Université de Montréal

Optimisation du processus de développement du médicament grâce à la modélisation PK et les simulations d'études cliniques

par
Philippe Colucci, M.Sc.

Sciences pharmaceutiques, Université de Montréal
Faculté de pharmacie

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.)
en Sciences pharmaceutiques
option pharmacologie

Juin 2011

Université de Montréal

Faculté des études supérieures

Cette thèse intitulée :

Optimisation du processus de développement du médicament grâce à la modélisation PK et les simulations d'études cliniques

présenté par :

Philippe Colucci

a été évaluée par un jury composé des personnes suivantes :

Line Labbé, Ph.D.
président-rapporteur, Faculté de Pharmacie, Université de Montréal

Murray P. Ducharme, Pharm.D.
directeur de recherche, Faculté de Pharmacie, Université de Montréal

Jacques Turgeon, Ph.D.
directeur de recherche, Faculté de Pharmacie, Université de Montréal

France Varin, Ph.D.
membre du jury, Faculté de Pharmacie, Université de Montréal

David Z. D'Argenio, Ph.D
examinateur externe, Department of Biomedical Engineering, University of Southern Calfornia

Annie Angers
représentant du doyen de la FES

RÉSUMÉ

Le développement d'un médicament est non seulement complexe mais les retours sur investissment ne sont pas toujours ceux voulus ou anticipés. Plusieurs médicaments échouent encore en Phase III même avec les progrès technologiques réalisés au niveau de plusieurs aspects du développement du médicament. Ceci se traduit en un nombre décroissant de médicaments qui sont commercialisés. Il faut donc améliorer le processus traditionnel de développement des médicaments afin de faciliter la disponibilité de nouveaux produits aux patients qui en ont besoin. Le but de cette recherche était d'explorer et de proposer des changements au processus de développement du médicament en utilisant les principes de la modélisation avancée et des simulations d'essais cliniques.

Dans le premier volet de cette recherche, de nouveaux algorithmes disponibles dans le logiciel ADAPT 5® ont été comparés avec d'autres algorithmes déjà disponibles afin de déterminer leurs avantages et leurs faiblesses. Les deux nouveaux algorithmes vérifiés sont l'itératif à deux étapes (ITS) et le maximum de vraisemblance avec maximisation de l'espérance (MLEM). Les résultats de nos recherche ont démontré que MLEM était supérieur à ITS. La méthode MLEM était comparable à l'algorithme d'estimation conditionnelle de premier ordre (FOCE) disponible dans le logiciel NONMEM® avec moins de problèmes de rétrécissement pour les estimés de variances. Donc, ces nouveaux algorithmes ont été utilisés pour la recherche présentée dans cette thèse.

Durant le processus de développement d'un médicament, afin que les paramètres pharmacocinétiques calculés de façon noncompartimentale soient adéquats, il faut que la

demi-vie terminale soit bien établie. Des études pharmacocinétiques bien conçues et bien analysées sont essentielles durant le développement des médicaments surtout pour les soumissions de produits génériques et supergénériques (une formulation dont l'ingrédient actif est le même que celui du médicament de marque, mais dont le profil de libération du médicament est différent de celui-ci) car elles sont souvent les seules études essentielles nécessaires afin de décider si un produit peut être commercialisé ou non. Donc, le deuxième volet de la recherche visait à évaluer si les paramètres calculer d'une demi-vie obtenue à partir d'une durée d'échantillonnage réputée trop courte pour un individu pouvaient avoir une incidence sur les conclusions d'une étude de bioéquivalence et s'ils devaient être soustraits d'analyses statistiques. Les résultats ont démontré que les paramètres calculer d'une demi-vie obtenue à partir d'une durée d'échantillonnage réputée trop courte influençaient de façon négative les résultats si ceux-ci étaient maintenus dans l'analyse de variance. Donc, le paramètre de surface sous la courbe à l'infini pour ces sujets devrait être enlevé de l'analyse statistique et des directives à cet effet sont nécessaires *a priori*. Les études finales de pharmacocinétique nécessaires dans le cadre du développement d'un médicament devraient donc suivre cette recommandation afin que les bonnes décisions soient prises sur un produit. Ces informations ont été utilisées dans le cadre des simulations d'essais cliniques qui ont été réalisées durant la recherche présentée dans cette thèse afin de s'assurer d'obtenir les conclusions les plus probables.

Dans le dernier volet de cette thèse, des simulations d'essais cliniques ont amélioré le processus du développement clinique d'un médicament. Les résultats d'une étude clinique pilote pour un supergénérique en voie de développement semblaient très encourageants. Cependant, certaines questions ont été soulevées par rapport aux résultats

et il fallait déterminer si le produit test et référence seraient équivalents lors des études finales entreprises à jeun et en mangeant, et ce, après une dose unique et des doses répétées. Des simulations d'essais cliniques ont été entreprises pour résoudre certaines questions soulevées par l'étude pilote et ces simulations suggéraient que la nouvelle formulation ne rencontrerait pas les critères d'équivalence lors des études finales. Ces simulations ont aussi aidé à déterminer quelles modifications à la nouvelle formulation étaient nécessaires afin d'améliorer les chances de rencontrer les critères d'équivalence. Cette recherche a apporté des solutions afin d'améliorer différents aspects du processus du développement d'un médicament. Particulièrement, les simulations d'essais cliniques ont réduit le nombre d'études nécessaires pour le développement du supergénérique, le nombre de sujets exposés inutilement au médicament, et les coûts de développement. Enfin, elles nous ont permis d'établir de nouveaux critères d'exclusion pour des analyses statistiques de bioéquivalence.

La recherche présentée dans cette thèse est de suggérer des améliorations au processus du développement d'un médicament en évaluant de nouveaux algorithmes pour des analyses compartimentales, en établissant des critères d'exclusion de paramètres pharmacocinétiques (PK) pour certaines analyses et en démontrant comment les simulations d'essais cliniques sont utiles.

Mots clés : ADAPT 5®; simulations d'essais cliniques; développement du médicament; demi-vie; MLEM; ITS.

SUMMARY

Drug development is complex with results often differing from those anticipated or sought after. Despite technological advances in the many fields which are a part of drug development, there are still many drugs that fail in the late stages of clinical development. Indeed, the success rate of drugs reaching commercialization is declining. Improvements to the conventional drug process are therefore required in order to facilitate development and allow new medications to be provided more rapidly to patients who require them. The aim of this Ph.D. project was to explore and propose ways to improve this inefficient drug development process with the use of advanced modeling and clinical trial simulations.

For the first part of this research, new algorithms available in ADAPT 5® were tested against other available algorithms in order to determine their potential strengths and weaknesses. The two new algorithms tested were the iterative two-stage and the maximum likelihood expectation maximization (MLEM) methods. Our results demonstrated that the MLEM algorithm was consistently better than the iterative two-stage algorithm. It was also comparable with the first order conditional estimate method available in NONMEM®, with significantly fewer shrinkage issues in the estimation of the variances. Therefore, these new tools were used for the clinical trial simulations performed during the course of this Ph.D. research.

In order to calculate appropriate noncompartmental pharmacokinetic parameter estimates during the drug development process, it is essential that the terminal elimination half-life be well characterized. Properly conducted and analyzed pharmacokinetic studies are essential to any drug development plan, and even more so for

generic and supergeneric (a formulation similar to the reference product, containing the same active ingredient; however differing from the original reference product it its delivery process) submission where they often are the only pivotal studies that need to be done to decide if a drug product is good or not. Thus, the purpose of the second part of the research was to determine if the pharmacokinetic (PK) parameters obtained from a subject whose half-life is calculated from a sampling scheme duration that is considered too short could bias the bioequivalence conclusions of a study and if these parameters should be removed from statistical analyses. Results demonstrated that subjects with a long half-life relative to the duration of the sampling scheme negatively influenced results when these were maintained in the analysis of variance. Therefore, these subjects should be removed from the analyses and guidelines to this effect are required *a priori*. Pivotal pharmacokinetic studies done within the drug development process should therefore follow this recommendation to make sure that the right decision be taken on a drug product formulation. This information was utilized with the clinical trial simulations that were subsequently performed in this research in order to ensure the most accurate conclusions.

Finally, clinical trial simulations were used to improve the development process of a nonsteroidal anti-inflammatory drug. A supergeneric was being developed and results from a pilot study were promising. However, some results from the pilot study required closer attention to determine if the test and reference compounds were indeed equivalent and if the test compound would meet the equivalence criteria of the different required pivotal studies. Clinical trial simulations were therefore undertaken to address the multiple questions left unanswered by the pilot study and these suggested that the test

compound would probably not meet the equivalence criteria. In addition, these results helped determine what modifications to the test drug would be required to meet the equivalence criteria. This research brought forward solutions to improve different aspects of the drug development process. Notably, clinical trial simulations reduced the number of studies that would have been done for a supergeneric, decreased the number of subjects unnecessarily exposed to a drug, lowered costs and helped established new criteria for the exclusion of subjects from analyses.

Research conducted during this Ph.D. provided concrete ways to improve the drug development process by evaluating some newly available tools for compartmental analyses, setting standards stipulating which estimated PK parameters should be excluded from certain PK analyses and illustrating how clinical trial simulations are useful to throughout the process.

Key words: ADAPT 5®, Clinical Trial Simulations, Drug Development, Half-life, MLEM, Iterative two-stage

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: INTRODUCTION

ARTICLE #1

ARTICLE #2

ARTICLE #3

LIST OF ABBREVIATIONS AND SYMBOLS

α	Individual PK parameter vector
AIC	Akaike information criterion
ANOVA	Analysis of variance
AUC0-t	Area under the curve from time 0 to the last measurable concentration
AUC0-τ	Area under the curve during the dosing interval
AUCinf	Area under the curve from time 0 to infinity
AUCt-inf	Area under the curve from time t to infinity
AUMC	Area under the moment curve
β	Variance model parameters
BA	Bioavailability
BE	Bioequivalence
BMSR	Biomedical simulation resource
C	Concentration
CI	Confidence intervals
CL	Total clearance
CL/F	Total apparent clearance
C_{Last}	Last measurable concentration
CLd	Distributional clearance
CLd/F	Apparent distributional clearance
Cmax	Maximum concentration
Cmin	Minimum concentration
CTS	Clinical trial simulations
CV	Coefficient of variation
ED_{50}	Dose providing 50% of maximal response
EM	Expectation maximization
Emax	Maximum effect
ε_i	Epsilon: error term describing difference between predicted and observed concentrations
ε_{ij}	Independent identically distributed random error
η_i	Eta: Difference between population and individual parameter value
F	Bioavailability
f	Function
FDA	Food and drug administration

FIH	First time in human
FO	First order
FOCE	First order conditional estimation
Frel	Relative bioavailability
g_i	Weight for predicted concentration
h	Hour
IR	Immediate release
ITS – IT2S	Iterative two-stage
IV	Intravenous
ka	Absorption rate constant
Kel	Apparent terminal elimination rate constant estimated noncompartmentally
λ_Z	Terminal elimination rate constant estimated compartmentally
MAP	Maximum *a posteriori* probability
MLEM	Maximum likelihood expectation maximization
MRT	Mean residence time
n	Number of total observations
NDS	New drug submission
NSAID	Non-steroidal anti-inflammatory drug
OGD	Office of generic drug
OLS	Ordinary least square
O_{OLS}	Ordinary least square objective function
O_{MAP}	Maximum *a posteriori* probability objective function
Ω	Variance covariance matrix
O_{NLL}	Negative log maximum likelihood objective function
O_{WLS}	Weighted least squares objective function
p	Vector of PK parameters of the model
PD	Pharmacodynamic
Pi	Vector of pharmacokinetic parameters for an individual
PHN	Postherpetic neuralgia
PK	Pharmacokinetic
POC	Proof of concept
q	Number of variance parameters in Akaike information criterion
R^2	Regression coefficient
R&D	Research and development
SD	Standard deviation
SERM	Selective estrogen receptor modulators
sh_η	η-shrinkage

σ^2	Variance
SR	Sustained release
STS	Standard two-stage
t	Time
t_j	J^{th} time point
T½	Terminal half-life
θ	System model parameters
T_{Lag}	Lag time prior to the start of absorption
Tmax	Time associated to the maximum concentration
TPD	Therapeutic Products Directorate
US	United States
var	Variance
Vc	Volume of central compartment
Vc/F	Apparent volume of central compartment
Vp	Volume of peripheral compartment
Vp/F	Apparent volume of peripheral compartment
Vss	Total volume of distribution
Vss/F	Apparent total volume of distribution
WLS	Weighted least squares
ω^2	Variance
X_i or X_{ij}	Vector of independent variables associated with y_i or y_{ij}
y_{ij}	i^{th} concentration for j^{th} individual
Z_i	i^{th} predicted concentration for an individual

DEDICATION

To my wife, Natali Carpentier, without whom this dream would never have been possible and to my son, Gabriel Colucci, whose passion, curiosity and thirst for knowledge give me reason to strive for more. I love you both with all my heart.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my director of research, Dr. Murray P. Ducharme, for his constant availability and unwavering generosity with both time and scientific knowledge. Each day, he taught me something new and inspired me to excel, and for that I am eternally grateful. As well as my mentor, he has become a dear friend whose insight I have come to value immensely.

I also would like to extend my gratitude to my other director of research, Dr. Jacques Turgeon, for allowing me to take up residence in the laboratory and exchange ideas with other colleagues. His lessons shared in the classroom and in laboratory meetings were always educational, motivating and thought-provoking and his comments on my work constructive and insightful.

My deepest appreciation also goes to the members of the jury for accepting to evaluate this thesis.

To my family and friends, who have been there throughout this process to encourage me on and give me comfort when I needed it most, thank you from the very bottom of my heart. Your support made this dream possible and put it within my reach.

Last, but not least, I wish to express my heartfelt appreciation to Corinne Seng Yue for the endless stream of scientific discussions and for serving as a constant source of motivation throughout my Ph.D.

GENERAL ORIENTATION OF THE THESIS

Despite technological advances in the many fields which are a part of drug development, there are still many drugs that fail in the late stages of clinical development. These failures negatively affect the numerous monetary and scientific resources that could have been best allocated to improve the development of other medications. The overall result is that the development of drugs is significantly slowed down. Reasons for the drug development failures include poorly designed studies, extracting insufficient information from the data collected and a generally poor understanding of the drug being developed. These problems highlight the inefficiencies of the conventional drug development process and indicate why improvements are essential. Positive changes would allow new medications to be provided more rapidly to patients who require them.

This research aimed at bringing forward solutions that may provide a more efficient drug development process. More specifically, this work focused on the use of advanced modeling and clinical trial simulations to achieve this objective. Clinical trial simulations have the potential to predict study outcomes, reduce unnecessary studies, decrease the number of volunteers exposed to drugs and lower development costs.

This thesis is divided into three sections, namely an introduction, a description of the research performed, and finally a general discussion and conclusion. The introduction discusses the cost of the drug development process, general pharmacokinetic concepts including noncompartmental and compartmental approaches for calculating parameters, their importance within the drug development process, and examples of the use of clinical trial simulations in the industry. It also includes a short summary of the different clinical

phases of the drug development process and explains how clinical trial simulations could improve it.

The second section presents the work accomplished through this research and includes three articles as the research was conducted in three main parts. The first aimed at determining if the new algorithms available in the software ADAPT 5® were adequate to be used for clinical trial simulations. The article is entitled "Performance of different population pharmacokinetic algorithms using clinical simulations". The second part of this research used clinical trial simulations to determine if an additional criterion should be required when estimating the noncompartmental apparent terminal elimination rate in pivotal pharmacokinetic studies performed during the drug development process. This second article is entitled "How critical is the duration of the sampling scheme for the determination of half-life, characterization of exposure and assessment of bioequivalence?". The last part of this research provides an example of how clinical trial simulations can help optimize and speed up the development of a new drug. The title of this last article is "Improved Drug Development and Bioequivalence Potential of a New Extended-Release Formulation Determined by Clinical Trial Simulations".

Finally, the thesis ends with a general discussion and conclusion discussing the main results of the research and its potential applicability as a whole to the drug development process.

CHAPTER I: INTRODUCTION

1. INTRODUCTION

The drug development process requires huge investments in order to produce a marketable drug. It is estimated that the cost to produce a new drug varies from 800 million to close to 2 billion United States (US) dollars.[1,2,3] In addition, costs are increasing at a higher rate than inflation further reducing profitability.[4] The above quoted cost includes approximately 400 million dollars in opportunity lost.[1] In addition, the cost includes research and development (R&D) costs for all failed drugs. The money invested in these failed drugs cannot be recovered. The average success rate of drugs entering the clinical phase of development is approximately 8 to 16%.[5,6] In order for a pharmaceutical company to be profitable, income from marketed drugs has to cover its production costs and the costs for the failed drugs. Because of that, pharmaceutical companies typically spend between 12 and 20% of their overall revenues in R&D. Despite this figure, the whole drug development process is at the present time anemic and inefficient. Besides the extremely high development cost, the time spent to produce a new drug is also on average disappointingly long. It generally takes 12 to 15 years for a new drug to arrive on the market.[7,8] Pharmaceutical companies in the US only benefit from exclusivity patents for 20 years, most of which is taken by the time to develop the drug, and a market exclusivity of a minimum of 5 year. The length of time to develop a drug is attributed to the fact that many different steps of development are needed must be very closely monitored and sometimes directly inspected by regulatory agencies to ensure public safety. It takes on average approximately 4 years for a drug to go through the pre-clinical stage and 7 years to go through the clinical stage until its regulatory submission, 1 to 2 years in Phase I, 2 years in Phase II, and 2 to 4 years in Phase III.[7,9] Regulatory

agencies may take after that 1.5 to 2.5 years to review the data and approve the submission.[7,10] The time and cost discussed above does not include any Phase IV post-marketing studies or activities that may be required.

Stage	Time (years)	Test Population
Pre-clinical	4	*In vitro* or animals
Phase I	1-2	Healthy volunteers
Phase II	2	Patient volunteers
Phase III	2-4	Patient volunteers
FDA Approval	1.5-2.5	
Total	10-15	

Table I: Timeline for the development of a new drug

In order for pharmaceutical companies to remain competitive as a business model, it is imperative that they make changes to their current practices. Based on the cost and time invested in developing a new drug, there are two ways that this can be accomplished. The first approach is that drugs that are not successful need to be abandoned as quickly as possible in the development process. There are still approximately 50% of drugs entering Phase III that fail to be marketed.[6,7,11] Phase III studies are the longest and include the most number of patients which are very expensive to enroll. Phase III should never fail due to foreseeable reasons such as having the wrong choice of regimen to prove efficacy. Thus, if a new compound reaches clinical testing, it should be known before Phase III if it has any chances of success. Phase I and II studies should be planned accordingly and data collected in the most efficient matter to extract all possible information from the drug before reaching Phase III. Reducing the development costs of failed drugs will decrease the overall cost and free funds to be invested in other compounds that have better chances of succeeding. In addition, the sooner a drug that is not marketable can be set aside, the faster scientists can spend their

time and energy on other potentially marketable products. The second way to improve pharmaceutical profitability is to reduce the number of unnecessary studies for successful drugs. Even if a drug is successfully marketed, the program may have conducted studies that were poorly planned or did not prove what they were intended to show. If the number of studies can be reduced or conducted more efficiently, this will lower costs and shorten the time period it takes to market the drug. Consequently, drugs that have a quicker turn around time between discovery and submission to regulatory agencies will benefit from a longer exclusivity time which in turn will increase profitability.

Pharmaceutical companies have successfully changed trends in the past. Previously, Phase I accounted for close to 40% of all clinical drug failures. This attrition level is now below 10%.[12] This was made possible because pharmaceutical companies recognized the problem and came up with a solution. The solution was that companies verified more diligently the pharmacokinetics (PK) of drugs including drug metabolism within the discovery and pre-clinical phases. Nonlinear products, compounds not well absorbed and drugs that may cause significant metabolite interactions are removed from the development process in the discovery or pre-clinical phases rather than in Phase I. However, the biggest problem right now is that drugs are failing mostly in the late stages of development. The main reasons for these failures are lack of efficacy and safety issues. The pharmaceutical industry is conscient of these problems and is trying to look for solutions.

The Food and Drug Administration (FDA) recognized the need to improve the drug development process and published a document entitled “Challenge and Opportunity on the Critical Path to Medical Products”.[5] In this paper, the FDA suggests

that pharmaceutical companies need to improve the development process and reduce failures. The FDA goes on to specify that tools such as pharmacokinetic and pharmacodynamic (PD) modeling could help bring a drug to market more efficiently.

In creating such a document, the FDA has shone a light on pharmacometrics, which is the science of applying mathematical and statistical methods to better understand and predict a drug's PK/PD behavior.[13] One aspect of pharmacometrics is the use of clinical trial simulations. With a better pharmacometrics-based understanding of a drug's PK/PD profile obtained as early as possible in the development process, it is anticipated that less drugs will fail in the late clinical stages. In addition, the more that is known about a drug, the greater its potential to benefit patient care as well as increase its likelihood to be efficiently pushed through the different stages of development.

In light of the challenges faced by those who attempt to bring a new drug on the market, the objective of this thesis is to propose solutions to improve the drug development process. One such solution is the more widespread application of pharmacometrics and the use of clinical trial simulations (CTS) during different stages of the drug process. To better understand how pharmacometrics can lead to more efficient drug development, a clear understanding of pharmacokinetic concepts is required and will be discussed hereafter. This will be followed by a short summary of the different clinical phases of drug development and how clinical trial simulations can improve the process.

1.1. PHARMACOKINETICS

Pharmacokinetics is the study of what the body does to a drug. It is a branch of pharmacology that studies the movement of the drug through the organism.[14] This movement through the body is generally divided into three main categories which are the absorption, distribution and elimination of the drug. Absorption is the process by which the drug enters the body; distribution is the description of how and where the drug will disperse throughout the body while the elimination process characterizes how the drug is finally cleared from the organism. Elimination can be further sub-divided into metabolism and excretion.

Pharmacokinetics is often studied in conjunction with pharmacodynamics. Pharmacodynamics is often referred to as what the drug does to the body. It relates the effects over time experienced by volunteers to the drug's PK. The characterization of the pharmacokinetics will first establish the systemic concentrations and consequently, the theoretical concentrations at the site of action. It is these concentrations at the site of action that are responsible for the PD effect.[15] If we assume that the systemic plasma concentrations have a relationship with those at the site of activity, then we can use these plasma concentrations to predict the theoretical concentrations at the site(s) of activity and build PK/PD relationships.[16-19] The PK/PD correlation is crucial in understanding the relationship between the systemic concentrations of the drug (exposure) and its effects on the body. The effect can be either wanted (beneficial) or unwanted (harmful). It is this relationship between the pharmacokinetic and pharmacodynamic properties of a drug that will help determine its level of activity.

Therefore, the first step in understanding a drug and its effect on the body is to describe its pharmacokinetics. There are two main approaches for determining the pharmacokinetics of a drug. The first is to use a noncompartmental approach[20-21] while the second is the compartmental approach.[22-23] Both have advantages and disadvantages and are not mutually exclusive.[24] When determining the pharmacokinetics of a drug, we can either choose to describe the PK parameters of an individual, known as individual PK, or those of a population which is known as population PK. The noncompartmental approach is better suited to the description of individual's pharmacokinetics while compartmental analyses are well suited for both. Each approach is useful during the drug development process, but to perform clinical trial simulations, compartmental models are required. However, study outcomes are sometimes based on noncompartmental analysis, so clinical trial simulations often simulate profiles which are then used to calculate noncompartmental parameters. The decision to use noncompartmental, compartmental or both approaches will depend on the purpose of the analyses as well as the available data collected. The following two sections will provide a general review of the two different approaches used to characterize the pharmacokinetics of a drug.

1.1.1. Noncompartmental analysis

The noncompartmental approach calculates pharmacokinetic parameters based on the graphical interpolation and extrapolation of concentrations over time. This approach is based on the theory of statistical moments which is a mathematical concept explaining the distribution of data.[25-29] Although statistical moments theory was used in other fields of research before being applied to pharmacokinetics, it was regularly used for PK analyses by the early 1980s. In PK, statistical moments are calculated from a set of concentration-time data and represent an estimate of the true moment. It is an estimate of the true relationship between concentration and time.

Typically, only the first two moments are used in PK.[30] The first moment defines the area under the concentration curve from time zero to infinity (AUCinf) and relates the exposure of the drug to the concentrations as defined in Equation 1.

$$AUC = \int_0^{\infty} C * dt \qquad (1)$$

Area under the curve to infinity is typically calculated using the trapezoidal method. Multiple trapezoidal methods exist such as the linear trapezoidal and the log-linear trapezoidal rules.[31] These methods consists of adding multiple small trapezoidal areas and an example of how this is calculated for the linear trapezoidal method in given in Equation 2.

$$\text{AUC0-t} = \sum_0^t (C_2 + C_1)/2 * (t_2 - t_1) \qquad (2)$$

AUCinf is defined in Equation 3 and is the sum of the area under the curve from time 0 to the last measurable concentration (AUC0-t) and the area under the curve that is

extrapolated beyond the last measurable concentration to infinity (AUCt-inf). The first term, AUC0-t, is calculated as per Equation 2. The second term, the extrapolated area, is calculated as the last measurable concentration (C_{Last}) divided by the apparent terminal elimination rate constant (Kel).

$$AUCinf = AUC0\text{-}t + (C_{Last}/Kel) \quad (3)$$

If a drug follows first-order elimination, Kel is calculated from the slope of the plot of the logarithm of concentration versus time. The slope has to be estimated during the apparent terminal phase of the profile.

The second statistical moment is involved in the measurement of the mean residence time (MRT) determined by the area under the moment curve (AUMC). AUMC is estimated by Equation 4.

$$AUMC = \int_0^\infty t * C * dt \quad (4)$$

AUMC has no physiological value and is simply a mathematical variable used to determine other pharmacokinetic parameters which have more useful physiological meaning. MRT is simply the AUMC divided by AUCinf for a bolus intravenous administration.[32-33] If a drug is administered intravenously via an infusion, then half the time of the infusion has to be subtracted from the MRT calculation. The third moment is the variance associated with the calculated parameter and is usually estimated with too much uncertainty to be useful.[29]

Using the graphical representation of the concentration versus time profile and statistical moment theory, other useful PK parameters of interest can be obtained. The first parameter is the observed maximum concentration (Cmax) and the time associated with this maximum concentration (Tmax). Both of these parameters are associated with

the rate of absorption and are taken directly from the profile without any interpolation of the data.

A crucial PK parameter is the clearance. It is a measure of the volume of blood or plasma from which the drug is removed per unit of time[25,26,34-35]. The total clearance (CL) of a drug is calculated by Equation 5.

$$CL = \frac{Dose * F}{AUC\,\mathrm{inf}} \tag{5}$$

Where bioavailability (F) is defined in Equation 6

$$F = \frac{Dose_{IV} * AUC_{oral}}{Dose_{oral} * AUC_{IV}} \tag{6}$$

Clearances are additive and therefore total clearance represents the sum of all the clearances from different organs, except for the lung as the blood supply to the lungs is in series rather than in parallel with other organs.[26] Another essential parameter is the total volume of distribution (Vss). This is a virtual volume and provides information on the extent to which the drug is distributed in the body. The formula is described in Equation 7.

$$Vss = CL * MRT \tag{7}$$

Clearance and volume of distribution are two independent parameters. The elimination half-life (T½) is a parameter dependent on these two PK parameters and it represents the time it takes the organism to eliminate half of the drug or reduce its concentrations by half. For drugs displaying first-order elimination, this half-life is independent of the amount of drug that is administered. It is estimated based on Kel and the calculation is shown in Equation 8.

$$T\,\tfrac{1}{2} = \frac{\ln(2)}{Kel} \tag{8}$$

The relationship between the elimination half-life, Vss and CL parameters is demonstrated in Equation 9, assuming a one compartment model.

$$T\,\tfrac{1}{2} = \frac{\ln(2) * Vss}{CL} \tag{9}$$

From Equation 9, the larger the volume of distribution, the longer the elimination half-life since a greater volume of distribution results in a lower blood or plasma concentrations. A lower concentration implies that a smaller amount of drug is reaching the eliminating organ so it will take longer to eliminate the drug. The opposite is true with clearance since the larger the clearance, the shorter the elimination half-life. This is evident as clearance represents the capacity of an organ to eliminate drug and the more efficient it is at eliminating the drug, the less time it will take to reduce the amount of drug by half.

Based on the noncompartmental equations, it is imperative that both the AUC0-t and Kel be well characterized to adequately calculate the noncompartmental parameters, as most of the noncompartmental PK parameters are dependent on both the AUC0-t and/or Kel. In order to properly characterize the AUC0-t parameters using the linear trapezoidal method, a sufficient number of samples are required. Generally, it is recommended that at least 15 samples be collected in each subject after a single dose administration.[36] These samples have to be collected at specific moments, with approximately 5 samples each taken during the absorption and distribution phase to properly characterize Cmax and Tmax, and 5 samples in the elimination phase to robustly characterize the Kel. The extrapolated portion of the AUCinf parameter is dependent on

both the last measurable concentration as well as the Kel parameter, so improper characterization of the Kel will lead to poor estimates of the extrapolated portion of the AUCinf. The CL, Vss and MRT parameters are dependent on AUCinf. Therefore, any poor estimate of AUCinf due to either a poor Kel or to improper sample selection for AUC0-t will lead to poor estimates of these parameters. Clearance is calculated using the AUCinf parameter and any error in the AUCinf estimation leads directly to the same error in the clearance estimation.[37] The error doubles for the estimated MRT and Vss parameters as these parameters are dependent on both AUCinf and AUMCinf (*i.e.*, both are dependent on extrapolation error).[21,27,29,37] In addition, the half-life is based directly on the Kel and thus a poor Kel estimate will lead to a poor T½ estimate.

