	4 items, 2 grades each	505 patients	1	No	Cmax, AUC of 3 active métabolites	none (LR test and SE)	NONMEM SAS	Identification of 3 toxicity vs exposure relationships		Gieschke 2002
Oxybutynin	Dry mouth	4	187 patients	2 to 6	Subject specific	Dose	Visual comparison of observed and predicted prob. vs dose	NONMEM	Probability of dry mouth is less with XR formulation	Placebo effect accounted for	Gupta 1999
Capecitabine	Hand and foot syndrome	4	595 patients	Each week, 30 weeks	Subject-specific	A(t) in biophase	PPC and data splitting	NONMEM	HFS score is related to drug amount in biophase	Markov component to account for previous observation	Henin 2007
Propofol	Sedation	6	20 patients	4 times / day 2 to 5 d	Subject specific on EC50	C(t)	Visual comparison of observed and predicted prob. vs conc	NONMEM	EC50 for each score		Knibbe 2002
Investiga-tional	Adverse event	4	811 patients	4	Subject specific	Dose	PPC	NONMEM	AE severity vs time, dose and age	Bimodality + Placebo effect accounted for	Kowalski 2003
Anti-IgE antibody	Rhinitis Sneezing scores	4	155 patients	Daily 12 weeks	Subject specific	Dose	VPC	WinBUGS	Score vs dose and seasonality	Bayesian framework. Markov component accounts for serial correlation	Lunn 2001
Ketorolac	Analgesia	5	522 patients	6	Subject specific	Ce(t)	VPC	NONMEM	Score vs dose and time. Percentage patients with adequate pain relief vs time	Placebo effect accounted for	Mandema 1996
Topotecan	Neutro-penia	6	438 patients	> 3	Subject specific	AUC	VPC	NONMEM	Optimal dose according to covariates		Mould 2002
Monoclonal antibody ATM-027	Receptor expression on T cells	3	73 patients	10	Subject specific	C(t)	Visual comparison of observed and predicted prob. vs concentration	NONMEM	Score vs time or concentration	Comparison with a continuous data model	Zingmark 2004
Investiga-tional	CNS side-effect	3 6 transi-tions	12 healthy subjects	Every 3 min, several hours	Subject specific	Ce(t)	PPC	NONMEM	Time course and frequency of the side effect.	Markov component accounts for serial correlation	Zingmark 2005

Table II: models for count data.

Drug	Response	Pop size	Nb occa-sions	Link model	Exposure measure	Validation	Software	Main result	Comment	Author year
Oxybutynin	Weekly urge urinary incontinence episodes	187 patients	2 to 6	Poisson	Dose	LR test and SE	NONMEM	Same efficacy of IR and XR formulations	Placebo effect accounted for. Evaluation of therapeutic index	Gupta 1999
Pregabalin	Monthly seizure frequency	1042 patients	3	Poisson	Dose	PPC	NONMEM	ED50, covariates for response	Mixture model. Placebo effect accounted for	Miller 2003 Frame 2003
Levetiracetam	Weekly seizure frequency	958 patients	3 to 6	Poisson	Dose	PPC	NONMEM	Dose-effect relationship	Placebo effect accounted for	Snoeck 2007

Table III: models for time-to-event data.

Drug	Event	Pop size	Nb occa-sions	Hazard function	Exposure measure	Validation	Software	Main result	Comment	Author year
Sildenafil	Time between sexual events	614 women	several	Weibull	dose	PPC	NONMEM	Median time between events		Claret 2006
Indinavir	Time to nephrolithiasis	282 HIV patients	1	Non-parametric (Cox)	Cmax, Cres, AUC, CR	cross-validation	NONMEM and SAS	Higher risk for high Cres		Collin 2007
Ondansetron	Time to emesis after ipecac	86 healthy	several	Time dependent function	C(t)	Residuals on time and probability scales + PPC	NONMEM	CE50 for reduction of emesis hazard		Cox 1999
Ciprofloxacin	Time to bacterial eradication	74 patients	1	Non-parametric (Kaplan-Meier)	2 ranges of AUC / MIC	none	ADAPT II	Lower time for AUC/CMI > 250		Forrest 1993
Grepafloxacin	Time to bacterial eradication	76 patients	1	Non-parametric (Kaplan-Meier)	3 ranges of AUC / MIC	none	ADAPT II	Lower time for AUC/CMI > 190		Forrest 1997
Capecitabine	Time to disease progression	481 patients	1	Non-parametric (Cox)	Cmax or AUC of 3 active metabolites	Not stated	SAS	Negative correlation with FU AUC	FU in plasma not relevant	Gieschke 2002
Ribavirin	Time to anemia	88 patients	1	Weibull	Dosing rate	PPC	NONMEM	Median time to anemia vs dose		Tod 2005
Docetaxel	Time to disease progression, death	151 patients	1	Weibull	Cumulative AUC	VPC and PPC	NONMEM	Identification of risk factors. Decision not to perform phase III	Decision based on CTS	Veyrat-Follet 2000
Cariporide	Time to death or myocardial infarction	2840 patients	1	Mixture of two Weibull functions	C(t)	Visual comparison of observed and predicted hazard rate vs dose	NONMEM	Estimation of the minimal effective concentration	Placebo effect accounted for	Weber 2002

Table IV: transitional models

Drug	Response	Nb of transitions	Pop size	Nb occa-sions	Statistical model	Exposure measure	Validation	Software	Main result	Comment	Author year
Temazepam	Sleep stage transitions	30 transitions 6 states non ordered	21 patients	Every 30 s during sleep	Probability of transitions	C(t)	PPC	NONMEM	Fine characterization of hypnotic effect	Placebo effect accounted for	Karlsson 2000
Sumatriptan Naratriptan	Migraine	4 grades 4 tran-sitions 3 states	Sum : 1180 episodes Nara : 1608	7 to 11 during 24 h	Hidden Markov : rate of transitions	C(t)	Visual comparison of observed and predicted prob. vs time	S-plus	Comparison of potency and maximal effect	Placebo effect accounted for	Maas 2006

CONCLUSION

This paper has reviewed the usual approaches used for modelling categorical, count and time-to-event data, as well as their applications to analyse discrete pharmacodynamic endpoints. Pharmacometrics is a rapidly growing science and new model extensions, numerical techniques for their estimation are introduced. Consequently, the number of published applications is growing and showing how modelling and simulation can be useful for identification of influential factors related to the clinical outcome, characterization and quantification of their impact, for making better informed predictions and clinical decisions, assessments of efficacy of therapeutic interventions, optimising the individual treatments and drug development studies.

REFERENCES

Aarons L, Graham G. Methodological approaches to the population analysis of toxicity data. Toxicol Lett. 2001; 120(1-3):405-10.

Agresti A. Modelling ordered categorical data: recent advances and future challenges. Stat Med. 1999; 18(2191-207).

Anderson JA, Philips PR. Regression, discrimination and measurement models for ordered categorical variables. Applied Statistics. 1981; 30:22-31.

Claret L, Cox EH, McFadyen L, et al. Modeling and simulation of sexual activity daily diary data of patients with female sexual arousal disorder treated with sildenafil citrate (Viagra). Pharm Res. 2006; 23(8):1756-64.

Collin F, Chêne G, Retout S, et al. Indinavir trough concentration as a determinant of early nephrolithiasis in HIV-1-infected adults. Ther Drug Monit. 2007; 29(2):164-70.

Cox EH, Veyrat-Follet C, Beal SL, et al. A population pharmacokinetic-pharmacodynamic analysis of repeated measures time-to-event pharmacodynamic responses: the antiemetic effect of ondansetron. J Pharmacokinet Biopharm. 1999; 27(6):625-44.

Dartois C, Freyer G, Michallet M, et al. Exposure-effect population model of inolimomab, a monoclonal antibody administered in first-line treatment for acute graft-versus-host disease. Clin Pharmacokinet. 2007; 46(5):417-32.

Diack C, Bois FY. Pharmacokinetic-pharmacodynamic models for categorical toxicity data. Regul Toxicol Pharmacol. 2005; 41(1):55-65.

Forrest A, Nix DE, Ballow CH, et al. Pharmacodynamics of intravenous ciprofloxacin in seriously ill patients. Antimicrob Agents Chemother. 1993; 37(5):1073-81.

Forrest A, Chodosh S, Amantea MA, et al. Pharmacokinetics and pharmacodynamics of oral grepafloxacin in patients with acute bacterial exacerbations of chronic bronchitis. J Antimicrob Chemother. 1997; 40 Suppl A:45-57.

Frame B, Miller R, Lalonde RL. Evaluation of mixture modeling with count data using NONMEM. J Pharmacokinet Pharmacodyn. 2003; 30(3):167-83.

Garrido MJ, Habre W, Rombout F, et al. Population pharmacokinetic/pharmacodynamic modelling of the analgesic effects of tramadol in pediatrics. Pharm Res. 2006; 23(9):2014-23.

Geldof M, Freijer J, van Beijsterveldt L, et al. Pharmacokinetic-pharmacodynamic modeling of the effect of fluvoxamine on p-chloroamphetamine-induced behavior. Eur J Pharm Sci. 2007; 32(3):200-8.

Gieschke R, Burger HU, Reigner B, et al. Population pharmacokinetics and concentration-effect relationships of capecitabine metabolites in colorectal cancer patients. Br J Clin Pharmacol. 2003; 55(3):252-63.

Gupta SK, Sathyan G, Lindemulder EA, et al. Quantitative characterization of therapeutic index: Application of mixed-effects modeling to evaluate oxybutynin dose–efficacy and dose–side effect relationships. Clin Pharmacol Ther. 1999; 65(6):672-84.

Hartzel J, Agresti A, Caffo B. Multinomial logit random effects models. Statistical Modelling. 2001; 1:81-102.

Henin E, You B, VanCutsem E, et al. A dynamic model of hand-and-foot syndrome in patients receiving capecitabine. Clin Pharmacol Ther. 2009; 85:418-25.

Karlsson MO, Schoemaker RC, Kemp B, et al. A pharmacodynamic Markov mixed-effects model for the effect of temazepam on sleep. Clin Pharmacol Ther. 2000; 68(2):175-88.

Knibbe CA, Zuideveld KP, DeJongh J, et al. Population pharmacokinetic and pharmacodynamic modeling of propofol for long-term sedation in critically ill patients: a comparison between propofol 6% and propofol 1%. Clin Pharmacol Ther. 2002; 72(6):670-84.

Kowalski KG, McFadyen L, Hutmacher MM, et al. A two-part mixture model for longitudinal adverse event severity data. J Pharmacokinet Pharmacodyn. 2003; 30(5):315-36.

Lee ET, Wang JW. Statistical Methods for Survival Data Analysis. 3rd edition. John Wiley & Sons, Inc., Hoboken, New Jersey.

Lindsey JK, Jones B, Ebbutt AF. Simple models for repeated ordinal responses with an application to a seasonal rhinitis clinical trials. Stat Med. 1997; 16:2873-82.

Liu I, Agresti A. The analysis of ordered categorical data: An owerview and a survey of recent developments. Sociedad de Estadistica e Investigacion Operativa Test. 2005; 14(1):1-73.

Lunn DJ, Wakefield J, Racine-Poon A.Cumulative logit models for ordinal data: a case study involving allergic rhinitis severity scores. Stat Med. 2001; 20(15):2261-85.

Maas HJ, Danhof M, Della Pasqua O. A model-based approach to treatment comparison in acute migraine. Br J Clin Pharmacol. 2006; 62(5):591-600.

Mandema JW, Stanski DR. Population pharmacodynamic model for ketorolac analgesia. Clin Pharmacol Ther. 1996; 60(6):619-35.

McCullagh P. Regression models for ordinal data (with discussion). Journal of the Royal Statistical Society, Series B. 1980; 42:109-42.

Meira-Machado L, de Uña-Alvarez J, Cadarso-Suárez C, et al. Multi-state models for the analysis of time-to-event data. Stat Methods Med Res. 2009; 18(2):195-222.

Miller R, Frame B, Corrigan B, et al. Exposure-response analysis of pregabalin add-on treatment of patients with refractory partial seizures. Clin Pharmacol Ther. 2003; 73(6):491-505.

Mould DR, Holford NH, Schellens JH, et al. Population pharmacokinetic and adverse event analysis of topotecan in patients with solid tumors. Clin Pharmacol Ther. 2002; 71(5):334-48.

Peterson B, Harrell Jr FE. Partial proportional odds models for ordinal response variables. Applied Statistics. 1990; 39(2):205-17.

Raman R, Hedeker D. A mixed-effects regression model for three-level ordinal response data. Stat Med. 2005; 24:3331-45.

Snoeck E, Stockis A. Dose-response population analysis of levetiracetam add-on treatment in refractory epileptic patients with partial onset seizures. Epilepsy Res. 2007; 73(3):284-91.

Tod M, Farcy-Afif M, Stocco J, et al. Pharmacokinetic/pharmacodynamic and time-to-event models of ribavirin-induced anaemia in chronic hepatitis C. Clin Pharmacokinet. 2005; 44(4):417-28.

Vermunt JK, Hagenaars JA. Ordinal longitudinal data analysis. Methods in Human Growth Research, 374-393. Ed. Hauspie, R. C., Cameron, N. & Molinari, L. Cambridge University Press. 2004.

Veyrat-Follet C, Bruno R, Olivares R, et al. Clinical trial simulation of docetaxel in patients with cancer as a tool for dosage optimization. Clin Pharmacol Ther. 2000; 68(6):677-87.
Weber W, Harnisch L, Jessel A, et al. Lessons learned from a phase III population pharmacokinetic study of cariporide in coronary artery bypass graft surgery. Clin Pharmacol Ther. 2002; 71(6):457-67.

Zingmark PH, Edenius C, Karlsson MO. Pharmacokinetic/pharmacodynamic models for the depletion of Vbeta5.2/5.3 T cells by the monoclonal antibody ATM-027 in patients with multiple sclerosis, as measured by FACS. Br J Clin Pharmacol. 2004; 58(4):378-89.

Zingmark PH, Kågedal M, Karlsson MO. Modelling a spontaneously reported side effect by use of a Markov mixed-effects model. J Pharmacokinet Pharmacodyn. 2005; 32(2):261-81.

Zingmark PH. Models for Ordered Categorical Pharmacodynamic Data, PhD thesis. Ed. Acta Universitatis Upsaliensis, Uppsala. 2005.

Cancer Chemother Pharmacol
DOI 10.1007/s00280-011-1714-9

ORIGINAL ARTICLE

Dose adaptation of capecitabine based on individual prediction of limiting toxicity grade: evaluation by clinical trial simulation

Ines Paule · Michel Tod · Emilie Hénin · Benoit You · Gilles Freyer · Pascal Girard

Received: 9 May 2011 / Accepted: 15 July 2011

Abstract
Purpose Anticancer drugs often show a narrow therapeutic index and high inter-patient variability, which can lead to the need to adjust doses individually during the treatment. One approach to doing this is to use individual model predictions. Such methods have been proposed to target-specific drug concentrations or blood cell count, both of which are continuous variables. However, many toxic effects are evaluated on a categorical scale. This article presents a novel approach to dose adjustments for reducing a graded toxicity while maintaining efficacy, applied to hand-and-foot syndrome (HFS) induced by capecitabine.
Methods A mixed-effects proportional odds Markov model relating capecitabine doses to HFS grades was individually adjusted at the end of each treatment cycle (3 weeks) by estimating subject-specific parameters by Bayesian MAP technique. It was then used to predict the risk of intolerable (grade ≥ 2) toxicity over the next treatment cycle and determine the next dose accordingly, targeting a predefined tolerable risk. Proof of concept was given by simulating virtual clinical trials, where the standard dose reductions and the prediction-based adaptations were compared, and where the therapeutic effect was simulated using a colorectal tumor inhibition model. A sensitivity analysis was carried out to test various specifications of prediction-based adaptation.
Results Individualized dose adaptation might reduce the average duration of intolerable HFS by 10 days as compared to the standard reductions (3.8 weeks vs. 5.2 weeks; 27% relative reduction) without compromising antitumor efficacy (both responder rates were 49%). A clinical trial comparing the two methods should include 350 patients per arm to achieve at least 90% power to show a difference in grade ≥ 2 HFS duration at an alpha level of 0.05.
Conclusions These results indicate that individual prediction-based dose adaptation based on ordinal data may be feasible and beneficial.

Keywords Capecitabine · Hand-and-foot syndrome · Dose adaptation · Cancer chemotherapeutics · Computer modeling and simulation · Pharmacometrics

I. Paule · M. Tod · E. Hénin · B. You · G. Freyer · P. Girard
Université de Lyon, 69622 Lyon, France

I. Paule (✉) · M. Tod · E. Hénin · B. You · G. Freyer · P. Girard
Université Lyon 1, Faculté de Médecine Lyon Sud, BP12, EMR3738 Ciblage Thérapeutique en Oncologie, 69921 Oullins, France
e-mail: ines.paule@gmail.com

M. Tod
Hospices Civils de Lyon, Hôpital de la Croix-Rousse, Pharmacie, 69317 Lyon, France

B. You · G. Freyer
Hospices Civils de Lyon, Centre Hospitalier Lyon Sud, Service d'Oncologie Médicale, 69310 Pierre-Bénite, France

Present Address:
P. Girard
Merck Serono, Merck Serono S.A., 9 chemin des Mines, 1202 Geneva, Switzerland

Introduction

The management of anticancer therapies is complicated by their narrow therapeutic indices (range between the minimum effective and toxic doses) and high inter- and intra-patient variability in pharmacokinetics (PK) and pharmacodynamics (PD). Dose scaling to body surface

Published online: 02 August 2011

area is used for most anticancer drugs, although this approach reduces some of the PK-variability for only a few drugs [1]. The overall objective for optimizing a dosing regimen is that each patient obtains the maximum possible anticancer effect without being subjected to an unacceptable risk of severe toxicities. To achieve this therapeutic goal, strategies for dose adaptation before and/or during the treatment are required. Because the dose reductions commonly applied after the occurrence of severe toxicity are suspected to be less than optimal, individualized dose adjustment alternatives are of high clinical relevance. The population model-based dose adaptation approach was first introduced by Sheiner [2]. The main idea is to individualize a population model relating dosage to a pharmacokinetic or pharmacodynamic outcome by Bayesian techniques, using data from the patient's previous responses to the drug. The optimal dosage is then determined based on individual response predictions given by the patient-specific model.

Bayesian dose adaptations are used widely to control the plasma concentrations of various classes of drugs or their active metabolites; a few examples have been developed for anticancer drugs too [3–5]. This is reasonable only if plasma concentration correlates strongly with the toxic outcomes, which is seldom the case for anticancer drugs [3]. In the absence of such direct relationships, PK-based control is not adequate. The alternative is to investigate the feasibility of dose adaptation based on toxicity predictions, without compromising the efficacy. Moreover, because toxicity is often measured on an ordinal and not a continuous scale, the usual model-based adaptive control techniques are not directly applicable and ordinal data-specific criteria must be determined.

Capecitabine (Xeloda, Roche) is an oral prodrug of 5-fluorouracil (5FU), a chemotherapeutic agent commonly used to treat solid tumors [6]. Because it is preferentially metabolized to the active molecule 5FU in tumor tissues, capecitabine is less toxic to healthy tissues [7], while having non-inferior efficacy as compared to intravenously administered 5FU (commonly given in conjunction with leucovorin (5FU/LV)) [8, 9]. However, hand-and-foot syndrome (HFS) is experienced by much more patients treated with capecitabine (60%) than with 5FU/LV (15%). This syndrome manifests as redness, swelling, numbness, or even painful blisters and desquamation of palm and sole skin, and often disturbs the daily activities of patients. HFS is measured on an ordinal scale of severity from grade 0 (none) to grade 3 (severe). According to the National Cancer Institute Common Terminology Criteria for Adverse Events v3.0 [10], grade 1 HFS is characterized by skin changes without pain, grade 2 by skin changes with pain not interfering with function, and grade 3 by skin changes with pain interfering with function. Grade 1 toxicity is considered acceptable within the context of cancer treatment and does not require any dosage modification. According to manufacturer's prescription guidelines [11], the occurrence of grade 2 or higher toxicity is considered intolerable and indicates treatment interruption until remission or at least decrease to grade 1 severity. Treatment is then reinitiated at a reduced dose (reduced by 25 or 50%) depending on the number of previous toxicity events. After the third event with grade 3 or the fourth with grade 2, treatment with capecitabine is terminated. However, the standard crude dose reductions may not be optimal and the therapeutic benefit might be improved by making patient-specific dose adaptations using model-based predictions. In addition, as with many anticancer drugs, no direct relationship between occurrence of HFS and plasma concentration of capecitabine or its metabolites has been detected. Pathophysiology of HFS has not been understood yet, and HFS is only suspected to be induced by the accumulation of the drug in the skin [12]. Therefore, the dosage of capecitabine should be based directly on the risk of intolerable grade HFS while maintaining at least the same efficacy in terms of tumor growth inhibition.

The present work aimed at developing an algorithm for individual prediction-based dose adaptation (IPBDA) based on a mixed-effects model for categorical endpoints. It is a direct application of the capecitabine-induced hand-and-foot syndrome model [13]. The method had to respect the requirement not to reduce the antitumor efficacy of the treatment, which was simulated using a colorectal tumor inhibition model [14]. The proof of concept was given by performing computer simulations of clinical trials that compared the standard and IPBDA methods.

Methods

Population hand-and-foot syndrome model

A mixed-effects proportional odds Markov model for an adverse effect called hand-and-foot syndrome (HFS) caused by the anticancer drug capecitabine has been described in [13], and we give only details relevant to the present simulation study.

While capecitabine is taken twice a day at 1,250 mg/m^2, this model assumes a total daily dose of 2,500 mg/m^2 is administered once a day. Because capecitabine is available in tablets of 150 and 500 mg and is taken twice per day, the manufacturer recommends to round the doses to 3,000; 3,300; 3,600; 4,000; 4,300; …; 5,600 mg per day, so that equal amounts can be taken in the morning and in the evening [11]. The drug was administered on a 3 weeks cycle basis: 2 weeks of treatment followed by 1 week of rest.

The severity of HFS is evaluated by grades from 0 (none) to 3 (severe). Since the frequency of grade 3 was low, grades 2 and 3 were combined into a single category

of grade ≥2 in the original model. HFS grade was evaluated once each week. The probabilities of grades are considered dependent on the preceding grade, i.e., a first-order Markov effect was incorporated in the population model. The model predicts that the overall distribution of grades in the treated population over 30 weeks (when standard dose adaptation is used) is approximately 66, 18 and 16% for grade 0, 1 and ≥2, respectively.

Due to the absence of pharmacokinetic data, the drug effect was quantified by a kinetic-pharmacodynamic (KPD) model [15], which relates the pharmacodynamic response to the administered drug doses without specifying a pharmacokinetic model. It is analogous to a PKPD model but the letter P is omitted to emphasize that PK is not measured. The principal idea of a KPD model is to assume a virtual compartment representing the biophase in which the concentration is in equilibrium with the observed effect. The elimination rate constant, K, from the virtual compartment governs the delay between the rate of dose administration and the observed effect. The parameter ED_{50} is the apparent in vivo potency of the drug at steady state, equal to the product of EC_{50}, the pharmacodynamic potency, and clearance, the PK "potency" at steady state.

The drug effect on the log-odds (logit) of the cumulative probabilities of grades is specified by an E_{MAX} function. The renal function (measured by creatinine clearance (CLcr) calculated using the Cockcroft–Gault formula [16]) at baseline was found to explain some of the interindividual variability and was included as an additive linear effect on the intercept [13], which corresponds to the idea that the better the renal function, the lower the patient's risk is. The full model is given below:

$$\frac{dQ(t)}{dt} = \text{Dose}(t) - K \cdot e^{\eta_{1i}} \cdot Q(t),$$

where $Q(t)$ is the accumulated drug amount in the virtual compartment at time t, Dose(t) is the amount of capecitabine administered at time t, 0 otherwise.

$$\text{logit}[P(Y_{it} \le 0 | Y_{it-1} = G^*)] = B_0^{G^*} - \frac{E_{\text{MAX}}^{G^*} \cdot (Q_{it} \cdot K \cdot e^{\eta_{1i}})}{ED_{50} + (Q_{it} \cdot K \cdot e^{\eta_{1i}})} + (\text{CLcr}_i - 75.5) \cdot \theta_{\text{CLcr}} + \eta_{2i}$$

$$\text{logit}[P(Y_{it} \le 1 | Y_{it-1} = G^*)] = B_0^{G^*} + B_1^{G^*} - \frac{E_{\text{MAX}}^{G^*} \cdot (Q_{it} \cdot K \cdot e^{\eta_{1i}})}{ED_{50} + (Q_{it} \cdot K \cdot e^{\eta_{1i}})} + (\text{CLcr}_i - 75.5) \cdot \theta_{\text{CLcr}} + \eta_{2i}$$

where Y_{it} is ith patient's grade at week t, G^* is the preceding grade (at week t–1); the baseline logit parameters, $B_0^{G^*}$ and $B_1^{G^*}$, as well as the maximum effect parameter, $E_{\text{MAX}}^{G^*}$, depend on G^* (i.e., each has three different values for $G^* = 0,\ 1,\ \ge 2$); η_{1i} and η_{2i} are individual-specific random effects, corresponding to the elimination rate constant K and the intercept, respectively. η_{1i} and η_{2i} follow a bivariate normal distribution with mean 0 and variance–covariance matrix Ω. η_{1i} adjusts the K value as an exponential multiplier ($K_i = K \cdot e^{\eta_{1i}}$,) while η_{2i} has an additive effect on the intercept.

The grade probabilities are obtained by the following transformations:

$$p_{it0} = P(Y_{it} = 0 | Y_{it-1} = G^*) = P(Y_{it} \le 0 | Y_{it-1} = G^*)$$
$$p_{it1} = P(Y_{it} = 1 | Y_{it-1} = G^*) = P(Y_{it} \le 1 | Y_{it-1} = G^*) - P(Y_{it} \le 0 | Y_{it-1} = G^*)$$
$$p_{it2} = P(Y_{it} = 2 | Y_{it-1} = G^*) = 1 - P(Y_{it} \le 1 | Y_{it-1} = G^*)$$

The model parameter estimates (as published in [13]) are given in Table 1.

Dose adaptation procedures

Standard dose adaptation was to reduce the initial dose by 25% after the second event with HFS grade >2 and by 50% after the third event.

Individual prediction-based dose adaptation (IPBDA) procedure consisted of: (1) estimating the individual random effects using both the data of the patient's past observations of HFS and the population model (the estimation step); (2) choosing the new dose so that the average risk of HFS grade ≥2 over the next 3 weeks would be closest to (but not greater than) the target risk (the dose calculation step).

The estimation step

Empirical Bayes' estimates (EBEs) of the individual random effects were obtained by the *maximum* a posteriori (MAP) method using the simplex optimization algorithm

Table 1 Hand-and-foot syndrome model parameter values [13]

Parameter	Value		
	$G^* = 0$	$G^* = 1$	$G^* = 2$
$B_0^{G^*}$	4.14	0.855	1.47
$B_1^{G^*}$	0.626	7.24	0.33
$E_{\text{MAX}}^{G^*}$	3.17	6.65	8.92
K		0.102	
ED_{50}		12,900	
θ_{CLcr}		0.0065	
ω_{η_1}		0.954	
ω_{η_2}		1.5	
$\text{corr}(\eta_1, \eta_2)$		0.67	

G^* HFS grade at previous week, ω_η standard deviations of random effects

Parameter significations are given in the text

[17]. Population parameters were not reestimated but fixed to their true value.

The distribution of $\eta_i = \begin{pmatrix} \eta_{1i} \\ \eta_{2i} \end{pmatrix}$ was assumed to be bivariate normal N(0, Ω), where

$$\Omega = \begin{bmatrix} \omega_{\eta_1}^2 & \rho\omega_{\eta_1}\omega_{\eta_2} \\ \rho\omega_{\eta_1}\omega_{\eta_2} & \omega_{\eta_2}^2 \end{bmatrix},$$

$$\rho = \mathrm{corr}(\eta_1, \eta_2) = \frac{\mathrm{cov}(\eta_1, \eta_2)}{\omega_{\eta_1}\omega_{\eta_2}}.$$

The likelihood function for the HFS grade at week t was denoted $\rho(Y_{it}|D_{it}, Y_{it-1}, \Theta, \mathrm{CLcr}_i, \eta_i)$, where $D_{it} = (\mathrm{dose}_{i1}, \ldots, \mathrm{dose}_{it})$ is the ith patient's dosing history, and the total set of population parameters was $\Theta = \left(B_0^0, B_0^1, B_0^2, B_1^0, B_1^1, B_1^2, E_{\mathrm{MAX}}^0, E_{\mathrm{MAX}}^1, E_{\mathrm{MAX}}^2, ED_{50}, K, \theta_{\mathrm{CLcr}}, \omega_{\eta_1}, \omega_{\eta_2}, \rho\right)$.

The indicator variable, z_{itg}, was used to indicate the toxicity grade of the ith patient at week t:

$$z_{itg} = \begin{cases} 1, & \text{if } Y_{it} = G, \\ 0, & \text{otherwise.} \end{cases} \text{with} \quad G = \{0, 1, \geq 2\}.$$

The likelihood of a single observation for the ith patient at week t, conditional on the individual parameter values η_i, was given by $p(Y_{it}|D_{it}, Y_{it-1}, \Theta, \mathrm{CLcr}_i, \eta_i) = \prod_{g=0}^{2} p_{itg}^{z_{itg}}$, with p_{itg} as defined in the model description.

The likelihood of all observations of the ith patient till week t was given by:

$$\prod_{j=1}^{t} p(Y_{ij}|D_{ij}, Y_{ij-1}, \Theta, \mathrm{CLcr}_i, \eta_i) = \prod_{j=1}^{t}\prod_{g=0}^{2} p_{ijg}^{z_{ijg}}.$$

The a posteriori distribution of η_i was obtained by applying Bayes' rule:

$$p(\eta_i|H_{it}, D_{it}, \Theta, \mathrm{CLcr}_i) = \frac{p(\eta_i) \cdot \prod_{j=1}^{t} p(Y_{ij}|D_{ij}, Y_{ij-1}, \Theta, \mathrm{CLcr}_i, \eta_i)}{p(H_{it})},$$

where $H_{it} = (Y_{i1}, \ldots, Y_{it})$ is the toxicity history for the ith patient.

The estimates of ith patient's random effects at week t were given by the mode of their a posteriori distribution:

$$\hat{\eta}_{it\mathrm{MAP}} = \mathrm{Arg}\left[\max_{\eta_i} \frac{p(\eta_i) \cdot \prod_{j=1}^{t} p\left(Y_{ij}|D_{ij}, Y_{ij-1}, \Theta, \mathrm{CLcr}_i, \eta_i\right)}{p(H_{it})}\right].$$

The calculation was performed using the simplex optimization algorithm [17], coded in Fortran 77 adapted from [18].

The dose calculation step

After the first appearance of HFS (any grade >0), the most appropriate dose for the next treatment cycle was considered to be the one with which the average predicted probability of HFS grade ≥2 over the next cycle (3 weeks) was closest to (but not higher than) the "target" risk of 4% (determined by sensitivity analysis). The lower limit of reduction was 50% of the nominal dose. The upper limit depended on the patient's HFS and tumor response history: if the patient was still in stable disease, doses could be increased up to 150% of the nominal dose after the first 4 cycles if the patient had no HFS at all or if grade 1 lasted for at least 6 consecutive weeks. If grade ≥2 had been observed at any time previously, the upper dose limit was the nominal dose (100%). Before reduction, possible daily doses were from 3,000 to 5,600 mg, with increments of 300 or 400 mg (3,000; 3,300; 3,600; 4,000; 4,300; …), according to prescription guidelines [11]. For adapted doses, the lower limit was 1,600 mg (closest to 50% of the minimal initial dose of 3,000 mg), the upper limit was 5,600 mg (the highest dose recommended in the prescription guidelines [11]). Once started, dose adaptation was made before starting each new cycle, after reestimation of the patient's HFS model random effects.

Colorectal tumor model

The tumor growth inhibition model was developed using phase II data capecitabine ($n = 34$) and phase III data of fluorouracil ($n = 252$) in advanced and/or metastatic colorectal cancer [14]. In our simulation study, we used only the parameter values estimated on capecitabine data. It describes tumor size (the sum of the longest diameters of target lesions) as a function of time and drug exposure. It accounts for the dynamics of natural tumor growth (k_g) and for the antitumor drug effect (k_d), as well as development of resistance to it (λ). The model is described by the following differential equation:

$$\frac{dy(t)}{dt} = k_g \cdot y(t) - k_d(t) \cdot \mathrm{Exposure}(t) \cdot y(t),$$
$$y(0) = \mathrm{baseline},$$

with $k_d(t) = k_{d,0} \cdot e^{-\lambda \cdot t}$, in which $y(t)$ is the tumor size at week t, Exposure(t) is the daily dose at week t. Inter-patient variability in the model parameters (k_g, k_d and λ) was assumed to be lognormally distributed. The values of parameters are given in Table 2.

Disease status

Dynamics of the disease was classified similarly to response evaluation criteria in solid tumors (RECIST) 1.1. [19] to:

Table 2 Colorectal tumor inhibition model parameter values [14]

Parameter	Value	Interindividual variance	CV%
Tumor growth rate k_g (week^{-1})	0.021	0.499	80
Drug cell kill rate k_d (week^{-1})	0.025	0.388	69
Resistance appearance rate λ (week^{-1})	0.053	1.260	159

CV coefficient of interindividual variability

- complete response (CR): observed sum of longest diameters <10 mm;
- partial response (PR): observed >30% reduction from baseline;
- progressive disease (PD): >20% and at least 5 mm increase above lowest observed value;
- stable disease (SD): all other cases.

In silico clinical trial protocol

One treatment cycle corresponded to 2,500 mg/m^2/day for 2 weeks, followed by 1 week rest. Treatment was interrupted in case of grade ≥2 HFS, until recovery to grade ≤1. Subsequent doses were modified according to the corresponding protocol. Fifty thousand patients per arm were simulated. Trial duration was 30 weeks. Patients were assumed to drop out of the trial if grade ≥2 HFS lasted more than 6 weeks or reoccurred for the 4th time, also if disease progression was observed. If patients had a complete response, they ended the treatment after 6 treatment cycles had been completed. HFS was monitored for 4 weeks after the end of treatment.

Simulation of the clinical trial data

The HFS grades were generated at the end of each week by simulation from the HFS model. Creatinine clearance (CLcr) and body surface area (BSA) values were randomly drawn from distributions estimated from a representative dataset of 595 patients used for HFS model building: a normal distribution with mean = 1.82 and standard deviation (SD) = 0.227 for BSA, and a log-normal distribution for CLcr with mean of log(CLcr) = 4.34 and SD of log(CLcr) = 0.349. The ranges were limited to [1.19, 2.5] for BSA and to [26.9, 218.5] for CLcr, according to the extreme values observed in the mentioned dataset. Individual random effects were drawn from a bivariate normal distribution with parameter values given in Table 1. The toxicity grades were obtained by random sampling according to the probability model. Tumor size baseline (base) observations were generated by sampling from a lognormal distribution with mean(log(base)) = 4.25 and SD(log(base)) = 0.5, with minimum accepted values equal 10 mm. This distribution corresponded to baseline data from the dataset used to estimate the tumor model. Subsequent observations were generated every 6 weeks by integrating the differential equation and then "adding" the measurement error: observation = true value * exp(error), with error sampled from a normal distribution with mean = 0 and SD = 0.25. Individual random effects were drawn from normal distributions with parameter values given in Table 2. The simulation of the trial was coded and performed in Fortran.

Criteria for comparing dose adaptation methods

Dose adaptation methods were compared in terms of toxicity related criteria: % of patients having (one and reoccurring) events with grade ≥2 HFS, average number of weeks with grade ≥2 HFS, average duration of reoccurring events with grade ≥2 HFS, % of patients who dropout due to HFS; as well as efficacy related criteria: % of patients having tumor response (PR + CR), % of patients who have progression of disease, relative change from baseline of tumor sizes.

Power analysis

In order to investigate the statistical power of the clinical trial for demonstrating the superiority of individualized dose adaptation versus the standard method, 100 replicate trials with 300, 350 and 600 patients per arm were simulated. The Wilcoxon rank sum test was used to test the difference in the total duration of time spent in HFS grade ≥2. The proportion of simulated studies with $P \leq 0.05$ gave the power for detecting a significant difference at $\alpha = 0.05$.