Advantages of the noncompartmental approach are that it is relatively simple, robust and almost model-independent (e.g., except for the extrapolation of the last trapeze which is based on a monoexponential decline). Because it is virtually model-independent, its results are not dependent on the scientist's ability at modeling data. In addition, the noncompartmental approach is usually not significantly influenced by experimental errors associated with each individual measured concentration as long as there are enough samples taken, as previously described. A certain experimental error is associated with each concentration, which includes the variability in the analytical analysis, dosing errors, collection errors and other clinical errors. However, the error associated with each concentration does not contribute to the overall variability of the AUC parameter. This is due to the fact that with numerous concentrations collected, the individual errors associated with each concentration cancel themselves out. More specifically, one concentration might be higher than expected while another

concentration may be lower than expected. The addition of all the overestimated and underestimated errors in the concentrations cancel each other out and the overall AUCinf is generally unaffected by the experimental errors.

Even though AUCinf determination is robust with respect to experimental errors, it is still influenced by the Kel precision. Previously published research suggests that to have a proper Kel value and consequently a proper AUCinf value, a minimum number of samples collected in the elimination phase is required, the predicted C_{Last} should be used rather than the observed C_{Last} and the extrapolated portion of the AUCinf should be maintained to a minimum (e.g., maximum 20%).[38-40,41] Based on prior experience, another criterion suspected to be important was the sample collection duration. Generally, a study is planned based on average PK parameters that are expected (from literature or previous studies). Sometimes the PK study is the first one ever conducted in humans, and therefore *a priori* data is not available. Due to this and to individual subject variability or unexpected PK values, it is our hypothesis that some subjects may not have the optimal sampling scheme to appropriately determine their PK parameters robustly. Consequently, it is important to understand how the sample collection duration may influence PK parameters values, and how subjects that may not have an optimal sampling scheme may influence the conclusions of a study. Therefore, a research study presented in this thesis aimed at determining the influence of the sample collection duration on the precision of PK parameter estimates. This work is presented in Chapter 2 (Article 2).

1.1.2. Individual compartmental analysis

The compartmental approach is the classical PK approach and is the foundation of the field of pharmacokinetics. The aim of compartmental analyses is to explain observed concentrations with the use of mathematical and statistical models. These models are comprised of hypothetical compartments representing the body and are used to explain how the drug reacts within the body. With empirical models, each compartment represents a group of tissues or organs with similar blood flow. Each compartment has a volume of distribution and all compartments together represent the extent to which the drug distributes in the body. The more compartments that are required for an analyte, the greater normally the volume of distribution a drug will have. Movement between compartments is comprised of rate constants. These are often labeled as K_{ij} (where i and j are different compartments) or as CLd. With the exception of physiologically-based pharmacokinetic models which attempt to reproduce the physiological aspects of the body, compartmental PK analyses attempt to find the simplest model to best explain the observed concentrations while still remaining true to being physiologically relevant.

Individual compartmental analyses have been around since the 1960s and made popular by many including Levy[42-47] in the 1960s and Sheiner in the early 1970s.[48-54] A basic model to explain the observed concentration from an individual can be written as in Equation 10.

$$y_i = f(P, X_i) + \varepsilon_i \tag{10}$$

Where y_i is the i^{th} concentration for an individual

P is the vector of pharmacokinetic parameters of the model

X_i is the vector of independent variables (such as time and dose) associated with y_i

ε_i is the statistical error that corresponds to the measurement error, the change in PK over time for the subject and also the model misspecifications.

In Equation 10, the predicted concentration at time i for an individual is the value determined by $f(P, X_i)$. Therefore, the difference between the predicted and observed concentrations is the error term ε_i in Equation 10 and is represented differently in Equation 11 and graphically in Figure I.

$$y_i = z_i + \varepsilon_i \tag{11}$$

Where z_i represents the predicted i^{th} concentration

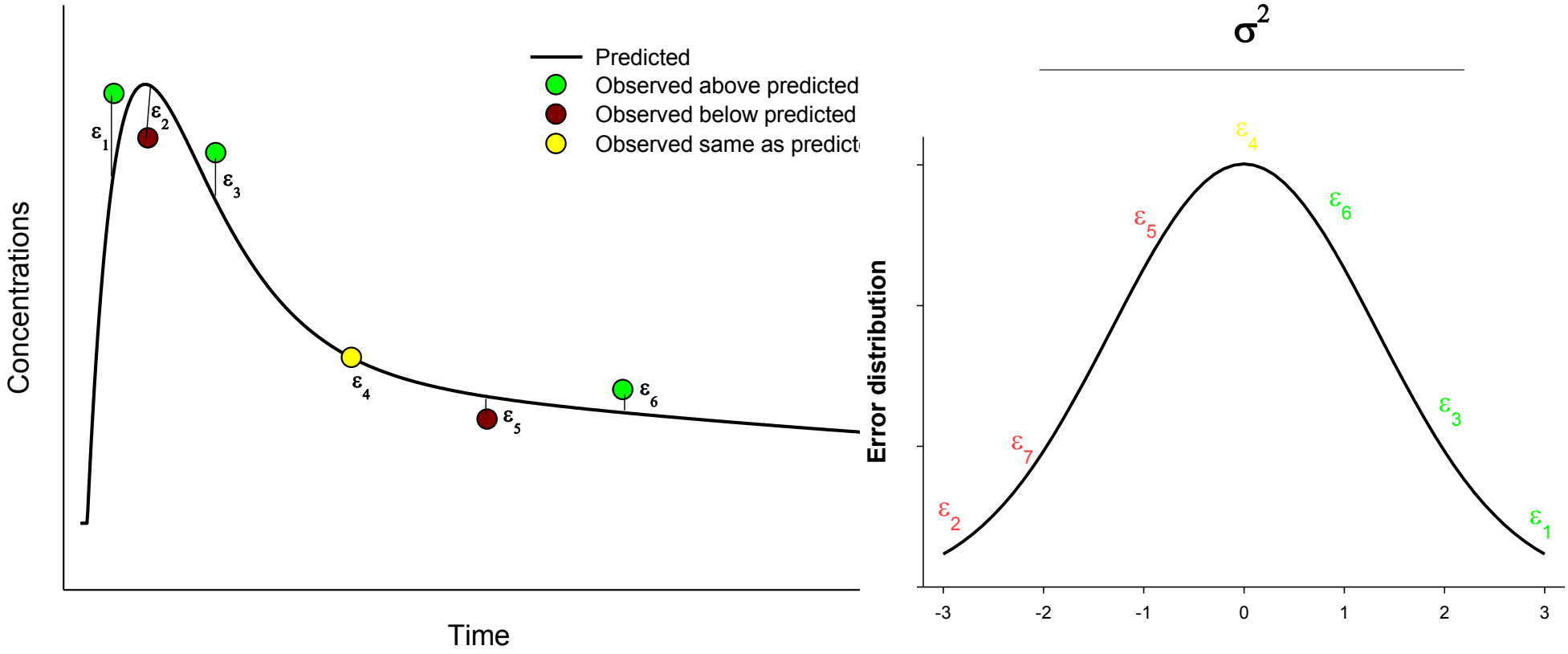


Figure I: Error distribution between predicted and observed concentrations

The errors ε_i are assumed to be independent and have a mean of 0 with a variance σ^2.[51,55-57] Unlike noncompartmental analyses which does not estimate the errors ε_i, compartmental analyses take these errors into consideration and try to find predicted concentrations to minimize the individual ε_i.

There have been many methods proposed to minimize the differences between the predicted and the observed concentrations. The first method, well known in statistics, is the ordinary least square (OLS) estimates.[58-60] The OLS function (O_{OLS}) minimizes the squared errors between the observed (y_i) and predicted (z_i) concentrations and is the value obtained in Equation 12.

$$O_{OLS} = \sum_{i=1}^{n} (y_i - z_i)^2 \qquad (12)$$

This value is relatively easy to obtain; however, it is not ideal if there is a wide range of concentrations. Concentrations in a profile vary greatly and may even span multiple logs (*i.e.*, Cmax might be 2 to 3 logs higher than the minimum concentration (Cmin)). Using Equation 12, high concentrations will have a greater impact on the OLS function than low concentrations. For example, a 10% error on a concentration value of 1000 is 100 while the same 10% error on a concentration of value of 1 is 0.1. Therefore the OLS function will minimize the distance between the higher predicted and observed concentrations and ignore this difference for low concentrations. Since we assume that all concentrations have the same percentage error, they should be considered equally as important when trying to determine the optimal PK parameters. Therefore, a weight is often added to this OLS function and the function becomes a weighted least squares

estimate (WLS). The weight most often used is the error variance[61] as shown in Equation 13.

$$O_{WLS} = \sum_{i=1}^{n} \frac{(y_i - z_i)^2}{\sigma^2} \qquad (13)$$

Where σ^2 is the variance

The two previous methods used to minimize the error between the predicted and observed concentrations are not the most efficient. A superior method to these functions is the maximum likelihood function. The observed data are approximated by a model function consisting of the parameters being estimated. This approach maximizes the probability of obtaining the observed data by estimating the best possible parameter estimates. In other words, the solution to the function is the best set of system and variance model parameters (θ and β) that renders the observed concentrations the most likely from any other estimates.[56] The maximum likelihood objective function (O_{NLL}) is described in Equation 14.

$$O_{NLL} = l \cdot m \cdot \ln(2\pi)/2 + \frac{1}{2} \sum_{i-1}^{l} \sum_{j=1}^{m} \left[\frac{(Z_i(t_j) - y_i(\theta, t_j))^2}{g_i(\theta, t_j, \beta)} + \ln g_i(\theta, t_j, \beta) \right] \qquad (14)$$

Where where i is the type of data (*e.g.*, plasma concentration and urinary output)

j is the number of samples

l and m are the total number of output data and total number of samples, respectively

$(Z_i(t_j) - y_i(\theta, t_j))^2$ is the square of distance between the predicted and observed concentrations

g_i is the weight for the predicted concentrations which is the variance var{v(t_j)}.

Another method to minimize the error between the predicted and observed concentrations is the generalized least squares (GLS) estimates.[62] With this method, the observed data are also approximated by a function consisting of system and variance model parameters (θ and β). Unlike the maximum likelihood approach, the system and variance parameters are estimated separately. In the first iteration, the system parameters are estimated using a least squares estimation. Then in a second iteration, the variance parameters are estimated using a maximum likelihood function and the parameters estimated in the first iteration. In a third iteration, the system parameters are re-evaluated using a weighted least squares estimate and the variance parameters estimated from the second iteration. The second and third iterations are repeated until convergence.

There is also a Bayesian method (*Maximum a Posteriori Probability* – MAP) that can be used to minimize the difference between predicted and observed concentrations.[60,62-64] The Bayesian method uses an objective function that takes into consideration the results of the individual and those from the population. The MAP objective function that is minimized is described in Equation 15.

$$O_{MAP} = \frac{1}{2}\sum_{i-1}^{l}\sum_{j=1}^{m}\left[\frac{(Z_i(t_j) - y_i(\alpha, t_j))^2}{g_i(\alpha, t_j, \beta)} + \ln g_i(\alpha, t_j, \beta)\right] + [\alpha - \mu]^T \Omega^{-1}[\alpha - \mu] \qquad (15)$$

Where $g_i(\alpha, t_j, \beta)$ is the weighting for the predicted concentration which is the variance var{v(t_j)}

α is the individual PK parameter vector

μ is the population PK parameter vector

T is the transpose of the matrix

Ω is the covariance matrix.

In this equation, the algorithm minimizes two distinct terms to obtain the individual PK parameters. The first term is the distance between the individual predicted concentrations and the observed concentrations while the second half of the equation represents the distance between the individual PK parameter estimates and the population PK parameter estimates. Therefore, if an individual has many observations, the equation add more weight to the individual's observations and the impact from the population parameter values will be minimal while the opposite is true when an individual has fewer observations. In this case, more weight will be given to the population PK parameter values and the subject's individual parameter estimates will tend to more closely resemble the population values.

1.1.3. Population compartmental analysis

Because of its superior robustness, the population approach is the preferred analysis when performing compartmental analyses.[39,65] This approach was first introduced in the 1970s to analyze sparse observational data collected from different clinical trials. In addition to estimating the mean PK parameters in a target population, the aim of population compartmental analysis is to determine the dispersion of these PK parameters (inter-individual variability) as well as the residual error (which includes intra-subject variability and measurement error). This is what differentiates this type of analysis from individual compartmental analysis which will only determine the PK parameters for each individual separately without considering any data from the other subjects in the analysis. Describing the variation of the PK parameters adds parameters to be estimated and contributes to the complexity of the analysis. In spite of these challenges, a proper population analysis not only predicts the results of the subjects that were analyzed but permits the user to make inferences on the population and future outcomes. Generally, decisions in drug development are based on the typical or average parameters of a drug in the population. However, knowledge of the typical concentration-time profile of the drug and how patients' profiles can vary is crucial to the regulatory agencies and the pharmaceutical companies to ensure efficient and safe administration of a drug.

In population PK analysis, to explain the observed data of a particular subject, Equation 10 is expanded to reflect the population in Equation 16.[60]

$$y_{ij} = f(P_j, X_{ij}) + \varepsilon_{ij} \tag{16}$$

Where y_{ij} is the ith concentration for the jth individual of the population analyzed

P_j is the vector of pharmacokinetic parameters for the jth individual

X_{ij} is the vector of independent variables (such as time and dose) associated with y_{ij}

ε_{ij} is the independent identically distributed random error with a mean of zero and a variance of σ^2.

Using a population analysis, P_j is further expanded to include every subject in the population as defined in Equation 17.

$$P_j = q(\theta, X_j, \eta_j) \tag{17}$$

Where q is a vector value function,

θ is the vector of the population PK parameters

X_j are the covariates that may influence P_i

η_j is the vector of independent identically distributed random error having means of zero and variances of ϖ^2. This is a covariance matrix often referred to as Ω.

It is this distribution of η_i for all the subjects around the mean PK parameter θ that provides information on the variability of the PK parameter θ. This variability is presented as the variance (ϖ_i^2) of θ. This variance represents the inter-subject variability. To clarify the above, we can use a simple 1-compartment model which has CL and volume of central compartment (Vc) as PK parameters. To explain an observation for a jth individual, such as a concentration, you have to determine the population PK parameters θ_{CL} and θ_{Vc} as well as the distance the subject's own PK parameter estimates

are from the population PK parameters. This distance between the population and individual value is presented as eta (η); therefore, the jth individual would have η_{jCL} and η_{jVc}. Using these individual parameters for the jth individual, a concentration can be predicted at any given time for this particular individual. To explain the observed concentration, an additional error term is required that accounts for the distance between the predicted concentration and the observed concentration and it is represented by epsilon ε_{ij}. It is the distribution of these ε_{ij} that provides information on the residual variability. As specified previously, this residual variability is a combination of the intra-subject variability (inter-occasion variability in the individual PK parameter), analytical error and model misspecification. Just like in an individual compartmental PK analysis, it is the magnitude of ε_{ij} that is minimized. However, minimization is for the population data and not for each individual separately.

An appropriate population analysis not only predicts the results of the subjects that were analyzed but enables the user to make inferences on other populations and future outcomes. Using all population PK parameters and their estimated variability, a typical concentration-time profile can be determined for a specific population. It can also be determined how this profile can vary within individuals. It is this variation in subjects' profiles that enable scientists to make appropriate decisions as to the acceptability of a compound to be developed. If the variation is too big and too many subjects are expected to have sub-efficacious or toxic concentrations, then the drug may be judged unacceptable. This information is crucial to regulatory agencies and pharmaceutical companies to ensure efficient and safe administration of a drug.

Compartmental pharmacokinetic analysis often uses non-linear equations to explain concentration-time profiles. As these equations are non-linear, there are no numerical solutions to the problem. Therefore, to provide solutions to the differential equations, numerous algorithms have been proposed. Some of these include the Livermore Solver for Ordinary Differential equations with Automatic method switching for stiff and nonstiff problems (LSODA) included with ADAPT® and NONMEM® algorithms which are based on work from the Lawrence Livermore Laboratory and modified by the NONMEM Project Group. In the latter, the user has to choose between stiff and nonstiff solutions.

Different methods exist to estimate the population PK parameters and their variances. In addition, new methods are being proposed in the quest to provide the most precise results possible. The following sub-sections will describe some of the methods available to perform population analyses.

1.1.3.1 Standard two-stage (STS)

The two-stage method is so called as it proceeds in two stages. The first step is to analyze all of the subjects individually. The second step is to calculate the population PK parameters (mean PK parameters, their variance and residual variability) directly from the individual results by simply taking the arithmetic mean and variances of the individual results. Using this method, no covariance matrix is obtained. Only the diagonal elements of omega are computed, so it is not possible to determine if there are any correlations that exist between parameters. In addition, the variability of the PK parameters is not a parameter estimated by the model. Therefore, it is not possible to

model the variability of a PK parameter using two terms or to estimate it using a log-normal distribution. Another drawback with this method is that there are no standard errors for the variability estimates. However, this method is easy, straightforward and uses techniques that are understood by most scientists.

This method is available in many software package including both ADAPT 5®[66] and its predecessor, ADAPT-II®[67]. ADAPT 5® was utilized for the STS analyses in this research.

1.1.3.2 Mixed effect modeling approaches

The first mixed effect modeling approach was introduced in the 1970s with the work of Stuart Beal and Lewis Sheiner.[51,53,60,68-71] Most nonlinear mixed effect approaches use maximum likelihood method to estimate the parameters. Different algorithms are available to estimate this maximum likelihood objective function and those used during this Ph.D. analyses are discussed hereafter.

1.1.3.3 NONMEM®

NONMEM®[72] stands for nonlinear mixed effect model. This was the first robust tool globally available for doing population analyses and has since been extensively used. It is often referred to as the gold standard for nonlinear mixed effect analyses and was demonstrated from the beginning by Sheiner and Beal to be superior to the standard two-stage approach.[51,53,65,70,73-77] Starting with the version IV, this tool proposed multiple algorithms besides the original first-order method. This is a first-order (FO) Taylor series

expansion around the mean mixed effects η_i and ε_i. It linearizes the random effects in the PK model. These random effects are independent normally distributed with a zero mean (*i.e.*, distributed around the population value) and a variance matrix Ω. The algorithm will simultaneously obtain estimates of the population parameters θ, population variance Ω as well as the residual variability σ^2. In order to obtain the most likely estimates, this algorithm will minimize an objective function which is the negative of twice the logarithm of the population likelihood as described in Equation 18.

$$-2\mathrm{LL} = \sum_{i=1}^{N} (\log(\det(C_i)) + (y_i - E_i)^T C_i^{-1} (y_i - E_i)) \qquad (18)$$

Where C_i represents the first derivative estimates of the model function with respect to η_i when η_i equal 0

E_i are the model predictions for y_i.

Individual results are obtained in a second step using a *post hoc* Bayesian analysis once the population parameters are estimated.

Although this is the original NONMEM® method and is still being used, it has been shown to have a potential to provide biased estimates especially if the inter-individual variability is large.[78-81] In addition it simplifies the PK parameter distribution to a normal one even when a log normal distribution is assumed in the equations. Therefore, two other algorithms were implemented in NONMEM® and are more efficient and provide less biased estimates. One of these is the first order conditional estimation (FOCE) method. The main difference between the FO and FOCE methods is that FOCE makes the expansion around the individual predicted values of η_i and ε_i and not around the population average predicted value (*i.e.*, zero). It also does not simplify the PK parameter distribution to a normal one. The other method is the

Laplacian method. The FOCE and Laplacian methods are both considered to be excellent and robust methods, but they obviously require more computing power and so the population analyses take longer to run and converge.

1.1.3.4 Iterative two-stage (IT2S®, ITS®)

This was a method derived from the work of Prevost and subsequently Steimer.[76] This approach has also been shown to be superior to the STS approach[76,82-87] at the same time as it was proposed by Steimer. The algorithm derives its results in an opposite manner from NONMEM®. As the name specifies, the calculations are completed in two steps. It first calculates the individual parameters for every patient. For the initial iteration, prior information of the population parameters is required. This can be obtained from the literature, previous studies or using the final results from STS. Individual PK parameters in the first iteration are determined using a maximum likelihood algorithm. All other iterations use a MAP Bayesian approach with a set of prior distribution estimates to determine the individual PK parameters. A Bayesian approach has no constraints on the number of samples each subject may have.

In the second step, population estimates for the population PK parameters, their variances as well as the population residual variability are determined from these newly calculated individual estimates. These population values are then used as prior distribution estimates for a subsequent iteration. This is repeated until the iterations have converged (*i.e.*, iterations are stable and population values fluctuate minimally).

This algorithm was first implemented in a fortran tool called IT2S® by Collins and Forrest and that was built using subroutines from ADAPT-II® release III

(1992).[82,83,88] The latest version of ADAPT-II called ADAPT 5® now directly provides an iterative two-stage algorithm (ITS). This algorithm is similar to the previous algorithm and functions in a similar manner. The main differences are that unlike its predecessor, ITS updates the residual variability parameter estimates automatically at each iteration and convergence is often achieved automatically. In the previous IT2S algorithm, the user had to update the variability parameters at random iterations and had to declare when the iterations were converged.

1.1.3.5 Maximum Likelihood Expectation Maximization

In the latest version of ADAPT-II (eg. ADAPT 5®), a maximum likelihood expectation maximization (MLEM) algorithm is implemented. This algorithm was based on the work done by Dempster, Laird and Rubin in 1977.[89] These authors proposed this algorithm to solve certain problems with maximum likelihood noted with linear mixed effects models. In 1995, Schumitzky used the expectation maximization (EM) algorithm to solve the nonlinear mixed effects maximum likelihood estimation problem.[90] Unlike FO and FOCE, this exact maximum likelihood solution to parametric population modeling does not require the linearization of the nonlinear equations. Instead, Schumitzky suggested the use of sampling-based methods (including importance sampling) to calculate the required integrals and to avoid a linearization approximation. The EM algorithm proceeds in two steps.[91] In the first step (estimation or E step), the conditional mean and covariance for each individual's parameters are estimated using the latest predicted parameter values and the observed data and Monte Carlo sampling. Integrals in the E-step are approximated by using a number of random samples known as

importance sampling, which provides an unbiased estimate of the integral. Basically, a normal density distribution near the mean is taken to reflect the posterior density distribution which may not be normally distributed. Then other normal density distribution samples are taken which are corrected by an importance sampling weight. The average of these samples provides an accurate estimate of the distribution curve. This allows population parameters to converge towards the position of exact maximum likelihood. The second step (Maximization-M step) updates the population mean, covariance and error variance parameters in order to maximize the log-likelihood function in the E-step. The new values are then reused for the subsequent iteration(s).

The proposed importance sampling is a practical solution to the challenging calculations of the conditional mean and covariance matrix required by the EM algorithm.[92-94] It is this EM algorithm with an importance based sampling that has been introduced in ADAPT 5® by Wang.

1.1.3.6 Testing of the new algorithms in ADAPT 5®

As described in Sections 1.1.3.4 and 1.1.3.5, two new algorithms have been implemented in the ADAPT 5® software. They include an iterative two-stage algorithm (ITS) as well as a maximum likelihood expectation maximization (MLEM) algorithm. Different algorithms (new and old) may provide different advantages and disadvantages. As part of the work performed for this thesis, the new methods available in ADAPT 5® were tested versus STS, FOCE, and IT2S to understand their strengths and weaknesses. In order to confidently use the new ITS and MLEM algorithms available in ADAPT 5®, it was necessary to verify if they were adequate to use in a clinical setting and if they

provided accurate (precision and bias)[95-96] results for population PK parameters, variances and residual variability. Therefore, prior to performing compartmental analyses and subsequent clinical trial simulations for this Ph.D., these algorithms were tested to ensure they were adequate and to research when and how they would be best used versus other tools routinely utilized at our laboratory such as NONMEM® version VI and IT2S®. This is critical to know so that the best tools are utilized in order to propose new ways of estimating the PKPD of drugs and optimize their development process. This is the work presented in Chapter 2 (Article 1).

During these analyses, the STS analyses were conducted using the software ADAPT 5®. For these analyses, maximum likelihood was used as no prior information was known for the PK parameters estimated (hypothetical drugs were simulated) and this method is considered as the best option when prior information is inexistent.[56] In order to ensure the best results, this analysis was always carried out in two steps. The first was to estimate the PK parameters for each individual by fixing the residual parameters to a low value (approximately 5%). The second step was to recalculate the individuals' PK parameters; however, the residual variability parameters (both additive and proportional) were simultaneously estimated with the PK parameters and the mean results from the first step was used as prior information for the second step. Using results from the first step allowed a more efficient estimation of the PK parameters and lowered the chances of obtaining results from a local minimum. Final parameter results were simply the mean and variance of the PK parameters and the mean of the residual variability obtained from the individuals.

1.1.3.7 Construction of PK models

With the use of empirical modeling, the first step in compartmental analyses is to build the structural model.[97] The characteristics of the medication as well as the observations collected will dictate the complexity of the model. A simple model, such as a 2-compartment linear model, may be used first to explain the observations. The structural model includes mean PK parameters, inter-subject variability parameters (variances) as well as residual variability parameters. Then, extra parameters are added to the population model in an attempt to better explain the observations. The added parameters can be PK parameters or different variance parameters. The more complicated models are tested to determine whether the additional parameters result in an observable improvement in goodness-of-fit graphs and/or statistically significant improvement in the description of the data. The final model chosen is the simplest model possible to explain the data. The assessment of the significance of the additional parameters is described in the following sections.

1.1.3.8 Objective function

Objective functions in both ADAPT® and NONMEM® are minimized using simplex algorithms. The lowest objective function is wanted. In NONMEM®, the assessment of the statistical significance of the additional parameters is based on the difference between the objective function values between two hierarchical models. In hierarchical or nested models, the more complex model can be reduced to the simpler model by making the additional terms equal 0. In theory, a more complicated model

should explain the observations at least as well as a simpler model. Therefore, to determine if the reduction in the objective function is significant for a more complex model, certain criteria are used. Sheiner had proposed that the difference may follow a chi-squared distribution and a reduction in the objective function of 3.8 and 10.8 would therefore represent a significant difference at levels of 5% and 1%, respectively.[98] This is true if the hierarchical models differ by one degree of freedom (*i.e.*, one parameter). If the model differs by more than one parameter, then the appropriate chi-squared value must be used. However, this method has been criticized by many and appears to overestimate the inclusion of covariates in the model (eg, covariates are too easily viewed as significantly improving the fit). Another test that is older and more conservative would be the Akaike information criterion (AIC).

1.1.3.9 Akaike information criterion

This criterion is built upon a sum of squares such as the objective function, but it imposes a penalty for the number of parameters included in the model.[99-100] This penalty is directly related to the number of parameters used in the model. This criterion is not associated with any p-values. As with the objective function and the sum of squares, this criterion has to be minimized. The AIC is calculated as described in Equation 19.

$$\mathrm{AIC} = 2 * \mathrm{O}_{NLL} + 2 * (p + q) \qquad (19)$$

Where p is the number of PK parameters, q is the number of variance parameters and O_{NLL} is the objective function. Based on this formula, if a more complex model does not significantly lower the objective function, the AIC will be higher due to the penalty

imposed by the increased number of parameters. Therefore, a model is chosen based on the smallest AIC value.

1.1.3.10 Goodness of fit figures

Although graphical goodness of fit is subject to individual's interpretation and is therefore subjective, they can often indicate problems with the model. Concentration-time profiles for the predicted versus the observed concentrations of each individual can be constructed, allowing the modeler to determine if there is a negative or bias trend in the fitting, such as a Cmax always being over or under estimated. A second plot that can be completed is a plot of all observed versus predicted concentrations. The figure should show a uniform spread of the results around the line of identity. If there is a spread of observations that is always below or above the line of identity, then this might mean that the model has difficulty explaining certain concentrations. Other figures prepared are residual by time plots and weighted residuals versus predicted concentrations. These allow the user to determine if there is any trend in how the observations are fitted over time and if there are any outlier concentrations that need to be further investigated. If there is a trend over time in the residuals, it may demonstrate that the PK changes over time due to a saturable or auto-induction process.

1.2. CLINICAL TRIAL SIMULATIONS

It is interesting to note that Levy, Jelliffe, Rodman, Sheiner, D'Argenio and others initially proposed advanced PK/PD analyses and nonlinear mixed effect modeling to improve patient care with drugs that were already commercialized.[57,69,101-112] By the early 1990's, several scientists proposed to expand its use to the drug development field and Carl Peck, Thomas Ludden and other regulators in the early 1990's have been key in this regard to push the population approach at the US FDA.[73,112-118] This eventually led to the development of the first US FDA guidance on population pharmacokinetics published in 1997 (draft) and 1999 (final).[119] Several key players in Europe, North America and other countries met in the late 1990's to push the population approach within the drug development process and resulted in a book and several paper.[117-118,120-124] Finally a key seminal paper published by Lewis Sheiner in 1997 proposed the "learn and confirm" paradigm within the drug development process to explain simply how modeling and simulations had to be done throughout the drug development process and not just within Phase III work.[125] The ultimate goal of modeling as described in Section 1.1 is to gain better knowledge of the drug, reduce development costs as well as decrease the number of volunteers exposed to the drug. This has led to the recent use of clinical trial simulations to optimize drug development.

Clinical trial simulations attempt to integrate important information already gathered to critically assess the known assumptions about the drug as well as the outcomes of a study being designed. This allows drug developers to determine the consequences if changes are made to the drug, the population or the study design. Clinical trial simulations allow the simulation of multiple different scenarios without

having to actually perform the studies, thereby avoiding costly clinical studies and unnecessarily exposing volunteers to the drug. Recent examples of the benefits of clinical trial simulations have been published;[126-130] however, the widespread use of this technique to improve drug development has yet to become mainstream. In most instances, clinical trial simulations are used to explain unexpected results or concerns from regulatory agencies. The consistent use of clinical trial simulations beginning in the early phases of development has the potential to reduce the number of studies that are done by the pharmaceutical industry. The possibility to reduce the number of studies is very real. For example, in the 1990s, there were on average 60 studies submitted in a new drug application or submission (NDA / NDS) and half of these studies were unnecessary as they failed to demonstrate efficacy in a statistically significant manner. The need to better design such studies is obvious. One way to achieve this goal is to simulate the study outcome before dosing and determine if the design and number of subjects meet the aims of the study. An example of a simple question that could be answered with clinical trial simulations is that one might wish to understand whether a drug could be as efficacious if administered once a day instead of twice a day and if a change in daily dose is required. Using data from previous studies, a PK/PD model can be built to incorporate both efficacy and toxicity endpoints. This would permit the developers to see if a once daily regimen could be as efficacious and if Cmax would attain unacceptable levels when the dose is administered as a single dose. The number of simulations and the different clinical scenarios possible are unlimited and may enable scientists to have a better understanding of the consequences of the proposed changes. Once an optimal study design is found based on simulations, then an actual study could

be performed to confirm the expected results of the clinical trial simulations. In order to have the best possible outcomes, it is important that simulations take into account what is known of the product and the uncertainty of the parameters. In order to make optimal use of clinical trial simulations, a multi-disciplinary approach is required. An added benefit is that this forces scientists from different divisions to discuss what information is available and what is still unknown about the drug being developed.

The following sections will describe the different clinical phases of a drug development program (DDP) and set the stage for the research to show how PK/PD modeling and clinical trial simulations could be used throughout these phases to optimize the DDP. Although modeling can be used earlier in the drug development process (pre-clinical), this Ph.D. research focused on the use of pharmacometrics in the clinical phases of development because of the drug products selected. However, we believe that for a new chemical entity the research shown here would need to be started directly from the preclinical stage.

1.3. CLINICAL DRUG DEVELOPMENT

To bring a new drug to market is expensive, time consuming and inefficient. Currently, new drugs take more than 12 years to be developed and cost between 500 million and 2 billion dollars.[131-132] Part of the cost to bring a new medication to the market includes costs for drugs that were unsuccessful and never marketed. In the 1980s and 1990s, the main reasons for drug failure were poor bioavailability, unacceptable pharmacokinetics from poor understanding of drug metabolism and transporters such as ABCB1, lack of efficacy and toxicity. The reasons for late stage failures have since changed with poor PK accounting for less than 10% of failures.[12,133] Pharmaceutical companies made conscious efforts to reduce drug attrition due to poor PK and have succeeded. More specifically, in order to avoid bringing drugs to the clinical stage that did not have a chance of succeeding for PK reasons, researchers focused their efforts on determining the PK of a drug earlier in development. Absorption, distribution, metabolism and elimination (ADME) studies are now performed in the discovery and pre-clinical stages of development process. By focussing earlier research on better understanding the PK properties of drugs, scientists are able to predict how the drug will perform in the clinical stage and avoid surprises. The net benefit is to discontinue poor drugs sooner and have a better understanding of the drug. However, drugs failing in late clinical stages remain an important problem. In addition to being costly for the pharmaceutical companies, the time and effort of scientists that are spent on drugs that fail could be used on other products. The main reasons that contribute to late stage failure have shifted from PK to safety, toxicity and lack of efficacy. Safety and toxicity

attrition rates have even increased. In fact, even when a drug is successfully marketed, it might be removed from the marketplace due to safety and toxicity concerns.

In order to reduce the number of drugs that fail in late stage, we can inspire ourselves from the work performed in the 1990s and early 2000s to reduce drugs that failed due to unacceptable PK characteristics. Similarly, knowledge of the drug's efficacy and safety is required earlier in drug development in order to understand the true potential of a compound before conducting expensive multi-center Phase III studies. We propose using modeling and simulations within Phase I instead of the later phases in order to optimize the drug development process and render it more efficient. We believe that by using modeling and simulations and performing clinical trial simulations as early as possible, the drug's dose-response and dose-safety characteristics can be better understood. This will lead to better designed studies, optimized dosing, fewer required studies, fewer patients who are exposed to unsafe or inefficient drugs and to a more efficient drug development process.

The following sections will describe the different clinical phases of a drug development program and set the stage for this research to show how PK/PD modeling and clinical trial simulations could be used throughout these phases to optimize the DDP. Although modeling can be used earlier in the drug development process (pre-clinical), this research focused on the use of pharmacometrics in the clinical phases of development because of the drug products selected. However, we believe that for a new chemical entity the research shown here would need to be started directly from the preclinical stage.

1.3.1. Phase I

These studies are the first time a drug is administered to humans. The typical mindset in drug development is that the purpose of studies performed in Phase I is to understand the tolerated doses, the adverse effects associated with increasing doses, the pharmacokinetics of the drug, more specifically the absorption, distribution, metabolism and elimination, and to gain initial knowledge of drug-exposure in humans. However, efficacy needs to be a clear motivator in Phase I if we are to optimize the drug development process. Any early evidence of effectiveness should be collected as this will increase the knowledge of the drug earlier in the process and improve clinical trial simulations to be performed.

The first study dosed in humans is known as the First Time in Human study (FIH).[134] Phase I studies will normally enroll a small number of volunteers. The drug is more often than not administered to healthy volunteers unless the anticipated adverse effects of the drugs are severe or if the drug is intended to be administered for life-threatening conditions. An example of a drug that would not be administered to healthy volunteers in a Phase I study would be an anticancer drug.

The initial drug dose administered to the volunteers in Phase I will depend on the information collected in the pre-clinical studies and may be chosen empirically or with the help of PK/PD modeling performed on the animal data collected.[135-138] However, any initial dose chosen will incorporate a safety factor to avoid any serious or permanent adverse effects. Therefore, studies in Phase I often incorporate a dose escalation process where the doses are increased in each subset of volunteers until the maximum tolerated

dosed is attained. This is performed in single dose and multiple ascending dose studies in order to have a better understanding of the drug's PK. It is in this stage, with the different strengths of the dose administered as well the single and multiple administrations, that it could be determined if a drug is linear and dose proportional. Other studies that might be conducted in Phase I include drug-drug interaction studies, food-effect studies and special population studies. Drug-drug interaction studies will verify the *in-vivo* impact of potential inhibitors or inducers that might have been demonstrated *in-vitro* or with drugs that are likely to be co-administered with the new drug. Special population studies often include a renal impaired study or a hepatic study. Another study that might be performed in Phase I is a radioactive study in which a small radioactive dose is administered in order to determine the metabolic pathways.

The design for Phase I studies is such that a wealth of information is collected. Blood or plasma samples are collected from most volunteers, if not all of them, using a rich sampling design with blood draws collected at set times throughout the collection period. Certain Phase I studies also collect other samples such as urine in order to determine with greater precision the pharmacokinetics of the drug. Samples are measured for the parent analyte and any major active metabolites.

To continue to later phases of development, scientists use the data collected in Phase I to determine if the drug had acceptable pharmacokinetic and safety profiles, if the expected efficacy concentrations were attained in the dose escalation studies and if the PK parameters show tolerable variability to avoid having volunteers with excessively high or low drug concentrations. Often, the maximum tolerated dose from Phase I is chosen to be administered in Phase II. This is inefficient and often leads to the selection

of a dose that is not optimal. The data should be used to determine a dose-response relationship in order to estimate the desired exposure in the target population and to choose the optimal dose to be developed. Thus, the use of PK/PD modeling is crucial in Phase I. Modeling in Phase I has the potential to reduce the number of subsequent studies required. Initial modeling should be developed linking drug exposure to response (safety and efficacy, if collected) using the different strengths of drug administered, the different sub-populations and some covariates such as gender, age, weight and others. This modeling permits a better dose selection for subsequent development of the drug.

1.3.2. Phase II

Phase II studies are normally conducted in patients. Unless the drug tested required patients in Phase I, studies in Phase II will be the first time the drug is administered to patients for whom the drug is indicated. Phase II studies are usually performed in a larger number of volunteers than in Phase I; however, the number of subjects remains limited to a maximum of a few hundred. As Phase II studies are usually the first studies conducted in patients, inclusion and exclusion criteria are usually very strict in order to limit external factors that would influence results and render interpretations harder to explain. The purposes of Phase II studies are to determine if the drug shows efficacy in the target population and to continue to collect safety data in a greater number of volunteers. Phase II is often divided into two different sections. A proof of concept (POC) study is performed in a small number of subjects and is considered a Phase IIa study. Unfortunately, the POC study often consists of

administering the maximum tolerated dose and determining if the effect is better than the placebo. This mindset should be avoided if at all possible as this will often lead to use of the wrong dose in subsequent studies and often lead to inacceptable toxicity. Instead, an optimal dose or set of doses could be established based on the data collected from Phase I studies and a PK/PD dose-response model could be developed. This would reduce adverse effects, maximize the possibility of success and avoid having to change the dose when administered in large Phase III studies. The proof of concept study is often a crucial step in the drug development process and is often used to decide if the development of the drug should be continued. Therefore, to increase drug development efficacy, it is imperative that the interpretation of the results at this step is done with the utmost precision. The purpose of Phase IIb studies is to repeat what was determined in the Phase IIa study with a larger set of volunteers. This study will also collect valuable efficacy and safety data.

In Phase IIa studies, there might still be a relatively large number of PK samples collected. However, the focus will be on the active analyte which is usually collected in blood or plasma. In addition to these PK samples, multiple efficacy measurements may be taken in order to construct a proper PK/PD model. Efficacy measurements can be taken directly from laboratory results (such as bacteria growth to test if an antibiotic is really effective or not) or they can be surrogate biomarkers when discussing long term diseases such as the LDL for cholesterol. The use of these PK and PD samples are of great help to build the structural PK/PD model. Phase IIb studies will usually have less PK samples collected, often using a sparse sampling design. However, PD data should

be collected often if the drug acts quickly and has a short duration of action. If changes in PD data are expected to be slow, then data can be collected more sporadically.

Prior to advancing to the next phase of development, scientists will use all of the data collected in Phases I and II. At this stage, the objective is to determine if the drug should go to Phase III where studies are expensive to conduct. If the drug is being developed for a new indication, it has to be safe and efficacious. However, if the drug is for an indication where there is already an acceptable treatment, then the new drug has to not only be safe and efficacious but also has to have an advantage over drugs already available. Modeling will help understand the potential of the drug at this stage and determine if Phase III studies are warranted. Data from Phase I and Phase II can be combined to increase the available data. At this stage, modeling can be performed to confirm results from the Phase I model and refine this model based on the collected data from Phase II. Ultimately, the objective is to determine the optimal dose for Phase III. Ideally, this dose will provide maximal efficacy while reducing expected adverse effects. For a drug competing with already accepted treatments, clinical trial simulations can determine if it is superior to the accepted treatments and if development is worth continuing.

1.3.3. Phase III

Phase III studies are pivotal for submission to regulatory agencies. The purpose of the Phase III studies is to prove that the drug is safe and efficacious in the target population. They usually consist of two studies performed at multiple clinical sites and include many patients, often in the thousands. Inclusion and exclusion criteria are often less stringent than Phase II studies which permits the drug to be administered under conditions that have not been tested until this phase such as patients with other conditions or diseases, and patients taking other drugs which could influence the PK of the drug being tested. The drug will most often be administered over a long period of time and adherence to the regimen is often not perfect.

Safety will be judged based on the drug's intended indication and different safety profiles will be acceptable depending on the benefits of the drug. If a drug is already on the market for the same indication, then the pivotal Phase III efficacy study will include the approved marketed drug. This is done in order to determine if the new drug is superior, has a better safety profile or is as safe and efficacious but with something novel such as a different route of administration (eg. oral formulation instead of a suppository). If no drug exists for the targeted indication, then the Phase III studies will contain a placebo as a control for the interpretation of the results. Typically, the drug used in Phase III trials will be the formulation that will be marketed. Information such as precautions that should be taken, contraindications and adverse effect frequency will be collected during Phase III studies and used for the labeling.

For these studies, it is recommended to still collect sparse blood samples and PD variables. This data may be useful to support the registration or help explain Phase III results that might differ from the previous phases. The cost of these studies in patients is extremely high and the cost of sparse sampling is minimal compared to the wealth of information it can provide, especially if the regulatory agencies demand explanation of some results. Contrary to the previous phases, the data collected in Phase III will be very sparse, will not be identical between subjects and often have a greater variability since the studies are being carried out in different sites with different personnel. In addition, the population being studied is less homogeneous.

The focus of modeling after Phase III serves a slightly different purpose than the modeling performed after Phases I and II. The main reason for modeling at this stage will be to confirm results and answer lingering questions. It will often bridge results from the different phases of development to show consistency between results. It will use the different covariates collected to establish possible dose adjustments for sub-populations. In addition, modeling and simulation can be used to respond to regulatory agencies' concerns that they might express over some results. For instance, they can be used to explain results that may at first appear unexpected.

1.3.4. Phase IV

If a Phase IV study is required, it occurs once the drug has been accepted by regulatory agencies and is marketed. Often, these studies are demanded by the regulatory agencies in order to improve the drug's safety profile, explore the treatment of other

conditions, confirm or add to the label claims or even determine its effect in special populations such as pediatrics. No matter how many volunteers and patients receive the drug during its development, this can never compare to the number of people who will take the drug once it is marketed. In addition, for drugs intended for chronic use, the effects of long time use are often unknown before a drug is available on the market. Rare complications or adverse events of the drug such as those that occur once in every 10000 or 100000 patients (ex: terfenadine, celecoxib or thioridazine) cannot be detected during the development process. The same can be true of carcinogenic products where effects can take many years to manifest. These will be detected with the everyday use of the medication in the population at large. These rare or long term use effects may change the risk/benefit ratio of a drug and regulatory agencies may need to determine if the drug is still acceptable for its marketed use.

Phase III study designs are usually controlled, randomized and double-blinded. However, designs for Phase IV studies will depend on the purpose of the study. It is possible to have a similar design as a Phase III study. However, other study designs may include a historical retrospective design, a longitudinal observational design, the use of a large computerized database, drug-drug interaction study, etc. Samples may or may not be collected depending on the actual design of the study. The application of modeling and simulation in Phase IV is not as well documented as with other phases of development.

1.3.5. Generic Drug Development

A generic drug product is a drug which is considered similar or identical to the brand (reference) product. A generic drug must contain the same active ingredient(s) as the original formulation. Generally, a generic drug is considered therapeutically equivalent to the original formulation if it has the same active ingredient(s), strength, dosage form, route of administration, quality, performance characteristics and therapeutic indications. Generic products can only be made available to the public once the patents of the original developer have expired. When generic products become available, the market competition often leads to substantially lower prices for both the original brand name product and the generic forms.

In order to meet the criteria established to consider therapeutic equivalence, most generic drug manufacturers will perform a bioequivalence study. According to the FDA, bioequivalence is "the absence of a significant difference in the rate and extent to which the active ingredient or active moiety in pharmaceutical equivalents or pharmaceutical alternatives becomes available at the site of drug action when administered at the same molar dose under similar conditions in an appropriately designed study".[135] Based on this definition, the relative bioavailability of the two compounds has to be similar in terms of rate and extent of exposure. Therefore, the two products are compared in terms of formulation performance. If the release and relative bioavailability of two products are equivalent, then the concentration-time profile between the two products should be equivalent. The focus of bioequivalence is to determine if two compounds have similar concentration-time profiles because if two products have similar systemic profiles, it is

assumed that they will have similar concentrations reaching the site of action either for efficacy or for adverse effects. It is believed that the greatest difference that can exist between two different formulations of the same active ingredient rests on how it is absorbed. The difference does not arise from the actual active ingredient as the two formulations are supposed to have very similar ingredients and *in-vitro* dissolution profiles. However, the two formulations will have different non-active ingredients which will influence the *in-vivo* release of the active ingredient and absorption parameters. The inter-occasion variability in the elimination process of a volunteer is considered negligible over short periods of time (*i.e.*, when the medical status of the volunteer does not change).

The most common study design to assess bioequivalence is a 2-way crossover study in which the parent drug concentrations are measured systemically in a biological fluid such as plasma or blood. Crossover studies are not always feasible, so it is possible that a parallel study is required to prove bioequivalence. However, a crossover study design is preferred as the two different formulations are administered to the same subjects after a suitable washout period. The highest formulated dose is normally used to perform bioequivalence studies. Unless there is an ethical issue, healthy volunteers are enrolled in these studies. Following drug administration, a concentration versus time curve is obtained for the two formulations. Typically, in a bioequivalence study, a noncompartmental approach is used to determine the PK parameters. The major parameters of interest are the observed Cmax, AUC0-t and AUCinf. These parameters are obtained directly from the concentration versus time curve for each formulation as depicted in Figure II. Other parameters of interest are Tmax, Kel and half-life.

Bioequivalence is declared to be met most of the time only when Cmax and AUC parameters meet the pre-defined criteria.

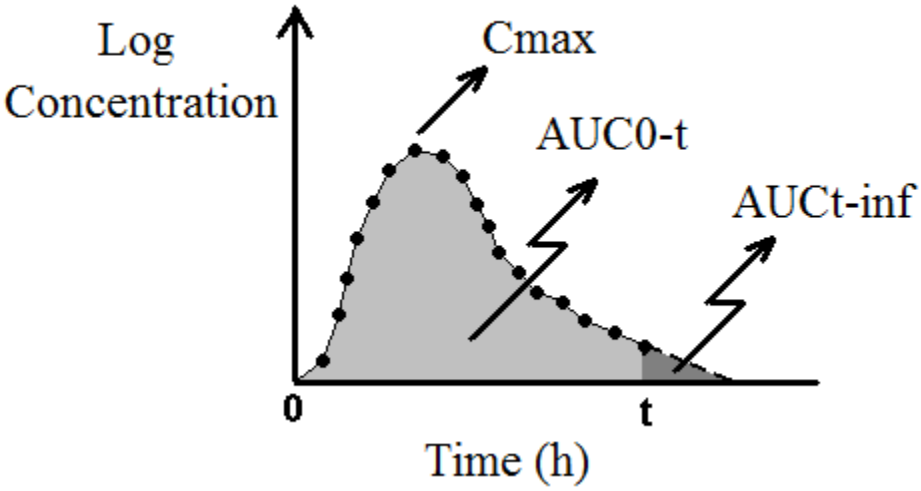


Figure II: Noncompartmental PK parameter determination

Bioequivalence studies are not only performed for generic compounds. Many bioequivalence studies are required by innovator companies in order to prove equivalence between formulations undergoing changes. For example, the formulation that is marketed by an innovator company is rarely, if ever, the same as the formulation that was used with the first studies in Phase I. In order to bridge the information gap between studies, bioequivalence studies are performed on the different formulations administered during the development process. In addition, any significant changes in production procedures or the relocation of a manufacturing plant entails bioequivalence studies. Regulatory agencies have to be assured that these changes do not affect the PK parameters of the drug. This is true for both innovator and generic companies.

Although bioequivalence based on systemic concentrations is always preferred, there are instances where this is not possible. One example is when systemic

concentrations of the parent or active compound cannot be measured because a sensitive analytical method does not exist. Another example is studies in which doses are administered locally, providing local responses, and systemic concentrations are not reflective of the concentrations at the site of action. These include studies involving topical cream for dermatological diseases, nasal sprays used for local response to allergies, ophthalmic drugs or even oral drugs not absorbed but acting locally in the gastrointestinal tract. Therefore, in these instances, the best approach to determine bioequivalence is to use pharmacodynamic endpoints or clinical endpoints. The choice of endpoints is crucial in demonstrating bioequivalence and should reflect the label indication of the product. These studies are usually blinded, may contain a placebo group, and are more variable than comparative studies based on systemic concentrations, which is often why more subjects are enrolled. In these studies, the dose which is chosen is important. For example, the onset, duration and strength of the PD response depend on the dose that is administered. In many, if not most instances, the response will follow a maximum effect (Emax) curve. The dose that is selected has to be in the linear portion of the Emax curve. Usually, the dose providing 50% of maximal response (ED_{50}) is the best choice and should be used in the study. That way, if the formulations are not bioequivalent, a significant change in response between the two formulations can be measured and distinguished. If doses are chosen in the upper region of the Emax curve, any difference in dose between formulations would only give rise to a small change in response. Therefore, it becomes practically impossible to distinguish two formulations that are not bioequivalent and the PD study is inappropriate. This is depicted in Figure III. Obviously, in order to perform such a study, knowledge of the Emax model for a

particular drug has to be known. If not, then using a low, middle and high dose of the drug plus a placebo should be used to build the Emax model and test bioequivalence. This approach is one that is recommended by the regulatory agencies (FDA) and has been used successfully on numerous occasions.[139-142]

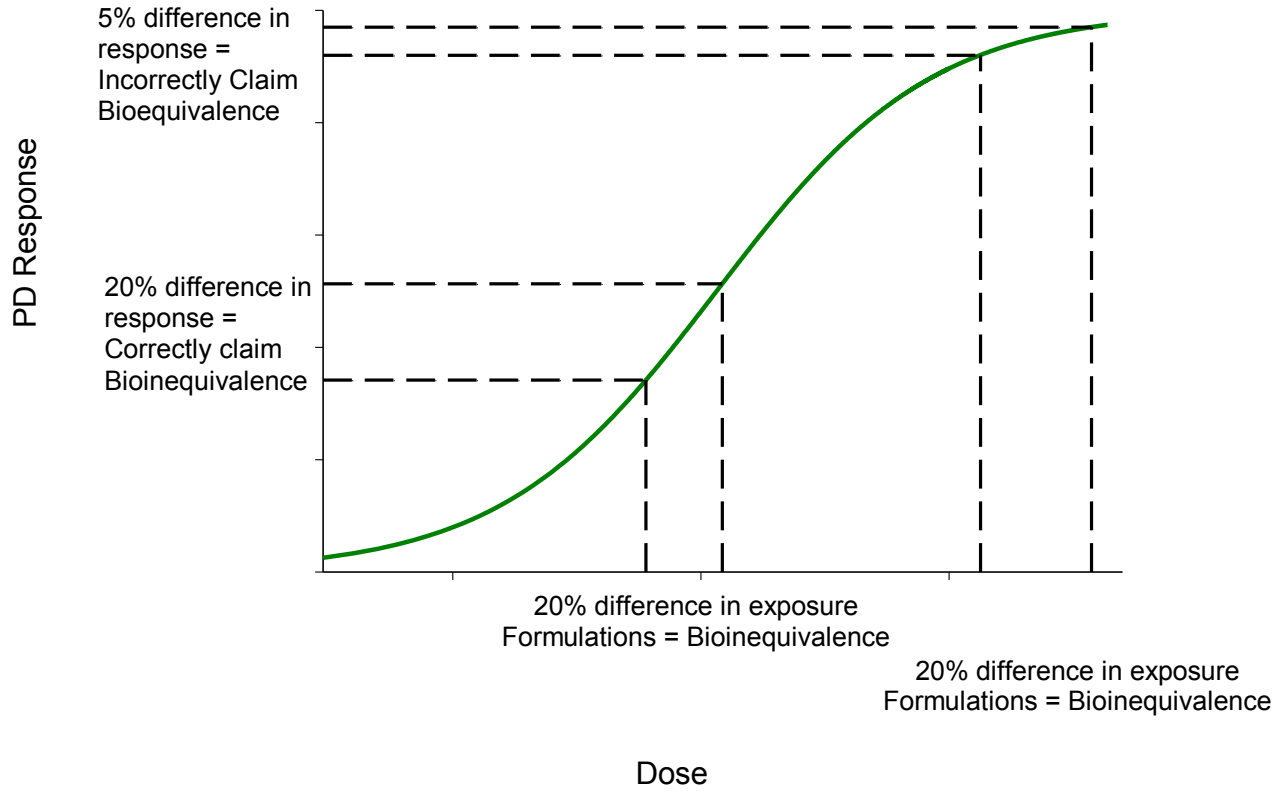


Figure III: Pharmacodynamic response versus dose on a log scale

In order to demonstrate bioequivalence, statistical tests have to be performed. These tests aim to show that the rate and extent of exposure for the compared formulations are equivalent. In a typical 2-way crossover, two-treatment study design, the statistical test performed is called a two one-sided test which was proposed by Schuirmann in the 1980s.[143] Basically, the average bioequivalence of the two products will be compared to show that the test formulation's rate and extent of absorption is equivalent with respect to the reference formulation. In order to be able to perform such a test, an acceptable equivalence limit has to be defined. The consensus is that for most drugs, a 20% difference in drug concentrations does not produce a significant change in clinical effects. Therefore, an analysis of variance (ANOVA) is performed on the test/reference ratio of the average Cmax and AUC parameters. The 90% confidence intervals (CI) for the ratios must be within 20%. Previously, non-transformed data was used and the ratio had to be between 80 and 120%. However, since the 1990s, the log-transformed Cmax and AUC data have been used and the 90% CI must be within 80 and 125%. This means that the full range of the 90% CI has to be within 80-125% in order for two products to be declared bioequivalent. As the 90% CI are constructed around the geometric mean ratio, the test to reference ratio will have to be close to 100% for the test formulation to be bioequivalent to the reference one using a "manageable" number of subjects. In a crossover study design, the 90% confidence intervals are dependent on the intra-subject variability and the range of the interval reflects the within-subject variability. If a parallel study design is used, the CI reflects the between-subject variability which will always be greater than within-subject variability. Therefore,

parallel studies will always require more subjects than a crossover design study. Figure IV shows the different ways in which a study can pass or fail.

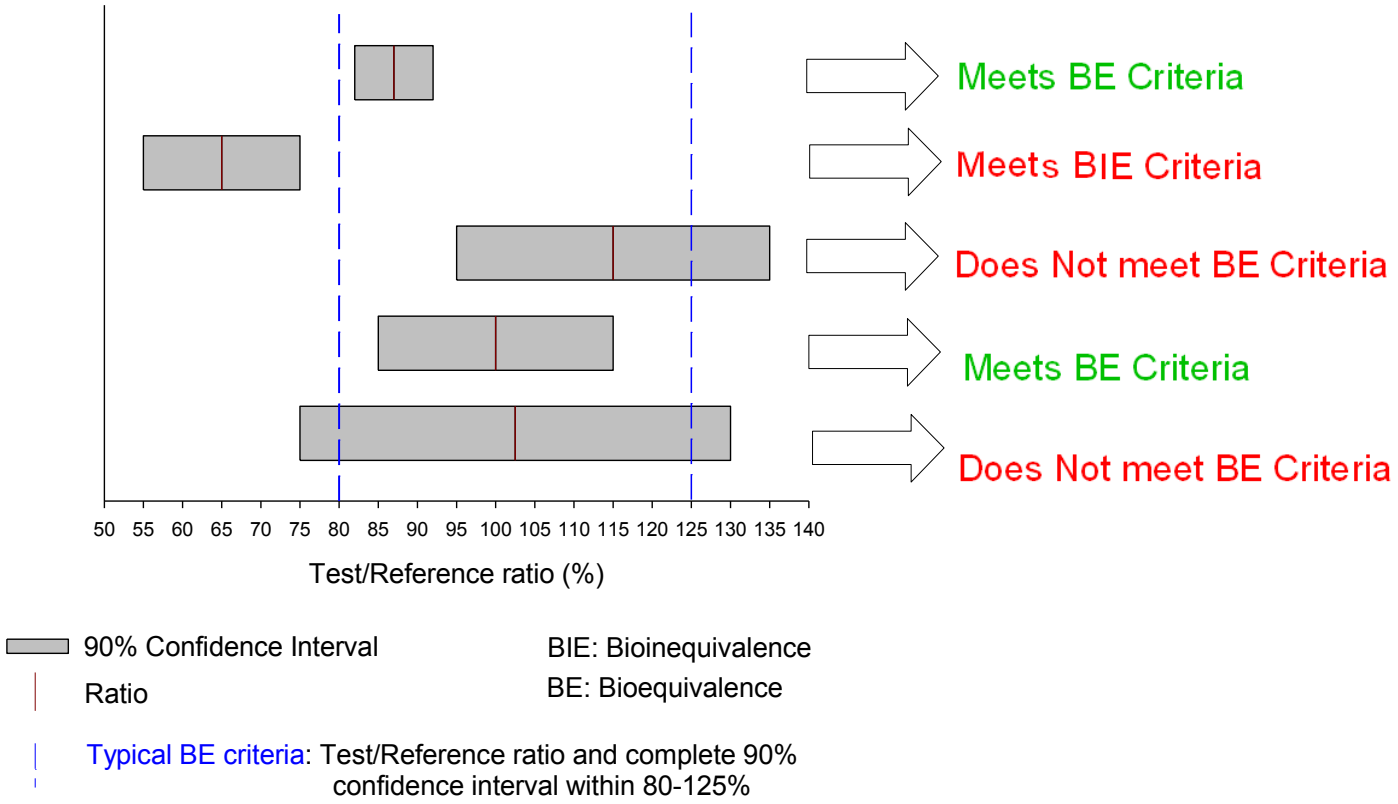


Figure IV: Possible bioequivalence results

Data are log-transformed prior to performing the ANOVA in order to follow a normal distribution, an important assumption for the statistical tests. The 90% CI are used as this allows a 5% alpha error on both sides of the intervals. When designing studies, the Type I error is set at 5% which means that there is a 5% chance that two formulations will be declared bioequivalent when in reality they are not. Type II error, which can be translated as the chance that two formulations are not declared

bioequivalent when they really are, is set by the manufacturer although it needs to be a maximum of 20% (1- 80% power). There is a price to pay to lower the Type II error. The lower a manufacturer sets this error, the greater the number of subjects required to perform the study. In other words, in order to decrease the risk of failing a study, a greater number of subjects is required to run the study.

Bioequivalence testing is the opposite of the classical statistical test to show that there are no differences. In classical statistics, H_0 is $\mu_1 = \mu_2$ and H_A is $\mu_1 \neq \mu_2$. For example, if you want to prove that a drug is better than the placebo, this would be the statistical alternate hypothesis to the null hypothesis of the study. In such a study design, enough subjects would be included in the study to try and prove a difference. Therefore, in classical statistics one aims to show that two products are different. This is accomplished by designing the study to prove that the two compounds are the same and if this hypothesis is rejected, the alternative hypothesis is that the two products are different. However, in bioequivalence testing, this would not be acceptable. The aim of the study is not to prove that the two products are the same but that the two products do not differ significantly one from the other. If the classical hypothesis testing were to be used, it would be easy to show two products are similar. All one would have to do is make sure there are insufficient subjects (*i.e.*, power) to show a difference and it would be concluded that the two products are bioequivalent. Thus, in a bioequivalence study, the hypothesis testing is done differently where H_0 is $\mu_1 \neq \mu_2$ and H_A is $\mu_1 = \mu_2$. Therefore, enough subjects must be included in the study to show that the two products are equivalent.