Results

Table 3 summarizes the main results of the virtual clinical trial, which compared the standard dose adaptation (reductions by 25 or 50%) and the individual prediction-based one, which uses the patient's HFS model to predict the future risk of grade ≥2 HFS and determines the next cycle dose accordingly. The overall dynamics of the

Table 3 Results of the two dose adaptation methods

Criteria	Standard	IPBDA
Toxicity (HFS)		
Average number of weeks with grade ≥2	5.2	3.8
Percentages of patients having grade ≥2	55.5	55.2
Percentages of patients having reoccurring events with grade ≥2	13.6	12.6
Duration of reoccurring events with grade ≥2 (weeks)	5.7	5.0
Percentages of patients who dropout due to HFS	23.2	21.6
Therapeutic efficacy		
Percentages of responders (CP + PR)	49.2	49.4
Percentages of patients who have disease progression	31.7	31.9
Relative change from baseline (median), %	−23.3	−23.1

IPBDA individual prediction-based dosage adaptation, *HFS* hand-and-foot syndrome, *CR* complete response, *PR* partial response

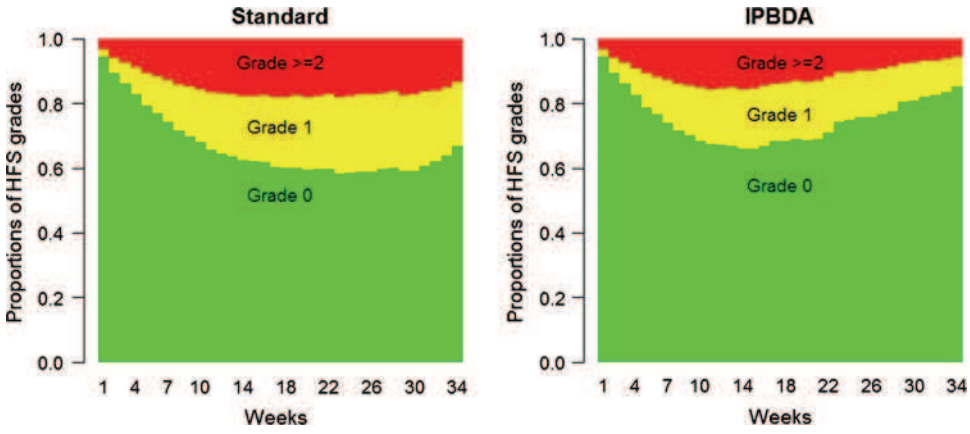

Fig. 1 Dynamics of the distributions of hand-and-foot syndrome grades with the standard dose adaptation (*left panel*) and the individual prediction-based one (IPBDA) (*right panel*)

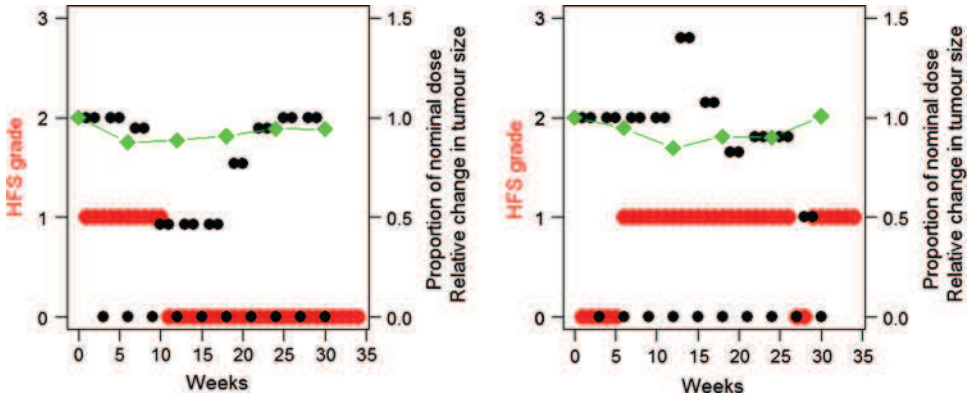

Fig. 2 Examples of individual changes in doses, HFS grades and tumor sizes. *Black dots* correspond to the dose amount relative to the initial dose; *red dots* correspond to the HFS grades; *green linked diamonds* correspond to tumor sizes relative to the initial size

distributions of grades in the two adaptation arms are shown in Fig. 1; a couple of examples of individual changes in the dose and HFS grade are shown in Fig. 2. The average total duration of grade ≥2 HFS was approximately 10 days shorter when applying the IPBDA as compared to the standard method (3.8 and 5.2 weeks, respectively; 27% relative reduction). This was achieved mainly by reduced frequency (from 13.6 to 12.6%) and duration (from 5.7 to 5 weeks) of reoccurring events with grade ≥2 HFS. Consequently, treatment discontinuation due to grade ≥2 HFS was 7% less frequent with IPBDA than with the standard approach (21.6 and 23.2%, respectively). This gain in terms of HFS toxicity was achieved with maintaining equivalent efficacy to the standard method with 49.4% of responders with individualized dosage regimen as compared to the 49.2% of the standard regimen (Table 3). A clinical trial comparing the IPBDA and the standard methods should include 350 patients per arm to achieve at least 90% power to show a difference in grade ≥2 HFS duration at an alpha level of 0.05.

Sensitivity analysis concerning IPBDA details has been performed, investigating the impact of different target risks

(TR) (4, 5, 6%), of the lower limit for dose (25, 50% of the nominal dose), of the upper limit for dose (100, 125, 150%). Different times of starting dose increases were tested: after 2 or 4 cycles. As it is observed by clinicians that if HFS remains at grade 1 for a long time, the risk to develop to a higher grade is very low, so it was tested if allowing dose increase after at least 6 consecutive weeks with grade 1 would be beneficial. Other tried variations were: lower TR for increases than for reductions (2, 3%), lower TR for reductions if patient has tumor response (2, 3%) (only if 95th percentile of predicted tumor size at the next measurement (in 6 weeks) does not correspond to disease progression). The results of all these variations were in smaller or bigger extent inferior to those of the presented specification of IPBDA (data not shown).

Discussion

This work aimed at developing and investigating the solutions for a rational individual adaptation of anti-cancer drug doses according to side-effects evaluated on a categorical scale and at evaluating the impact of dose adjustments in terms of antitumor efficacy. In particular, the study dealt with the hand-and-foot syndrome induced by capecitabine and its tumor growth inhibition action as efficacy biomarker. An alternative to the standard empirical dose reduction protocol that would optimize the dosages individually based on toxicity risk predictions given by a longitudinal patient-specific HFS model was investigated. Virtual clinical trials were simulated to compare the different adaptation approaches and demonstrated that individual prediction-based dose adaptation (even using only grade-type data) was a feasible strategy for improving toxicity control during chemotherapy, without compromising the therapeutic efficacy. In comparison with the PK-guided methods [3–5], the presented method of dose adaptation based on toxicity grade has the advantages of neither requiring costly and constraining invasive measurements nor the assumptions about the relationships of plasma concentrations and the observed toxic outcome.

In this particular example, the benefit was rather small, but clinically significant. The statistical significance ($\alpha = 0.05$) is reached in at least 90% of studies with 350 patients per arm. The main obstacle for a stronger impact was the lack of sensitivity of the grade probabilities to the dose changes. HFS develops and reduces much slower than doses change. The sensitivity analysis of the model parameters indicated that due to the assumed accumulation of capecitabine (through the KPD model), the doses received during the preceding 3 weeks has a much lower impact on toxicity risk than do the total accumulated doses received, previous HFS grade, and individual random effects. Therefore, some patients are predisposed to experience the HFS and the therapeutic index shrinks quickly for them. In these cases, monitoring the toxicity risk serves to indicate the time when treatment with capecitabine should be discontinued and replaced by another therapy.

As dose adaptation can only be started after appearance of HFS, IPBDA could only reduce the frequency of reoccurring events with grade $\geq$2 HFS, but not of the first event. When a patient has grade 1 HFS, the predicted risk of HFS developing to the grade $\geq$2 is lower than when in grade 0, so dose reduction is less likely to be done.

In general, the rather low quality of empirical Bayes' estimates (EBEs) of individual random effects may reduce the efficacy of IPBDA. In studying this issue [20], it was shown that unbiased and precise EBEs can be obtained only in particularly favorable conditions, which were impossible to meet in the HFS case. In particular, effect saturation must be reached to correctly identify the $E_{\max}$ function parameters, within-subject distribution of grades should be close to uniform, and a high number of observations are needed as categorical data is information-poor. The quality of EBEs influences the quality of prediction of severe toxicity risk, and therefore the decisions concerning the dose adaptation. In extreme cases of the HFS example, the error in dose due to imprecision of EBEs could be 10-fold (data not shown). However, due to low sensitivity of HFS to dose changes, the impact of poor EBEs on the performance of dose adaptation was minimal (data not shown).

Some other limitations imposed by the employed HFS model may include not considering some possibly influential information about the pharmacokinetics, such as the activity of dihydropyrimidine dehydrogenase (DPD), the rate-limiting enzyme for 5FU degradation [21]. Therefore, the approach presented here concerns only patients having no DPD deficiency.

Polymorphisms in genes coding for enzymes involved in transformation of capecitabine into 5FU have been associated with increased risk of severe HFS (cytidine deaminase [22]) and with better response and time to progression (carboxylesterase 2 [23]). Thymidylate synthase is a target enzyme for 5FU (and capecitabine, as its oral prodrug), polymorphisms in its gene have been associated with response to capecitabine in advanced colorectal cancer [24]. Methylenetetrahydrofolate reductase is a key enzyme in folate metabolism; polymorphisms in its gene have been related to toxicity rates with capecitabine [25]. Statistical associations were observed between polymorphisms in the transporter gene ATP-binding cassette B1 and a lower risk of HFS in capecitabine-treated patients [26]. Moreover, different amounts of folates in food were suggested as a possible explanation of regional differences in capecitabine tolerability [27]. The identified possible factors of

variability in efficacy and tolerability of capecitabine need further large well-designed studies to better determine their impact, and then could possibly be taken into account in dose adaptation schemes.

To enable extrapolation of the model-based dose adaptation algorithm to other dosing schedules, other cancer types and specific populations (prior therapy, performance status, age, sex, genotype, co-medications, etc.), mechanistic (physiologically based) models are needed.

The drawback of the Markov proportional odds model is that it is only a statistical description of dependence of the current week's toxicity grade on the previous week's toxicity grade; it does not rely on a physiological ground.

Furthermore, taking into account the possible uncertainty in grading might be suggested for future investigations in categorical data modeling, for example using hidden Markov chain model [28].

The adaptation method presented here determined the "best" dose according to a single HFS toxicity. Clinical benefit could be further improved by extending the method to address all frequent dose-limiting toxicities once their models are developed, and/or combined treatment protocols. The highest benefit of dose adaptations is expected to be achieved with reversible toxicities that have rapid kinetics, such as gastrointestinal ones.

This work constitutes the first application of an individualized dose adaptation of an anticancer drug based on a longitudinal mixed-effects dose-toxicity model for categorical observations. The proof of concept by clinical trial simulations represents the first step in the general approval in clinical practice of such an approach. Such virtual trials are particularly useful for assessing the possible impact of an investigational protocol and can be a valuable aid in decision-making concerning launching large, prospective, randomized clinical trials to validate the approach.

This work demonstrated the feasibility of individualized dose adaptation based on categorical endpoints and assessed its potential superiority to the standard empirical dose reductions via *in silico* clinical trials. Even in the case of the rather inert hand-and-foot syndrome induced by capecitabine, this new approach may enable clinically significant reduction in the duration of intolerable toxicity, as well as earlier detection of patients intolerant to the drug, without compromising drug efficacy on tumor shrinkage.

Acknowledgments We thank Novartis AG for providing financial support to Ines Paule's PhD. Pascal Girard was funded by INSERM (Institut National de la Santé et de la Recherche Médicale), France.

References

1. Miller AA (2002) Body surface area in dosing anticancer agents: scratch the surface!. J Natl Cancer Inst 94:1822–1823
2. Sheiner LB, Rosenberg B, Melmon KL (1972) Modelling of individual pharmacokinetics for computer-aided drug dosage. Comput Biomed Res 5:411–459
3. de Jonge ME, Huitema AD, Schellens JH et al (2005) Individualised cancer chemotherapy: strategies and performance of prospective studies on therapeutic drug monitoring with dose adaptation: a review. Clin Pharmacokinet 44:147–173
4. Zandvliet AS, Schellens JH, Beijnen JH, Huitema AD (2008) Population pharmacokinetics and pharmacodynamics for treatment optimization in clinical oncology. Clin Pharmacokinet 47:487–513
5. Rousseau A, Marquet P (2002) Application of pharmacokinetic modelling to the routine therapeutic drug monitoring of anticancer drugs. Fundam Clin Pharmacol 16:253–262
6. Hoff PM, Cassidy J, Schmoll HJ (2001) The evolution of fluoropyrimidine therapy: from intravenous to oral. Oncologist 6(Suppl 4):3–11
7. Blesch KS, Gieschke R, Tsukamoto Y et al (2003) Clinical pharmacokinetic/pharmacodynamic and physiologically based pharmacokinetic modeling in new drug development: the capecitabine experience. Invest New Drugs 21:195–223
8. Hoff PM, Ansari R, Batist G et al (2001) Comparison of oral capecitabine versus intravenous fluorouracil plus leucovorin as first-line treatment in 605 patients with metastatic colorectal cancer: results of a randomized phase III study. J Clin Oncol 19:2282–2292
9. Van Cutsem E, Twelves C, Cassidy J et al (2001) Oral capecitabine compared with intravenous fluorouracil plus leucovorin in patients with metastatic colorectal cancer: results of a large phase III study. J Clin Oncol 19:4097–4106
10. National Cancer Institute. Common terminology criteria for adverse events. http://ctep.cancer.gov/protocolDevelopment/electronic_applications/docs/ctcaev3.pdf. Accessed 2 May 2011
11. Capecitabine prescribing information. http://www.gene.com/gene/products/information/xeloda/pdf/pi.pdf. Accessed 2 May 2011
12. Kara IO, Sahin B, Erkisi M (2006) Palmar-plantar erythrodysesthesia due to docetaxel-capecitabine therapy is treated with vitamin E without dose reduction. Breast 15:414–424
13. Henin E, You B, Van Cutsem E et al (2009) A dynamic model of hand-and-foot syndrome in patients receiving capecitabine. Clin Pharmacol Ther 85:418–425
14. Claret L, Girard P, Hoff PM et al (2009) Model-based prediction of phase III overall survival in colorectal cancer on the basis of phase II tumor dynamics. J Clin Oncol 27:4103–4108
15. Jacqmin P, Snoeck E, van Schaick EA et al (2007) Modelling response time profiles in the absence of drug concentrations: definition and performance evaluation of the K-PD model. J Pharmacokinet Pharmacodyn 34:57–85
16. Cockcroft D, Gault MD (1976) Prediction of creatinine clearance from serum creatinine. Nephron 16:31–41
17. Nelder JA, Mead R (1965) A simplex method for function minimization. Comput J 7:308–313
18. Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1992) Numerical recipes in fortran 77: the Art of scientific computing. Cambridge University Press, Cambridge
19. Response Evaluation Criteria In Solid Tumors. http://www.recist.com. Accessed 2 May 2011
20. Paule I, Girard P, Tod M (2011) Empirical Bayes estimation of random effects of a mixed–effects proportional odds Markov model for ordinal data. Comput Methods Programs Biomed Published online 26-05-2011
21. Milano G et al (1999) Dihydropyrimidine dehydrogenase deficiency and fluorouracil-related toxicity. Br J Cancer 79:627–630
22. Caronia D, Martin M, Sastre J et al (2011) A polymorphism in the cytidine deaminase promoter predicts severe capecitabine-induced hand-foot syndrome. Clin Cancer Res 17(7):2006–2013

23. Ribelles N, López-Siles J, Sánchez A et al (2008) A carboxylesterase 2 gene polymorphism as predictor of capecitabine on response and time to progression. Curr Drug Metab 9(4):336–343
24. Park DJ, Stoehlmacher J, Zhang W et al (2002) Thymidylate synthase gene polymorphism predicts response to capecitabine in advanced colorectal cancer. Int J Colorectal Dis 17(1):46–49
25. Sharma R, Hoskins JM, Rivory LP et al (2008) Thymidylate synthase and methylenetetrahydrofolate reductase gene polymorphisms and toxicity to capecitabine in advanced colorectal cancer patients. Clin Cancer Res 14(3):817–825
26. Gonzalez-Haba E, García MI, Cortejoso L et al (2010) ABCB1 gene polymorphisms are associated with adverse reactions in fluoropyrimidine-treated colorectal cancer patients. Pharmacogenomics 11(12):1715–1723
27. Midgley R, Kerr DJ (2008) Capecitabine: have we got the dose right? Nat Clin Pract Oncol 6:17–24
28. Delattre M, Savic R, Miller R et al (2010) Estimation of mixed hidden Markov models with SAEM. Application to daily seizures data. PAGE 19 Abstr 1694. http://www.page-meeting.org/?abstract=1694. Accessed 2 May 2011

COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE XXX (2011) XXX–XXX

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Empirical Bayes estimation of random effects of a mixed-effects proportional odds Markov model for ordinal data

Inès Paule[a,b,*], *Pascal Girard*[a,b], *Michel Tod*[a,b,c]

[a] *Université de Lyon, Lyon, France*
[b] *EA3738 Ciblage thérapeutique en oncologie, Faculté de Médecine Lyon-Sud, Université Lyon 1, Oullins, France*
[c] *Hôpital Croix-Rousse, Hospices Civils de Lyon, Lyon, France*

ARTICLE INFO

Article history:
Received 20 December 2010
Received in revised form
29 March 2011
Accepted 27 April 2011

Keywords:
Ordered categorical data models
Mixed-effects models
Empirical Bayes estimation
Individual random effects
Optimization algorithm comparison

ABSTRACT

The objective of this work was to investigate the factors influencing the quality of empirical Bayes estimates (EBEs) of individual random effects of a mixed-effects Markov model for ordered categorical data. It was motivated by an attempt to develop a model-based dose adaptation tool for clinical use in colorectal cancer patients receiving capecitabine, which induces severe hand-and-foot syndrome (HFS) toxicity in more than a half of the patients. This simulation-based study employed a published mixed-effects model for HFS. The quality of EBEs was assessed in terms of accuracy and precision, as well as shrinkage. Three optimization algorithms were compared: simplex, quasi-Newton and adaptive random search. The investigated factors were amount of data per patient, distribution of categories within patients, magnitude of the inter-individual variability, and values of the effect model parameters. The main factors affecting the quality of EBEs were the values of parameters governing the dose–response relationship and the within-subject distribution of categories. For the chosen HFS toxicity model, the accuracy and precision of EBEs were rather low, and therefore the feasibility of their use for individual model-based dose adaptation seemed limited.

1. Introduction

For anti-cancer therapies, the occurrence of dose-limiting toxicities is a common event that often requires dose changes during the course of treatment. Adapting the dose for each individual, based on predictions of toxicity risk from a model describing the patient's toxicity–time course, is expected to be superior to the standard method of reducing doses in steps of 25% once severe toxicities are observed. A model-based method requires obtaining the toxicity model for each individual. Empirical Bayes estimates (EBE) of the individual random effects (RE) are obtained by the relative weighting of the patient's toxicity data collected during treatment, and the population toxicity model, for which the parameter values are known.

Many non-haematological adverse effects are evaluated on a graded scale with typically 4 or 5 ordered categories. Such data are frequently described by ordered categorical models, including the proportional odds model [1–3]. Categorical data contain less information than continuous data, making it more difficult to estimate each patient's toxicity

* *Corresponding author at*: EA3738 Ciblage thérapeutique en oncologie, Faculté de Médecine Lyon-Sud, Université Lyon 1, Oullins, France. Tel.: +33 426235959.
E-mail addresses: ines.paule@gmail.com, ines.paule@recherche.univ-lyon1.fr (I. Paule).
0169-2607/$ – see front matter © 2011 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cmpb.2011.04.006

parameters. To our knowledge, there has been no study published on the feasibility to obtain unbiased and precise estimates of individual parameters of ordinal categorical data models.

The objective of this work was to investigate the estimation of individual random effects of a mixed-effects Markov ordered categorical data model, and the factors impacting on the quality of the estimates. Insight into this issue should help determine the feasibility of individual dose adaptation based on ordinal categorical data models for adverse effects.