Summary

There is an urgent need to improve the drug development process. We believe that Clinical trial simulations and advanced modeling and simulations are one way to respond to this need. The potential would be there to prepare better study designs, lower the number of unnecessary studies and individuals receiving the drug, better predict study outcomes and improve knowledge of the drug being tested. This should be true for both innovator and generic drug compounds. The objective of this thesis is therefore to propose ways to hopefully improve the development success of drugs by using advanced modeling and simulation techniques. To this end, three articles are presented. The first presents the results of the assessment of new tools available to perform population compartmental PK analyses, which are the foundation of proper PK/PD characterization of drugs and a better understanding of their safety and efficacy. The second article presents a novel technique to ensure that the PK of drugs is correctly characterized in Phase I studies by focusing on the better characterization of the terminal half life, a key parameter used to derive other key parameters such as the total exposure (AUCinf) of a drug. Without the correct characterization of the PK of a drug, its development process cannot be optimized. The last article provides a concrete example of how modeling and simulations can be used to simulate clinical trials and how we proposed and used it successfully to improve and optimize the development process of an innovator drug.

CHAPTER II: RESEARCH ARTICLES

1. ARTICLE #1

Performance of different population pharmacokinetic algorithms using clinical simulations

1.1. INTRODUCTION

The main objective for the first part of this research project was to determine if the new algorithms available in ADAPT 5® are adequate for population compartmental analyses. Compartmental population analysis has been evolving since its conception over 40 years ago. New tools are available and researchers require the knowledge of their strengths and weaknesses in order to understand when their use is appropriate.

Population analyses are useful as they not only permit to predict the individual's data but allow quantification of the inter-individual and intra-individual variability. This information is crucial to predict results for future subjects receiving the drug as well as predicting outcomes of studies. However, an improper estimation of the variability could lead to inaccurate predictions and wrong outcome conclusions. It was therefore important to test if the two new methods available in ADAPT 5® (ITS and MLEM) would provide more accurate results of the variability estimates than the algorithms often used in the past. A better accuracy in these estimates would provide greater confidence in predictions.

The two new ADAPT 5® non-linear mixed effect algorithms are described in Sections 1.1.3.4 and 1.1.3.5 on pages 27-29. The following article compared these methods to established ones. They are first compared to a standard two-stage method as this method is considered to be less precise in the estimates of variability parameters than mixed effect methods. Therefore, these new methods are expected to provide better results than the STS method. They are also compared to an iterative two-stage method (IT2S) available in IT2S® and FOCE in NONMEM®. The iterative two stage algorithm

available in ADAPT 5® is expected to perform better than the previous IT2S method due to the changes in the residual variability calculation as discussed on page 28 of Section 1.1.3.4. Unlike FOCE which approximates the maximum likelihood estimates, the new MLEM algorithm offers an exact estimation of the maximum likelihood estimates. Although similar MLEM algorithms already exist, these were dependent on Monte Carlo parametric maximization. MLEM algorithms based on importance sampling like the one offered in ADAPT 5® has only been offered recently. It is unclear if this method would perform differently than the FOCE method which is known to be an excellent population algorithm.

Thus, the following article presents the results of the comparisons between the previously described algorithms.

1.2. ARTICLE

Title

Performance of different population pharmacokinetic algorithms using clinical simulations

Authors: Philippe Colucci, M.Sc.[1,2]; Julie Grenier, PhD[3]; Corinne Seng Yue, M.Sc.[1,2]; Jacques Turgeon, PhD[1,4]; and Murray P. Ducharme; PharmD[1,2]

Institutional affiliation:

1) Faculté de Pharmacie, University of Montreal, Montreal, Canada
2) Cetero Research, Cary, North Carolina
3) Celerion, Montreal, Canada
4) CRCHUM, Centre Hospitalier de l'Université de Montréal, Montreal, Canada

Key Words: NONMEM, ADAPT 5, IT2S, ITS, MLEM, Shrinkage

Conflict of Interest: There are no personal financial or non-financial conflicts of interest

Abstract

Background There has been an increased focus on population pharmacokinetics (PK) by pharmaceutical companies since the FDA documented the essential need to improve the drug development process in their "Critical Path paper". The increased interest in population PK analyses has given rise to additional algorithms to perform these analyses.

Objectives The main purpose of this exercise was to compare the new algorithms Iterative Two Stage (ITS) and Maximum Likelihood Expectation Maximization (MLEM) available in ADAPT 5® with other methods established in the industry. A secondary purpose was to determine if the results were impacted when reducing the number of samples taken in the expectation maximization step of the MLEM algorithm.

Methods A total of 29 clinical trials were simulated using different types of drug administration, number of subjects, number of samples per subject and number of PK parameters. Drug concentrations were simulated in plasma or plasma and urine. Using the simulated concentrations, different algorithms were used to fit the data and estimated parameters were compared to the true values. The algorithms ITS and MLEM were compared to Standard-Two-Stage (STS) analyses and to other nonlinear mixed effect modeling approaches including Iterative-2-Stage (IT2S) method in the IT2S® package and First Order Conditional Estimate (FOCE) method in NONMEM Version VI®. Precision and bias in the population PK parameters and their variances as well as the individual PK parameter estimates were used to compare the different methods. The scientist performing the fitting was blinded to the true values.

Results Population PK parameters were well estimated and bias was low for all nonlinear mixed effect modeling approaches. These approaches were superior to the STS analyses.

The MLEM algorithm was better than IT2S and ITS to predict the PK and variability parameters. Residual variability was better estimated using MLEM and FOCE. A difference in the estimation of the variance exists between FOCE and the other methods. Variances estimated with FOCE often had shrinkage issues while MLEM in ADAPT 5® had practically no shrinkage problems. Using MLEM, a reduction from 3000 to 1000 samples in the expectation maximization (EM) step had no impact on the results.

Conclusion The new algorithm MLEM in ADAPT 5® was consistently better than IT2S and ITS in its prediction of PK parameters, their variances as well as the residual variability. It was comparable to the FOCE method available in NONMEM with fewer shrinkage issues in the estimation of the variances. The number of samples used in the EM step with MLEM did not influence the results.

Introduction

Pharmacokinetic (PK) modeling and simulations have been part of drug development for decades.[1] However, it is only recently that its use has become widespread in an attempt to improve the process and reduce the cost associated with bringing a new drug to market. Two papers have helped bring a conscious effort to improve and strengthen model-based drug development. The first, by Sheiner[2], discussed the use of multiple cycles of predictive and confirmatory PK modeling (learn and confirm) throughout the drug program while the second, a "Critical Path paper" published by the Food and Drug Administration (FDA)[3], emphasized the importance of improving the process in order to expedite the development of new drugs and reduce associated costs. The consequence has been to increase the use of population PK analyses over the years where practically every new drug submitted to the FDA has population PK analyses included in the submission.[4] The increased use of population PK analyses has given rise to new tools and algorithms to perform these analyses. In order to make the best use of these new algorithms, one has to better understand their strengths and limitations.

One new tool that was made available in 2009 was ADAPT 5®.[5] This is an updated version of the ADAPT-II® software created by the Biomedical Simulations Resource (BMSR) at the University of Southern California. Within this software, two new nonlinear mixed effect algorithms were implemented. The first is an iterative two-stage (ITS) algorithm and the second is a maximum likelihood expectation maximization (MLEM) algorithm.

Objectives

The main purpose of this exercise was to compare the new algorithms ITS and MLEM available in ADAPT 5® with other methods established in the industry, such as the standard two-stage (STS), the iterative two-stage (IT2S) in the IT2S® program[6] and the First Order Conditional Estimate (FOCE) in NONMEM version VI®.[7] Different study designs were used to determine if any of the algorithms performed differently under varying conditions. A secondary purpose was to determine if the results were impacted when reducing the number of samples taken in the expectation maximization step of the MLEM algorithm of ADAPT 5®.

Hypotheses

Based on our previous work with all of these methods, it was hypothesized that the new available algorithms in ADAPT 5® would be better than STS and IT2S. We believed that ITS would be better than IT2S for population PK parameters, variances and residual variability because residual variability parameters with ITS are updated automatically with every population iteration steps whereas with IT2S they are updated manually by the scientist at random iterations determined by the modeler. Based on the literature and our previous experience, it was expected that MLEM would be a good alternative to the FOCE method.

Methods

Algorithms

The new ADAPT 5® algorithms were compared against different approaches of obtaining population parameters. It was compared to STS, IT2S and FOCE methods.

The first approach used as a reference was the STS approach. In this approach, PK parameters as well as the unexplained variability (residual variability) are determined for each subject individually and the population values are directly obtained from the individual results by simply taking the mean and variances of the individual results. This method served as a control as it was expected that mixed effect population analyses would procure more precise and less biased results. Indeed, it has been already documented by many that a mixed effect population analysis is superior to STS.[8-13] The individual analyses methods implemented in ADAPT 5® are identical to the ones found in the previous version, ADAPT-II® release IV. Some of the important new features in ADAPT 5® are the population algorithms and the automation of the creation of tables listing the STS results.

The first new algorithm tested from ADAPT 5® was ITS. This iterative two-stage algorithm is a modification of the algorithm as implemented in IT2S®. This approach is based on the work done by Prevost and Steimer and implemented by Collins and Forrest in IT2S® program.[6,14] In both ITS and IT2S, the first iteration uses a standard STS approach as previously described. However, multiple individual and population iterations are performed after this first iteration using a maximum a posteriori (MAP) analysis. More specifically, the average PK parameters and their variances from the previous iteration are used as new prior information in a subsequent iteration. These updated

iterations continue until stable population results are obtained and convergence is achieved. The main difference between the two algorithms is how the residual variability is obtained. In the previous IT2S® program, the residual variability(ies) was(were) updated periodically by the user until it was(were) found to be stable. In contrast to the IT2S analyses, the ITS algorithm in ADAPT 5® updates the residual variability parameters automatically at every population iteration using maximum likelihood. In addition, ITS stops its iterations (*i.e.*, converges) when all subjects' results are converged and no additional iterations are required whereas the user had to declare convergence manually in IT2S®.

The second new algorithm available in ADAPT 5® is a new MLEM approach. A MLEM algorithm was first proposed by Schumitzky and Walker in the mid 1990s[15] based on work published by Dempster, Laird and Rubin in the 1970s.[16] Unlike other parametric algorithm methods such as FOCE in NONMEM® which use an approximate maximum likelihood to solve the nonlinear problem, MLEM uses an exact maximum likelihood solution. For all methods, maximum likelihood (ML) estimates model parameters that are most likely to explain the observed data. In the MLEM algorithm, ML is combined with an expectation maximization (EM) algorithm. The EM algorithm consists of two steps, the E-step and the M-step. In the E-step, parameter variables are estimated using the latest predicted parameter values and the observed data. In the M-step, parameter values are updated in order to maximize the log-likelihood function in the E-step. These new values are then reused for the subsequent iteration. To avoid a linearization approximation, an importance sampling-based estimation is proposed. Integrals in the E-step are approximated by using a number of random samples known as

importance sampling, which provides an unbiased estimate of the integral. In ADAPT 5®, the number of random samples can be set from 1000 to 3000. Similarly to the IT2S algorithm, convergence using this algorithm in ADAPT 5® has to be determined by the user.

Finally, the last mixed effect algorithm evaluated was the FOCE algorithm. NONMEM® and its available algorithms are based on the work by Beal and Sheiner.[7] FOCE uses a different approach than the previous algorithms discussed. Unlike the previous approaches, the subject's individual results are obtained in a *post hoc* step using a MAP-Bayesian algorithm once the population parameter values are obtained. In order to obtain these population values, all individual data are simultaneously fitted. An approximate maximum likelihood method is employed and FOCE uses a first-order Taylor expansion of the nonlinear model around the variance estimators (known as η in NONMEM®).

Simulation of different studies

To replicate the different types of clinical data one typically has to analyze, a total of 29 studies were simulated. These simulated studies varied in the number of subjects from low to high (n = 6 to 200), different sampling strategy (sparse vs. rich), route of drug administration (oral and intravenous), number of compartments to describe the PK (1, 2 and 3), and the matrix in which the concentrations were simulated (plasma with or without urine). For simplicity purposes, population PK parameters and their variances were assumed to be normally distributed. The simulated variances corresponded to an

inter-individual coefficient of variation (CV%) for PK parameters ranging from 7 to 250%.

Studies with sparse and rich sampling were simulated. When there were fewer samples than the number of parameters to be estimated, the sampling was considered to be sparse. These studies contained between 2 and 5 plasma samples and between 2 and 3 urine samples. This corresponded to a ratio of concentrations to parameters of 0.6 to 0.9. Sparse sampling studies either had a fixed sampling scheme for each subject or different collection time points for the subjects. In either case, samples were collected so that all phases of the PK profile could be characterized. Simulated studies having a rich sampling had a ratio of concentrations to parameters ranging between 1.6 and 5.7. Again for simplicity purposes, absorption and elimination followed first order linear processes. Table 1 summarizes the different study designs used for the analyses.

Simulations were performed by an independent scientist who did not participate in the fitting of the data. The studies were simulated using ADAPT-II® release IV. Concentrations were directly simulated using a known residual variability that was different for each study. All residual errors were simulated using a proportional and additive error model.

Fitting of the simulated data

The scientist performing the fitting of the data was blinded to the true values of the parameters, however was informed of the structural model and the distribution of the PK parameters (*i.e.*, normal distribution). In addition, a proportional and additive error

model was used to estimate all residual variabilities. This was done to avoid comparing results obtained from different structural models.

The first step in the fitting process was to obtain results using an STS approach with ADAPT 5®. No prior knowledge of the parameter values was provided. Therefore, in order to obtain the best STS estimations, the analysis was performed in two steps. The first was to set the priors of the PK parameters to 1 and fix the residual variability to a small number such as 5% and estimate the PK parameters. The second step was to use the median results from the first step as priors for the second STS analysis. In this second step, the residual variability was not fixed. Final parameter results were simply the mean and variance of the PK parameters and the mean of the residual variability obtained from the individuals. Results from the STS analyses were then used as prior information for all mixed effect analyses.

For all IT2S analyses, the proportional and additive components of the residual variability were updated with maximum likelihood at Iterations 10, 30, 60, 100 and 150. The residual variability parameters were updated automatically in all other algorithms.

In the work conducted for this exercise, results from the IT2S and MLEM analyses were considered converged when population estimates varied by no more than 1% for at least 200 consecutive population iterations. From the iterations where estimates varied less than 1%, the optimal iteration was the iteration representing the median estimates of these converged iterations.

For MLEM analyses, the maximum number of samples was used, which was 3000. As a secondary objective, the MLEM analyses were redone with the number of

samples set to 1000 to determine the impact on the PK results when reducing the number of samples taken in the expectation maximization step.

For each simulated study, FOCE was permitted to converge and results were only used if optimization was concluded successfully with significant digits being greater than 3. If FOCE was not able to converge, then prior parameter values were modified first by using the median STS values instead of the average STS values, then by increasing the residual variability and then using the FO algorithm to obtain different population priors.

Precision and Bias

Algorithms were compared based on their estimates of population PK parameters, population PK variances, individual PK parameters and residual variability. The true values for these estimations were determined *a priori* by an independent scientist, as previously described. For each simulated study, an overall precision and bias of the population PK parameter estimates with respect to the true values was assessed using the following formulas:

$$\text{Precision} = \text{mean}_{i=1}^{n}\left[\frac{\text{Absolute}(\text{estimated value}_i - \text{true value}_i)}{\text{true value}_i}\right]*100$$

$$\text{B}ias = \text{mean}_{i=1}^{n}\left[\frac{\text{estimated value}_i - \text{true value}_i}{\text{true value}_i}\right]*100$$

where i is the i^{th} population PK parameter estimate and n is the total number of population PK parameters in a study. Overall precision and bias for the population PK variances, plasma residual variability and urine residual variability were estimated in the same manner.

Similarly, in each study an overall precision and bias of the individual PK parameter estimates was calculated with respect to the true values. However, the median was used instead of the mean as the mean values were occasionally inflated by the difficulties in estimating some individual values. This was true for all algorithms. For the remainder of the text, the median or mean precision and bias will be referred to as precision and bias.

Variance shrinkage

It has been reported in the literature that variances may be underestimated during the modeling process. Underestimated variances have been described in detail by Karlsson.[17] This issue has been labeled as η-shrinkage (sh_η). The author proposed a formula in order to determine which parameters have too small of a variance associated with them. The formula is presented below.

$$sh_\eta = 1 - \frac{SD(\eta_{EBE})}{\omega}$$

In this formula, $SD(\eta_{EBE})$ is the standard deviation of the individual estimates of η for each parameter and ω is the estimate of the standard deviation of the estimated population variance. A shrinkage value close to zero corresponds to minimal shrinkage while a value close to 1 corresponds to severe shrinkage. As these were simulated studies, the true variances were known and the exact percentage by which a variance was underestimated were also calculated. For the purposes of this paper, a severely underestimated variance was defined as a percentage bias value between -95 and -100%.

Results

Bias

Mean and median biases for each of the different algorithms are presented in Table 2. The median bias for the population PK parameters, their variances as well as individual PK parameters estimates were all close to zero.

Precision

Mean and median precision values are also presented in Table 2 for the population PK parameters, their variances, individual PK parameters estimates as well as the residual variability for plasma and urine. To better understand the different algorithms' overall precision results, box plots are presented showing the 25^{th}, 50^{th}, 75^{th} and the 95^{th} percentiles of the precision. These precision results are discussed in details in the following subsections.

Population PK Parameters Estimates

Of all parameters estimated, the population PK parameters were the parameters that were most precisely estimated by all algorithms. Figure 1A represents the box plots for the mean precision in PK parameters of the different studies for each of the algorithms. The 50^{th} percentile of the mean precision was 9 and 10% for FOCE and MLEM, respectively. These methods were followed by the ITS, IT2S and STS methods at 14, 22 and 36%, respectively. When looking at the dispersion of the full box plot, both FOCE and MLEM had smaller ranges than the rest of the methods. STS was the least precise method.

Variance Estimates

Figure 1B represents the box plots for the mean precision in variance estimates of the different studies for each algorithm. These parameters were not as precisely estimated as the population PK parameters. The results varied greatly from one study to the next as well as from one algorithm to the other. However, all mixed effect modeling fared better at estimating the variances than the STS method. Again, FOCE and MLEM algorithms were better at estimating the population variances with the 50^{th} percentile of the mean precision being 76 and 208% respectively. These methods were followed by IT2S, ITS and STS methods at 557, 580 and 7064%, respectively. However, if the box plot was compared as a whole and not just the 50^{th} percentile, FOCE and MLEM fared better than the other methods. In addition, ITS had less variable estimates as a whole than IT2S.

In the 29 simulated studies that were analyzed, a total of 158 variances for PK parameters were estimated. Of these 158 estimated variances, the FOCE algorithm produced a severely underestimated variance in 38 instances or for 24% of the parameters. In contrast, the MLEM algorithm only had five instances of a severely underestimated variance corresponding to 3% of the parameters. More importantly, of the 29 studies that were simulated and fitted, the FOCE method severely underestimated the variance for either the elimination clearance or the volume of the central compartment or both in 9 studies. The MLEM algorithm never severely underestimated the variances of these parameters.

Applying the shrinkage formula proposed by Karlsson to the results of our 29 studies revealed that 28 of the 38 variance parameters that had a bias between -95 and -100% had a shrinkage value of 0.95 and greater. All ten other severely underestimated

variances had a shrinkage value > 0.9. None of the MLEM estimated values had a shrinkage value > 0.95. In fact, of the 158 variances estimated with MLEM, only 13 sh_{η} values were greater than 0.5 with almost half of these coming from 1 study while FOCE had 70 estimated variances with a sh_{η} value above 0.5.

Residual Variability

Figure 1C and 1D represent the box plots for the mean precision in plasma and urine residual variability. The precision in estimating the true residual variability for plasma and urine varied depending on the algorithm chosen. The best algorithm in determining the residual variability was FOCE with the 50th percentile of the mean precision being 5 and 10% for plasma and urine, respectively. This was closely followed by the MLEM algorithm which had a 50th percentile of 9 and 15% for plasma and urine, respectively. The width of the boxes was very similar for FOCE and MLEM concerning the plasma residual variability while the precision for urine was slightly better with FOCE. All three other methods fared poorly against these two methods. The 50th percentile for plasma and urine were 11 and 33%, 12 and 30% and 16 and 38% for IT2S, ITS and STS, respectively. Looking at the whole box plot, ITS did not fare any better than the previous IT2S algorithm nor the STS method.

Individual PK Parameter Estimates

Box plots of the precision for the individual estimates can be found in Figure 2. The 50th percentile of the median precision was 9 and 10% for FOCE and MLEM, respectively. These methods were followed by the ITS, IT2S and STS methods at 12, 19

and 24%, respectively. When looking at the dispersion of the full box plot, both FOCE and MLEM had a smaller range than the rest of the methods. The others were all very similar.

In order to determine if certain methods performed better than others under different conditions, other graphs were created to assess the precision in individual PK estimates with studies having rich or sparse sampling and less than 24 subjects or 24 subjects and more. These box plots are in Figure 3A to 3D. Under a rich sampling design, all methods performed reasonably well with FOCE and MLEM having the best precision. ITS was slightly better than IT2S and STS. However, the box plots representing the results under sparse sampling showed that MLEM was the most precise algorithm followed by FOCE. The other three methods were not as precise, with the new ITS algorithm having the best results of the three. The number of subjects appeared to have a great impact on the precision of the individual PK parameter estimates. With only 6 or 12 subjects in a study (Figure 3D), all mixed effect algorithms had a 50^{th} percentile at 7%; however, the box plot was narrower for ITS and FOCE followed by MLEM. When a study had many subjects (24 and above), FOCE and MLEM had the best precision while ITS seemed to have difficulties at estimating the individual PK parameters for some studies.

MLEM Population Estimates Using 3000 Samples versus 1000 Samples

The population PK parameter estimates and the population PK variance estimates obtained from the MLEM analyses using 1000 samples were compared to the results obtained from the analyses using 3000 samples. Excluding the one study in which

iterations were prematurely terminated (*i.e.*, crashed), the mean difference between the population PK parameter estimates using 1000 samples was less than 1% of those obtained from the analyses using 3000 samples. However, there was a greater difference in the estimation of the variances if 1000 or 3000 samples were used. The mean difference was 7% (excluding the same study).

Discussion

Comparisons between different algorithms were made using 29 simulated clinical studies where the mean and variance values for the PK parameters as well as the residual variability and the individual PK parameters were known to an independent scientist. Having the fitting of the data performed by a scientist blinded to the true values removed any potential preferred bias this scientist had toward a particular algorithm. Including multiple studies permitted the testing of the algorithms under varying conditions such as different types of drug administration (oral vs. IV), number of subjects (6 to 200), number of samples per subjects (sparse vs. rich) and number of PK and variability parameters (6 to 21).

For all comparisons, mixed effect modeling approaches provided better results than a standard method of STS. Mixed effect modeling provided better population PK parameters, variances, residual variability as well as individual PK parameters. This was not surprising and was in agreement with previously published papers demonstrating this fact.[8-13] This was true regardless of the condition of the study such as sparse or rich sampling, little or many subjects, oral versus IV administration and plasma or plasma with urine collection. The most profound difference between mixed effect modeling and

a STS approach was the fact that STS grossly over estimated the variances of the population PK parameters which was in accordance with the literature.[8-13] Therefore, the values obtained from STS analyses should not be used as final results but rather as starting priors for mixed effect modeling algorithms.

The new MLEM algorithm performed better than the IT2S and ITS algorithms in the estimation of the population PK parameter estimates, their corresponding variances, residual variability and individual PK estimates. For individual PK parameters, ITS was slightly more precise for studies with less than 24 subjects. However, the difference was marginal. MLEM was more precise than the other two algorithms for all parameter estimates for all the different study designs.

Of all the mixed effect algorithms tested, MLEM was the only one that used an exact estimation of the maximum likelihood estimates instead of a linear approximation of the maximum likelihood. In order to avoid a linearization of the model, MLEM uses a sampling-based method implemented in the expectation step. In ADAPT 5®, the number of samples used for the estimation of the parameter values is set by the user and can vary from 1000 to 3000 samples. Wang et al demonstrated that accuracy to the second significant digit was obtained for all parameters using 1000 samples when compared to estimation obtained with 2000 and 3000 samples.[18] However, the author only demonstrated this finding using 1 study. In our analyses, we simulated 29 studies with different study designs. The use of 1000 samples versus 3000 had little impact in population PK parameter estimates. There was a greater difference in the estimation of the variances if 1000 or 3000 samples were used. However, in our analyses, choosing 3000 samples instead of 1000 samples did not guarantee better variance estimates. In

fact, the precision in estimates even worsened in some cases. Considering that the analysis time was approximately 50 to 60% faster using 1000 samples and that the estimates were no worse than when 3000 samples were used, 1000 samples could safely be used for population analyses.

In our analyses, FOCE was considered as a control analysis to benchmark the new algorithms, as this algorithm is used extensively by the scientific community and is a widely accepted population PK tool. Of the new algorithms available in ADAPT 5®, MLEM was the best algorithm. It compared favorably with FOCE in all estimates. Individual PK estimates fared better with a sparse sampling design using MLEM although only seven of the studies used in the analyses utilized a sparse sampling). This was a surprising result considering how the two algorithms function. FOCE in NONMEM® analyzes the population parameters first and then uses a *post hoc* Bayesian analysis to obtain its individual PK parameters. This is contrary to MLEM in ADAPT 5® which computes the individual PK parameters to then update the population parameters at each iteration. With a sparse sampling design, it was expected that it may be harder to obtain individual results with precision when determined as a first step as in ADAPT 5®. In contrast, FOCE might have a slight edge when there are many subjects included in the analyses.

The main difference between FOCE and the other algorithms was its estimation of the variances. A word of caution is required when looking at the variance results for FOCE. With the exception of FOCE, all methods over-inflated the variance when the variance parameter was poorly estimated while FOCE under estimated and often estimated a value very close to zero for a variance parameter that was difficult to

estimate. FOCE was the only algorithm to have a negative median bias of the variances. That means that the median variances estimated for a study was underestimated in more than 50% of the studies. This had a consequence on the precision comparisons. When a variance parameter was under estimated (*i.e.*, near zero), the precision obtained for that estimate will be close to 100% as determined by the following formula: absolute(0-True value)/True*100 value ≈ 100. On the other hand, a precision value toward infinity can be obtained for a variance parameter that was over estimated as seen in the next formula: absolute(High value – True value)/True*100 ≈ very high value. In our exercise, 24% of population PK parameters had variances estimated by NONMEM that were very close to zero. This problem has been well described by Karlsson and is known as η-shrinkage.[17] This author specified that a level of shrinkage above 20 to 30% should be viewed with caution. Using published data, Karlsson suggested that shrinkage occurred in approximately one third of the clearances estimated in studies. This was in line with our simulated studies where clearance or volume of the central compartment had variances with a severe shrinkage issue in approximately one third of the studies (9 out of 29). In fact, in the 29 simulated studies, almost half of the variances estimated by FOCE had shrinkage issues ($sh_\eta > 0.3$) while MLEM had shrinkage issues with less than 10% of the estimated variances. Often, the primary goal of the population pharmacokinetic analysis is to estimate or predict individual profiles and PK parameters. If variances are close to zero, then it becomes difficult to estimate the differences between subjects. Thus, if shrinkage occurs, it may be wise to switch to another algorithm especially if the model will be used to simulate individual results or profiles.

The new ITS algorithm in ADAPT 5® is a modification of the iterative two-stage algorithm found in IT2S®. In general, ITS fared slightly better than IT2S with population PK parameters, their corresponding variances and with the estimation of the individual PK parameters. However, it had difficulties with the residual variability estimates. It had a tendency to shrink the additive error parameter very quickly preventing some analyses from converging. To avoid this problem, the additive error had to be fixed in those analyses to a small value after a few iterations. In ITS, residual variability parameters can be fixed to different values by the user at any time if it is judged that the values were converging to a local minimum. The problem is that it can never be known if the residual variability estimates are acceptable or converging to incorrect values.

The major advantage of the new ITS algorithm in ADAPT 5® was the speed at which results were obtained. In most instances, convergence was obtained automatically and after few iterations. In addition, iterations did not need to be interrupted by the user to update the residual variability parameters. Therefore, ITS can be a useful tool depending on the required analyses. However, to ensure more precise results, MLEM or FOCE are recommended.

To have the most success with population PK analyses, one should exploit the strengths of every existing tool in order to ensure the most accurate results. In population analyses, it is always important to keep in mind the purpose of the analyses. It is possible that STS results fulfill the purpose of the analyses. In fact, STS is very useful for structural model discrimination and for obtaining prior information for mixed effect modeling. If more precise analyses are required, then mixed effect modeling should be favored over a STS method regardless of the study design or the number of subjects.

At the time at which these analyses were performed, NONMEM 7® was not available. The NONMEM 7 ® package offers ITS and MLEM algorithms, as well as a STS methodology which were not previously part of NONMEM VI®. Therefore, further work will be required to determine if the MLEM algorithm in the latest NONMEM® version will reduce the number of shrinkage issues.

Conclusions

The algorithms offered in the new ADAPT 5® software fared well compared with other established methodologies. In addition, the number of samples used in the expectation maximization step of the MLEM algorithm in ADAPT 5® did not influence the obtained results.

The FOCE and MLEM algorithms offered the most precise estimations of the population PK parameters, their variances, residual variability and individual PK estimates. The MLEM algorithm in ADAPT 5® was slower than the FOCE algorithm in NONMEM VI®. However, it had less shrinkage issues for the variance estimates.

All methods of analyses have their strengths and weaknesses and no method is clearly superior under all conditions. Scientists should use all available tools at their disposal to provide the best results.