A simulation study comparing estimation methods for ordered categorical data was published in [4]. They compared estimation methods that employed different approximations of the integral of the likelihood (the Laplacian and the Gaussian Quadrature (GQ) methods). Savic et al. [5] have used the same simulation study to investigate and compare the quality of population parameter estimates given by Stochastic Approximation Expectation Maximization (SAEM) method. In [4], the distributions of the simulated REs and of the EBEs were compared. The authors observed that the EBE distributions are highly non-normal and non-symmetric when the observations are unevenly distributed in the categories. This leads to bias in population parameter estimates when using the Laplacian method (without centering option [6]) in NONMEM and SAS NLMIXED, which conditions on EBEs of REs without taking into account their uncertainty. Since GQ methods allow uncertainty around EBEs and SAEM method does not use them at all, they produce population parameter estimates without appreciable bias.

Estimation methods for binary data models were investigated in [7] and the authors have also observed that the Laplacian method produced biased population parameter estimates when the distribution of categories was highly uneven.

Kowalski et al. [3] looked at the distribution of EBEs in which 82% of the individuals had only grades 0 and observed a highly bimodal distribution, where EBEs of individuals with only grades 0 were very close to 0 and the EBEs of the remaining individuals where large positive values. They proposed a two-part mixture model (where random effects were specified only for individuals with toxicity) to improve the distribution of EBEs, but did not perform a simulation study to allow comparison of true (simulated) and estimated values of random effects.

Individual parameter estimates of PKPD mixed effects models for continuous data were investigated in [8], where the authors bring to attention the shrinkage of EBEs when there is little data per individual. Categorical data contain less information than continuous data; therefore it is suspected that EBEs' quality might be an even bigger issue than in continuous data models.

This work presents a simulation study to assess the factors influencing the quality of EBEs of random effects in ordinal data mixed-effects models. The model used in this simulation exercise was a mixed-effects proportional odds Markov model for the hand-and-foot syndrome toxicity [9].

Several estimation algorithms were compared, and a sensitivity analysis was carried out, investigating the possible impact on the EBE quality of several factors, namely the amount of data per patient, the distribution of categories within individuals, the magnitude of the true variability of random effects, and the values of population parameters governing the dose–response relationship. The tested estimation methods were: two local optimization algorithms, quasi-Newton (as implemented in NONMEM software [6]) and simplex [10]; a global optimization algorithm, adaptive random search (ARS) [11]; and MCMC sampling from the *a posteriori* RE distribution (using WinBUGS software [12]). The main reason to use 4 different algorithms was to make sure that there is not some purely numerical problem specific to this model. It is known that optimization algorithms are not equivalently efficient for different problems. Global optimization method ARS was tested to ensure that there was no problem of local optimums in the estimates of simplex and quasi-Newton. Estimation of means of *a posteriori* distributions using WinBUGS was used to check if it would be a better estimator than the mode. The performance of the methods was investigated in terms of bias and precision of the estimates, as well as approximate run time.

2. Methods

2.1. Population hand-and-foot syndrome model

A mixed-effects proportional odds Markov model for an adverse effect called hand-and-foot syndrome (HFS) caused by the anticancer drug capecitabine has been described [9] and this was the model used for our own simulation study.

While capecitabine is taken twice a day at 1250 mg/m^2, this model assumes a total daily dose of 2500 mg/m^2 is administered once a day. The drug is taken intermittently: 2 weeks of treatment followed by 1 week of rest, which corresponds to one cycle of 3 weeks.

The severity of HFS is evaluated by grades from 0 (no toxicity) to 3 (severe toxicity). However, because of the low frequency of grade 3, grades 2 and 3 were combined into a single category of grade $\geq$2 in the original model and used in the same way in our simulations from it. The toxicity grade was evaluated once each week. The probabilities of grades are dependent on the preceding grade, i.e. a first-order Markov model was incorporated in the population model. The model predicted overall distribution of grades in the treated population, over 30 weeks and when standard dose adaptation is used, is approximately 66%, 18%, 16% for grade 0, 1 and $\geq$2 respectively.

Due to the absence of pharmacokinetic data, the drug effect was quantified by a kinetic–pharmacodynamic (KPD) model [13], which relates the pharmacodynamic response to the administered drug doses without specifying a pharmacokinetic model. It is analogous to a PKPD model but the letter P is omitted to emphasize that PK is not measured. The principal idea of a KPD model is to assume a virtual compartment representing the biophase in which the concentration is in equilibrium with the observed effect. The elimination rate constant, K, from the virtual compartment governs the delay between the rate of dose administration and the observed effect. The parameter ED_{50} is the apparent in vivo potency of the drug at steady state, equal to the product of EC_{50}, the pharmacodynamic potency, and clearance, the PK "potency" at steady state.

Please cite this article in press as: I. Paule, et al., Empirical Bayes estimation of random effects of a mixed-effects proportional odds Markov model for ordinal data, Comput. Methods Programs Biomed. (2011), doi:10.1016/j.cmpb.2011.04.006

The drug effect on the log-odds (logit) of the cumulative probabilities of grades is specified by an E_{max} function. The renal function (measured by creatinine clearance (CLcr) calculated using the Cockcroft–Gault formula [14]) at baseline was found to explain some of the interindividual variability and was included as an additive linear effect on the intercept, which corresponds to the idea that the better the renal function, the lower the patient's risk is. The full model is given below:

$$\frac{dQ(t)}{dt} = Dose(t) - K \cdot e^{\eta_{1i}} \cdot Q(t),$$

where $Q(t)$ is the accumulated drug amount in the virtual compartment at time t, Dose(t) is the amount of capecitabine administered at time t, 0 otherwise.

$$\text{logit}[P(Y_{it} \le 0 | Y_{it-1} = G_*)] = B_0^{G_*} - \frac{E_{max}^{G_*} \cdot (Q_{it} \cdot K \cdot e^{\eta_{1i}})}{ED_{50} + (Q_{it} \cdot K \cdot e^{\eta_{1i}})} + (\text{CLcr}_i - 75.5) \cdot \theta_{\text{CLcr}} + \eta_{2i}$$

$$\text{logit}[P(Y_{it} \le 1 | Y_{it-1} = G_*)] = B_0^{G_*} + B_1^{G_*} - \frac{E_{max}^{G_*} \cdot (Q_{it} \cdot K \cdot e^{\eta_{1i}})}{ED_{50} + (Q_{it} \cdot K \cdot e^{\eta_{1i}})} + (\text{CLcr}_i - 75.5) \cdot \theta_{\text{CLcr}} + \eta_{2i}$$

where Y_{it} is ith patient's grade at time t, G_* is the preceding grade (at time $t-1$); the baseline logit parameters, $B_0^{G_*}$ and $B_1^{G_*}$, as well as the maximum effect parameter, $E_{max}^{G_*}$, depend on G_* (i.e. each has three different values for $G_* = 0, 1, \ge 2$); η_{1i} and η_{2i} are individual-specific RE, corresponding to the elimination rate constant K and the intercept respectively. η_{1i} and η_{2i} follow a bivariate normal distribution with mean 0 and variance–covariance matrix Ω. η_{1i} adjusts the K value as an exponential multiplier ($K_i = K \cdot e^{\eta_{1i}}$), while η_{2i} has an additive effect on the intercept.

The grade probabilities are obtained by the following transformations:

$$p_{it0} = P(Y_{it} = 0 | Y_{it-1} = G_*) = P(Y_{it} \le 0 | Y_{it-1} = G_*)$$
$$p_{it1} = P(Y_{it} = 1 | Y_{it-1} = G_*) = P(Y_{it} \le 1 | Y_{it-1} = G_*) - P(Y_{it} \le 0 | Y_{it-1} = G_*)$$
$$p_{it2} = P(Y_{it} = 2 | Y_{it-1} = G_*) = 1 - P(Y_{it} \le 1 | Y_{it-1} = G_*)$$

The model parameter estimates (as published in [9]) are given in Table 1.

Table 1 – HFS model parameter values (G_*: HFS grade at $t-1$; ω_η: standard deviations of RE).

Parameter	Value		
	$G_* = 0$	$G_* = 1$	$G_* = 2$
$B_0^{G_*}$	4.14	0.855	1.47
$B_1^{G_*}$	0.626	7.24	0.33
$E_{max}^{G_*}$	3.17	6.65	8.92
K		0.102	
ED_{50}		12,900	
θ_{CLcr}		0.0065	
ω_{η_1}		0.954	
ω_{η_2}		1.5	
corr(η_1, η_2)		0.67	

2.2. Simulation of data

The data were generated by simulation from the HFS model described in Section 2.1. Creatinine clearance and body surface area (BSA) values were randomly drawn from distributions estimated by HFS model authors from a representative dataset of 600 patients used for model building: a normal distribution with mean = 1.82 and standard deviation (SD) = 0.227 for BSA, and a log-normal distribution for CLcr with mean of log(CLcr) = 4.34 and SD of log(CLcr) = 0.349. The ranges were limited to [1.19, 2.5] for BSA and to [26.9, 218.5] for CLcr, according to the extreme values observed in the mentioned dataset. Individual REs were drawn from a bivariate normal distribution with parameter values given in Table 1. The toxicity grades were obtained by random sampling according to the likelihood.

In the simulations, the starting daily dose was 2500 mg/m^2, with an intermission of 1 week after every 2 weeks – the approved standard regimen. If a patient experienced severe toxicity (grade $\ge$2), treatment was stopped until recovery to at least grade 1 and then continued with doses reduced according to an individual model-based prediction of toxicity risk. This calculation was performed in the following order: (1) estimate η_i from the current individual data; (2) calculate the dose for the next treatment cycle so that the predicted probability of grade $\ge$2 at the end of the following 2 weeks is equal to 1%. If the estimated "optimal" dose is lower than 50% of the starting dose, no drug is given at all. Doses are not increased above the starting dose. Seeking to prevent severe toxicity, such reassessment of the "optimal" dose was also performed before starting each 3-week cycle, once a grade 1 toxicity had been observed. If the treatment had to be interrupted for more than 6 weeks, it was definitely stopped. However, given the model, there is still drug remaining in the body and the toxic effect may continue. Therefore, HFS observations after the treatment has been stopped were included in the EBE estimation. 1500 virtual patients were simulated for 29 weeks, corresponding to 10 standard treatment cycles. There was one observation of toxicity grade per week. The simulation was performed in Trial Simulator 2 [15].

2.3. Estimation of individual parameters

An empirical Bayes approach was used to estimate the individual RE estimates, given the patient's simulated response data during treatment and the population parameter values. For estimation of individual parameters, the population parameters were fixed to their true value (i.e. population parameters were not estimated).

2.3.1. Distribution of random effects

The distribution of $\eta_i = (\eta_{1i}, \eta_{2i})$ was assumed to be bivariate normal $N(0, \Omega)$, where:

$$\Omega = \begin{bmatrix} \omega_{\eta_1}^2 & \rho\omega_{\eta_1}\omega_{\eta_2} \\ \rho\omega_{\eta_1}\omega_{\eta_2} & \omega_{\eta_2}^2 \end{bmatrix}, \quad \rho = \text{corr}(\eta_1, \eta_2) = \frac{\text{cov}(\eta_1, \eta_2)}{\omega_{\eta_1}\omega_{\eta_2}}.$$

2.3.2. Likelihood function

The likelihood function for the HFS grade at time t was denoted $p(Y_{it} | D_{it}, Y_{it-1}, \Theta, \text{CLcr}_i, \eta_i)$, where $D_{it} = (\text{dose}_{i1},$

Please cite this article in press as: I. Paule, et al., Empirical Bayes estimation of random effects of a mixed-effects proportional odds Markov model for ordinal data, Comput. Methods Programs Biomed. (2011), doi:10.1016/j.cmpb.2011.04.006

..., $dose_{it}$) is the ith patient's dosing history, and the total set of population parameters is denoted by $\Theta = (B_0^0, B_0^1, B_0^2, B_1^0, B_1^1, B_1^2, E_{max}^0, E_{max}^1, E_{max}^2, ED_{50}, K, \theta_{CLcr}, \omega_{\eta_1}, \omega_{\eta_2}, \rho)$.

The indicator variable, z_{itg}, was defined to indicate the toxicity grade of the ith patient at week t:

$$z_{itg} = \begin{cases} 1, & \text{if} Y_{it} = G, \\ 0, & \text{otherwise.} \end{cases} \quad \text{with} \quad G = \{0,\ 1,\ \geq 2\}$$

The likelihood of a single observation for the ith patient at time t, conditional on the individual parameter values η_i, was given by:

$$p(Y_{it} | D_{it}, Y_{it-1}, \Theta, CLcr_i, \eta_i) = \prod_{g=0}^{2} p_{itg}^{z_{itg}},$$

with p_{itg} as defined in Section 2.1.

The likelihood of all observations of the ith patient till time t was given by:

$$\prod_{j=1}^{t} p(Y_{ij} \big| D_{ij}, Y_{ij-1}, \Theta, CLcr_i, \eta_i) = \prod_{j=1}^{t} \prod_{g=0}^{2} p_{ijg}^{z_{ijg}}$$

2.3.3. *A posteriori distribution of individual random effects*

The *a posteriori* distribution of η_i was obtained by applying Bayes' rule:

$$p(\eta_i | H_{it}, D_{it}, \Theta, CLcr_i) = \frac{p(\eta_i) \cdot \prod_{j=1}^{t} p(Y_{ij} \big| D_{ij}, Y_{ij-1}, \Theta, CLcr_i, \eta_i)}{p(H_{it})},$$

where $H_{it} = (Y_{i1}, \ldots, Y_{it})$ is the toxicity history for the ith patient.

2.3.4. *Empirical Bayes estimates*

2.3.4.1. *Maximum a posteriori estimates.* Using the method of *maximum a posteriori* (MAP), the RE estimates were given by the mode of their *a posteriori* distribution:

$$\hat{\eta}_{itMAP} = Arg \left[\max_{\eta_i} \frac{p(\eta_i) \cdot \prod_{j=1}^{t} p(Y_{ij} \big| D_{ij}, Y_{ij-1}, \Theta, CLcr_i, \eta_i)}{p(H_{it})} \right].$$

As the denominator $p(H_{it})$ does not depend on η_i, it may be neglected in the maximization procedure for $\hat{\eta}_{itMAP}$, so that

$$\hat{\eta}_{itMAP} = Arg \left[\max_{\eta_i} \left\{ p(\eta_i) \cdot \prod_{j=1}^{t} p(Y_{ij} \big| D_{ij}, Y_{ij-1}, \Theta, CLcr_i, \eta_i) \right\} \right].$$

Taking the log of the *a posteriori* distribution, and given the RE distribution function $p(\eta_i) = \frac{1}{(2\pi|\Omega|^{1/2})} \exp\left(-\frac{1}{2}\eta_i^T \Omega^{-1} \eta_i\right)$, the MAP estimate is given by

$$\hat{\eta}_{itMAP} = Arg \left[\max_{\eta_i} \left\{ \log \left(\frac{1}{2\pi|\Omega|^{1/2}} \exp\left(-\frac{1}{2}\eta_i^T \Omega^{-1} \eta_i\right) \prod_{j=1}^{t} \prod_{g=0}^{2} p_{ijg}^{z_{ijg}} \right) \right\} \right]$$

The function to be maximized may be further simplified to $-\frac{1}{2}\eta_i^T \Omega^{-1} \eta_i + \log\left(\prod_{j=1}^{t} \prod_{g=0}^{2} p_{ijg}^{z_{ijg}}\right)$, demonstrating that $\hat{\eta}_{itMAP}$ is a balance between the fit to the individual data and the population fit.

2.3.4.2. *Mean of a posteriori random effects distribution.* The mean of the *a posteriori* RE distribution can also be estimated as an alternative to the mode. Here, the Bayesian modelling software WinBUGS v1.4. [12] was used to draw samples from the *a posteriori* distribution of η for each individual. Means of these samples were compared to the modes obtained with simplex.

The used prior distribution for $\eta_i = (\eta_{1i}, \eta_{2i})$ was a bivariate normal distribution with parameter values as given in the published model for the interindividual variability matrix (Table 1). The total number of iterations was 100,000, with first 4000 discarded as "burnin", therefore the sample size per chain was 96,000. Convergence was checked by the Gelman–Rubin statistic [16], defined as $\sqrt{\hat{R}} = \sqrt{1 - (1/n) + (1/n)(B/W)}$, where W is parameter within-chain variance, B is parameter between-chain variance, n is the number of chains. It requires several chains to be run from different initial values. The convergence is said to be achieved when the value of $\sqrt{\hat{R}}$ is close to 1, which corresponds to within-chain variances being very close to between-chain variance. If $\sqrt{\hat{R}}$ is not close to 1 then the chains have not converged to the posterior distribution of the parameter. This statistic is based on the assumption that the target distribution is normal.

The dispersion of the *a posteriori* distribution for each individual η indicates the uncertainty in the estimate of each individual's η. To summarise the dispersion of the *a posteriori* distribution of the individual ηs, the median and the 90% interval of SDs across the population (with at least one grade >0) was given. The same summary statistics of standard errors (SE) of individual ηs given by NONMEM (PsN function [17]) were also shown.

2.4. *Optimization algorithms for MAP estimation*

2.4.1. *Simplex*

This local optimization algorithm, which does not require function derivatives, was introduced in [10]. Here it was coded in Fortran 77 adapted from [18].

2.4.2. *Quasi-Newton*

This local optimization method, which uses the Hessian matrix (second derivatives) of the objective function, was used as implemented in NONMEM software [6].

2.4.3. *Adaptive random search (ARS)*

This global optimization algorithm is based on random sampling with constant readjustment of sampling spaces. Here it was adapted and coded in Fortran 77 according to a scheme given in [11].

Please cite this article in press as: I. Paule, et al., Empirical Bayes estimation of random effects of a mixed-effects proportional odds Markov model for ordinal data, Comput. Methods Programs Biomed. (2011), doi:10.1016/j.cmpb.2011.04.006

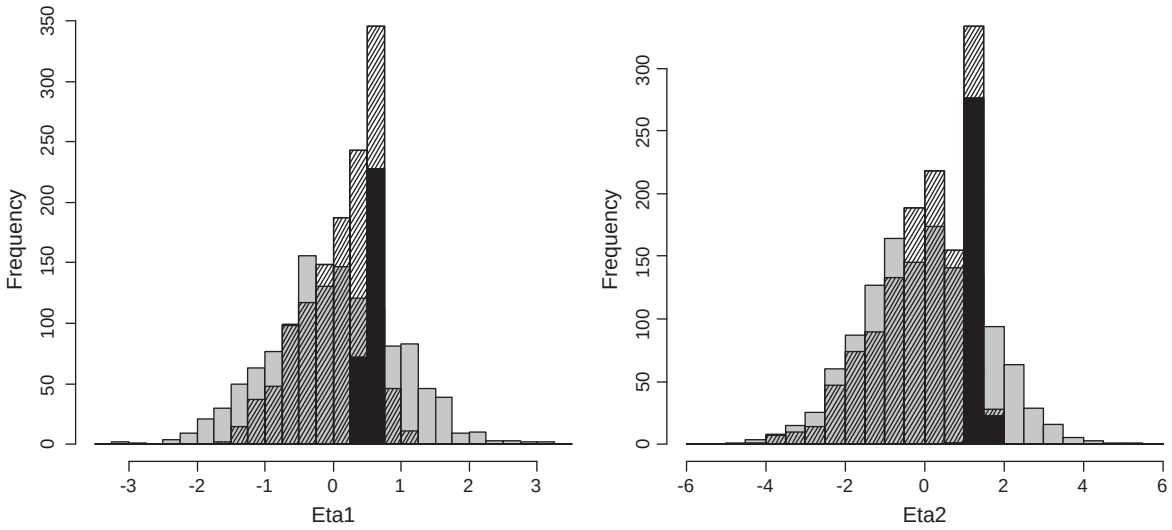

Fig. 1 – Histograms of simulated η values (grey) and EBEs (striped). Black areas represent the EBEs of the subgroup of patients with only grades 0.