References

1. Sheiner LB, Rosenberg B, Marathe VV. Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. J Pharmacokinet Biopharm. 1977 Oct;5(5):445-79.
2. Sheiner, LB. Learning versus confirming in clinical drug development. Clin Pharmacol Ther. 1997 Mar;61(3):275-91.
3. Challenge and Opportunity on the Critical Path to Medical Products. U.S. Department of Health and Human Services Food and Drug Administration. March 2004. Available at http://www.fda.gov/downloads/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm113411.pdf
4. Joga Gobburu. How EOP2A Meetings with FDA can Cut Down on Phase Attrition Rate by 50%? AAPS Webinar; July 15, 2010.
5. D'Argenio, D.Z., A. Schumitzky and X. Wang. ADAPT 5 User's Guide: Pharmacokinetic/Pharmacodynamic Systems Analysis Software. Biomedical Simulations Resource, Los Angeles, 2009.
6. Collins, D. and A. Forrest, IT2S User's Guide. 1995, State University of New York at Buffalo: Buffalo.
7. Beal, S.L., Sheiner, L.B. & Boeckmann, A. NONMEM Users Guide (1989-2006) Version VI. Icon Development Solutions, Ellicott city, MD, 2008.
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.
9. Ludden TM. Population pharmacokinetics. J Clin Pharmacol. 1988 Dec;28(12):1059-63. Review.
10. White DB, Walawander CA, Tung Y, Grasela TH. An evaluation of point and interval estimates in population pharmacokinetics using NONMEM analysis. J Pharmacokinet Biopharm. 1991 Feb;19(1):87-112.
11. Steimer JL, Mallet A, Golmard JL, Boisvieux JF. Alternative approaches to estimation of population pharmacokinetic parameters: comparison with the nonlinear mixed-effect model. Drug Metab Rev. 1984;15(1-2):265-92.
12. Yu DK, Hutcheson SJ, Wei G, Bhargava VO, Weir SJ. A comparison of population and standard two-stage pharmacokinetic analyses of vigabatrin data. Biopharm Drug Dispos. 1994 Aug;15(6):473-84.
13. Vermes A, Math t RA, Van der Sijs IH, Dankert J, Guchelaar HJ. Population pharmacokinetics of flucytosine: comparison and validation of three models using STS, NPEM, and NONMEM. Ther Drug Monit. 2000 Dec;22(6):676-87.
14. Forrest A, Hawtor J, Egorin MJ. Evaluation of a new program for population PK/PD analysis – applied to simulated Phase I data. Clin Pharm Ther 1991; 49(2):153.
15. Walker S. An EM algorithm for nonlinear random effects models. Biometrics 1996; 52(3):934-944.
16. Dempster AP;Laird NM; Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society. Series B (Methodological) 1977 39(1):1-38.
17. Savic RM, Karlsson MO. Importance of shrinkage in empirical bayes estimates

for diagnostics: problems and solutions. AAPS J. 2009 Sep;11(3):558-69.
18. Wang X, Schumitzky A, D'Argenio DZ. Nonlinear Random Effects Mixture Models: Maximum Likelihood Estimation via the EM Algorithm. Comput Stat Data Anal. 2007 Aug 15;51(12):6614-6623.

Figures

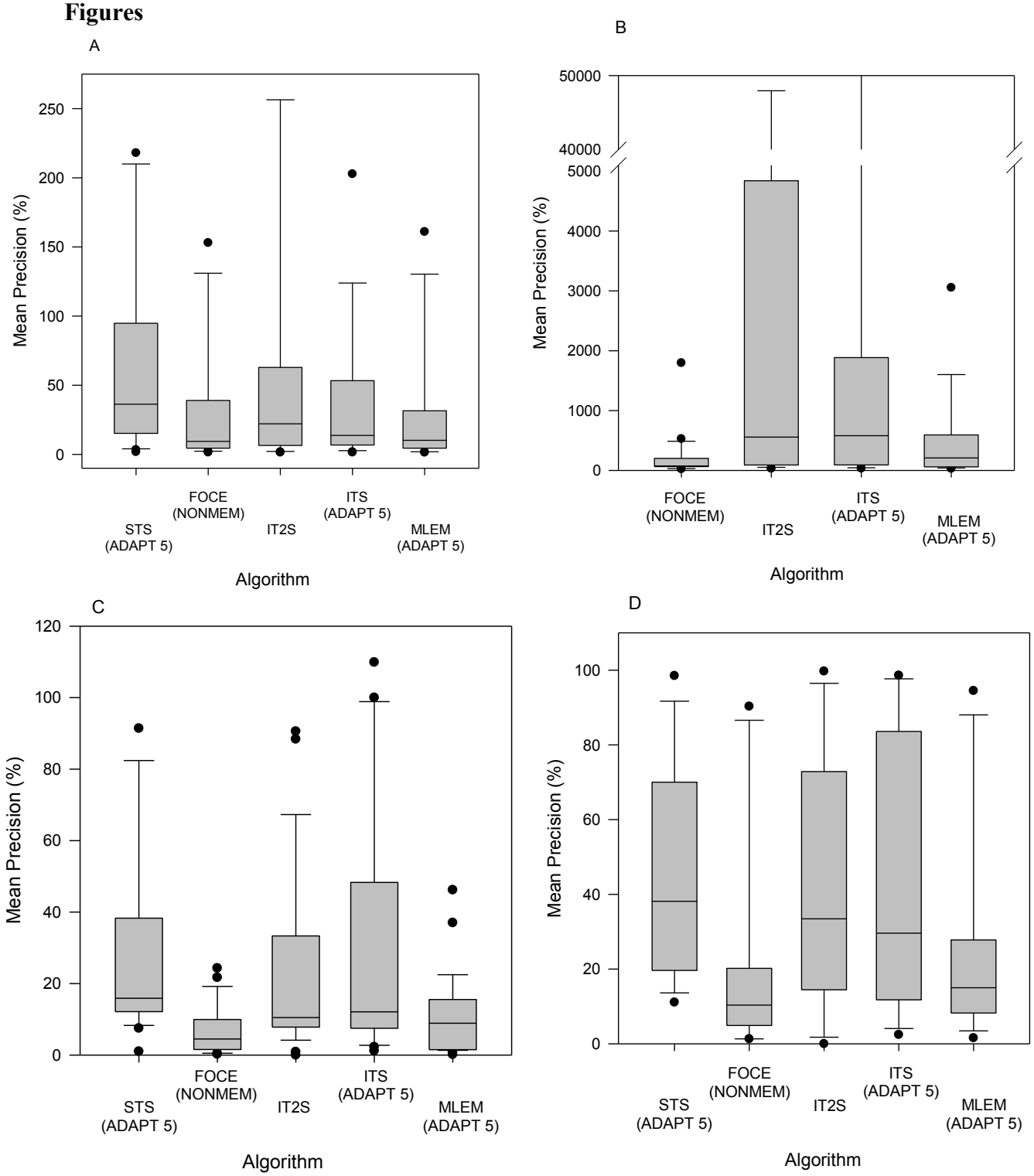


Figure 1 Box plot of mean % precision on estimates for the different algorithms. A) population PK parameters; B) population PK variances; C) plasma residual variability; D) urine residual variability. The bar in the box represents the 50th percentile, the box boundaries are the 25th and 75th percentiles, and the limits of the whiskers are the 10th and 90th percentiles. The dots are outliers.

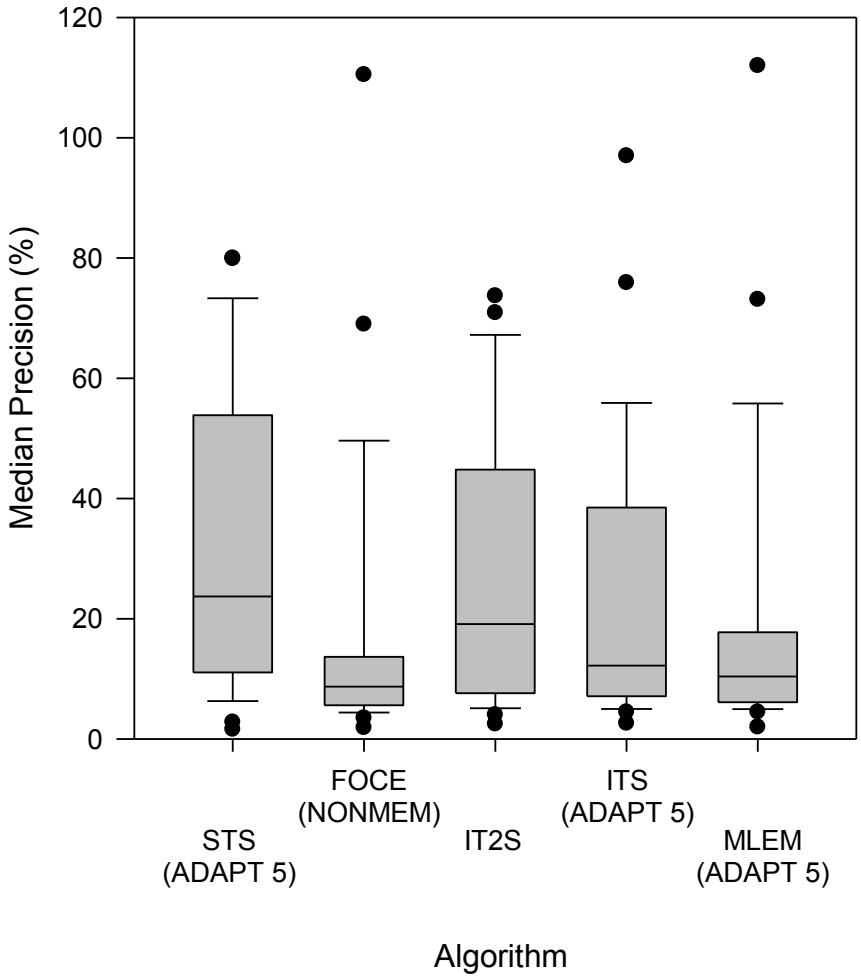


Figure 2 **Box plot of median precision on individual estimates for the different algorithms including all studies.** The bar in the box represents the 50th percentile, the box boundaries are the 25th and 75th percentiles, and the limits of the whiskers are the 10th and 90th percentiles. The dots are outliers.

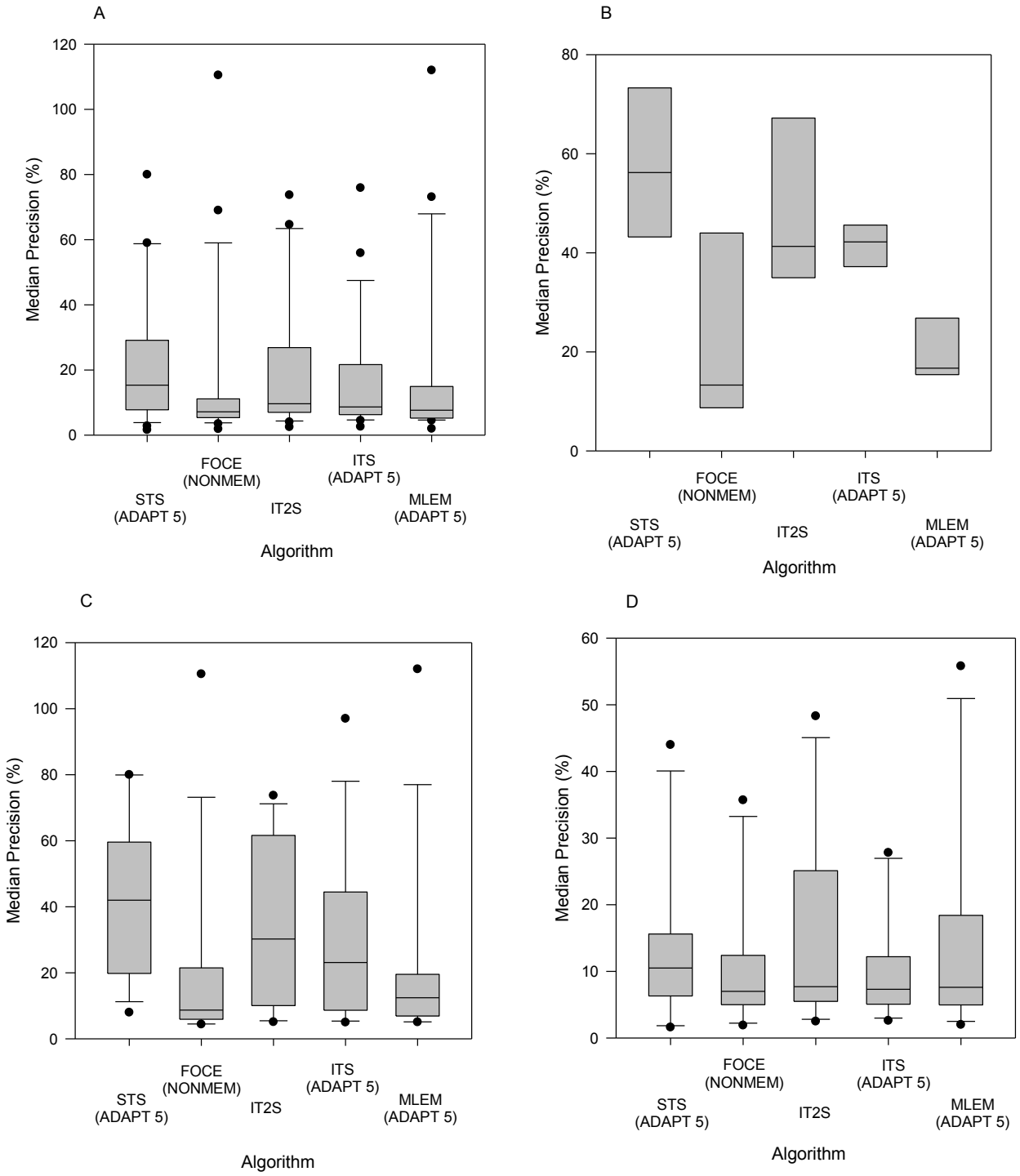


Figure 3 **Median precision on individual estimates for the different algorithms.** A) Rich sampling; B) Sparse sampling; C) ≥ 24 subjects; D) < 24 subjects. The bar in the box represents the 50th percentile, the box boundaries are the 25th and 75th percentiles, and the limits of the whiskers are the 10th and 90th percentiles. The dots are outliers.

Tables

Table 1: Study design for the different simulated studies

Study #	# of subjects	# of compartments	Route of Administration	Type of sampling	Matrix
1	6	3	IV	Rich	Plasma + Urine
2	6	3	IV	Rich	Plasma
3	6	2	IV	Rich	Plasma + Urine
4	6	2	IV	Rich	Plasma
5	6	1	IV	Rich	Plasma + Urine
6	12	3	IV	Rich	Plasma
7	12	3	IV	Rich	Plasma + Urine
8	12	2	IV	Rich	Plasma
9	12	2	IV	Rich	Plasma + Urine
10	16	2	IV	Rich	Plasma + Urine
11	18	2	Oral	Rich	Plasma
12	24	2	IV	Rich	Plasma + Urine
13	24	1	Oral	Rich	Plasma
14	30	2	IV	Rich	Plasma
15	30	2	Oral	Rich	Plasma
16	30	2	IV	Rich	Plasma + Urine
17	30	2	Oral	Rich	Plasma + Urine
18	30	2	IV	Sparse	Plasma
19	30	2	Oral	Sparse	Plasma
20	30	2	IV	Sparse	Plasma + Urine
21	30	2	Oral	Sparse	Plasma + Urine
22	50	3	IV	Rich	Plasma + Urine
23	200	2	IV	Rich	Plasma
24	200	2	IV	Rich	Plasma + Urine
25	200	2	IV	Sparse	Plasma + Urine
26	200	2	Oral	Rich	Plasma
27	200	2	Oral	Rich	Plasma + Urine
28	200	2	Oral	Sparse	Plasma
29	200	2	Oral	Sparse	Plasma + Urine
Total	1850	2x1cpt 22x2cpt 5x3cpt	19 IV 10 Oral	22 rich 7 sparse	13 plasma 16 plasma + urine

Table 2: Mean and median precision and bias for the population PK parameters, their variances, plasma residual variability, urine residual variability and individual parameters for each algorithm

		Precision					Bias		
		Population		Individual	Residual Variability		Population		Individual
		Parameter	Variance	Parameter	Plasma	Urine*	Parameter	Variance	Parameter
FOCE	**Mean**	38	197	37	6	22	20	116	14
	Median	7	84	9	5	10	-1	-5	-1
IT2S	**Mean**	63	14817	81	23	43	37	14796	52
	Median	10	104	19	11	33	0	94	0
MLEM	**Mean**	47	4887	51	11	26	29	4860	31
	Median	8	87	10	9	15	1	45	1
ITS	**Mean**	45	7981	50	31	43	17	7965	18
	Median	9	111	12	12	30	0	119	-1
STS	**Mean**	88	176470	102	33	47	58	176459	64
	Median	18	1817	24	16	38	4	1817	-2

* n = 16

2. ARTICLE #2

How critical is the duration of the sampling scheme for the determination of half-life, characterization of exposure and assessment of bioequivalence?

2.1. INTRODUCTION

In drug development, noncompartmental PK analysis is often the preferred choice by the regulatory agencies especially in the early phases when many samples are collected. For instance, noncompartmental analysis is used for bioequivalence studies. It is simple, robust and not dependent on the scientist's ability at modeling data. However, as discussed in Chapter 1, Section 1.1.1, AUCinf has to be well characterized in order to produce appropriate results. Previous experiences with bioequivalence studies have shown that subjects with a long half-life relative to the duration of the sampling scheme possibly influenced study conclusions when these were maintained in the ANOVA analyses. Therefore, the purpose of this part of the research was to determine and confirm if a subject whose half-life is calculated from a sampling scheme duration that is considered too short could affect bioequivalence conclusions of a study and if these parameters should be removed from statistical analyses.

One tool that can be used to answer the above questions is clinical trial simulations. Indeed, these can be used to answer multiple "what if" questions or to predict study outcomes, since the number of simulations that can be conducted is limitless. Because so many simulated profiles can be simulated, this method is ideal for testing our hypotheses since some simulated individual concentration-time profiles may be associated with longer half-lives and thus influence the predicted outcomes.

In order to perform clinical trial simulations, subjects' concentration-time profiles are required. Therefore, to test the objectives described above using clinical trial simulations, subjects' concentration-time profiles were simulated with different sampling

scheme durations in order to calculate noncompartmental PK parameters from these profiles. It is the same principle when efficacy and safety clinical trials are simulated. Efficacy and safety of a drug are often correlated with the concentration-time profiles and it is these profiles that are simulated. Monte Carlo simulations, which form the basis of clinical trial simulations, will thus provide PK parameters for subjects based on the population PK parameters and their expected variability (variance). The next step to obtain simulated individual subject concentration-time profiles is to use these individual PK parameters obtained from the Monte Carlo simulations to produce the expected profile for each subject. However, the individual concentrations are simulated with a variability error, which reflects analytical and clinical errors.

To answer the objectives of this part of the research, subject profiles were therefore simulated and the sampling scheme was progressively shortened to determine the impact on the estimated noncompartmental PK parameters. In addition, bioequivalence studies were simulated to determine the impact of maintaining subjects with a long half-life relative to the sampling scheme in the ANOVA results. Results are presented in the following article.

2.2. ARTICLE

Title

How critical is the duration of the sampling scheme for the determination of half-life, characterization of exposure and assessment of bioequivalence?

Authors: Philippe Colucci, M.Sc.[1,2]; Turgeon J, PhD[1,3]; and Murray P. Ducharme; PharmD[1,2]

Institutional affiliation:

1) Faculté de Pharmacie, University of Montreal, Montreal, Canada
2) Cetero Research, Cary, North Carolina
3) CRCHUM, Centre Hospitalier de l'Université de Montréal, Montreal, Canada

Key Words: AUCinf, half-life, simulations, bioavailability, bioequivalence, ADAPT-5

Conflict of Interest: There are no personal financial or non-financial conflicts of interest

Abstract

In noncompartmental analysis, poor characterization of the terminal elimination rate constant (Kel) will lead to biased results for half-life and total exposure (AUCinf), providing incorrect relative bioavailability and bioequivalence conclusions. We set out to determine if the sampling scheme duration was crucial for proper half-life and AUCinf determination. Profiles for 1000 subjects were simulated with a sampling scheme covering five half-lives. Concentrations were gradually removed from the end of the profile to determine if precision and bias in the half-life and AUCinf values were affected. Additionally, 30 bioequivalence studies were simulated to determine the influence of unreliable AUCinf PK parameter on BE conclusions. Precision and bias became unacceptable for AUCinf and half-life if Kel was not determined with a sampling scheme covering at least 2 and 4 half-lives, respectively. Bioequivalence conclusions also deteriorated if unreliable PK parameters were maintained. Sampling scheme duration is important when calculating noncompartmental parameters. In conclusion, sampling scheme duration should be at least 4 times the average measured half-life in order to have confidence in the reported half-life values. In a bioequivalence setting, individual subject's pharmacokinetic parameter AUCinf should be removed from the pivotal statistical ANOVA analysis when their associated calculated half-life is longer than half of the total sampling interval.

Introduction

Bioequivalence (BE) studies usually assess and compare the rate and extent to which the active ingredient or active moiety becomes available into the systemic circulation.[1] Theoretically, if two formulations of the same active product produce similar systemic concentration-time profiles, they will also produce similar concentration-time profiles at the site of action and therefore produce comparable pharmacological effects.[1,2]

To properly characterize relative bioavailability (BA) or establish BE in terms of extent of exposure, total exposure is measured by the PK parameter area under the curve from time zero to infinity (AUCinf).[3] This parameter is calculated as the area under the curve from time zero to the last measurable concentration (AUC0-t) plus the last measurable drug concentration divided by the terminal elimination rate constant (λ_z if calculated compartmentally or Kel if calculated noncompartmentally).[4,5] Therefore, in order to properly characterize AUCinf, the terminal rate constant and terminal elimination half-life (T½) must be properly determined.

For a drug displaying linear PK properties, the terminal phase seen in a graphical log-concentration versus time profile decreases in a straight line, independently of concentrations, and represents either the true elimination or the absorption of the drug (e.g., in the case of flip-flop pharmacokinetics). In noncompartmental analyses, Kel is estimated by linear regression from the slope of the terminal log-linear portion of the drug concentration versus time curve. The terminal elimination half-life is calculated from this constant as ln(2)/Kel.

An improper characterization of the PK parameters Kel and T½ will lead to a poor determination of the total exposure (*i.e.*, AUCinf), an unacceptable characterization

of the relative BA and therefore may lead to an incorrect BE conclusion. As such, certain criteria are useful in the determination of the Kel parameter. These include having at least three measurable concentration time-points in the terminal phase,[3] excluding the maximum observed concentration (Cmax) from the regression analysis and having a regression coefficient of at least 0.8.[6-9] In addition to these common criteria, an additional important criterion that should be considered is the sampling scheme duration.

Simulations were thus undertaken to determine how the duration of a sampling scheme could affect the proper determination of the extent of exposure. The objectives of the work presented in this article are first to assess how the sampling duration affects the noncompartmental half-life or AUCinf determination and secondly to assess if these parameters should be excluded from the overall pivotal statistics based on the sampling duration to ensure adequate BA calculations and BE conclusions.

Methods

Simulations

Concentration data for one thousand (1000) subjects receiving a single 100 mg dose of a fictitious drug were simulated by Monte Carlo techniques in ADAPT 5®.[10] The model used to perform these simulations exhibited linear absorption (Ka) and elimination characteristics, included a central compartment (Vc/F), a peripheral compartment (Vp/F), a clearance from the central compartment (CL/F) and a distributional clearance (CLd/F). Figure 1 depicts the model used. The simulated concentration data were based on the following hypothetical PK population mean and variability values:

Table 1: Hypothetical population PK and variability parameters

Parameter	Population mean	Inter-individual CV%
Ka (h^{-1})	1	60
CL/F (L/h)	2.65	50
Vc/F (L)	8.5	30
CLd/F (L/h)	1.75	75
Vp/F (L)	16.7	75
T½* (h)	12 hours	

* T½ was calculated based on the population parameters. Average T½ presented.

The simulated concentration-time profiles included the following 21 sample time points: 0 (pre-dose), and 0.25, 0.5, 1, 1.5, 2, 3, 4, 5, 6, 8, 10, 12, 14, 16, 18, 21, 24, 30, 36, 48 and 60 hours after dosing. Simulations were performed using a residual variability of 10%. To determine the impact of a shorter sampling scheme, the concentration data at the end of the simulated profiles were progressively removed so that the duration of the entire sampling profile varied from 1 to 5 half-lives (i.e., 12 to 60 hours).

Noncompartmental PK Analysis

The simulated subjects' data were then analyzed using a standard noncompartmental approach with SAS® to obtain Kel, T½ and AUCinf. These noncompartmental PK parameters were re-calculated for each different sampling scheme. Kel, T½, and AUCinf were only set to missing if Kel was positive. Kel was calculated as the apparent first-order elimination rate constant calculated from a semi-log plot of the plasma concentration versus time curve. The parameter was calculated by linear least-squares regression analysis to the terminal log-linear phase.

Precision and Bias

True and known values for AUCinf and T½ for each subject were those obtained from the simulation in ADAPT 5® (*i.e.*, ADAPT values were the true values). True AUCinf for each subject was calculated by dividing dose by CL/F and true T½ was calculated from the parameters CL/F, CLd/F, Vc/F and Vp/F.[4] To determine how the duration of a sampling scheme affected the robustness of the calculated noncompartmental parameters, bias and precision as described by Sheiner[11] were evaluated for the noncompartmental parameters AUCinf and T½. Precision of the noncompartmental AUCinf and T½ results with respect to the true results were assessed using the following formula: 100*ABS(noncompartment value – ADAPT value)/ADAPT value. Bias of the noncompartmental AUCinf and T½ results with respect to the true results were evaluated using the following formula: 100*(noncompartment value – ADAPT value)/ADAPT value. Precision and bias judged acceptable were set *a priori* to 10% and 5%, respectively.

Simulated subjects with an absorption half-life that was longer than the terminal elimination half-life (*i.e.*, flip-flop profiles) were excluded from the precision and bias calculation. This was done in order to avoid mixing different half-life types. The apparent terminal noncompartmental half-life would be a reflection of the absorption half-life and comparing this value to the known elimination half-life from the compartmental analysis would be wrong and would affect the precision and bias.

Bioequivalence studies

In addition to verifying the robustness (precision and bias) of the noncompartmental parameters, the impact of including or not data from subjects with unreliable noncompartmental AUCinf estimates on bioequivalence conclusions was evaluated using simulated studies. Using the same population PK parameters and variability parameters, thirty 2-way crossover studies each having 24 subjects were simulated. Ten studies per sampling scheme were derived based on sampling schemes that covered on average 2 half-lives (24 hours), 2.5 half-lives (30 hours) and 3 half-lives (36 hours). This resulted in some subjects in each study with potentially unreliable half-life estimates. Noncompartmental parameters were obtained in the same way as previously described and comparisons were made between the true values obtained from ADAPT 5® and the noncompartmental estimates. Noncompartmental results were considered the test results while the ADAPT results were considered the reference results.

AUCinf were determined to be reliable or not using two different methods. The first method (Method 1) considered an AUCinf to be unreliable if it did not meet both the acceptable *a priori* precision and bias of 10% and 5%, respectively. For this first method, an AUCinf was considered to be unreliable if the associated T½ was longer than half of the duration of the sampling scheme (as demonstrated in the results section). The second method (Method 2) ignored the overall length of the sampling scheme as it considered only the duration of the sampling scheme over which Kel was determined. For this second method, an AUCinf was considered to be unreliable if the associated T½ was longer than the time span over which Kel was estimated as proposed by Purves.[5]

Analysis of Variance

Analysis of variance (ANOVA) with Proc Mixed model in SAS® (Version 9.1.3) was performed to calculate the ratio and 90% confidence intervals (CI) on the ln-transformed AUCinf parameter (noncompartmental vs. true value). Three different ratios and 90% confidence intervals were calculated for each of the 30 simulated bioequivalence studies. Firstly, with all AUCinf included (*i.e.*, without the removal of any unreliable AUCinf); secondly, with the removal of unreliable AUCinf as determined by Method 1 and thirdly, with the removal of unreliable AUCinf as determined by Method 2. These ANOVA were performed to determine the consequences on bioequivalence conclusions if unreliable results were included or not.

Results

Precision & Bias

The precision and bias of the noncompartmental method to estimate the true PK parameters AUCinf and T½ are presented in Figure 2 A and B. Results are presented for sampling schemes covering 1 to 5 half-lives (i.e., 12 to 60 hours). As the sampling scheme duration was shortened from five half-lives to one half-life, the precision and bias in the noncompartmental PK parameters AUCinf and T½ deteriorated. The median precision and bias for AUCinf changed from 2.1% and -0.1% to 14.0% and -13.4%, respectively. Similarly, the precision and bias for T½ deteriorated from 5.0% and -1.4% to 65.5% and -64.7%, respectively.

Results indicated that to have both acceptable precision and bias in the noncompartmental AUCinf or T½ estimates, a sampling scheme that spanned ≥ 2 and 4

half-lives was required, respectively. Sampling schemes covering less than 2 half-lives led to unreliable AUCinf estimates. In other words, AUCinf were judged to be unreliable with Method 1 when they were calculated using half-life values observed to be longer than half of the overall sampling scheme.

Based on the bias results from all sampling schemes, the noncompartmental method estimated a shorter half-life than the true value (negative bias) and the shorter the sampling scheme, the shorter the noncompartmental half-life that was obtained.