2.5. *Estimation quality criteria*

Accuracy and precision of the estimation methods were compared using bias and mean absolute error (MAE).

$$\text{Bias} = \frac{1}{N}\sum_{i=1}^{N}(\hat{\eta}_i - \eta_i) \quad \text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{\eta}_i - \eta_i\right|$$

In order to facilitate the interpretation of the absolute values of bias and MAE, relative measures were obtained by dividing those criteria by the true SD of ηs (and then giving the result in %). The approximate time needed was also recorded.

Shrinkage shows how much the individual parameter estimates are "shrunk" to the population mean, due to the low amount of data (per patient) to inform about the true individual values. It was calculated as $1 - (SD(\hat{\eta}_i)/SD(\eta_i))$ [8].

2.6. *Sensitivity analysis*

2.6.1. *Magnitude of the true interindividual variability*

The impact of the magnitude of the true interindividual variability (IIV) was assessed by simulating datasets from the model with modified IIV: 10%, 50%, 200% and 300% of the original SD. 5000 patients were simulated for 30 weeks applying the individual model-based dose adaptation. Empirical Bayes estimation was performed using the simplex method.

2.6.2. *Amount of data per subject*

To investigate the impact of the amount of data per patient, 10,000 virtual patients were simulated for 200 weeks applying the individual model-based dose adaptation. REs were estimated using the simplex method after 29, 100 and 200 weeks for those patients who had at least one non-zero grade over the corresponding periods.

2.6.3. *Parameters of the effect model*

In order to investigate the impact of the effect model parameters, the values of parameters K and ED_{50} were altered. 14 combinations of modified parameter values were tested, the range of altered values was from *5 to *20 for K and from *1/2 to *1/30 for ED50. For each case, 1500 virtual patients were simulated for 29 weeks applying the individual model-based dose adaptation. EBEs of patients with at least one non-zero grade were obtained by simplex.

2.6.4. *Within-patient distribution of categories*

In order to assess a possible impact of the individual distribution of grades, grades were counted for each individual and related to the error of their EBEs. This analysis used a dataset of 7652 virtual patients with at least one non-zero grade simulated for 29 weeks applying the individual model-based dose adaptation.

3. Results

The histograms of RE estimates (obtained by ARS) showed a high peak in a narrow interval for η_1 and η_2 (Fig. 1). It was found that this corresponded to patients who had no toxicity (only grades 0). Their response data bring the same information, therefore their EBEs are close to each other, varying only due to differences in doses and in CLcr. However, this is not a concern from the clinical point of view, since these individuals do not require dose reductions. Therefore, in order to have a more precise evaluation of estimation quality for the patients at risk, only the data of the first 1000 patients with at least one grade higher than 0 were used in the subsequent analysis.

3.1. *Comparison of optimization algorithms for MAP*

All the three tested optimization algorithms gave similar results (Table 2), those of the simplex and ARS were almost identical, with a big advantage of simplex in terms of run time. The EBEs had some bias and low precision (Fig. 2), especially η_1 corresponding to the elimination rate constant K.

Please cite this article in press as: I. Paule, et al., Empirical Bayes estimation of random effects of a mixed-effects proportional odds Markov model for ordinal data, Comput. Methods Programs Biomed. (2011), doi:10.1016/j.cmpb.2011.04.006

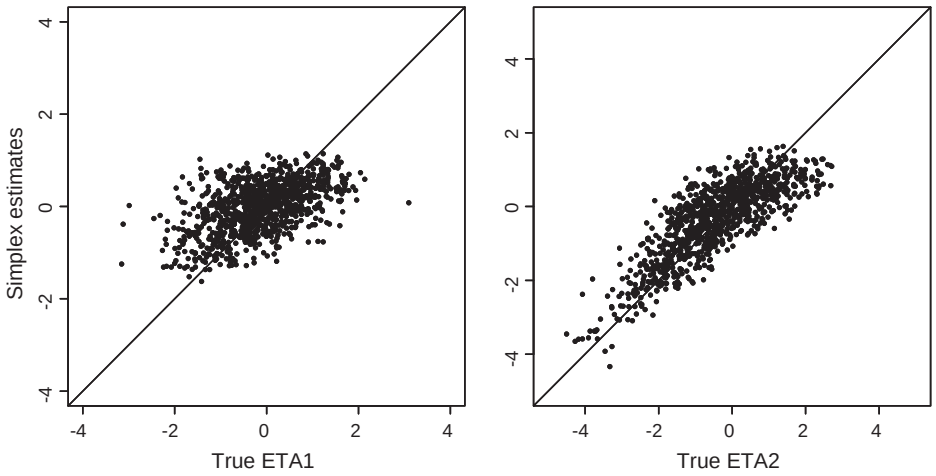

Fig. 2 – Simplex estimates vs. true values of η_i. n = 1000 patients with at least one grade > 0.

3.2. Comparison of estimators (mode and mean)

For 161 of 1000 patients WinBUGS did not work well: there was no convergence according to the $\sqrt{\hat{R}}$ criterion, the chains almost did not move from the initial values. Nothing particular in the data of these patients could be found. The distributions of their simplex MAP estimates were comparable to those of other patients. Estimation with "over relax" did not improve the result. "Over relax" option generates multiple samples at each iteration and then selects one that is negatively correlated with the current value. The within-chain correlations should be reduced and hence fewer iterations may be necessary [12].

The mean of the *a posteriori* distribution was a less biased estimator than the mode (Table 3, the results are given for 839 patients for whom the estimation in WinBUGS worked correctly).

Table 2 – Estimation quality of different optimization algorithms. In parentheses: bias or MAE divided by the true SD of η. n = 1000 patients with at least one grade >0.

	Simplex	NONMEM	ARS
Bias η_1	0.120 (13%)	0.102 (11%)	0.120 (13%)
Bias η_2	0.086 (6%)	0.098 (7%)	0.085 (6%)
MAE η_1	0.592 (62%)	0.608 (64%)	0.592 (62%)
MAE η_2	0.595 (41%)	0.607 (40%)	0.595 (41%)
Time	5 s	21 s	5 min 40 s

Table 3 – Estimation quality of the mean and the mode. In parentheses: bias or MAE divided by the true SD of η.

	Mean (WinBUGS)	Mode (simplex)
Bias η_1	−0.029 (3%)	0.101 (11%)
Bias η_2	−0.014 (1%)	0.076 (5%)
MAE η_1	0.587 (62%)	0.586 (61%)
MAE η_2	0.596 (40%)	0.605 (40%)
Time	7 h 52 min	4 s

3.3. Uncertainty of EBEs

The credible interval of the *a posteriori* distribution (from WinBUGS) of each individual η was typically wide: the median SD of the 839 individual's *a posteriori* distributions for η_1 was 0.75 (the 90% interval of SDs of each individual's *a posteriori* distribution was (0.54, 0.97)), the median SD of η_2 was 0.7 with 90% interval (0.55, 1.01). Comparison of these values to the size of population variance (SD(η_1) = 0.954, SD(η_2) = 1.5) clearly indicates a high uncertainty of the EBEs.

The SEs of EBEs (median and 90% interval) from NONMEM were close to the SDs from WinBUGS: 0.85 (0.51, 1.14) for η_1 and 0.8 (0.54, 1.08) for η_2.

3.4. Impact of the magnitude of the true interindividual variability on the EBE quality

There was no clear tendency of the impact of the true IIV on the quality of EBEs. However, some differences were observed: with lower IIV of η_1, its bias tended to be lower, shrinkage higher, MAE of η_2 slightly smaller; with lower IIV of η_2, its shrinkage tended to be higher, MAE of η_1 slightly larger.

3.5. Impact of the amount of data on the EBE quality

The results in Table 4 showed that more individual data improved the EBEs in terms of precision, and of bias to some extent, but did not reduce it to 0 for η_1. As expected, shrinkage was decreasing with increasing amount of individual data.

3.6. Impact of the effect model parameter values on the EBE quality

The best EBEs were obtained with ED_{50} divided by 20 and K multiplied by 10. A comparison of the results with these parameter values and with the original values in Table 5 and Fig. 3 showed that the effect model parameter values have a high impact on both bias and precision of EBEs.

Please cite this article in press as: I. Paule, et al., Empirical Bayes estimation of random effects of a mixed-effects proportional odds Markov model for ordinal data, Comput. Methods Programs Biomed. (2011), doi:10.1016/j.cmpb.2011.04.006

Table 4 – Estimation (simplex) quality with increased amount of data: 29, 100 and 200 weeks. In parentheses: bias or MAE divided by the true SD of η.

	29 weeks	100 weeks	200 weeks
Bias η_1	0.159 (17%)	0.068 (7%)	0.061 (6%)
Bias η_2	0.102 (7%)	0.003 (0.2%)	−0.011 (0.8%)
MAE η_1	0.616 (65%)	0.513 (54%)	0.498 (52%)
MAE η_2	0.583 (39%)	0.383 (26%)	0.313 (21%)
Shrinkage η_1 (all patients)	45%	24%	20%
Shrinkage η_2 (all patients)	21%	8%	6%
Shrinkage η_1 (excluding patients without toxicity)	40%	23%	20%
Shrinkage η_2 (excluding patients without toxicity)	18%	8%	5%

Table 5 – Estimation (simplex) quality with a more reactive model and with the original one. In parentheses: bias or MAE divided by the true SD of η.

	Model with $ED_{50}/20$ and K^*10	Original model
Bias η_1	0.005 (0.5%)	0.120 (13%)
Bias η_2	−0.041 (3%)	0.086 (6%)
MAE η_1	0.234 (25%)	0.592 (62%)
MAE η_2	0.531 (35%)	0.595 (41%)
Shrinkage η_1 (all patients)	9%	43%
Shrinkage η_2 (all patients)	15%	22%
Shrinkage η_1 (excluding patients without toxicity)	7%	39%
Shrinkage η_2 (excluding patients without toxicity)	17%	20%

3.7. *Impact of the within-patient distribution of categories on the EBE quality*

Results summarised in Table 6 show that the higher the proportion of grades 1 and 2 (the closer to a uniform distribution), the better the RE estimates.

Table 6 – Counts of grades according to the error of EBEs (in patients with at least one grade >0 over 29 weeks).

	Mean count of weeks in grade 2	Mean count of weeks in grade 1	Number of patients
η_1 error			
[min = −3.2,−2)	2.1	2.7	24
[−2,1)	2.4	3.7	416
[−1,0)	4.6	5.1	2908
(0,1]	5.0	5.0	3241
(1,2]	4.2	4.7	957
(2, max = 3.08]	3.5	4.5	106
η_2 error			
[min = −3.91,−2)	0.7	1.1	38
[−2,1)	2.8	3.1	436
[−1,0)	4.5	5.0	2855
(0,1]	5.0	5.2	3503
(1,2]	4.1	4.5	784
(2, max = 2.96]	2.8	4.8	36

4. Discussion

In this work, simulations were conducted to investigate what were the potential factors influencing the quality of EBEs of random effects in mixed-effects models for ordinal data. Comparison of the optimization methods showed that there was almost no difference in the quality of the MAP estimates in this study. The simplex method, a local optimization algorithm, gave almost identical estimates to those of a global

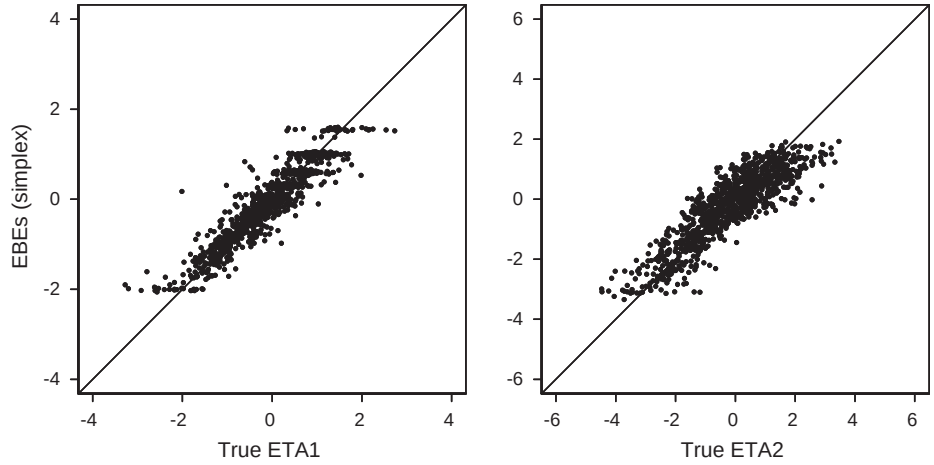

Fig. 3 – Simplex estimates vs. true η_i with altered values of K and ED_{50}.

Please cite this article in press as: I. Paule, et al., Empirical Bayes estimation of random effects of a mixed-effects proportional odds Markov model for ordinal data, Comput. Methods Programs Biomed. (2011), doi:10.1016/j.cmpb.2011.04.006

optimization method, ARS, which ran approximately 60 times longer.

The mean of the *a posteriori* random effects distribution was a less biased estimate than the mode, which is consistent with estimation theory [19]. However, the mean had the same low precision as the mode.

EBEs obtained with larger amount of individual data had better accuracy and precision. However, the bias of η_1 did not reduce to 0 with 200 observations per subject, but for nonlinear models it may not do so even with infinite number of observations [20]. It is well known that the precision of estimates of parameters depends on the amount of data per individual [21], the same is true for shrinkage [8]. From a practical perspective, the large amount of data required to improve the quality of EBEs could not be gathered in clinical practice. Hence this is not a viable way to overcome the problem.

It was shown that the quality of EBEs depends highly on the population parameter values. EBEs obtained from the same amount of data but generated from a model with altered parameter values (K*10 and ED_{50}/20) had much better accuracy and precision; their shrinkage was also lower. In this model, the effect-driving variable is Q*K, so increased *K* corresponds to faster kinetics (faster increase but also faster elimination, shorter time to reach steady state, larger peak to trough ratios), as well as higher average levels. Reduced ED_{50} means that maximal achieved Q*K is much higher than ED_{50} so that the effect reaches a plateau. In order to estimate the parameters of the E_{max} model well, data covering the whole predictor-effect curve (from the quasi-linear increase to the plateau) must be available. Dutta et al. [22] have reported that "when the highest measured effect intensity was less than 95% of E_{max}, E_{max} and EC_{50} were poorly estimated". Concentrations leading to 95% of E_{max} are equivalent to concentrations 19 times higher than EC_{50}. In the HFS model used in our work, the effect-driving variable is not concentration but Q*K; its value leading to 50% of the maximal effect is the ED_{50}. With the original values of *K* and ED_{50}, the Q*K values ranged approximately from 1000 (8% of ED_{50}) to 12,000 (93% of ED_{50}). With the *K* value multiplied by 10 and the ED_{50} value divided by 20, the median Q*K values ranged approximately from 14 (2% of ED_{50}) to 23,000 (3600% ED_{50}). With the original parameter values, the saturation of the effect is not achieved and this may prevent the parameters from being estimated correctly.

A major determinant for the quality of EBEs is the individual distribution of categories. The problem of estimating the REs for individuals with highly uneven distribution of grades was observed in [4,7] and was also visible in our results. In the context of dose adaptations to avoid severe toxicity (grade $\geq$2), an approximately even distribution of grades might happen only in exceptional cases (where there is a failure to control severe adverse effects of the treatment). So in "clinically normal" cases, the distribution of grades would not be "optimal" for estimating the individual REs.

Throughout all the scenarios studied in the sensitivity analysis, η_2 (the individual effect on the intercept) is better estimated than η_1. It may be due to the specific feature of this model which is that η_2 has a much higher impact on the probabilities of grades than η_1. Moreover, due to the correlation between the two ηs (67%), the estimate of η_2 might "pull" the estimate of η_1 to the values of the same sign. This is particularly "visible" in the case of patients without any toxicity (only grades 0). One would expect ηs to be estimated to the values that correspond to the toxicity risk being the lowest: highly positive η_2 (increased logit of grades $\leq$1) and highly negative η_1 (lowest value of Q*K). But the result is that both ηs are estimated to positive values (Fig. 1). As explained in the first paragraph of results section, these patients without any toxicity have been excluded from the analysis; but as the distribution of grades in most of the patients is dominated by grade 0 (Table 6), the same effect may explain the positive bias of EBEs.

In terms of impact on the clinical decision concerning the next dose to take after some toxicity has been observed, the impact is variable but considerable. In extreme cases, the impact of incorrect EBEs on the predicted risk of grade $\geq$2 can be 4-fold, and in the corresponding doses calculated according to true and estimated risk prediction can differ 10-fold (data not shown).