ANOVA - Bioequivalence studies

The mean ratio and 90% confidence intervals comparing the noncompartmental ln-transformed AUCinf to the true ln-transformed AUCinf for the different sampling schemes are presented in Table 2. Bioequivalence results from the 30 simulated studies are presented in Appendix 1. These results are presented with and without the unreliable AUCinf data based on the two already mentioned methods. Results indicated that if unreliable AUCinf were removed from within each BE study, the ratios were closer to 100% indicating that the noncompartmental estimates were closer to the true values. A longer sampling scheme reduced the number of unreliable AUCinf and decreased the magnitude of the improvement associated with removing these unreliable parameters. When 10 different BE studies were simulated per sampling scheme covering 2 (24 hours), 2.5 (30 hours) and 3 half-lives (36 hours), removing the unreliable AUCinf improved mean BE ratio by 6% (from 93.6% to 99.7%), 4% (from 95.3% to 99.1%) and 3% (from 96.9% to 99.6%), respectively (Table 2). The ln-transformed AUCinf mean confidence interval ranges (Upper CI – Lower CI) for each sampling scheme including or

not the unreliable values are depicted in Figure 3. On average, the 90% CI were narrower if the unreliable AUCinf were not included in the ANOVA.

Discussion

Pharmacokinetic parameters calculated using noncompartmental methods do not require much technical *a priori* expertise, and are considered to be the gold standard approach to be used for the vast majority of BE studies. It is generally appreciated though, that this method will only be reliable if the terminal half-life has been properly characterized. The absence of any formal guidance in this regard led us to perform simulations to determine the exact impact of the sampling scheme duration on the robustness of the half-life and AUCinf parameters. One thousand subjects were simulated with a sampling scheme covering 5 half-lives. Then, for these 1000 subjects, sampling schemes of shorter duration were created by progressively removing concentration data at the end of the simulated profiles. This approach avoided having to simulate different subjects for each sampling scheme which would have added variability to the comparisons. By removing concentrations from the end of the profile, it was possible to directly compare the effect a shorter sampling scheme had on the precision and bias of the Kel and T½ determinations.

The noncompartmental method estimated a shorter half-life than the true value. The difference between the estimated and true half-life was more pronounced as the sampling scheme was shorten. This was expected, because with fewer samples collected during the elimination phase, the half-life determined from the noncompartmental

method incorporated concentrations in the distribution phase instead of concentrations from only the elimination phase.

A shorter sampling scheme had a greater impact on the determination of the noncompartmental half-life than on the AUCinf parameter. This was also expected as AUCinf is dependent on both the half-life determination as well as AUC0-t. Therefore, any error on the half-life can only affect the extrapolated portion of AUCinf. With a sampling scheme spanning two half-lives, the extrapolated portion of AUCinf is only 25%. Consequently, even a 20% bias of the half-life would only lead to a 5% error of the total AUCinf. The extrapolated portion is even smaller as the duration of the sampling scheme increases. This limits the impact a poor half-life determination could have on the AUCinf calculation.

For a pivotal PK study, a sampling scheme spanning greater than or equal to 4 half-lives was required to adequately characterize the half-life of a drug in order to maintain the bias and precision below 5% and 10%, respectively. This is in line with the office of generic drugs of the US FDA requirements, but stricter than what is required in Canada or Europe (e.g., 80% of the AUCinf has to be observed, which is equivalent to a sampling duration spanning 2.3 half-lives).[12,13] Therefore, in drug development, one can have confidence in the half-life estimates of a drug if it was obtained from studies in which the sampling scheme duration spanned the equivalent of 4 half-lives or more.

In a bioequivalence setting, AUCinf precision and bias quickly deteriorated if the sampling scheme did not cover at least 2 half-lives, with bias being greater than +/-5%. When designing a BE study, the sample size of a study is often based on a predicted ratio that is ± 5%. Therefore, a bias greater than 5% in the calculation of a PK parameter may

add uncertainty and lead to a study that will be under-powered to prove BE. However, if subjects with AUCinf calculated using an unreliable half-life (eg, based on a value that is greater than half the sampling scheme or if the calculated T½ is less than the time span used in the calculation of Kel) were removed from the ANOVA results, ratios of ln-transformed AUCinf improved and were closer to the true ratio of 100%. In addition, the average confidence intervals for ln-transformed AUCinf tightened if the unreliable subject's data were removed from the ANOVA. Improvement in the ratios and confidence intervals for ln-transformed AUCinf was noted even with the loss of degrees of freedom due to the removal of data from subjects with AUCinf considered unreliable. Therefore, the loss of degrees of freedom was more than offset by the removal of variability caused by keeping these unreliable values.

Increased bias in the AUCinf PK parameter with decreasing time span for Kel determination was previously noted by Purves.[5] This author reported that the variance of the extrapolation in the AUC estimates rapidly increased if the time span for Kel determination was less than the half-life and suggested that the Kel determination should be obtained from a regression that spanned at least two half-lives. Other authors also determined the impact of large bias in the extrapolation portion of the AUCinf or area under the moment curve (AUMC) parameters.[14,15] However, their work focused on the impact on the mean residence time (MRT). The analyses presented in this paper aimed to determine the duration of a sampling scheme that was required to obtain robust half-life and AUCinf results. In addition, this work demonstrated the negative consequences of maintaining unreliable AUCinf on BE conclusions.

As previously mentioned, many scientists in industry and academia use different criteria to avoid having an improper characterization of the PK parameters Kel, T½ and AUCinf. These typically include determining Kel with at least three terminal time points without including the Cmax, and having a regression coefficient (R^2) of at least 0.8. [6-9] To precisely determine the impact of the sampling scheme duration on the PK parameters Kel, T½ and AUCinf, these criteria were not used. However, it is possible that by adding these criteria, the influence of the sampling scheme duration may be further reduced. Therefore, the noncompartmental parameters were recalculated with these criteria added to determine if the criterion of the sampling scheme had the same impact on the conclusions. Although precision and bias were slightly improved when these criteria were used in addition to sampling scheme duration, conclusions remained identical.

Two methods were used to identify the unreliable AUCinf values and verify the impact of removing them on the 90% confidence intervals of bioequivalence studies. Based on the BE results, both methods were comparable and helped to obtain results that were closer to the truth and reduce the overall uncertainty. A difference that exists between the two methods is that the criterion of half the sampling scheme (Method 1) is identical for all subjects within a study making it easier to use as a criterion, while the time span of the Kel (Method 2) is different for every subject.

Conclusions

The results of this study suggest that the sampling duration of any pivotal PK study be at least 4 times the average measured half life in order to have confidence in the reported half-life values and thereby in the extent of exposure parameter AUCinf. The analyses also suggest that individual subject's PK parameters AUCinf, kel and T1/2 should be removed from the pivotal statistical ANOVA analysis (i.e. such as in a BE study) when their associated calculated half-life is longer than half of the total sampling interval, as this will reduce the overall uncertainty and provide results closer to the true values.

References

1. Bioavailability and bioequivalence requirements. Code of Federal Regulations Title 21, Volume 5, Pt. 320.1. Revised April 1, 2009. Access May 6, 2010.
2. Niazi, S.K. Handbook of bioequivalence testing. Informa Healthcare, New York, NY, 2007.
3. Guidance for Industry. Bioavailability and bioequivalence studies for orally administered drug products — general considerations. U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER). March 2003.
4. Gibaldi M and Perrier D. Pharmacokinetics. 2nd ed (revised and expanded). Informa Healthcare, New York, NY, 2007.
5. Purves RD. Bias and variance of extrapolated tails for area-under-the-curve (AUC) and area-under-the-moment-curve (AUMC). J Pharmacokinet Biopharm. 1992 Oct;20(5):501-10.
6. D'Argenio, D.Z., A. Schumitzky and X. Wang. ADAPT 5 User's Guide: Pharmacokinetic/Pharmacodynamic Systems Analysis Software. Biomedical Simulations Resource, Los Angeles, 2009.
7. Guidance for Industry. Bioavailability and bioequivalence studies for orally administered drug products — general considerations. U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER). March 2003.
8. Guidance for Industry. Conduct and Analysis of Bioavailability and Bioequivalence Studies - Part A: Oral Dosage Formulations Used for Systemic Effects. Canada Health Products and Food Branch Guidance Document. 1992.
9. Pecking M, Montestruc F, Marquet P, Wodey E, Homery MC, Dostert P. Absolute bioavailability of midazolam after subcutaneous administration to healthy volunteers. Br J Clin Pharmacol. 2002 October; 54(4):357–362.
10. Damle B et al. Effect of food on the oral bioavailability of UFT and leucovorin in cancer patients. Clin Cancer Res. 2001 Mar;7(3):517-23.
11. Sheiner LB, Beal SL. Some suggestions for measuring predictive performance. J Pharmacokinet Biopharm. 1981 Aug;9(4):503-12.
12. Guidance for Industry. Conduct and analysis of bioavailability and bioequivalence studies - Part A: Oral dosage formulations used for systemic effects. Canada Health Products and Food Branch Guidance Document. 1992.
13. Draft guideline on the investigation of bioequivalence. European Medicines Agency; Committee for Medicinal Products for Human use (CHMP). London, 24 July 2008.
14. Riegelman S, Collier P. The application of statistical moment theory to the evaluation of in vivo dissolution time and absorption time. J Pharmacokinet Biopharm. 1980 Oct;8(5):509-34.
15. Purves RD. Numerical estimation of the noncompartmental pharmacokinetic parameters variance and coefficient of variation of residence times. J Pharm Sci. 1994 Feb;83(2):202-5.

Figures

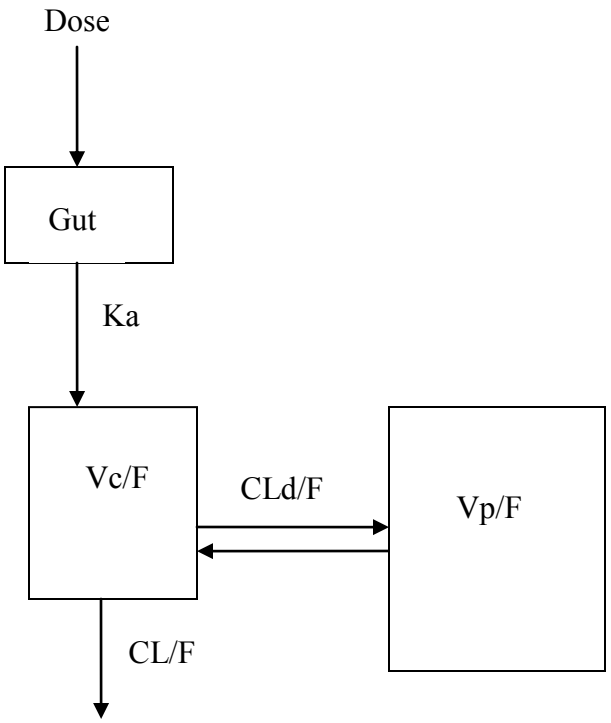


Figure 1 Model used to simulate concentrations

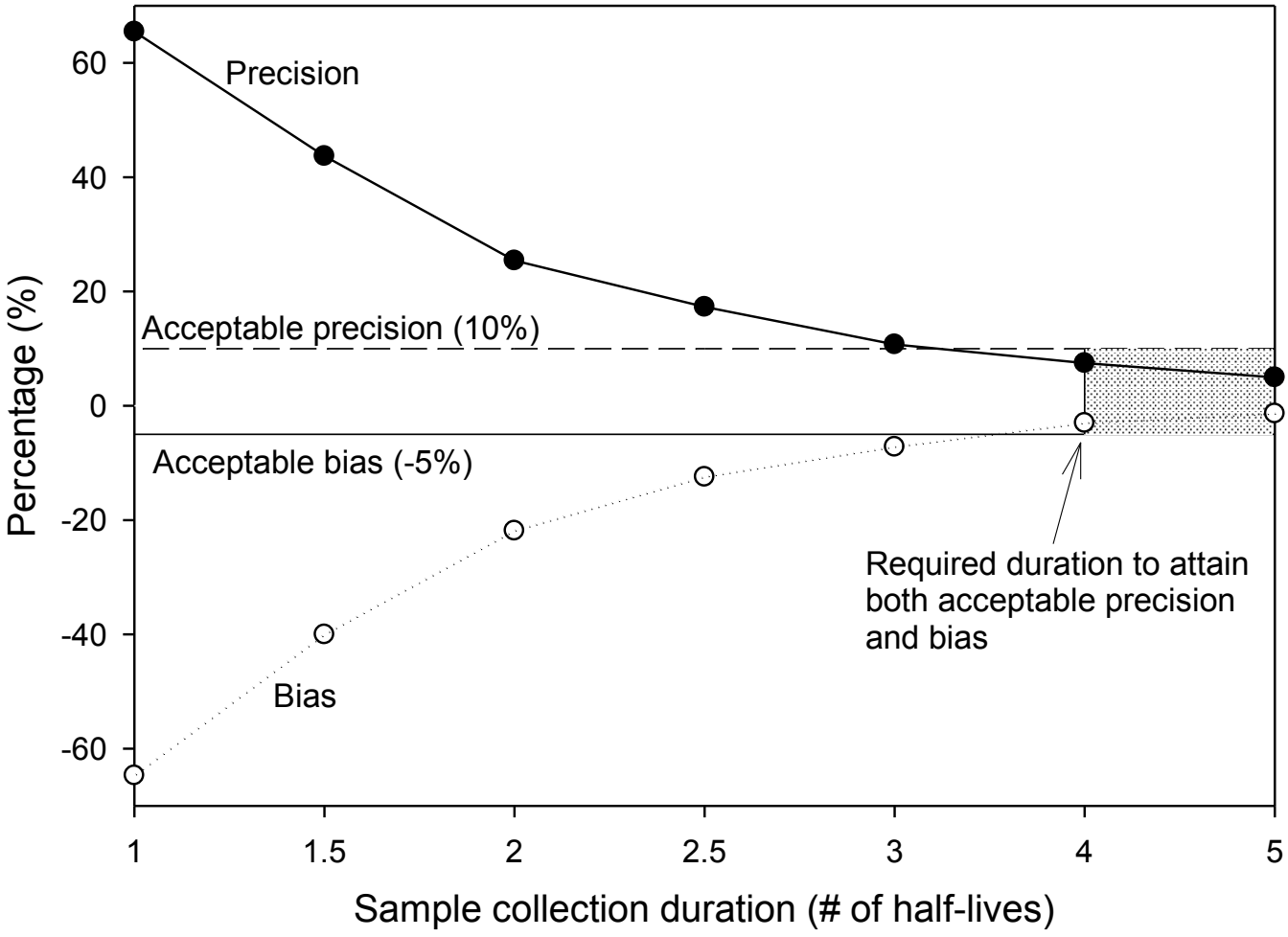


Figure 2A Bias and precision of the half-life PK parameter

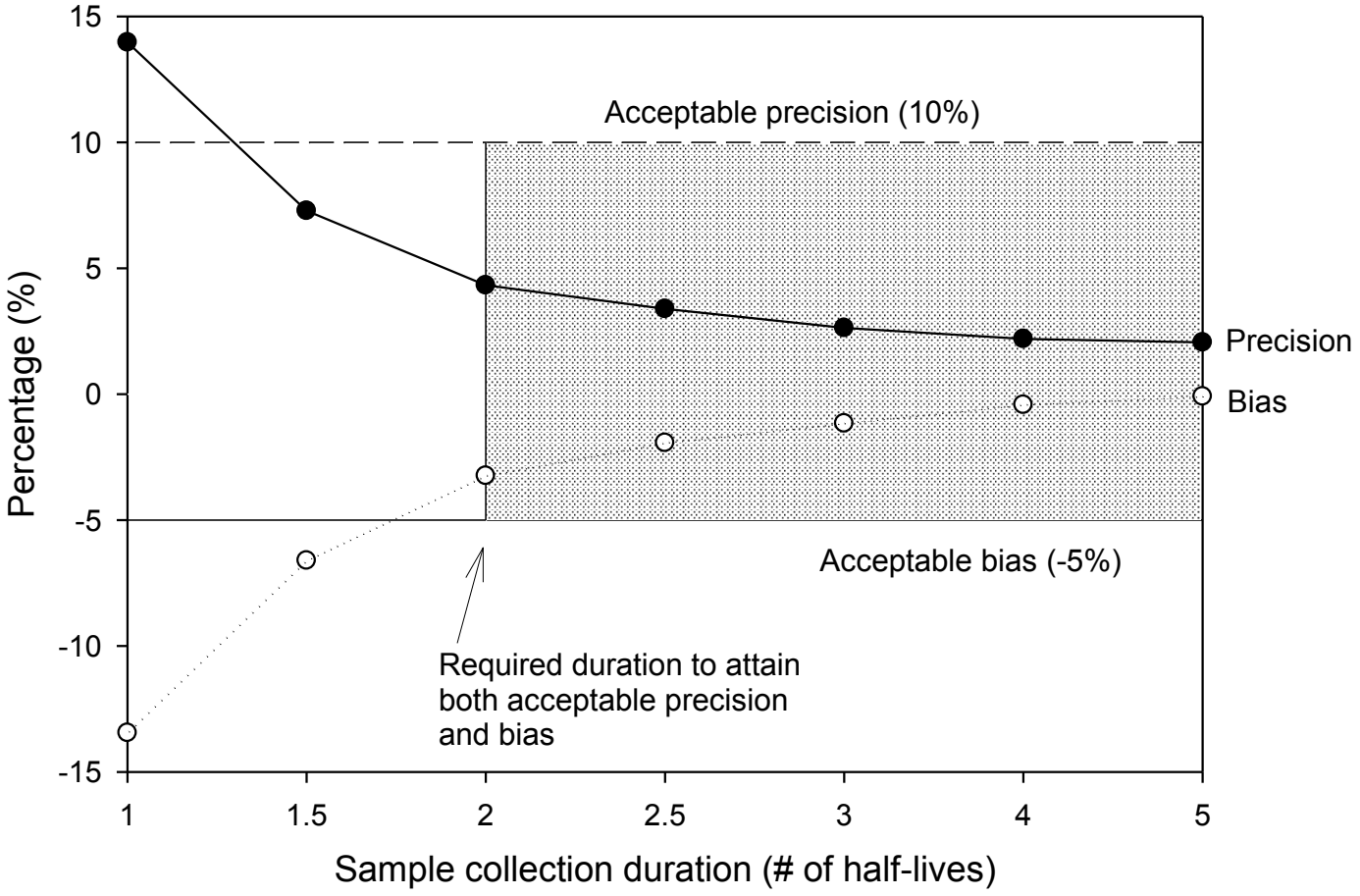
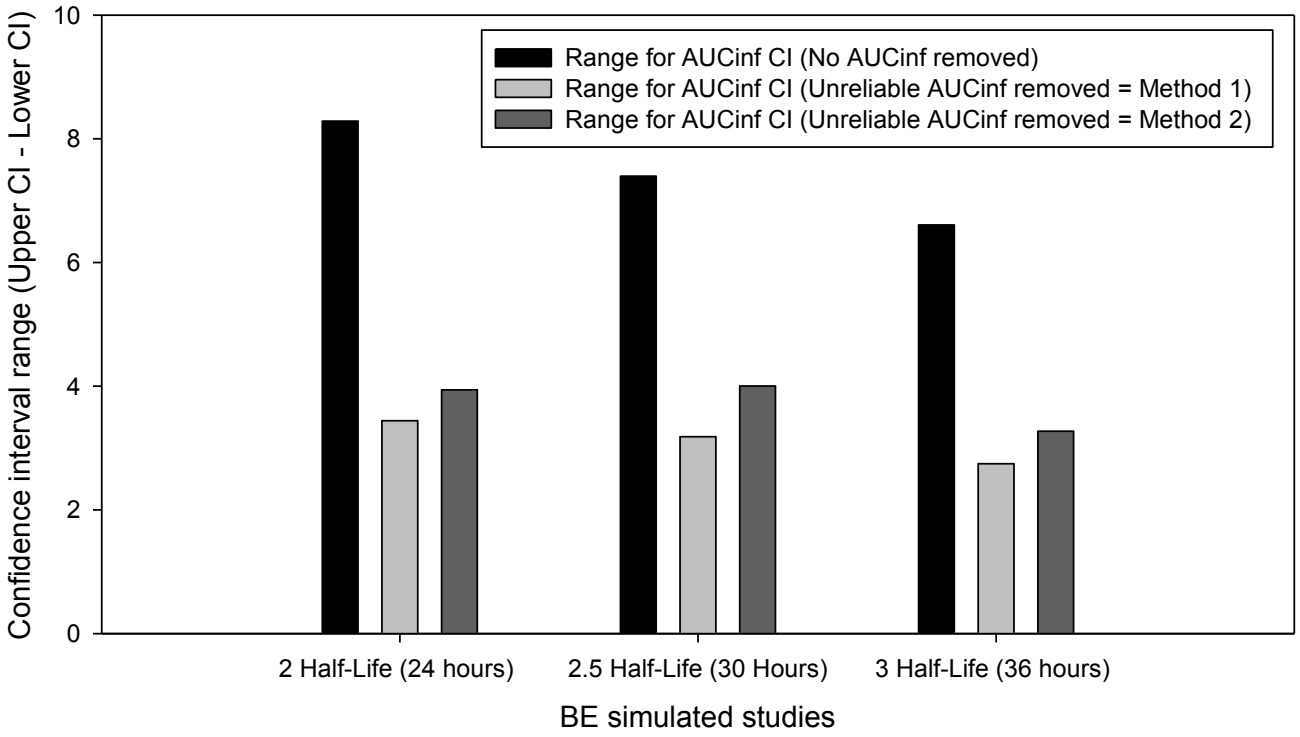


Figure 2B Bias and precision of the AUCinf PK parameter



Note1: For Method 1, AUCinf based on half-lives greater than half of the sampling scheme duration was considered unreliable.

Note2: For Method 2, AUCinf based on half-lives greater than the sampling duration over which Kel was determined was considered unreliable

Figure 3 Ln-transformed AUCinf confidence interval range for the different studies with and without unreliable AUCinf values

Tables

Table 2: Average ln-transformed AUCinf BE results from 10 studies (n=24 subjects in each study) per sampling scheme that covered only 2 half-lives (24 hours), 2.5 half-lives (30 hours) and 3 half-lives (36 hours) with and without unreliable AUCinf removed from analyses

	All AUCinf values			Without unreliable AUCinf values (Method 1)			Without unreliable AUCinf values (Method 2)		
Mean	**Ratio (%)**	**Lower CI (%)**	**Upper CI (%)**	**Ratio (%)**	**Lower CI (%)**	**Upper CI(%)**	**Ratio (%)**	**Lower CI (%)**	**Upper CI(%)**
24h sampling scheme	93.55	89.52	97.80	99.72	98.02	101.46	99.13	97.18	101.12
30h sampling scheme	95.29	91.67	99.07	99.08	97.51	100.69	98.79	96.71	100.71
36h sampling scheme	96.90	93.67	100.28	99.57	98.20	100.95	99.27	97.65	100.93

Appendix 1: Ln-transformed AUCinf BE results from 30 studies (n=24 subjects in each study) with and without unreliable AUCinf removed from analyses

24h sampling scheme

Study	All AUCinf values			Without unreliable AUCinf values (Method 1)			Without unreliable AUCinf values (Method 2)		
	Ratio (%)	**Lower CI (%)**	**Upper CI (%)**	**Ratio (%)**	**Lower CI (%)**	**Upper CI(%)**	**Ratio (%)**	**Lower CI (%)**	**Upper CI(%)**
1	92.87	87.49	98.57	100.50	98.12	102.93	98.86	96.60	101.18
2	95.91	92.84	99.08	99.84	97.51	102.22	99.05	96.88	101.28
3	93.51	90.19	96.95	99.50	98.55	100.46	99.18	97.77	100.61
4	95.80	93.40	98.27	101.76	98.95	104.65	100.48	98.12	102.89
5	88.67	84.20	93.38	99.45	97.93	101.00	99.41	97.21	101.66
6	96.91	93.25	100.71	99.41	97.89	100.96	98.54	96.46	100.67
7	92.58	90.10	95.13	97.47	96.40	98.55	97.39	95.23	99.60
8	96.89	93.21	100.71	100.24	98.73	101.77	99.86	98.69	101.04
9	88.28	80.07	97.33	99.00	97.04	100.99	99.45	97.78	101.14
10	94.09	90.43	97.90	100.05	99.05	101.05	99.09	97.07	101.15
Mean	93.55	89.52	97.80	99.72	98.02	101.46	99.13	97.18	101.12

30h sampling scheme

Study	All AUCinf values			Without unreliable AUCinf values (Method 1)			Without unreliable AUCinf values (Method 2)		
	Ratio (%)	**Lower CI (%)**	**Upper CI (%)**	**Ratio (%)**	**Lower CI (%)**	**Upper CI(%)**	**Ratio (%)**	**Lower CI (%)**	**Upper CI(%)**
11	96.06	91.39	100.98	99.56	97.65	101.51	99.27	96.62	102.00
12	95.91	92.84	99.08	98.98	96.98	101.03	97.93	95.42	100.51
13	93.59	88.60	98.86	99.20	98.03	100.38	99.20	98.03	100.38
14	98.84	97.23	100.47	99.97	98.59	101.37	99.00	97.39	100.63
15	93.29	89.43	97.32	98.73	96.62	100.89	99.69	96.16	101.28
16	95.85	92.36	99.48	99.03	97.44	100.64	99.02	97.69	100.38
17	95.07	92.92	97.28	97.38	96.28	98.49	96.67	95.21	98.15
18	97.55	94.78	100.39	99.07	97.33	100.84	97.93	94.71	101.26
19	91.78	85.93	98.03	99.05	97.23	100.90	99.16	96.88	101.48
20	94.94	91.24	98.80	99.86	98.90	100.84	100.00	98.98	101.04

Mean	95.29	91.67	99.07	99.08	97.51	100.69	98.79	96.71	100.71

36h sampling scheme

	All AUCinf values			Without unreliable AUCinf values (Method 1)			Without unreliable AUCinf values (Method 2)		
Study	**Ratio (%)**	**Lower CI (%)**	**Upper CI (%)**	**Ratio (%)**	**Lower CI (%)**	**Upper CI(%)**	**Ratio (%)**	**Lower CI (%)**	**Upper CI(%)**
21	96.10	91.94	100.44	99.26	97.51	101.04	99.45	97.65	101.29
22	95.91	92.84	99.08	99.30	97.33	101.30	99.58	97.56	101.65
23	95.39	91.21	99.77	100.33	99.35	101.32	100.00	99.01	101.01
24	99.42	97.39	101.50	100.37	99.29	101.45	99.90	98.49	101.33
25	96.71	92.71	100.88	99.37	97.50	101.28	99.73	97.90	101.60
26	98.01	94.50	101.65	99.73	98.68	100.79	100.04	98.95	101.16
27	97.40	96.17	98.66	98.31	97.14	99.50	97.83	96.85	98.81
28	97.84	95.79	9.92	100.13	98.80	101.49	98.44	96.03	100.91
29	94.42	89.73	99.35	99.04	97.55	100.55	98.12	95.57	100.74
30	97.82	94.43	101.54	99.83	98.88	100.78	99.64	98.53	100.76
Mean	96.90	93.67	91.28	99.57	98.20	100.95	99.27	97.65	100.93

3. ARTICLE #3

Improved Drug Development and Equivalence Potential of a New Extended-Release Formulation Determined by Clinical Trial Simulations

3.1. INTRODUCTION

The article presented in this section provides an example of how drug development can be improved by using clinical trial simulations, which were performed using information acquired from the first two articles. A pharmaceutical company was developing a supergeneric drug. The new compound was formulated to reduce the frequency of administration of an oral suspension non-steroidal anti-inflammatory drug (NSAID) from four times a day to two times a day. A pilot study was performed and the 90% confidence intervals were within the typical equivalence boundaries. However, there were some issues with the design of the pilot study and a potential food effect was observed for the reference product. Clinical trial simulations were undertaken to determine the equivalence potential of the two products with and without the observed food effect while correcting for the study design flaws of the pilot study. Modeling and simulation have a greater chance of success at attaining their objectives when all knowledge of the drug is used. Thus, literature data was available for the NSAID compound and was used as prior information for the PK model. The innovator drug was an immediate release compound. It is quickly absorbed with a Tmax at approximately 1 hour and bioavailability of over 80%. Protein binding is above 98%. Exposure increases in a dose dependent fashion. The drug is eliminated by CYP2C9 followed by glucuronidation. It is mostly eliminated in the urine as metabolites. Food effect in the literature was unclear. Some studies reported a reduction in exposure while others did not.

Unlike simple generics which are formulations that are manufactured to copy as closely as possible the brand products, supergenerics try to improve on the innovator compound. In that aspect, companies manufacturing supergenerics can be regarded as innovator companies. In fact, patents are sought after and market exclusivity is often awarded to companies that succeed in developing a supergeneric. These can differ from the original product in formulation or method of delivery. However, if a supergeneric compound can show similar concentration-time profiles to the reference product, the company can file an abbreviated new drug application based on the aforementioned reference product and avoid many pre-clinical and clinical studies. It is therefore advantageous for a company to try and avoid expensive studies and produce drugs that are an improvement to the older formulations while still respecting the principles of generic drug development and bioequivalence. This was another reason for which clinical trial simulations were performed for this study. The aim was to help the formulation experts understand the differences in absorption processes between the test and reference products and determine if changes in the supergeneric formulation were required to meet the equivalence criteria.

The absorption of orally administered drugs is often affected by external factors such as co-administration of drugs or food. These can lead to significant changes in the concentration-time profiles of the subjects. This can also cause some difficulties in the comparison of a supergeneric to an innovator compound if these compounds are administered at different times and under different conditions. For example, with the study presented in the article, the supergeneric was administered twice a day making it

easier to administer it relatively to meals versus the reference compound that had to be administered four times a day.

Effects on absorption pharmacokinetics can be divided into different categories. First, a delay in absorption can occur. This will translate into longer lag times (greater Tmax) and smaller Cmax. When modeling, this change in PK is reflected in longer lag times (Tlag) and smaller absorption rate constants (Ka). A delayed absorption may also be associated with a reduction in the bioavailability of a formulation. A decrease in bioavailability results in lower AUCinf and Cmax values, which can be observed with the co-administration of food. With a decrease in bioavailability, the PK model would require a parameter that lowers the bioavailability of an administered dose following the consumption of food. The opposite is also possible with food intake. Food can increase absorption, thereby increasing the bioavailability. This will increase AUCinf and Cmax and possibly lower the lag time prior to the start of absorption. When modeling, this would translate into an increased Ka value and a parameter increasing bioavailability after a meal. These changes in PK parameters are obvious on a concentration-time profile. However, when performing clinical trial simulations, changes in the inter-subject variability of the PK parameters are also important, because between subject variability can also be affected by food intake. This is less obvious from looking at the profiles. Clinical trial simulations require both the population PK parameters and their variances. Therefore, it is important to also model the changes in the variance parameters to have accurate predictions. Absorption that is significantly increased by food tends to have lower variability while the opposite is usually true with variability in PK parameters increasing when food decreases significantly the absorption.[147,148]

As will be demonstrated in the following article, clinical trial simulations were used together with previous knowledge to answer questions pertaining to a drug's bioequivalence potential and food effect.