REFERENCES

[1] G. Graham, S. Gupta, L. Aarons, Determination of an optimal dosage regimen using a Bayesian decision analysis of efficacy and adverse effect data, Journal of Pharmacokinetics and Pharmacodynamics 29 (1) (2002) 67–88.
[2] D.R. Mould, N.H.G. Holford, J.H.M. Schellens, J.H. Beijnen, P.R. Hutson, H. Rosing, W.W. ten Bokkel Huinink, E.K. Rowinsky, J.H. Schiller, M. Russo, G. Ross, Population pharmacokinetic and adverse event analysis of topotecan in patients with solid tumors, Clinical Pharmacology and Therapeutics 71 (5) (2002) 334–348.
[3] K.G. Kowalski, L. McFadyen, M.M. Hutmacher, B. Frame, R. Miller, A two-part mixture model for longitudinal adverse event severity data, Journal of Pharmacokinetics and Pharmacodynamics 30 (5) (2003) 315–336.
[4] S. Jönsson, M.C. Kjellsson, M.O. Karlsson, Estimating bias in population parameters for some models for repeated measures ordinal data using NONMEM and NLMIXED, Journal of Pharmacokinetics and Pharmacodynamics 31 (4) (2004) 299–320.
[5] R.M. Savic, F. Mentré, M. Lavielle, Implementation and evaluation of the SAEM algorithm for longitudinal ordered categorical data with an illustration in pharmacokinetics–pharmacodynamics, The AAPS Journal 13 (1) (2011) 44–53.
[6] S. Beal, L.B. Sheiner, A. Boeckmann, R.J. Bauer, NONMEM User's Guides (1989–2009), Icon Development Solutions, Ellicott City, MD, 2009.
[7] V. Murphy, A. Dunne, Mixed effects versus fixed effects modelling of binary data with inter-subject variability, Journal of Pharmacokinetics and Pharmacodynamics 32 (2) (2005) 245–260.
[8] R.M. Savic, M.O. Karlsson, Importance of shrinkage in empirical Bayes estimates for diagnostics: problems and solutions, The AAPS Journal 11 (3) (2009) 558–569.
[9] E. Hénin, B. You, E. Van Cutsem, P.M. Hoff, J. Cassidy, C. Twelves, K.P. Zuideveld, F. Sirzen, C. Dartois, G. Freyer, M. Tod, P. Girard, A dynamic model of hand-and-foot syndrome in patients receiving capecitabine, Clinical Pharmacology and Therapeutics 85 (4) (2009) 418–425.
[10] J.A. Nelder, R. Mead, A simplex method for function minimization, Computer Journal 7 (1965) 308–313.
[11] T. Ye, S. Kalyanaraman, A recursive random search algorithm for network parameter optimization, ACM

Please cite this article in press as: I. Paule, et al., Empirical Bayes estimation of random effects of a mixed-effects proportional odds Markov model for ordinal data, Comput. Methods Programs Biomed. (2011), doi:10.1016/j.cmpb.2011.04.006

SIGMETRICS Performance Evaluation Review 32 (3) (2004) 44–53.
[12] D. Spiegelhalter, A. Thomas, N. Best, D. Lunn, WinBUGS Version 1.4 User Manual, MRC Biostatistics Unit, Cambridge, UK, 2003.
[13] P. Jacqmin, E. Snoeck, E.A. van Schaick, R. Gieschke, P. Pillai, J.L. Steimer, P. Girard, Modelling response time profiles in the absence of drug concentrations: definition and performance evaluation of the K-PD model, Journal of Pharmacokinetics and Pharmacodynamics 34 (1) (2007) 57–85.
[14] D. Cockcroft, M.D. Gault, Prediction of creatinine clearance from serum creatinine, Nephron 16 (1) (1976) 31–41.
[15] Pharsight Trial Simulator v2.1.2 User's Guide, Pharsight Corporation, Mountain View, CA, 2001.
[16] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Rubin, Bayesian Data analysis, Chapman and Hall, Boca Raton, 1995.
[17] L. Lindbom, P. Pihlgren, E.N. Jonsson, PsN-Toolkit – a collection of computer intensive statistical methods for non-linear mixed effect modeling using NONMEM, Computer Methods and Programs in Biomedicine 79 (3) (2005) 241–257.
[18] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, Numerical Recipes in Fortran 77: the Art of Scientific Computing, Cambridge University Press, 1992.
[19] P. Congdon, Bayesian Statistical Modelling, 2nd edition, Wiley, 2006.
[20] D.M. Bates, D.G. Watts, Nonlinear Regression Analysis and its Applications, Wiley-Interscience, 2007.
[21] S. Retout, S. Dufful, F. Mentré, Development and implementation of the population Fisher information matrix for the evaluation of population pharmacokinetic designs, Computer Methods and Programs in Biomedicine 65 (2001) 141–151.
[22] S. Dutta, Y. Matsumoto, W.F. Ebling, Is it possible to estimate the parameters of the sigmoid E_{max} model with truncated data typical of clinical studies? Journal of Pharmaceutical Sciences 85 (2) (1996) 232–239.

Please cite this article in press as: I. Paule, et al., Empirical Bayes estimation of random effects of a mixed-effects proportional odds Markov model for ordinal data, Comput. Methods Programs Biomed. (2011), doi:10.1016/j.cmpb.2011.04.006

Paule et al. Orphanet Journal of Rare Diseases 2011, 6:30
http://www.ojrd.com/content/6/1/30

ORPHANET JOURNAL OF RARE DISEASES

RESEARCH Open Access

Population pharmacokinetics and pharmacodynamics of hydroxyurea in sickle cell anemia patients, a basis for optimizing the dosing regimen

Ines Paule[1,2], Hind Sassi[3], Anoosha Habibi[4], Kim PD Pham[3], Dora Bachir[4], Frédéric Galactéros[4], Pascal Girard[1,2], Anne Hulin[3] and Michel Tod[1,2,5*]

Abstract

Background: Hydroxyurea (HU) is the first approved pharmacological treatment of sickle cell anemia (SCA). The objectives of this study were to develop population pharmacokinetic(PK)-pharmacodynamic(PD) models for HU in order to characterize the exposure-efficacy relationships and their variability, compare two dosing regimens by simulations and develop some recommendations for monitoring the treatment.

Methods: The models were built using population modelling software NONMEM VII based on data from two clinical studies of SCA adult patients receiving 500-2000 mg of HU once daily. Fetal hemoglobin percentage (HbF %) and mean corpuscular volume (MCV) were used as biomarkers for response. A sequential modelling approach was applied. Models were evaluated using simulation-based techniques. Comparisons of two dosing regimens were performed by simulating 10000 patients in each arm during 12 months.

Results: The PK profiles were described by a bicompartmental model. The median (and interindividual coefficient of variation (CV)) of clearance was 11.6 L/h (30%), the central volume was 45.3 L (35%). PK steady-state was reached in about 35 days. For a given dosing regimen, HU exposure varied approximately fivefold among patients. The dynamics of HbF% and MCV were described by turnover models with inhibition of elimination of response. In the studied range of drug exposures, the effect of HU on HbF% was at its maximum (median I_{max} was 0.57, CV was 27%); the effect on MCV was close to its maximum, with median value of 0.14 and CV of 49%. Simulations showed that 95% of the steady-state levels of HbF% and MCV need 26 months and 3 months to be reached, respectively. The CV of the steady-state value of HbF% was about 7 times larger than that of MCV. Simulations with two different dosing regimens showed that continuous dosing led to a stronger HbF% increase in some patients.

Conclusions: The high variability of response to HU was related in part to pharmacokinetics and to pharmacodynamics. The steady-state value of MCV at month 3 is not predictive of the HbF% value at month 26. Hence, HbF% level may be a better biomarker for monitoring HU treatment. Continuous dosing might be more advantageous in terms of HbF% for patients who have a strong response to HU.

Trial Registration: The clinical studies whose data are analysed and reported in this work were not required to be registered in France at their time. Both studies were approved by local ethics committees (of Mondor Hospital and of Kremlin-Bicetre Hospital) and written informed consent was obtained from each patient.

* Correspondence: michel.tod@chu-lyon.fr
[1]Université de Lyon, Lyon, France
Full list of author information is available at the end of the article

Disease name

Sickle cell anemia.

Definition

Sickle cell anemia is an autosomal recessive genetic blood disorder, caused by a mutation in the hemoglobin gene and characterized by rigid sickle-shaped red blood cells. Sickling decreases the cells' elasticity and leads to vaso-occlusion which may result in various complications, such as acute painful crises, ischemia and damage of various organs, acute chest syndrome or stroke.

Background

The antineoplastic agent hydroxyurea (hydroxycarbamide) (HU) is the first approved pharmacological treatment of sickle cell anemia (SCA). It inhibits the production of the hemoglobin S that causes SCA and favors the reactivation of fetal hemoglobin (HbF) expression [1]. In fact, a variety of mechanisms are believed to be involved in HU beneficial effects in SCA, including increased HbF synthesis by erythroid regeneration, NO-related increases in soluble guanylate cyclase activity and cyclic guanidine monophosphate (cGMP) that stimulate *HBG* expression [2]. Other mechanisms may be myelosuppression with a reduction of circulating neutrophils, increased erythrocyte water content, modified erythrocyte endothelial cell interactions and altered vascular tone by increasing NO bioavailability [3]. Recently, Bartolucci *et al.* reported that HU could reduce abnormal sickle cell adhesion to the vascular wall by regulating the activation state of adhesion molecules [4]. HU treatment reduces the rate and severity of painful attacks [5] and was shown to possibly increase survival time [6].

The usual dosing of this oral treatment is daily doses of 15-35 mg/kg (or less if there is renal insufficiency) [7]. The most adequate individual dose is determined by starting with 15 mg/kg and monitoring full blood cell counts every two weeks. If after twelve weeks no cytopenia has developed, the dose is increased by 5 mg/kg. Once their maximal tolerable dose is determined, the patient can continue the treatment life-long, if no serious toxicities manifest or other issues arise.

Despite the widespread use of HU, only a few studies have been reported in the literature, especially concerning its use in the indication of SCA. Little is known about the relationship between drug exposure and efficacy, evaluated by fetal hemoglobin (HbF) and mean corpuscular volume (MCV) measurements. Although a number of genetic polymorphisms have been found to be associated with response to HU [8-10], the variability of this response remains poorly characterized. The optimal dosing schedule, the best strategy for monitoring and adjusting the treatment, and the impact that may have the prior determination of candidate genotypes on HU dosing remain open to discussion. Part of these questions may be addressed through simulations from pharmacokinetic (PK) and pharmacodynamic (PD) models of HU. Therefore, this study aimed to develop population PK-PD models for HU in order to characterize the exposure-efficacy relationships and their variability. These models were then used (1) to compare two dosing regimens (one continuous daily and the other with interruptions of 2 days after every 5 days) by simulation, and (2) to develop some recommendations for monitoring the treatment.

Methods

Two datasets from two studies with different designs were used in this analysis: one from a PK-PD study with up to 9 samples taken per patient over a period of up to 30 months, and the other from a PK study with 10 samples per patient taken over 24 hours after drug administration.

PKPD study design (sparse sampling)

Study design and patient population

This 30-month, open-label, noncomparative, prospective, observational study was conducted in 81 adult patients with sickle cell anemia at the *Centre de référence pour les syndromes drépanocytaires majeurs*, AP-HP, GH H. Mondor, *Université Paris Est-Créteil*, France from 2007 to 2010. It focused on the benefits and risks of HU, in particular the side effects in the short and medium term, as well as the need for regular hematological monitoring. The protocol was approved by the local ethics committee of Mondor hospital, and written informed consent was obtained from each patient. The recruited SCA patients were of 18 years or older and with Hb genotype *HbSS* (homozygous sickle Hb). In all patients, SCA diagnosis was documented by standard methods [7]. Patients who received erythropoietin or a blood transfusion at a time that can interfere with the results were excluded.

Treatment

HU treatment was started at a dose of 20 mg/kg orally unless there was renal insufficiency (in this case it started at 10 and 15 mg/kg). Since hematotoxicity is the dose-limiting toxicity, the hematologic control was performed every 15 days during the first month and then every month; if needed, doses were adjusted to maintain neutrophil counts higher than 3×10^9/L. In most cases, the final dose of HU did not exceed 30 mg/kg.

Pharmacokinetics protocol

Blood was drawn from most patients on day 0 (D0), D15, after 1 month (M1), M2, M4 and M6, as well as at various later timepoints more sparsely (up to M30). It was collected in heparinized tubes and centrifuged at

2000 g for 10 minutes at room temperature to obtain plasma. The plasma was then stored at -20°C until the samples were assayed. Plasma samples were assayed using high performance liquid chromatography (HPLC) coupled with UV-detection at 449 nm [11]. The analytical method was linear between 5 to 1000 μM, precise (coefficients of variation ranging from 1.7 to 9.9%), accurate (97.7 to 103.9%). The limit of quantification (LOQ) was 7 μM (0.532 mg/L).

The following biologic variables were measured during this study: creatinine, urea, lactate dehydrogenase (LDH), hemoglobin (Hb), fetal hemoglobin (HbF), mean corpuscular volume (MCV), mean corpuscular hemoglobin (MCH), ferritin, bilirubin, aspartate transaminase (AST), alanine transaminase (ALT), neutrophils (PMN) and platelets. Body weight, age and sex data were also documented.

PK study design (rich sampling)

These rich PK data come from a bioequivalence study of standard hydroxyurea capsules and a new formulation of 1000 mg coated breakable tablets in adults with homozygous SCA or S/β-thalassemia. The complete study protocol is given in a publication of the noncompartmental analysis of the PK data [12]. The assay validation parameters were the same as for the sparse PK data study [11].

These data contain 10 blood samples per patient, from 16 patients who took hydroxyurea doses ranging from 1000 mg to 2000 mg. The samples were taken at baseline and 45, 90, 120, 150, 180, 240, 360, 480 minutes after HU administration at the study center, as well as trough samples after 24 hours.

Population PK-PD models

Population analyses were performed using NONMEM software (version VII) [13]. PK model parameters were estimated using the second order conditional estimation (Laplacian) method with interaction between interindividual and residual variabilities. This estimation method is more accurate than the standard first order conditional estimation (FOCE) method. In this PK analysis it was needed to correctly handle the concentrations below the limit of quantification (BLQ) (see below). PD models parameters were estimated using the FOCE method with interaction between interindividual and residual variabilities. Confidence intervals (CI) of parameter estimates were obtained by nonparametric bootstrap (n = 1000) with stratification by study and dose.

Population pharmacokinetic model

First, the structural PK model was built using both sparse and rich datasets. One, two and three-compartment models with first-order absorption and elimination were tested, as well as with saturable (Michaelis-Menten) elimination. The most appropriate model was chosen on the basis of the objective function value (OFV) and simulation-based diagnostics such as normalised predictive discrepancy errors (NPDE) [14]. The NPDE analysis was performed with BLQ points excluded.

In the two-compartment model, the parameters were the elimination clearance (CL/F), the volume of the central compartment (V_c/F), the rate constants for transfer from central to peripheral (k_{cp}) or peripheral to central compartment (k_{pc}), the rate constant for absorption (k_a).

The BLQ measurements of the sparse data were included in the data and modelled as censored observations using "method 3" described in [15].

The interindividual variability (IIV) for the PK parameters was described using exponential models: $P_i = \theta \cdot e^{\eta_i}$, where η is a random variable with normal distribution, zero mean and variance to be estimated, θ is a typical value. The IIV was added in a stepwise manner, firstly to clearance and central volume of distribution. The interindividual random effects were kept in the model if their inclusion significantly reduced the OFV and if their relative SE was <50%. A full interindividual variance-covariance matrix was estimated to assess if there was any significant covariance in the IIV structure.

To find the most appropriate residual error model, additive, proportional and mixed error models were tested. As data were collected from two quite differently designed studies, separate residual error models for each study were tested.

The variables that were investigated for their ability to explain IIV in the PK of HU were body weight, creatinine, age and sex. Scaling by body weight (for clearance and central volume of distribution) as $V_c = \theta_V \cdot \frac{WGT}{70}$ and $CL = \theta_{CL} \cdot \left(\frac{WGT}{70}\right)^{0.75}$, as well as $CL = \theta_{CL} + f(covar),\ f(covar) = \frac{(140 - age) \cdot bodyweight \cdot (1 - coeff_{sex} \cdot sex)}{creatinine}$ (with sex = 0 for men and 1 for women, $coeff_{sex}$ estimated), similar to the Cockroft-Gault formula for creatinine clearance [16], were the tested forms of relationships. The conditions for the variable to be retained as a covariate were to be biologically plausible, and to decrease the OFV by at least 5 units, corresponding to a p-value less than 2.5%.

Missing values of body weight were predicted by the model: if the patient's weight was known at other time-point(s), it was predicted by using an interpolated or adjacent value and an additive individual random effect whose variance was fixed to the observed intraindividual variance of the body weight of the dataset (1.5); if no data of patient's weight were available, it was predicted by using the average weight of the dataset (62 kg for

women, 65 kg for men) and an additive individual random effect whose variance was fixed to the observed interindividual variances of the dataset (92 for women, 40 for men).

The time to reach 95% of HU concentration steady-state was determined by simulation of a typical patient with 1000 mg dose every day for 6 months.

Population pharmacodynamic models

The PD responses to be described by the model were HbF percentage and MCV. The observed level of both these responses depends on the ratio of a production rate to an elimination rate. Therefore, turnover models [17] were chosen to fit the data. Models assuming either stimulation of the production or inhibition of the elimination of response by HU were tested. The metrics of HU exposure used as the input into the PD model was mean drug concentration at steady state, calculated using the individual posthoc estimates of HU clearance. Therefore, a sequential modelling approach was used. The uncertainty in the estimate of clearance was taken into account by using the IPPSE method [18]. Linear, Emax, sigmoid and power functions were tested for the effect of the drug. The final model was chosen on the basis of the OFV and diagnostic plots, stability and precision of parameter estimates. IIV and residual error models were constructed in the same way as for the PK model.

In addition to demographic variables, baseline measurements of pharmacodynamic variables, as well as their previous values, rate of change per day from baseline to the previous value ($\Delta V = \frac{V_{t-1} - V_0}{days}$), rate of change per day between the two previous values ($\Delta V = \frac{V_{t-1} - V_{t-2}}{days}$) were investigated as covariates. They were included as additive or proportional to the drug effect, production rate or elimination rate. Missing values were predicted by the model by using the data average and an individual random effect whose variance was fixed to the value calculated from the data. The conditions for covariate inclusion were the same as for the PK model.

Model evaluation

The quality of models was assessed by goodness-of-fit plots and simulation-based methods (using 1000 simulations): visual predictive check (VPC) and NPDE. The mean of prediction error distribution was compared to zero by a Wilcoxon signed rank test, while its variance was compared to unity by a Fisher test. For the VPC, prediction corrections were used so that data of all dose levels could be used in one plot [19]. BLQ points in the observed and simulated PK data sets were assigned values equal to LOQ/2. The VPC plots showed 80% prediction intervals (PI) and medians of the observed and of the predicted data, as well as 95% confidence areas around the percentiles. For the PK, a VPC plot in log scale was also given.

Simulation of alternative dosing regimens

The two simulated dosing regimens were: 1000 mg daily doses 7 days a week (7/7) and 1000 mg daily doses 5 consecutive days a week (5/7) for the duration of 12 months; 10000 patients were simulated in each arm. For each covariate, the average observed value was used. The results were compared graphically by representing the median profile and the 90% prediction interval of the HbF% and MCV. In order to determine the steady-state values of the HbF% and of the MCV and time to reach 95% of them, such simulations with 7/7 dosing were extended for 48 months.

Results

Pharmacokinetic data analysis

A summary of PK related patient characteristics is given in Table 1. In the sparse data, 78% of last doses before the concentration measurements were 1000 mg, 15% were 1500 mg, and the rest were 500 mg, 1250 mg or 2000 mg. In the rich dataset, 44% were 1000 mg, 31% were 1500 mg, and the rest were either 1250 mg or 2000 mg. The median number of measurements per patient in the sparse dataset was 4 (range: 1 - 9); in the rich dataset, 10 measurements were available for each patient.

The PK profiles were best described by a two-compartment model (with first-order absorption and elimination). The NPDE diagnostics indicated that one-compartment model could not adequately describe these data. The OFV of the three-compartment model was not lower than that of the two-compartment model. No significant nonlinearity in absorption or elimination could be detected in these data.

Significant interindividual variability was found for V_c, CL, k_a and k_{cp}. Correlations were significant among the individual values of V_c, CL and k_{cp}. The mixed residual error model was best for both datasets. The OFVs of PK models with and without weight-scaling of V_c and CL (as shown in methods section) were nearly significant, but this covariate was kept in the model for the sake of

Table 1 Summary of patient characteristics

	PKPD dataset	Rich PK dataset
Characteristics	**Median (range)**	**Median (range)**
Age	30 (18 - 54)	32 (24 - 52)
Men/women	24/57	5/11
Weight (kg)	60 (45 - 163)	63 (42 - 71)
Creatinine (μM)	65 (27 - 558)	72 (47 - 129)
Urea (μM)	3.1 (0.8 - 10.5)	3.4 (2.4 - 7.7)

coherence with clinical practice and possible application of the model to children (where the effect of weight would be much more perceptible). The allometric model for clearance with power 0.75 was better than the model with power 1 (p = 0.0016).

The final model estimates are given in Table 2. As the bioavailability was not estimated here, the reported estimates for CL and V_c represent apparent pharmacokinetic parameters CL/F and V_c/F. Their values in Table 2 are given for a patient of 70 kg weight, which is the scaling base. To obtain values for patients of different weight, the population value should be multiplied by weight/70 for V_c and by $(weight/70)^{0.75}$ for CL. The estimate of the rate constant of transfer from the peripheral to the central compartment (k_{pc}) was very small and unstable, so its value was fixed to 0.004 (h^{-1}) (its best estimate) and this resulted in lower estimate of IIV of k_{cp} and better stability of the model. With this model, the time to reach 95% of the steady-state was typically about 35 days.

Simulation-based diagnostics VPC (Figure 1) and NPDE indicated that the model adequately described the measured data. The mean of normalized prediction errors was significantly different from zero (mean = 0.12, p = 0.016), but translated into HU concentrations, the mean difference between observed and predicted concentrations was only 0.08 mg/L. This difference was not considered as relevant from a clinical point of view. The variance of normalized prediction errors was not significantly different from unity (0.893, p = 0.12).

Pharmacodynamic data analysis

The median number of HbF% measurements per patient was 5 (range: 1 - 10); the total number of patients and measurements was 77 and 391 respectively. The median number of MCV measurements per patient was 6 (range: 3 - 10); the total number of patients and measurements was 80 and 439 respectively. 43% of patients with HbF% measurements and 49% of patients with MCV measurements were followed for 6 months (median (range) for HbF%: 6 months (9 days - 30 months); for MCV: 6 months (1 - 30 months)). Summary statistics of the pharmacodynamic variables at the beginning and after 6 months in the study are given in Table 3. The treatment with HU induced an increase in HbF, its percentage (Figure 2 shows its maximums) and in MCV, as well as decreases in bilirubin and LDH, which were indicative of decrease in the rate of hemolysis. Decreases in neutrophils (PMN) and platelets were mild, not reaching below normal levels.

Population PD model of the percentage of fetal hemoglobin

The turnover model with inhibition of the elimination rate was found to describe best the HbF% data. However, in this dataset, no relationship between HU concentration and HbF% could be identified, because all patients were estimated to have the maximum drug effect. The rate of change per day between the two last MCV observations (ΔMCV) was found to be a significant covariate for the production rate, K_{in} (p < 0.00001). The median (range) ΔMCV was 0.16 (-0.27 to 0.87). The final model was (cf. parameter estimates in Table 4):

$$\frac{dHbF\%}{dt} = K_{in,i} - K_{out,i} \cdot \left(1 - I_{\max,i}\right) \cdot HbF\%$$

where $K_{in,i} = K_{in,TV} \cdot e^{\eta_{K_{in},i}} \cdot e^{\theta_{\Delta MCV} \cdot \Delta MCV}$, $K_{out,i} = K_{out,TV} \cdot e^{\eta_{Kout,i}}$, $I_{\max,i} = \frac{e^{L_{I\max,TV} + \eta_{I\max,i}}}{1 + e^{L_{I\max,TV} + \eta_{I\max,i}}}$ and L_{Imax} is the logit-transformed Imax.

Table 2 Parameter estimates of the population PK model

Parameters	Typical values (95% CI)	Standard deviations of IIV (95% CI)	Interindividual CV
V_c/F (L) (for a patient of 70 kg)	45.3 (38.9 - 50.5)	0.34 (0.23 - 0.46)	35%
Cl/F (L/h) (for a patient of 70 kg)	11.6 (10.4 - 12.9)	0.29 (0.22 - 0.40)	30%
k_a (h^{-1})	3.29		
θ_{ka} (h^{-1}) ($k_a = \theta_{ka} \cdot e^{\eta_{ka}} + \alpha$)	3.02 (2.25 - 4.19)	1.34 (1.16 - 1.65)	224%
k_{cp} (h^{-1})	0.027 (0.021 - 0.037)	0.57 (0.43 - 0.95)	62%
k_{pc} (h^{-1}) (fixed)	0.004	-	-
SD of the additive component of the residual error (mg/L)			
- for densely sampled data	0.319 (0.197 - 0.492)		
- for sparsely sampled data	0.353 (0.257 - 0.522)		
SD of the proportional component of the residual error			
- for densely sampled data	0.12 (0.083 - 0.154)		
- for sparsely sampled data	0.435 (0.349 - 0.506)		
Correlation (η_{Vc}, η_{Cl})	0.71		
Correlation (η_{Vc}, η_{kcp})	-0.26		
Correlation (η_{CL}, η_{kcp})	0.37		

CI: confidence intervals, obtained by bootstrap (n = 1000), IIV: interindividual variability, CV: the apparent coefficient of variation of interindividual variability.

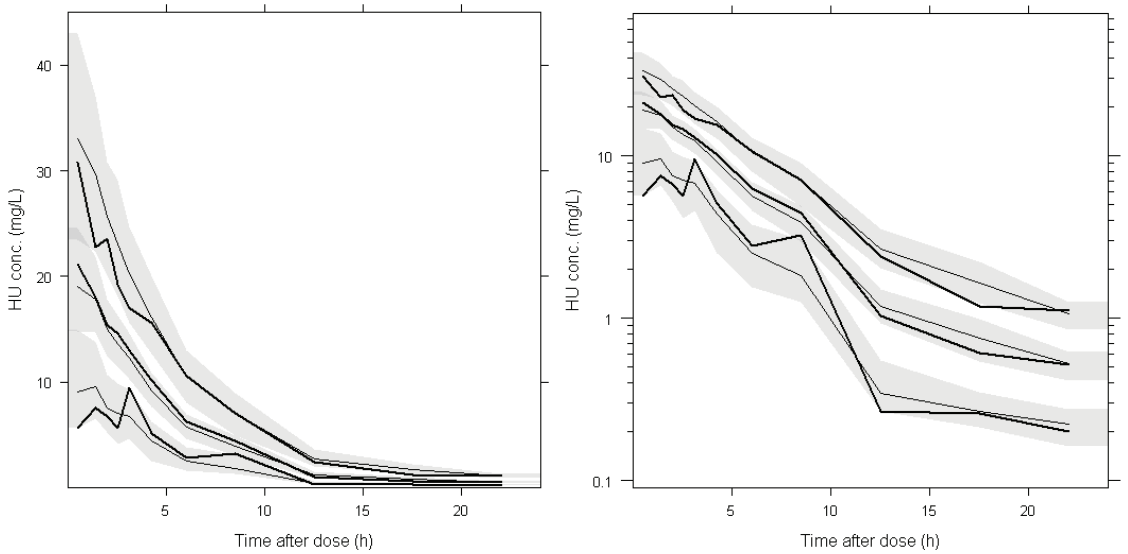

Figure 1 PC-VPC for HU concentrations. The lines show the 10th, 50th and 90th percentiles of observed data (thick lines) and of predictions (thin lines); the grey areas represent the 95% confidence areas around the percentiles. PC-VPC: prediction-corrected visual predictive check. Left panel: linear scale, right panel: log scale.

Significant interindividual variability was found for K_{in}, K_{out} and L_{Imax}, with correlation between K_{out} and L_{Imax}. A proportional residual error model was selected.

Simulation-based diagnostics VPC (Figure 3) and NPDE indicated that the model adequately described the observed HbF% data. The mean of NPDE was 0.03 (p = 0.3), the variance was 0.937 (p = 0.38). Concerning the VPC, for the first 300 days, the difference between the medians of HbF% observations and of its predictions was approximately 2%, which was not clinically significant. Only 14 patients out of 81 continued the treatment longer than 300 days, therefore the percentiles at later times may be imprecise. The NPDE versus time plots did not indicate any prediction deficiencies (data not shown).

In the simulation, median HbF% at steady-state was about 18.6%, 95% of it was reached in about 26 months.

Population PD model of the mean corpuscular volume

The turnover model with inhibition of the elimination rate was found to describe best the MCV data. The inhibition was best described by a power function of average concentration. The rate of change per day between the two last HbF% observations (ΔHbF%) was found to be a significant covariate on the parameter β ($p < 0.00001$).

Table 3 Summary statistics of PD variables

PD variables	At baseline Median (range)	Number of patients	After 6 months of treatment Median (range)	Number of patients
HbF%	6.3 (0.6 - 30.7)	65	15.7 (3.9 - 41.6)	46
HbF (g/dL)	0.48 (0.04 - 2.7)	63	1.59 (0.34 - 4.04)	46
Hemoglobin (g/dL)	8.8 (6.3 - 11.9)	73	9.6 (6.9 - 14.4)	55
MCV (fL)	90 (68 - 113)	74	111 (81 - 131)	55
MCH (pg)	30 (21 - 36)	73	37 (25 - 44)	55
PMN (10^9/L)	56 (25 - 80)	71	49 (26 - 86)	53
Platelets (10^9/L)	428 (122 - 995)	74	316 (109 - 528)	55
Bilirubin (μM)	43 (9 - 96)	75	30 (6 - 113)	53
LDH (UI/L)	355 (155 - 800)	73	317 (168 - 766)	52
Ferritin (μg/L)	346 (16 - 4500)	72	275 (14 - 2940)	52
AST (UI/L)	32 (17 - 79)	75	31 (12 - 81)	53
ALT (UI/L)	22 (7 - 84)	75	21 (7 - 83)	53
Creatinine (μM)	65 (38 - 142)	75	64 (35 - 137)	52
Urea (μM)	2.9 (1.2 - 13.3)	75	3.3 (1.5 - 9)	52

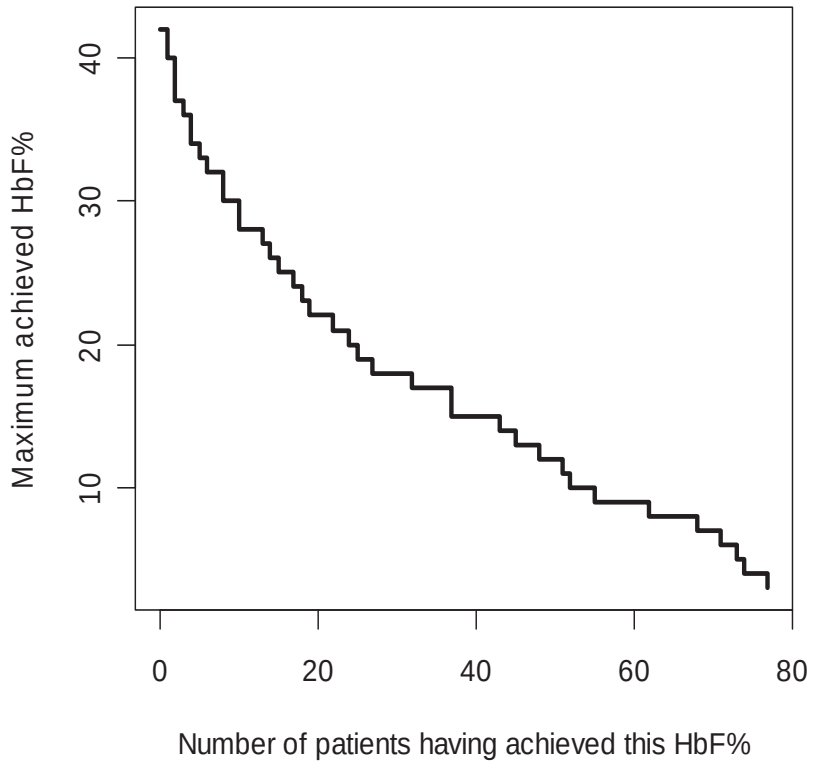

Figure 2 Maximum achieved HbF% and corresponding cumulative number of patients.

The median (range) ΔHbF% was 0.047 (-0.278 to 0.653). The final model was (cf. parameter estimates in Table 5):

$$\frac{dMCV}{dt} = K_{in,i} - K_{out,i} \cdot \left(1 - \beta_i \cdot Conc^{\gamma}\right) \cdot MCV$$

where $K_{in,i} = K_{in,TV} \cdot e^{\eta_{K_{in},i}}$, $K_{out,i} = K_{out,TV} \cdot e^{\eta_{Kout,i}}$, $\beta_i = \beta_{TV} \cdot e^{\eta_{\beta,i}+\theta \cdot \Delta HbF\%}$

Significant interindividual variability was found for K_{in}, K_{out} and β, with correlations between all three parameters. A proportional residual error model was selected.

Simulation-based diagnostics VPC (Figure 4) and NPDE indicated that the model adequately described the observed MCV data. The mean of NPDE was -0.02 (p = 0.87), the variance was 0.9 (p = 0.15). In the VPC, the difference between the medians of MCV observations and of its predictions was approximately 5 fL, which was not clinically significant. In the simulation, median MCV level at steady-state was about 104 pL, 95% of it was reached in about 90 days.

Simulation of alternative dosing regimens

The simulated HbF% and MCV profiles with the two dosing regimens are shown in Figures 5 and 6 respectively. For MCV, the difference was very small. For HbF %, continuous dosing led to more significantly stronger response, especially for patients reaching the highest levels of HbF%. It can be observed that HbF% required

Table 4 Parameter estimates of the population HbF% model

Parameters	Typical values (95% CI)	SD of IIV (95% CI)	Interindividual CV
K_{in} (%/day)	0.071 (0.055 - 0.094)	0.585 (0.472 - 0.681)	63%
K_{out} (day^{-1})	0.013 (0.010 - 0.019)	0.486 (0.334 - 0.602)	52%
L_{Imax} (unitless)	0.276 (-0.081 - 0.644) (I_{max} = 0.569 (0.48 - 0.656))	1.44 (1.07 - 1.97)	27%
$\theta_{\Delta MCV}$ (day^{-1})	1.37 (0.95 - 1.76)		
SD of proportional residual error	0.142 (0.119 - 0.162)		
Correlation (η_{Kout}, η_{Imax})	0.892		

CI: confidence intervals, obtained by bootstrap (n = 1000), SD: standard deviation, IIV: interindividual variability, CV: the apparent coefficient of variation of interindividual variability.

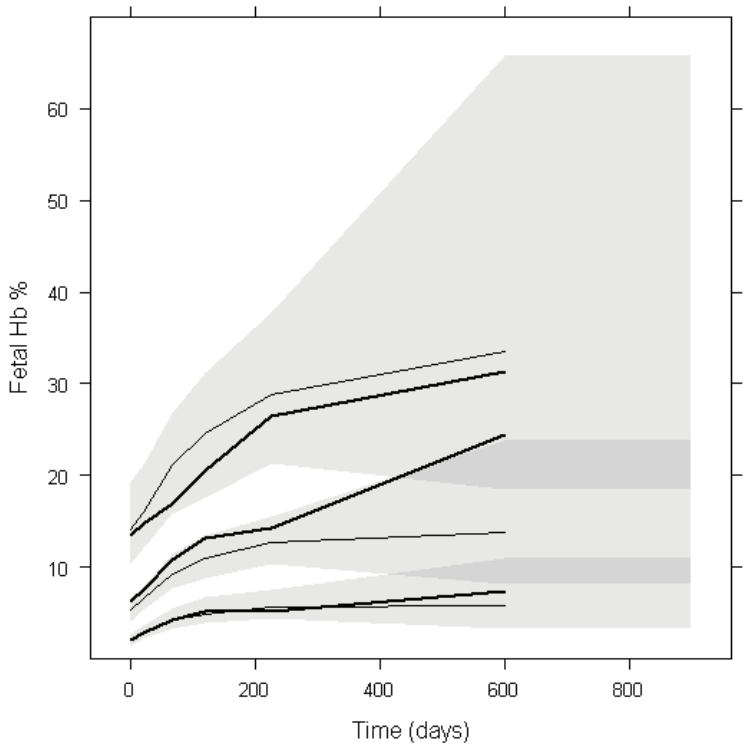

Figure 3 PC-VPC for HbF%. The lines show the10th, 50th and 90th percentiles of observed data (thick lines) and of predictions (thin lines); the grey areas represent the 95% confidence areas around the percentiles. PC-VPC: prediction-corrected visual predictive check.

a much longer time than MCV to reach the steady-state (approximately 26 and 3 months for 95% of steady-state levels respectively). The inter-individual variability of steady-state of HbF% was higher than that of MCV: the ratios of the 95th to the 5th percentile were approximately 10 and 1.5 respectively.

Discussion

In this study, population PK-PD models were developed for the first time, in order to characterize the exposure-efficacy relationships of HU and its variability.

The pharmacokinetics of HU was found to be linear. Other studies in rats and in humans using doses ranging between 10 to 800 mg/kg in patients with malignancies identified parallel linear renal and saturable non-renal elimination [20]. The latter could not be detected in the studies reported here, probably because the doses administered were not high enough to reach saturation of non-renal elimination pathways (10 to 35 mg/kg *per os* in SCA). Otherwise, the presented model was consistent with the results of previously reported studies.

Table 5 Parameter estimates of the population MCV model

Parameters	Typical values (95% CI)	SD of IIV (95% CI)	Interindividual CV
K_{in} (%/day)	3.71 (3.13 - 4.30)	0.191 (0.083 - 0.401)	19%
K_{out} (day^{-1})	0.042 (0.035 - 0.048)	0.186 (0.044 - 0.415)	19%
β (L. mg^{-1})$^{1/\gamma}$	0.099 (0.064 - 0.135)	0.457 (0.336 - 0.599)	48%
γ (unitless)	0.19 (0.02 - 0.46)		
$\theta_{\Delta HbF\%}$ (day^{-1})	1.22 (-0.07 - 2.21)		
SD of proportional residual error	0.036 (0.030 - 0.040)		
Correlation (η_{Kin}, η_{Kout})	0.87		
Correlation (η_{Kin}, η_{β})	-0.98		
Correlation (η_{Kout}, η_{β})	-0.95		

CI: confidence intervals, obtained by bootstrap (n = 1000), SD: standard deviation, IIV: interindividual variability, CV: the apparent coefficient of variation of interindividual variability.