3.2. ARTICLE

Title

Improved Drug Development and Equivalence Potential of a New Extended-Release Formulation Determined by Clinical Trial Simulations

Authors: Philippe Colucci, M.Sc.[1,2]; Turgeon J, PhD[1,3]; and Murray P. Ducharme; PharmD[1,2]

Institutional affiliation:

1) Faculté de Pharmacie, University of Montreal, Montreal, Canada
2) Cetero Research, Cary, North Carolina
3) CRCHUM, Centre Hospitalier de l'Université de Montréal, Montreal, Canada

Key Words: Modeling, Clinical Trial Simulations, Bioequivalence; Food Effect

Abstract

Background: The benefits of modeling and simulation in drug development is well documented, however its use in generic and supergeneric drug development is still in its early stages. The purpose of this work is to illustrate the benefits of modeling in the supergeneric drug development. More precisely, a model will be developed to describe the PK of two formulations of a drug and using this model, clinical trial simulations will be performed to assess the equivalence potential of the new formulation versus a reference one. Based on simulation results, recommendations will be made pertaining to study design or formulation.

Methods: A compartmental model was built using data from a pilot study. Using this model, clinical trial simulations were performed under different conditions in order to determine if a supergeneric compound would be considered equivalent to a reference product in pivotal single-dose and steady-state studies. Individual subject concentration data were obtained from these simulations to determine noncompartmental pharmacokinetic (PK) parameters. Analyses of variance were determined for each study and predicted confidence intervals for pivotal PK parameters were obtained.

Results: The products were well-described by a 2-compartment model with different absorption processes for each formulation. Simulations suggested that it would be difficult to prove equivalence between the two formulations at steady-state and possibly at single dose. Based on the modeling, the absorption half-life that corresponded to the sustained-release portion of the test formulation was too long. Modifications to the absorption process of the test formulation were postulated and further simulations

indicated that these modifications would increase the chance of meeting the equivalence criteria.

Conclusion: Concentrations for both formulations were well described by the final model and improved the understanding of the drug formulations. Results of the clinical trial simulations played a critical part in the plans to modify the test formulation, thereby avoiding unnecessary studies and reducing overall development costs.

Introduction

Currently, a low percentage of medications is successfully commercialized in the United States, in stark contrast with the last decade.[1] The Food and Drug Administration (FDA) has recognized this problem and has published a document entitled "Challenge and Opportunity on the Critical Path to Medical Products"[2] asking stakeholders to improve the development process to help bring new medications more efficiently to patients who require them. One method to improve drug development suggested in the document was to utilize tools such as pharmacokinetic (PK) and pharmacodynamic (PD) modeling. Sheiner had previously published on the use of multiple cycles of predictive and confirmatory PK modeling (learn and confirm) throughout the drug development process.[3]

Occasionally in drug development, results are interpreted using a "check box" approach (i.e., do the results answer the regulatory criteria). This will often lead to an inadequate understanding of the drug and to surprising and most times unwanted results in later studies. In contrast, multiple cycles of PK modeling allow us to use all available relevant information in order to have a better understanding of the drug being developed. Available data can be obtained from a current study, previous studies or the literature. Once models have been established, they can predict future outcomes with clinical trial simulations, help better understand the drug, avoid uninformative studies, minimize unnecessary exposure to subjects and reduce drug development costs.

The use of multiple cycles of predictive and confirmatory PK modeling has been demonstrated to be useful in optimizing drug development.[4,5,6] Furthermore, PK or PD modeling is not only limited to innovator products but can also be useful for the

development of generic products. However, few published examples of the value of PK/PD modeling exist for generic or supergeneric compounds. Most examples are for drugs for which noncompartmental analyses have limited usefulness such as data in patients or when systemic blood draws are not collected.[7,8,9] However, modeling could and should be used throughout generic and supergeneric drug development to ensure a better understanding of the drug being developed.

This paper illustrates the benefits of PK modeling in the development of a supergeneric drug. A supergeneric of an NSAID was being developed to be administered every 12 hours. In contrast, the reference product is administered every 6 hours. Noncompartmental results from a pilot study were promising as ratios and 90% confidence interval (90% CI) for ln-transformed AUC0-t, AUCinf and Cmax were within the FDA equivalence criteria of 80-125%. Therefore, the normal development process would have been to plan for the definitive studies using these results and this formulation. However, upon closer inspection of the results of the pilot study, it was determined that some information/results required further attention to determine the impact on the future development of the drug. The first issue was that concentrations after the second administration of the reference formulation were lower than after the first dose, implying a potential food effect which influenced the absorption process even though lunch was served two hours prior to the second administration. The second observation was that there were a total of six additional blood draws taken within the first two hours of administration at Time 0h in comparison with the administration at Time 6h. It was possible that the Cmax and AUC of this second dose administration were not properly

characterized. Lastly, 24 hours of sampling for the test formulation was insufficient (i.e., AUC0-t/AUCinf < 80%).

Although the results of the study were positive from a relative bioavailability perspective, it was unclear if the test product would be equivalent to the reference product under pure fasting (without a meal given at 4 hours or food effect from 4-hour meal removed) or fasting/fed conditions (conditions similar to the pilot study where concentrations after second reference administration are lower), and whether or not steady-state equivalence conclusions would be similar to the pilot single dose results. To better understand the pharmacokinetic behavior of the drug, modeling and simulations were therefore undertaken with data from the pilot study. The objectives of the modeling and simulations were:

1) Develop a model to describe the PK of the test and reference formulations
2) Perform clinical trial simulations of different scenarios using the final model to determine if the two products would be equivalent in a single-dose and steady-state study design under fasting/fed or true fasting conditions
3) Help design definitive studies if clinical trial simulations show that the two products are probably equivalent
4) Determine what modifications could be made to the test formulation to improve chances of meeting the equivalence criteria if simulations show that the two products are probably not equivalent

Methods

Brief Study Design of Pilot

The design of the pilot study was a single-dose, randomized 2-way crossover fasted equivalence study where a dose of a supergeneric extended-release oral suspension was compared in terms of relative bioavailability to the same equivalent dose of the reference formulation (eg given twice over 12 hours). The first treatment was a single extended-release oral suspension dose (equivalent to twice the reference dose) of a test formulation (supergeneric) administered at Time 0 after an overnight fast. The second treatment consisted of two immediate-release doses of a reference product. The first dose was administered at Time 0 and the second administration was 6 hours after the first which corresponded to two hours after the start of lunch. The total daily dose administered was identical for both formulations and the washout between periods was seven days. Plasma samples for the parent compound were collected over 24 hours (0, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20 and 24 hours). A total of 11 healthy subjects (men and women) completed both periods and data from these subjects were used in the modeling and clinical trial simulations.

Model Buildup and Discrimination

All modeling and simulations were performed using ADAPT 5®.[10] Individual analyses were first conducted using maximum-likelihood. Model discrimination was based on the minimization of the average Akaike information criterion and the Negative Log Likelihood score as well as the maximization of the R^2 of the fits. Visual inspection

of figures was used to confirm the appropriateness of the model by determining that there were no trends in the observed versus predicted concentrations or in the residual graphs and by ensuring that Cmax, distribution and the elimination phases were properly characterized in individual graphs. A population PK analysis was then performed with the final model by the iterative two-stage method (ITS) using priors obtained from the individual maximum-likelihood analysis in order to obtain the most accurate population PK parameters, variance, residual variability and individual results.

Compartmental PK analyses were performed using the plasma parent data from all available subjects in the pilot study. Concentrations below the limit of quatitation were set to missing. All systemic concentrations were modeled using a weighting procedure of $W_j=1/S_j^2$ where the variance S_j^2 was calculated for each observation using the equation $S_j^2=(a+b*Y)^2$. The parameters a and b are the intercept and slope of the variance model. The slope is the residual variability proportional to each concentration and the intercept is the additive component of the error. Three different models were created: a) one for the reference formulation, b) one for the test formulation and c) one combining both formulations.

Model for reference formulation

A model was first constructed based only on the reference data. Doses and concentration data were simultaneously modeled. Basic 1-, 2- and 3-compartment models were tested for their ability to predict the observed concentrations. Although this product was an immediate-release formulation that was absorbed quickly, multiple absorption peaks occurred following dosing. Therefore, different absorption models were tested

(e.g., single first-order absorption rate_Ka, multiple Ka which were identical after each administration or multiple Ka which were different after each administration). Concentrations after dosing at Hour 6 were lower than the concentrations measured after the initial dose which suggested a food-effect. Therefore, a change in bioavailability ($F_{2nd\ dose}$) was also included as a parameter to describe the second dose.

Model for test formulation

The parent concentration data was also fitted after the test formulation was administered. Like the model for the reference formulation, basic 1-, 2- and 3-compartments were also evaluated for the test formulation. Because this was an extended-release formulation, different absorption rates were tested for inclusion in the model (single to multiple absorption peaks using the same Ka and multiple absorption peaks using different Ka). Only one dose of this formulation was administered; therefore, no parameter explaining a change in bioavailability was required.

Model combining both formulations

In order to achieve results consistent with a crossover design, the test and reference models were combined. Combining the models ensured that all subjects had the same parameters for volume of distribution and elimination regardless of the formulation. However, each formulation had a different absorption process. This model reflected the actual conditions of the study.

Population analysis was performed on this final combined model.

Clinical Trial Simulations

Using the PK parameters and variances from the final combined model, clinical trial simulations were performed using a Monte Carlo technique in Adapt 5® for the following study designs: single-dose under true fasting conditions (ignoring the food effect after the second reference administration), steady-state under true fasting conditions, single-dose under similar conditions as the pilot (fasting/fed for reference formulation) and steady-state under similar conditions as the pilot (fasting/fed for reference formulation).

Concentrations for the single-dose study were collected over 24 hours. The times chosen corresponded to the times that would be used in a single-dose equivalence study between the two products, with the same number of plasma samples after each administration of the reference formulation. For the steady-state design, identical concentration time points up to and including the 12-hour time point were simulated. Concentrations were only simulated for a 12 hour period as this corresponds to the dosing interval for the test formulation.

For each study design, 10 studies varying from 20 to 42 subjects were simulated in order to determine a range of expected equivalence criteria results.

Noncompartmental PK Analysis

Using simulated concentrations, standard noncompartmental analysis with SAS® (Version 9.1.3) was performed to determine the area under the curve from time 0 to 24 hours post-dose (AUC 0-t for single-dose), area under the curve to infinity (AUCinf for single-dose), area under the curve during the dosing interval (AUC0-τ for steady-state), maximum concentration (Cmax single-dose and steady-state) and minimum

concentration (Cmin for steady-state) for each subject. Using these noncompartmental PK parameters, an analysis of variance (ANOVA) with the Proc Mixed model in SAS® was performed for each study to calculate the ratio and 90% confidence intervals (CI) for the primary PK parameters (AUC0-t, AUCinf, AUC0-τ, Cmax and Cmin). The ANOVA model included formulation as a fixed effect and subject as a random effect. ANOVA were conducted to determine if the 90% confidence intervals would meet the FDA equivalence criteria of 80-125%. This was also performed for simulated studies under true fasting conditions (i.e., food effect after the second reference administration was removed).

Results

The ratios for AUC0-t, AUCinf and Cmax from the original noncompartmental analysis of the pilot study were 97.7, 112.4 and 103.5%, respectively. The estimated terminal half-lives for the reference and test formulations were 3 and 11 hours, respectively. Intra-subject variability was low and all 90% confidence intervals fell within the 80-125% acceptance criteria.

Compartmental analyses were performed on the available data from the pilot study. Based on the model discrimination process, a 2-compartment model was chosen. This was not surprising as the parent compound exhibited a biphasic profile following the administration of both the reference and test formulations. The final model included a central compartment and a peripheral compartment with a clearance from the central compartment. It had the following PK parameters: volume of the central compartment

(Vc/F), volume of the peripheral compartment (Vp/F), total clearance (CL/F) and distributional clearance (CLd/F).

For the reference compound, the absorption process included a rate of absorption (Ka) that was different after each dose; however, three absorption peaks appearing at different times were required to properly fit the data. The addition of multiple absorption peaks to the model reduced the Akaike's Information Criterion value from 135.8 to 94.0 and the residual variability from 18.7% to 8.0%. Each absorption peak was associated with a different lag time (time before the start of absorption for each peak) and fraction of the dose. Therefore, the model had the following PK parameters: 2 absorption rate constants (Ka1 and Ka_6h), 6 different lag times (Tlag1, Tlag2, Tlag3, Tlag1_6h, Tlag2_6h and Tlag3_6h) and a change in bioavailability for the second dose ($F_{2nd\ dose}$). The $F_{2nd\ dose}$ represented the relative bioavailability of the reference product under fed conditions versus its bioavailability under fasting conditions.

As is often the case with extended-release formulations, different absorption rates were required to fit the multiple absorption peaks of the test formulation. The absorption process included 3 different rates of absorption (Ka1, Ka2 and Ka3) each associated with a different lag time (Tlag1, Tlag2 and Tlag3). A parameter for relative bioavailability (Frel) of the test formulation (with respect to the reference formulation) was also required. The final combined modeled is depicted in Figure 1.

Goodness of fit plots are presented in Figures 2 and 3. As demonstrated by these plots, the model adequately describes all observed concentrations. Predicted versus observed concentrations were randomly scattered around the line of identity. No trends were observed with respect to the standardized weighted residuals. In addition, the

residual variability was only 8.4%, further demonstrating that the model adequately described the observed concentrations.

The estimated PK parameters for the final model are presented in Table 1. The relative bioavailability for the test reference was 129% which would make total exposure outside of the usual accepted equivalence range of 80-125%. The main difference between the two compounds was in the absorption process. The absorption rate constants are summarized in Table 2. The absorption of the test formulation was characterized by two different absorption rates; the absorption half-life for one of these absorption rate constants was approximately 15 hours. Fifty percent (50%) of the dose was absorbed with a sustained-release (SR) rate (*i.e*, slower rate) while the remaining 50% of the dose was absorbed with an immediate-release (IR) rate (*i.e.*, faster rate). The immediate-release portion was described by two peaks (one Ka, two lag times) with the first peak accounting for 40% and the second for 10% of total dose. Both peaks started within 1.7 hours of administration and lasted approximately 2 hours. The sustained-release peak started at 0.7 h and would theoretically last 108 hours.

Clinical trial simulations were conducted for single-dose and steady-state equivalence studies under true fasting conditions and with the same conditions as the pilot study (*i.e.*, fasting/fed for reference formulation). Average ANOVA results are tabulated in Table 3. When comparing the single dose simulation under the same conditions as the pilot study, the ratios for AUC0-t, AUCinf and Cmax were similar to the original noncompartmental results; however Cmax was slightly lower. If the food effect was removed, then the ratio for both simulated AUC0-t and Cmax decreased. Although the ratio for AUC 0-t under a fasting/fed single-dose design was approximately

100%, the fasting/fed steady-state ratio for AUC 0-τ was significantly above 100%. Therefore, simulations suggested that it would be difficult to prove equivalence between the two formulations at steady-state (Cmin) and possibly at single dose for Cmax (true fasting condition) or AUCinf (fasting/fed conditions).

Based on these simulation results and the differences in the absorption process that existed between the formulations, further clinical trial simulations were undertaken to determine the likelihood of meeting the BE requirements if modifications were made to the test formulation absorption process. Two options were recommended to increase the chances of meeting the equivalence criteria. The first option was to delay the sustained release portion of the formulation by 2 hours and to increase its rate by 6.5 fold. The second option was to have a portion of immediate-release being absorbed at 6 hours, to delay the absorption of the sustained release portion of the formulation by 2 hours and to increase its absorption rate by 2.3 fold. Figures 4 and 5 depict the predicted mean test concentrations for these two options versus the observed reference concentrations.

Discussion

Although the drug name was not specified for confidentiality purposes, the objective of this article was still fulfilled, which was to demonstrate the utility of modeling and simulations in order to better develop generic and supergeneric formulations. Furthermore, the PK results reported for this NSAID agent are consistent with what is known about its pharmacokinetic characteristics.

When two formulations of the same drug are given to the same subjects on two different occasions, only the rate and extent of absorption differ between the two products

to possibly cause significant changes to a profile, while the distribution and elimination process remain constant because they are drug and not formulation specific processes.[7,11] Therefore, to avoid having different mean population and variance parameters for the distribution and elimination processes between the two formulations, the final model fitted both formulations simultaneously. This reduced the variability in the results, allowed the use of a relative bioavailability parameter for the test formulation and reflected the actual conditions of the study. However, prior to reaching this step, the two formulations were fitted separately. This was done to permit a better understanding of the absorption process for each formulation.

Clinical trial simulations conducted under the conditions of the pilot study (i.e., change in bioavailability after the second dose for reference) reflected what was observed in the pilot study. The original noncompartmental ratio for AUC0-t and Cmax were approximately 100% and the ratios estimated by compartmental analysis were also close to 100%. However, if the change in relative bioavailability for the reference formulation was not taken into consideration, then the ratios for both AUC0-t and Cmax decreased. This is understandable, as observed concentrations under true fasting conditions (first dose) were greater than under fed conditions (second dose). Therefore, removal of $F_{2nd\ dose}$ parameter increased the simulated reference concentrations after the second dose, leading to a greater Cmax and AUC for the reference formulation. Conversely, increased reference parameters caused a reduction in the ratios (i.e., ratio = test/reference).

Although the ratio for AUC0-t under a fasting/fed single-dose design was approximately 100%, the simulated fasting/fed steady-state ratio for AUC0-τ was significantly above 100%. The ratio of approximately 117% for AUC0-τ at steady-state is

similar to the single dose AUCinf ratio. The AUCinf ratio was greater than the AUC0-t ratio due to significant amount of AUC extrapolated for the test formulation. This demonstrated that there was still a significant amount of the test formulation still being absorbed at the end of the profile (24 hours post-dose) while the reference product was completely absorbed. For the test formulation, this translates into accumulation and a high ratio for AUC0-τ.

The absorption half-life corresponds to the time for half of the administered dose to be liberated and absorbed into the systemic circulation. The absorption half-life for the reference formulation was short under both fasting and fed conditions although multiple absorption peaks were included in the model. Food slightly delayed the lag times of these absorption peaks and increased the absorption half-life. However, the absorption half-life for the reference formulation was still short even after lunch and no flip-flop kinetics was observed with the modeling. On the other hand, three different absorption rates were required to model the absorption process for the test product with two absorption rates similar to those observed with immediate-release reference formulation and a third absorption rate that corresponded to the sustained portion of test formulation which had a half-life that was approximately 20 to 135 times longer than the reference absorption half-life. This led to flip-flop kinetics for the test formulation which explained why the apparent terminal half-life was much longer for the test formulation using the noncompartmental analysis. Furthermore, approximately 50% of the test product was liberated at this rate and after 24 hours, approximately 20% of the formulation was still being absorbed.

An advantage of clinical trial simulations is the possibility of determining the impact that a change to an aspect of the study will have on the study conclusions.[12,13] For example, clinical trial simulations have been used to determine the impact on a drug's profile if renally impaired volunteers are included instead of healthy volunteers.[14,15] In this study, clinical trial simulations showed that in order to optimize the chances of meeting bioequivalence requirements, the test formulation would require some modification. Based on the compartmental analyses, the absorption of the test dose associated with the sustained release rate was the main difference between the two products. Simulations of different proposed modifications to the test product were performed to optimize chances of meeting the equivalence criteria. With the formulation scientists' input on the feasibility of the proposed changes, two proposed modifications were retained. Both the proposals required a delay in the start of the sustained-release rate and an increase in this rate. Clinical trial simulations using the proposed changes to the test product suggested that it would meet all equivalence criteria under both a single dose and steady-state design under fasting/fed conditions. For these clinical trial simulations, the fasting/fed study design was used based on literature studies performed with the reference formulation which showed a constant food effect.

A limitation to this work is the lack of knowledge on the effect of food on the relative bioavailability of the test formulation administered at steady-state. Clinical trial simulations discussed in this paper were performed assuming no food effect on the test formulation at steady-state. If a food effect is observed with the test formulation in future multiple dose studies, it is assumed that the AUC0-t, AUCinf and Cmax would be lower; therefore lowering the ratio and 90% CI for these parameters. Another limitation is the

lack of concentration data collected after 24 hours for the test formulation. This did not allow the proper characterization of the slow absorption rate for the test formulation. Physiologically, it is possible that absorption stopped prior to the 108 hours predicted by the modeling. If absorption would cease prior to the complete absorption of the test formulation, the ratio for AUCinf would be decreased. It is therefore possible that the situations described here could be associated with lower ratios for PK parameters, however results would still indicate that changes need to be brought about in order to meet BE criteria. Thus, none of these limitations affect the conclusions about the differences in the profiles and absorption process between the two formulations. It is clear that in order to achieve more comparable profiles, the test formulation would require changes.

Conclusions

Concentrations for both formulations were well described by a 2-compartment model with 3 absorption phases. The results of the modeling were similar to the results of the noncompartmental analysis which supports the selection of the final model. Modeling and simulations permitted a better understanding of the potential food effect and the differences between the two formulations.

Clinical trial simulations suggested that it would be difficult to prove equivalence between the two formulations under a fasting single-dose design or fasting/fed steady-state design. To improve the chances of meeting the equivalence criteria, the test formulation had to be modified to be more similar to two doses of the reference product.

Results of these clinical trial simulations played a critical part in the plans to modify the test formulation. Once the formulation is modified, *in vitro* dissolution information will be used to predict *in vivo* bioequivalence potential.

References

1. Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov. 2004 Aug;3(8):711-5.
2. Challenge and Opportunity on the Critical Path to Medical Products. U.S. Department of Health and Human Services Food and Drug Administration. March 2004. Available at http://www.fda.gov/downloads/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm113411.pdf
3. Sheiner, LB. Learning versus confirming in clinical drug development. Clin Pharmacol Ther. 1997 Mar;61(3):275-91.
4. Mandema JW, *et al*. Model-based development of gemcabene, a new lipid-altering agent. AAPS J. 2005 Oct 7;7(3):E513-22.
5. Lockwood P, Ewy W, Hermann D, Holford N. Application of clinical trial simulation to compare proof-of-concept study designs for drugs with a slow onset of effect; an example in Alzheimer's disease. Pharm Res. 2006 Sep;23(9):2050-9.
6. Karlsson MO, *et al*. Pharmacokinetic/pharmacodynamic modelling in oncological drug development. Basic Clin Pharmacol Toxicol. 2005 Mar;96(3):206-11.
7. Colucci P, Marier JF, Ducharme MP. 'Chapter 8 – Population Pharmacokinetic Approaches for Assessing Bioequivalence' in the book 'Population PK approaches for assessing bioequivalence. In; Kanfer I, Shargel, eds. Generic Drug Product Development – Bioequivalence'. Informa Healthcare; Ney York, New York, 2008.
8. Fradette C, Lavigne J, Waters D, Ducharme MP. The utility of the population approach applied to bioequivalence in patients: comparison of 2 formulations of cyclosporine. Ther Drug Monit. 2005 Oct;27(5):592-600.
9. Tsai JC, Cheng CL, Tsai YF, Sheu HM, Chou CH. Evaluation of in vivo bioequivalence methodology for topical clobetasol 17-propionate based on pharmacodynamic modeling using Chinese skin. J Pharm Sci. 2004 Jan;93(1):207-17.
10. D'Argenio, D.Z., A. Schumitzky and X. Wang. ADAPT 5 User's Guide: Pharmacokinetic/Pharmacodynamic Systems Analysis Software. Biomedical Simulations Resource, Los Angeles, 2009.
11. Guidance for Industry. Bioavailability and bioequivalence studies for orally administered drug products — general considerations. U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER). March 2003.
12. Holford NH, Kimko HC, Monteleone JP, Peck CC. Simulation of clinical trials. Annu Rev Pharmacol Toxicol. 2000;40:209-34.
13. Bonate PL. Clinical trial simulation in drug development. Pharm Res. 2000 Mar;17(3):252-6.
14. Turpie AG, Lensing AW, Fuji T, Boyle DA. Pharmacokinetic and clinical data supporting the use of fondaparinux 1.5 mg once daily in the prevention of venous thromboembolism in renally impaired patients. Blood Coagul Fibrinolysis. 2009 Mar;20(2):114-21.
15. Piotrovsky V, Van Peer A, Van Osselaer N, Armstrong M, Aerssens J.

Galantamine population pharmacokinetics in patients with Alzheimer's disease: modeling and simulations. J Clin Pharmacol. 2003 May;43(5):514-23.

Figures

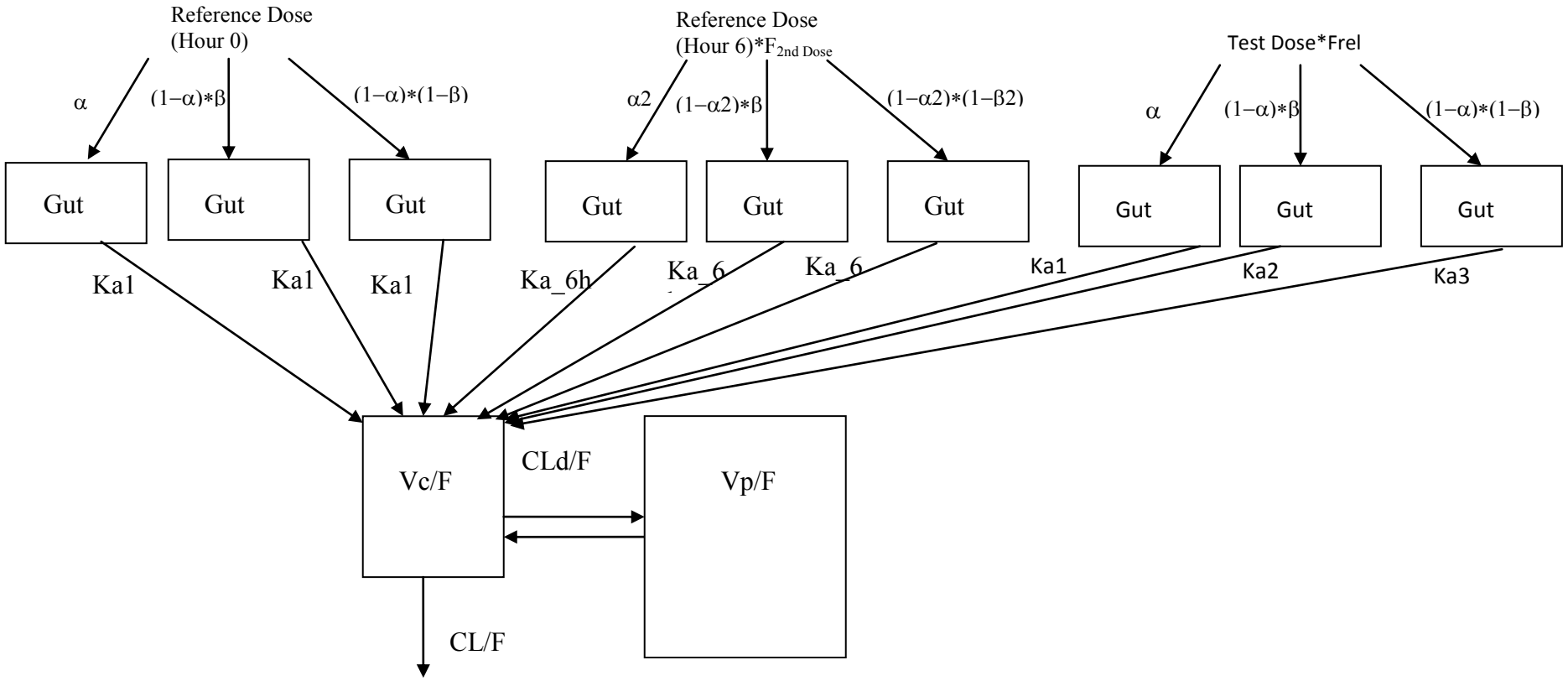


Figure 1 Final model used to simulate concentrations

Legend: Ka is an absorption rate constant; α and β are the portion of the available dose associated with an absorption process; F_{2nd} Dose is the relative bioavailability of the 2nd reference dose; Frel is the relative bioavailability of the test versus the reference product; Vc/F is the volume of the central compartment (reference); Vp/F is the volume of the peripheral compartment (reference); CL/F is the total clearance (reference); CLd/F is the distributional clearance (reference)