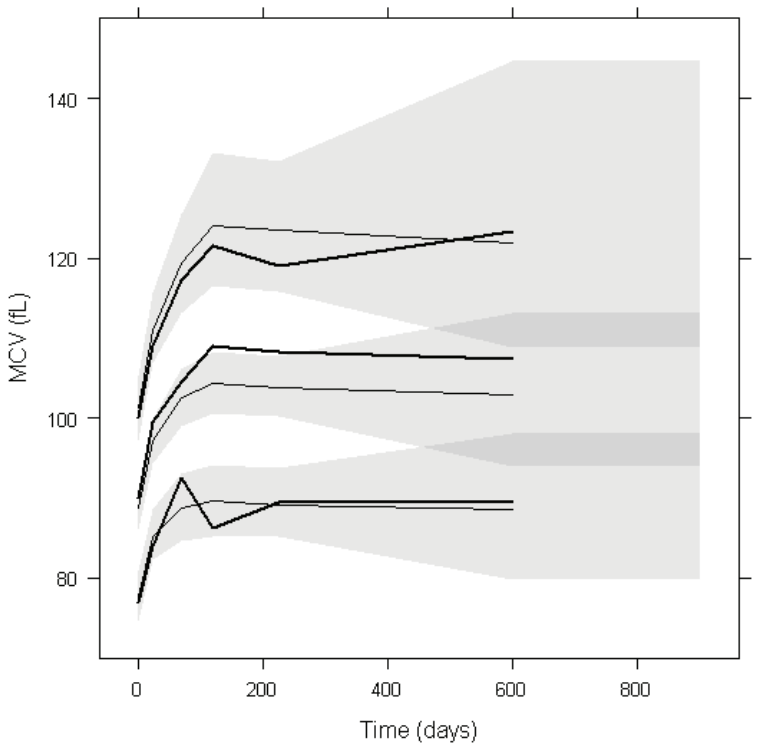

Figure 4 PC-VPC for MCV. The lines show the 10th, 50th and 90th percentiles of observed data (thick lines) and of predictions (thin lines); the grey areas represent the 95% confidence areas around the percentiles. PC-VPC: prediction-corrected visual predictive check.

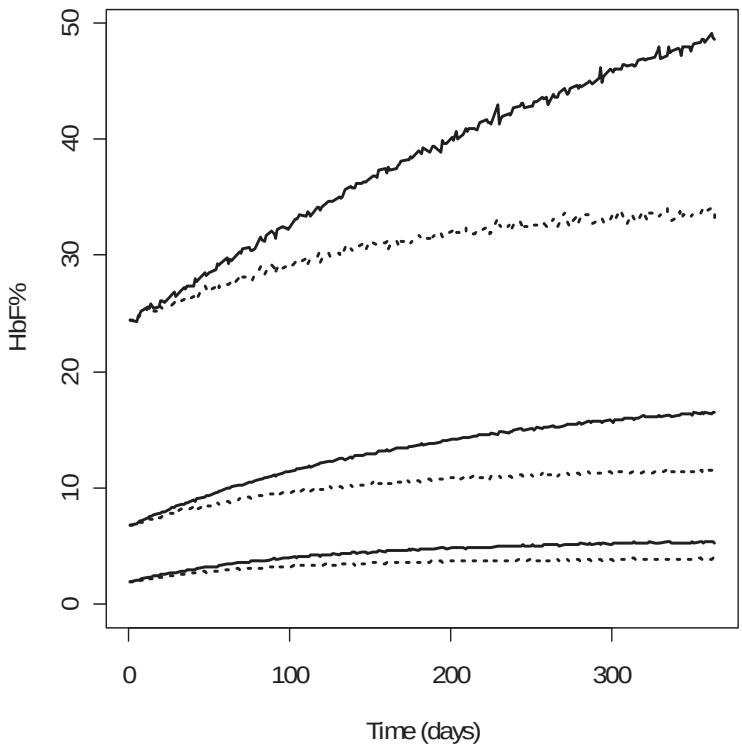

Figure 5 Profiles of simulated HbF% with two dosing regimens. 90% prediction intervals and medians of simulated HbF% with HU 7/7 (solid lines) and 5/7 (dotted lines) (n = 10000).

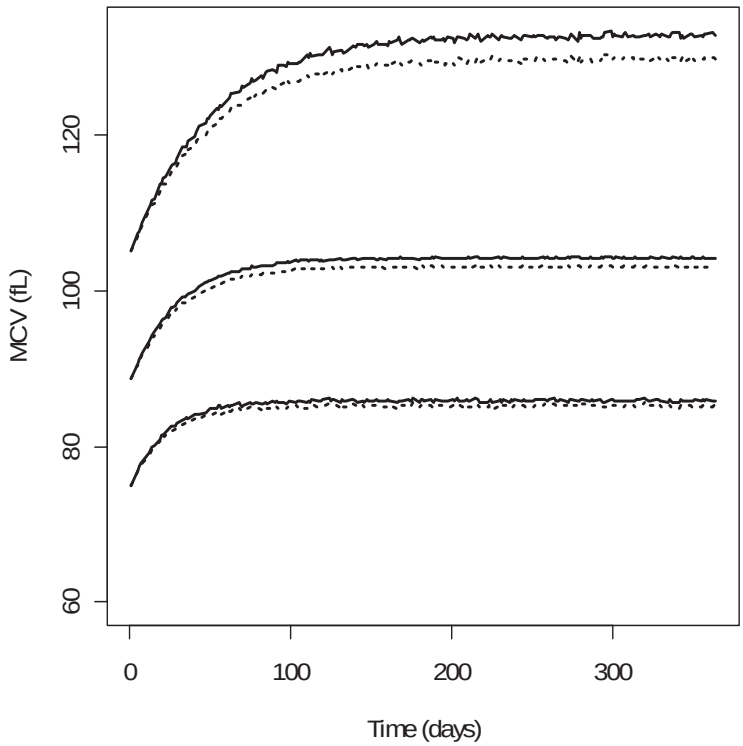

Figure 6 Profiles of simulated MCV with two dosing regimens. 90% prediction intervals and medians of simulated MCV with HU 7/7 (solid lines) and 5/7 (dotted lines) (n = 10000).

For a given dosing regimen, HU exposure varied approximately fivefold among patients. Part of the variability of apparent clearance and apparent volume of the central compartment was related to body weight. Clearance was better correlated with body weight at a 0.75 power, according to allometric scaling laws [21]. Because the maintenance dose of a drug to reach a desired average concentration is determined solely by its clearance, this allometric relationship implies that the HU dosing rate should be calculated with respect to body weight to the power of 0.75, or equivalently, to free fatty mass [22], in order to decrease the interindividual variability in HU exposure. Finally, the time to reach 95% of the pharmacokinetic steady-state was typically 35 days, in contrast with the delay to reach the maximal effect of HU, as discussed below.

The haematological results obtained in this study are compatible with those previously reported [23]. Our study brought further insight on the relationship between exposure and efficacy.

First, from a kinetic point of view, if we expressed the estimated K_{out} parameters as half-lives and then multiplied them by 5 to obtain approximate times to reach steady-state before the drug is taken, we could see that they are around 265 and 83 days for HbF% and MCV respectively. HU is assumed to reduce K_{out} and therefore extend this time to steady-state. The simulations under a constant dosing rate at 1000 mg per day show that 95% of the steady-state levels of HbF% and MCV need 26 months and 3 months to be reached, respectively. If the dosing regimen is modified, the same delay is required to reach a new steady-state. Hence, the variation of MCV is more rapid than that of HbF%. The 3 month delay for MCV is certainly related to the life span of RBC of 120 days and corresponds to the time needed to renew three quarters of RBCs.

Second, the effect of HU on HbF% was estimated to be at its maximum independently of the exposure, in the dose range of our study (500 to 2000 mg/day). However, the intensity of the effect (I_{max}) varied among patients, with a typical value of 0.57 and a coefficient of variation of 27%. None of the demographic and biological indices was correlated with these variations. Part of this variability might be explained by genetic polymorphisms in genes regulating HU metabolism or transporters, HbF expression and erythroid progenitor proliferation [8-10]. These polymorphisms might modulate the patient response to HU. In addition, the HU-inducible small guanosine triphosphate-binding protein, secretion-associated and RAS-related (SAR) protein has been demonstrated to play a key role in *HBG* induction and erythroid maturation by causing cell apoptosis and

G1/S-phase arrest [24]. Some genetic polymorphisms related to this pathway have been described such as *sar1a* promoter polymorphisms [10] and may also contribute to variability. Finally, patient compliance to treatment might also be a source of variability in response, but no information on compliance was available in this study.

Third, HU increased HbF% by reducing HbF elimination rate constant by 57% (for a typical patient). Absolute values of HbF per RBC (medians) at baseline and after 6 months were 1.9 pg and 5.6 pg respectively, which confirms that HU leads to a real increase in HbF per cell. Theoretically, a full inhibitor could reduce the elimination rate further, leading to a higher increase of HbF%. Hence, there is room for improvement, e.g. by looking for stronger inhibitors, or combining HU with other drugs to be discovered.

Fourth, a relation between HU exposure and effect on MCV could be identified, but this relation was flat as in the studied range of drug exposure the effect was close to its maximum. When the average HU concentration was 2 or 9 mg/L (the extremes of this study), the MCV decay rate constant (K_{out}) was multiplied by 0.88 or 0.84 respectively, with an interindividual coefficient of variation of 49%. Hence the inhibition of MCV "elimination" by HU was less potent than that of HbF, and the interindividual variability was greater.

Regarding simulations, a close inspection of Figures 5 and 6 reveals that the interindividual variability of the steady-state values of HbF% and MCV are different, the ratio of the 95^{th} to 5^{th} percentile being approximately 10 and 1.5 respectively. Although the effects of HU on MCV and HbF% variations are correlated, the steady-state value of MCV at month 3 is not predictive of the HbF% value at month 26. Hence, HbF% level, which is also directly related to the relief of sickle cell disease symptoms, may be the best biomarker for monitoring HU treatment.

No dose-limiting toxicity occurred in this study, which prevented a toxicity model from being developed. Nevertheless, cytopenia may occur during HU treatment, leading to dose reduction. We compared by simulation two dosing regimens, one continuous daily and the other with interruptions of 2 days after every 5 days. The difference was very small regarding the MCV profile, but larger for the HbF% profile, particularly for simulated patients in the last quartile of HbF% distribution. For these patients, continuous dosing may induce a clinically relevant increase of HbF% compared with the discontinuous schedule. The limits of this simulation exercise are that genetic polymorphisms were not accounted for, and some other biomarkers (arginase, NO enzymes, activated adhesion molecules, phosphatidylserine externalization [25,26]) were not evaluated.

Conclusions

The mode of action of HU on two clinically relevant biomarkers of its efficacy was established. The high variability of response to HU was related in part to pharmacokinetics (HU exposure varied approximately fivefold among patients), and to pharmacodynamics. The steady-state of HbF% and MCV levels need 26 months and 3 months to be reached, respectively, and the interindividual variability of the steady-state values of HbF% is much greater than that of MCV. As a result, the steady-state value of MCV at month 3 is not predictive of the HbF% value at month 26. Hence, HbF% level may be a better biomarker than MCV for monitoring HU treatment. Simulations showed that continuous dosing led to a stronger response than intermittent dosing (5 days out of 7), especially for patients reaching the highest levels of HbF%. Hence, a continuous dosing should be prescribed. Finally, an exciting perspective suggested by the model is that HbF could be further increased by more potent drugs or by drug combinations. In future studies, the model may allow to describe quantitatively the impact of relevant polymorphisms on the variability of response to HU, in order to refine the simulations and to yield specific recommendations for each genotype or haplotype.

Acknowledgements and Funding

We thank Emilie Hénin for advice concerning VPC plots and Christine Fauroux for collecting the data. The PK study was funded by OTL-Pharma Laboratories which participated solely to the study design. The PKPD study was funded by the *Centre de référence pour les syndromes drépanocytaires majeurs* and the *Laboratoire de Pharmacologie*, AP-HP, GH H. Mondor to which belong most authors.

Author details

[1]Université de Lyon, Lyon, France. [2]EMR3738 CTO, Faculté de Médecine Lyon-Sud, Université Lyon 1, Oullins, France. [3]Laboratoire de Pharmacologie, AP-HP, GH H. Mondor, Université Paris Est-Créteil, Créteil, France. [4]Centre de référence pour les syndromes drépanocytaires majeurs, AP-HP, GH H. Mondor, Université Paris Est-Créteil, Créteil, France. [5]Hôpital Croix-Rousse, Hospices Civils de Lyon, Lyon, France.

Authors' contributions

Study design: MT, AHulin and FG; inclusion and follow-up of patients: AHabibi, DB, FG; measurement of HU, monitoring of biological data: HS, KPDP and AHulin; statistical analysis, modelling and simulations: IP, PG and MT, writing of the manuscript: IP, MT, PG, AHulin and FG. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Received: 21 December 2010 Accepted: 28 May 2011
Published: 28 May 2011

References

1. Rodgers GP, Dover GJ, Noguchi CT, Schechter AN, Nienhuis AW: **Hematologic responses of patients with sickle cell disease to treatment with hydroxyurea.** *N Engl J Med* 1990, **322(15)**:1037-1045.
2. Cokic VP, Smith RD, Beleskin-Cokic BB, Njoroge JM, Miller JL, Gladwin MT, Schechter AN: **Hydroxyurea induces fetal hemoglobin by the nitric oxide-**

dependant activation of soluble fetal hemoglobin by the nitric oxide-dependant activation of soluble guanylyl cyclase. *J Clin Invest* 2003, **111(2)**:231-239.
3. Okpala IE: **New therapies for sickle cell disease.** *Hematol Oncol Clin North Am* 2005, **19(5)**:975-987.
4. Bartolucci P, Chaar V, Picot J, Bachir D, Habibi A, Fauroux C, Galactéros F, Colin Y, Le Van Kim C, El Nemer W: **Decreased sickle red blood cell adhesion to laminin by hudroxyurea is associated with inhibition of Lu/BCAM protein phosphorylation.** *Blood* 2010, **116(12)**:2152-2159.
5. Charache S, Terrin ML, Moore RD, Dover GJ, Barton FB, Eckert SV, McMahon RP, Bonds DR, the investigators of the multicenter study of hydroxyurea in sickle cell anemia: **Effect of hydroxyurea on the frequency of painful crises in sickle cell anemia.** *N Engl J Med* 1995, **332**:1317-1322.
6. Steinberg MH, Barton F, Castro O, Pegelow CH, Ballas SK, Kutlar A, Orringer E, Bellevue R, Olivieri N, Eckman J, Varma M, Ramirez G, Adler B, Smith W, Carlos T, Ataga K, DeCastro L, Bigelow C, Saunthararajah Y, Telfer M, Vichinsky E, Claster S, Shurin S, Bridges K, Waclawiw M, Bonds D, Terrin M: **Effect of hydroxyurea on mortality and morbidity in adult sickle cell anemia.** *JAMA* 2003, **289**:1645-1651.
7. Platt OS: **Hydroxyurea for the treatment of sickle cell anemia.** *N Engl J Med* 2008, **358(13)**:1362-1369.
8. Collins FS, Green ED, Guttmacher AE, Guyer MS: **A vision for the future of genomics research.** *Nature* 2003, **422(6934)**:835-847.
9. Ma Q, Wyszynski DF, Farrell JJ, Kutlar A, Farrer LA, Baldwin CT, Steinberg MH: **Fetal hemoglobin in sickle cell anemia: genetic determinants of response to hydroxyurea.** *Pharmacogenomics J* 2007, **7**:386-394.
10. Kumkhaek C, Kumkhaek C, Taylor JG, Zhu J, Hoppe C, Kato GJ, Rodgers GP: **Fetal haemoglobin response to hydroxycarbamide treatment and sar1a promoter polymorphisms in sickle cell anaemia.** *Br J Haematol* 2008, **141**:254-259.
11. Bachir D, Hulin A, Huet E, Habibi A, Nzouakou R, El Mahrab M, Astier A, Galacteros F: **Plasma and urine hydroxyurea levels might be useful in the management of adult sickle cell disease.** *Hemoglobin* 2007, **31(4)**:417-425.
12. De Montalembert M, Bachir D, Hulin A, Gimeno L, Mogenet A, Bresson JL, Macquin-Mavier I, Roudot-Thoraval F, Astier A, Galactéros F: **Pharmacokinetics of hydroxyurea 1000 mg coated breakable tablets and 500 mg capsules in pediatric and adult patients with sickle cell disease.** *Haematologica* 2006, **91(12)**:1685-1688.
13. Beal S, Sheiner LB, Boeckmann A, Bauer RJ: *NONMEM User's Guides (1989-2009)* Ellicott City, Icon Development Solutions; 2009.
14. Brendel K, Comets E, Laffont C, Laveille C, Mentré F: **Metrics for external model evaluation with an application to the population pharmacokinetics of gliclazide.** *Pharm Res* 2006, **23**:2036-2049.
15. Beal SL: **Ways to fit a PK model with some data below the quantification limit.** *J Pharmacokinet Pharmacodyn* 2001, **28(5)**:481-504.
16. Cockcroft D, Gault MD: **Prediction of creatinine clearance from serum creatinine.** *Nephron* 1976, **16(1)**:31-41.
17. Krzyzanski W, Jusko WJ: **Mathematical formalism and characteristics of four basic models of indirect pharmacodynamic responses for drug infusions.** *J Pharmacokinet Biopharm* 1998, **26(4)**:385-408.
18. Lacroix BD, Friberg LE, Karlsson MO: **Evaluating the IPPSE method for PKPD analysis [abstract].** *Population Approach Group in Europe: 8-11 June 2010; Berlin* [http://www.page-meeting.org/default.asp?abstract=1843], Abstract Nr. 1843.
19. Bergstrand M, Hooker AC, Wallin JE, Karlsson MO: **Prediction-corrected visual predictive checks for diagnosing nonlinear mixed-effects models.** *AAPS J* , Published online 08-02-2011..
20. Gwilt PR, Tracewell WG: **Pharmacokinetics and pharmacodynamics of hydroxyurea.** *Clin Pharmacokinet* 1998, **34(5)**:347-358.
21. West GB, Brown JH, Enquist BJ: **A general model for the origin of allometric scaling laws in biology.** *Science* 1997, **276(5309)**:122-6.
22. Green B, Duffull SB: **What is the best size descriptor to use for pharmacokinetic studies in the obese?** *Br J Clin Pharmacol* 2004, **58(2)**:119-33.
23. Lanzkron S, Strouse JJ, Wilson R, Beach MC, Haywood C, Park H, Witkop C, Bass EB, Segal JB: **Systematic review: Hydroxyurea for the treatment of adults with sickle cell disease.** *Ann Intern Med* 2008, **148(12)**:939-955.
24. Tang DC, Zhu J, Liu W, Chin K, Sun J, Chen L, Hanover JA, Rodgers GP: **The hydroxyurea-induced small GTP-binding protein SAR modulates gamma-globin gene expression in human erythroid cells.** *Blood* 2005, **106**:3256-3263.
25. Darghouth D, Koehl B, Madalinski G, Heilier JF, Bovee P, Xu Y, Olivier MF, Bartolucci P, Benkerrou M, Pissard S, Colin Y, Galacteros F, Bosman G, Junot C, Roméo PH: **Physiopathology of sickle cell disease is mirrored by red blood cells metabolome.** *Blood* 2010, prepublished online December 6.
26. Yasin Z, Witting S, Palascak MB, Joiner CH, Rucknagel DL, Franco RS: **Phosphatidylserine externalization in sickle red blood cells: associations with cell age, density, and hemoglobin F.** *Blood* 2003, **102(1)**:365-370.

doi:10.1186/1750-1172-6-30
Cite this article as: Paule *et al.*: Population pharmacokinetics and pharmacodynamics of hydroxyurea in sickle cell anemia patients, a basis for optimizing the dosing regimen. *Orphanet Journal of Rare Diseases* 2011 **6**:30.