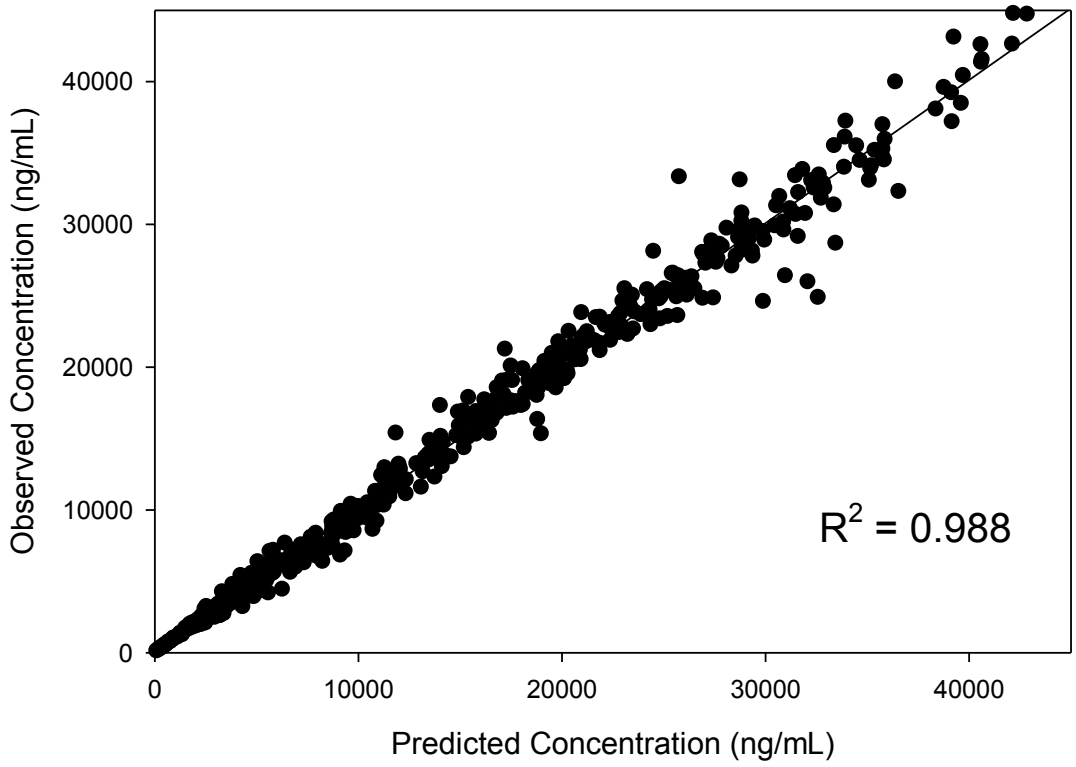


Figure 2 Observed versus predicted concentrations

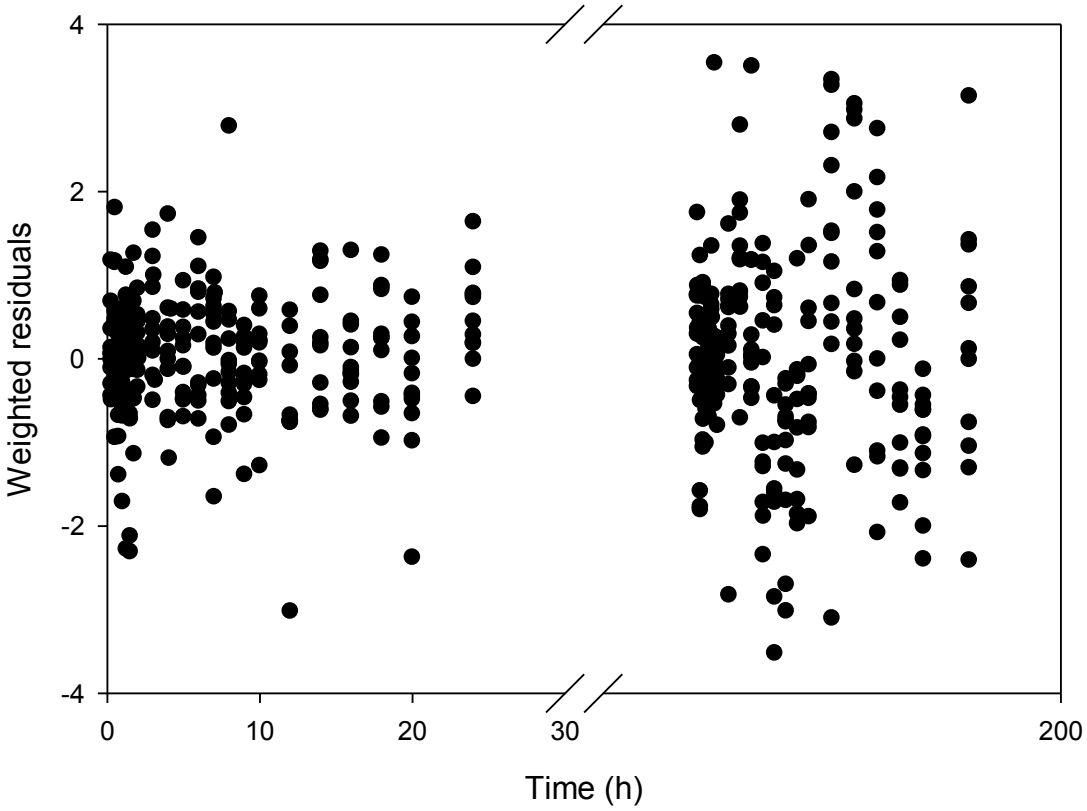


Figure 3 Weighted residual versus time

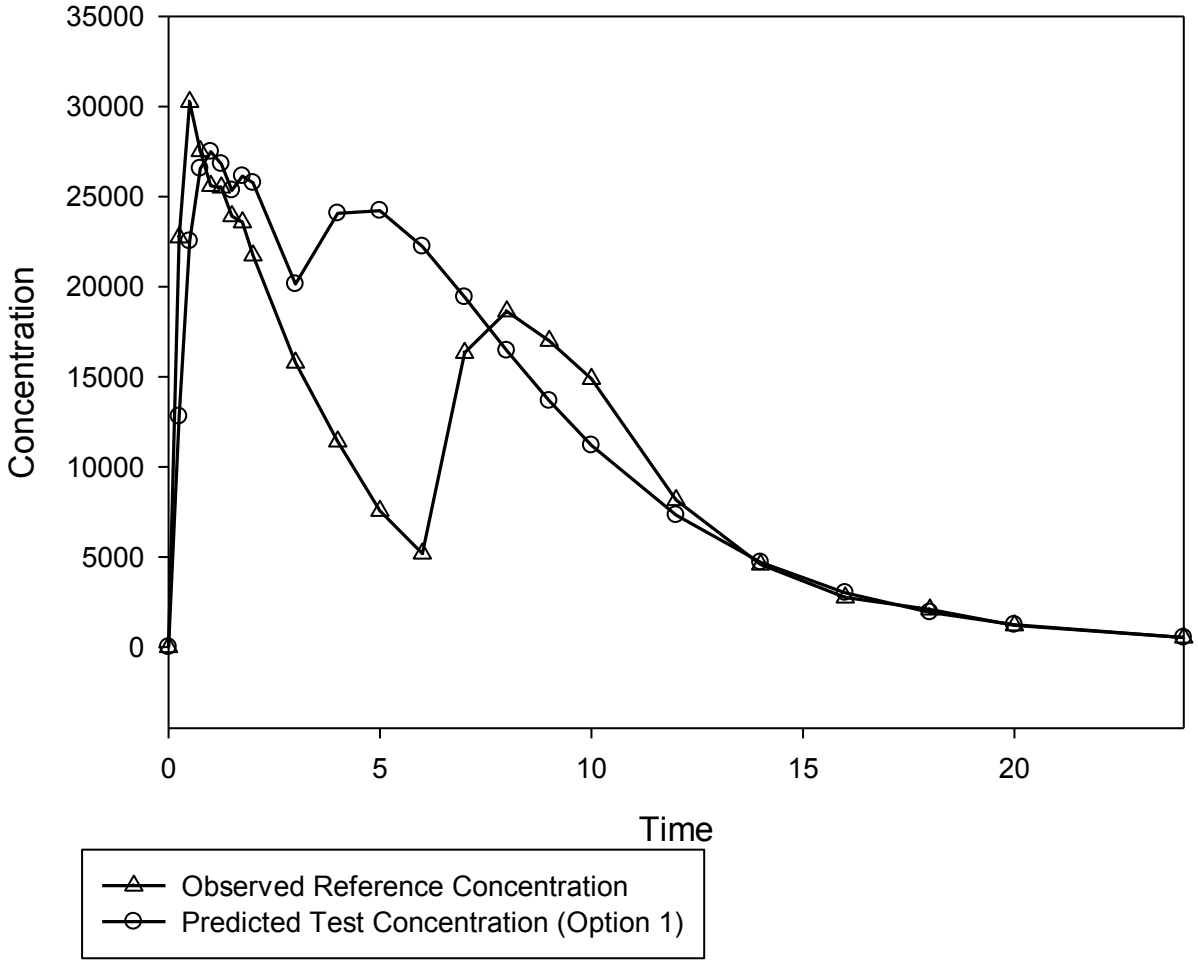


Figure 4 Simulated test formulation profile using Option 1 modification versus observed reference concentrations

Note: This simulation is for a single-dose study. The test formulation is simulated as an administration at Time 0 hour while the reference formulation is simulated as an administration at Times 0 hour and 6 hour.

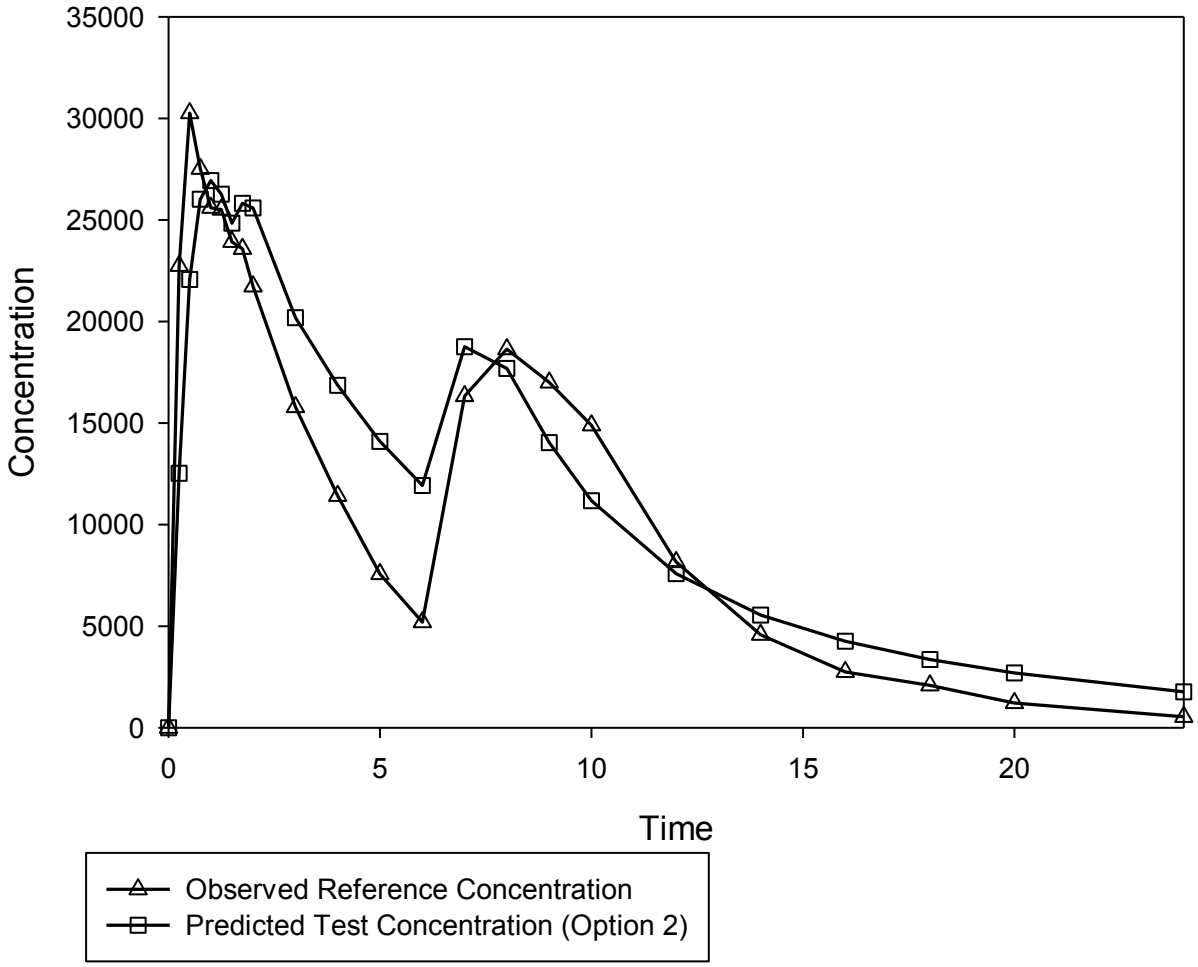


Figure 5 Simulated test formulation profile using Option 2 modification versus observed reference concentrations

Note: This simulation is for a single-dose study. The test formulation is simulated as an administration at Time 0 hour while the reference formulation is simulated as an administration at Times 0 hour and 6 hour.

Tables

Table 1: Estimated PK parameters

	Reference Fasting Model			Test Model		
Parameters	Mean	Median	Variance	Mean	Median	Variance
CL/F (L/h)	2.76	2.86	1.12	Same	Same	Same
Vc/F (L)	6.88	6.88	4.69	Same	Same	Same
CLd/F (L/h)	0.28	0.31	0.02	Same	Same	Same
Vp/F (L)	1.31	1.41	0.38	Same	Same	Same
Ka1 (h^{-1})	6.34	6.34	7.50	2.09	2.37	0.48
Ka1 for 6 hour reference dose (h^{-1})	0.93	0.92	0.07	-	-	-
Ka2 (h^{-1})	-	-	-	0.043	0.039	0.0003
Ka3 (h^{-1})	-	-	-	1.31	1.29	0.53
Tlag1 (h)	0.076	0.069	0.003	0.059	0.048	0.004
Tlag2 (h)	1.11	1.04	0.044	0.74	0.74	0.028
Tlag3 (h)	1.90	1.88	0.034	0.64	0.68	0.11
Tlag1 for 6 hour reference dose (h)	0.076	0.069	0.003	-	-	-
Tlag2 for 6 hour reference dose (h)	1.11	1.04	0.044	-	-	-
Tlag3 for 6 hour reference dose (h)	1.90	1.88	0.034	-	-	-
$F_{2nd\ Dose}$	0.92	0.93	0.011	-	-	-
Frel	-	-	-	1.29	1.34	0.026

Legend: Ka is an absorption rate constant; α and β are the portion of the available dose associated with an absorption process; Tlag is the time before the start of an absorption rate constant; $F_{2nd\ Dose}$ is the relative bioavailability of the 2nd reference dose; Frel is the relative bioavailability of the test versus the reference product; Vc/F is the volume of the central compartment (reference); Vp/F is the volume of the peripheral compartment (reference); CL/F is the total clearance (reference); CLd/F is the distributional clearance (reference)

Table 2: Estimated absorption PK parameters for the reference and test formulations

		Parameters associated with absorption process	**Reference**	**Test**
Before lunch	1st Absorption peak	Lag time before start (h)	0.06	0.07
		% associated to peak (%)	85.0	40.8
		Absorption half life (h)	0.11	0.31
	2nd Absorption peak	Lag time before start (h)	1.0	0.71
		% associated to peak (%)	12.0	50.2
		Absorption half life (h)	0.11	14.9
	3rd Absorption peak	Lag time before start (h)	2.7	1.68
		% associated to peak (%)	3.0	9.0
		Absorption half life (h)	0.11	0.47
After lunch	4th Absorption peak	Lag time before start (h)	0.08	-
		% associated to peak (%)	77.7	-
		Absorption half life (h)	0.74	-
	5th Absorption peak	Lag time before start (h)	1.3	-
		% associated to peak (%)	12.5	-
		Absorption half life (h)	0.74	-
	6th Absorption peak	Lag time before start (h)	2.9	-
		% associated to peak (%)	9.7	-
		Absorption half life (h)	0.74	-

Note: For the 4th to 6th absorption peaks, the lag times does not include the dosing time (6h)

Table 3: Average ratio and 90% confidence intervals for each dosing regimen

Predicted PK Parameters	SD Fasting	SS Fasting	SD Fasting/Fed	SS Fasting/Fed
Ln AUC0-t	91.7 (88.5 - 95.1)	NA	103.1 (100.6 - 105.6)	NA
Ln AUCinf	102.9 (99.0 - 106.9)	NA	117.0 (113.6 - 120.4)	NA
Ln Cmax	89.4 (84.1 - 95.1)	106.3 (100.5 - 112.4)	99.9 (93.9 - 106.4)	105.6 (98.4 - 113.4)
Ln AUC0-τ	NA	103.8 (100.0 - 107.6)	NA	117.7 (114.2 - 121.4)
Ln Cmin	NA	124.3 (114.9 - 134.5)	NA	139.0 (129.4 - 149.2)

Note: 10 studies per dosing regimens were simulated.

CHAPTER III: DISCUSSION

Modeling and simulations rely heavily on compartmental PK/PD analyses. To this end, new tools are made available in order to constantly improve analyses. One such new tool is ADAPT 5® with its new nonlinear mixed effect algorithms MLEM and ITS. It was to better understand the new tools available in ADAPT 5® that the first part of this thesis was performed. It was important to determine if these new available algorithms would perform better and/or provide better predictions than other tools that are already utilized, such as NONMEM® and IT2S®. In order to test these algorithms, twenty-nine studies were simulated. The use of simulated data instead of real clinical data allowed parameter estimations of the different algorithms to be compared directly to the true values of the parameters. Indeed, the main objective of using a nonlinear mixed effect modeling approach is to characterize with the utmost precision the population PK parameters and the inter-individual and intra-individual variability of these parameters. This information allows researchers to obtain accurate predictions to correctly anticipate future study outcomes. Thus, the precision and bias of each method were obtained by comparing estimated values against the true values. The use of precision and bias to compare method has been used in the past.[31,53,70,144] To the author's knowledge, this version of the software has never been tested in such a way before.

As expected and demonstrated by others,[65,73,76,145] results showed that all population PK analyses were superior to the standard two-stage analyses in determining the population PK parameters and their variabilities. PK estimates from this last method should therefore never be used to provide predictions results of future studies. Similarly, it was clear based on the results of this study that the new algorithms as implemented within ADAPT 5®, ITS and MLEM, were superior to IT2S® and we therefore

recommend the use of either MLEM or ITS rather than IT2S®. Based on our analyses, the two best methods to estimate variances were the MLEM and FOCE algorithms, implemented within ADAPT 5® and NONMEM® respectively. However, some differences existed between the two methods. The FOCE method often underestimated variance parameters. This shrinkage of the variance estimates has been well documented of late and is not a new issue.[146] Based on the literature, standard errors of the PK parameter estimates were often considered too small. It is for this reason that standard errors from NONMEM® are not often used to build confidence intervals around the estimates. Instead, bootstrapping is the method of choice to correctly provide confidence intervals of the estimates when NONMEM® is used.[137-150] In our analyses, the new MLEM algorithm almost never had any issues with variance shrinkage and in the 29 simulated studies it never severely underestimated the clearance or volume of the central compartment. Therefore, MLEM is an interesting alternative to FOCE in NONMEM® in this regard. Overall, results showed that of all the nonlinear mixed effect algorithms tested, MLEM in ADAPT 5® and FOCE method implemented within NONMEM® were the most precise. MLEM is comparable to FOCE with fewer shrinkage issues and its use in the drug development process should be increased.

The research presented in this thesis focused on compartmental PK analyses and clinical trial simulations used in the context of the Phase I drug development process. These studies enrolled healthy volunteer subjects selected from a generally homogeneous population. Therefore, no covariate analyses were required, although both NONMEM® and ADAPT 5® are capable of performing such analyses. Both continuous and discrete variables are easily modeled in NONMEM® and the difference in the objective function

serves as comparison between models with and without covariates.[151,152] Continuous variables are also easy to model with ADAPT 5®. However, in the beta version of this software, an option for the modeling of discrete covariates did not exist. Therefore, all 29 studies in the analyses were not designed to include testing of covariates. In the later stages of clinical drug development, covariates are important and may influence both the expected PK estimates and the study design. Clinical trial simulations for late clinical phases almost always include covariates in the model. It would be interesting to test this new module in ADAPT 5® to determine if it is a valuable tool for covariate testing and if it compares favorably to NONMEM®. This would facilitate the use of the software throughout the drug development process.

In terms of computational time for each of the nonlinear mixed effect algorithms, ITS was the fastest followed by FOCE, IT2S and MLEM. When modeling, it is always crucial to remember the purpose of the analyses. This will help guide the choice of the algorithm required for the analyses. If an ITS analysis suffices, then this analysis will save much time. The time difference between algorithm analyses can reach up to several days for a complex model containing many subjects.

Overall, the new algorithms available in ADAPT 5® performed very well and this new tool was used for all compartmental analyses and clinical trial simulations performed for the research projects undertaken within this thesis. In addition to confirming the choice of algorithm and software to be used for subsequent research, the results presented in the first part of the thesis will be useful to scientists applying these new algorithms and will help them better understand their strengths and weaknesses.

Despite the availability of new methodologies such as those explored in the first article, noncompartmental analysis remains a useful tool for the pharmaceutical industry and regulatory agencies to characterize the pharmacokinetics of drugs. Conclusions for many clinical studies depend on this approach. This is the case for the vast majority of Phase I studies including those having a bioequivalence or equivalence endpoint.[119] However, for noncompartmental PK parameters to be considered robust, the terminal elimination half-life is a critical parameter that must be properly calculated, as it is essential to the calculation of the overall exposure of a drug (e.g., AUCinf), total clearance, and total volume of distribution. Therefore, in these studies, the sampling scheme duration is important and regulatory agencies have set criteria to ensure proper analysis. For example, the Therapeutic Products Directorate (TPD) of Health Canada stipulates that the extrapolated portion of the AUCinf cannot be more than 20% of the total value,[41] while the office of generic drugs (OGD) of the US FDA recommends a value less than or equal to 12.5%.[36] These criteria are based on the mean results of a study and are meant to ensure the validity of the entire study. Criteria for individual results are not, however, clearly defined. To our knowledge, there was a gap in the understanding of what makes a half-life result robust or not in an individual subject. Based on past experiences, we believed that the conclusion of certain studies may have been biased or affected by individual results in which the PK parameters may not have been well estimated because of a sampling scheme duration that may have been too short (e.g., for these individuals' particular PK characteristics). We hypothesized that the exclusion of such inaccurate individual subject data would improve the reliability of the overall conclusions. Thus, the purpose of our analyses presented in the second article

was to determine if individual PK estimates judged unreliable negatively influenced the confidence intervals and if these individual subjects' kel, half-life and AUCinf should be removed from the ANOVA analyses.

Results from the analyses presented in the second research article of this thesis clearly showed that it was better to remove results from individuals whose PK parameters were estimated from a sampling scheme considered to be too short from ANOVA analysis. More specifically, the criterion that we proposed was to remove any subject's half-life value from the pivotal dataset in the ANOVA analysis when it was longer than half of the overall sampling duration. Maintaining subjects with unreliable estimates in the ANOVA adds uncertainty to the results while removing these subjects' parameters improves the accuracy of the overall results. This improvement outweighed the loss in degrees of freedom due to the exclusion of subjects. Although Purves[153] had previously stated that to have an acceptable precision in the half-life estimation, the Kel parameters had to be estimated from an appropriate terminal phase, his specific recommendation is not easy to implement, and the impact on the 90% confidence intervals of pivotal PK parameters had never been investigated. In contrast, our simulations determined the overall impact of including or excluding data judged unreliable on the 90% confidence interval for pivotal analyses.

Our results do not attempt to salvage poorly designed bioequivalence studies. It is not acceptable for a bioequivalence study to have a sampling scheme that spans only 2 half-lives on average. However, all studies are based on expected mean values and a certain amount of inter-subject and intra-subject variability. Occasionally, results will be unexpected or an individual subject's results will fall outside the anticipated limit of

variability. Our results suggest that the additional suggested criterion should be set *a priori* even if the mean results (*e.g.*, AUCt/AUCinf ratio) are within accepted criteria set by the regulatory agencies. This will provide confidence intervals that are closer to the truth.

The results also suggested that the sampling duration should be at least 4 times the average measured half-life in order to have confidence in the reported half-life value. This is true for both noncompartmental and compartmental analyses. This result is not surprising and further specifies the required duration of the sampling scheme for appropriate half-life estimation. Such information is crucial in clinical trial simulations for a new drug. Often, Phase II and III studies are simulated based on Phase I results. However, Phase I studies are the first time the drugs are administered in humans. These studies are sometimes based on short sampling schemes resulting in a half-life that might be underestimated, and consequently a clearance that might be overestimated. If clinical trials are simulated based on results from Phase I studies that have inaccurate estimated PK parameters, simulations will be incorrect and the actual study outcomes will be different than those expected. Therefore, before making any predictions, one must always ensure that the PK parameters are based on sampling schemes that span at least four half-lives.

Our research study did not compare the actual degree of bias in the alpha and beta error when we implemented this recommendation in pivotal Phase I studies to situations when the criterion was not applied. This would be an interesting avenue of research that could be investigated. One possible way to achieve this would be to retrospectively look back at a database of studies and determine how many studies would have had different

conclusions based on the new sampling scheme duration criterion. If a study has a proper design and results are those which are expected, the chance of a wrong conclusion would be inside the usual alpha and beta errors of 5% and 20%, respectively. If the same study were reproduced with clinical trial simulations, many subjects and study designs (with different sampling schemes) could be simulated. These designs would occasionally be less than optimal due to the different scenarios tested. Therefore, the chances of having an individual subject's data judged unreliable are greater due to the different study designs simulated.

Finally, when noncompartmental analysis is used during the drug development process, we recommend that the criterion to remove a subject's half-life value from the pivotal dataset when it is longer than half of the overall sampling duration be implemented in all pivotal Phase I studies (bioequivalence studies) to ensure that results are as accurate as possible. This is also true for clinical trial simulations as to have precise predictions by the simulations, the criteria to remove data for some individuals should be the same for the simulations as for the actual study that is performed. Therefore, we suggest that the added criterion be applied to both actual bioequivalence studies as well as any clinical trial simulations.

The use of new tools and the definition of subject inclusion criterion described in the first two articles lay the foundation for other innovative approaches used in the last article presented in this thesis. Because the current drug development process has been referred to be inefficient by many,[5,154] new approaches are required to extract meaningful information from collected data. The use of clinical trial simulations is one such new approach which has much potential to improve the drug development process. It is our

hypothesis that clinical trial simulations throughout the drug development plan will improve the efficiency of the process. In the last article presented in this thesis, clinical trial simulations were used to improve the drug development process for one compound. With ADAPT 5®, the new tool tested in the first article, and the new criteria established in the second article, population compartmental analyses and clinical trial simulations were used to extract valuable information for a new formulation of a NSAID which would have otherwise been impossible to determine using typical noncompartmental approaches. Only additional Phase I trials would have provided this information, which would have taken considerable time and money.

A pilot study was dosed in which a new sustained-release twice a day suspension was compared to an immediate-release four times a day reference suspension. Unfortunately, the pilot study did not properly capture the absorption phase of the second dose of the reference drug and the sampling scheme was somewhat short for proper characterization of the total exposure of the test product. In addition, a food effect was observed after the second dose of the reference drug. Using modeling and simulations, it was possible to characterize the absorption process of the reference product, establish the reduction in bioavailability for the reference product due to the food effect and calculate the absorption half-life for the two different products. The terminal half-life for the test formulation was representative of the absorption rate constant instead of the elimination rate constant due to flip-flop kinetics. The simulations also revealed that it would be difficult to demonstrate equivalence of the products in single-dose and multiple-dose pivotal studies. Therefore, although the pilot study demonstrated promising results with the 90% confidence intervals for the noncompartmental parameters AUC0-t, AUCinf and

Cmax being within the equivalence criteria, the drug development of this product was put on hold because of the modeling and simulation results, and until a new formulation be designed.

Clinical trial simulations permit researchers to answer many "what if" questions that are left unanswered by previously performed studies.[154,155] For example, Zhang *et al* have shown that clinical trial simulation could be used to answer certain questions.[126] Importantly, these questions can be answered without requiring the conduct of new studies. This was the case with the pilot study presented in the last article which had many unresolved issues. The first question was to determine if the observed food effect would impact the equivalence results. It was crucial to know whether or not conclusions would be different if a study was completed with a true fasting state throughout dosing for the reference product. Simulations demonstrated that conclusions would be different and that it would be important to know if all doses could be administered under true fasting conditions, especially for a multiple-dose study. When performing clinical trial simulations, it is critical to gain as much information as possible on the drug being simulated in order to provide the best predictions. In addition to the simulation results, literature data on studies conducted with the active ingredient demonstrated an observable food effect.[156,157] The food effect was similar to the one observed in the pilot study. In addition, one study showed that the effect was present for two of the four administered daily doses. This exemplifies how the use of information from two different sources (a previously dosed study and literature data) can help establish the most likely expected exposure. Thus, based on the literature data and simulation results, a food effect would most likely be present for the second administration of the reference

product. It is not uncommon to try and extract information from different sources. Lockwood presented an example of the use of previous data, either from literature or other drugs in the same class to improve the results of clinical trial simulations.[158]

When interpreting the data from the pilot study presented in the last part of this thesis, a second question that was answered by the clinical trial simulations concerned the impact of a change in the absorption rate constant of the test formulation. The formulation scientists had planned to use a slower sustained-release portion in the hopes that this would increase chances of meeting the equivalence criteria. The plan was to make these changes to the formulation and run another pilot study. Using clinical trial simulations, it was obvious that this would not help meet the equivalence criteria and would actually lower the chances of being equivalent at steady state. Using this tool made it easier for scientists with different backgrounds and expertise to comprehend the consequences of proposed changes. Discussions ensued, feasible changes to the formulation were proposed and clinical trial simulations were performed to determine which changes would have the best chance of helping the product meet equivalence criteria. These simulations helped the formulation scientists understand what changes should be made to the test formulation to increase chances of meeting equivalence criteria. This also reduced the number of hit and miss changes they would have done to their formulation and pilot studies that would have been performed.

Answering these questions and performing clinical trial simulations avoided the unnecessary dosing of additional studies. If the company had continued with a normal development plan based solely on the pilot results, pivotal studies would have been performed. These include one single dose study under fasting conditions, one single dose

study under fed conditions and two steady-state studies under fasting and fed conditions. These studies would most likely have failed to meet the equivalence criteria. Thus, potentially up to four unsuccessful studies were avoided thanks to clinical trial simulations. In addition, once it was determined from clinical trial simulations that the product would probably not meet the equivalence criteria, scientists had planned to make modifications to the product that they believed were necessary and dose another pilot study. As these changes were very different from those dictated by the clinical trial simulations, this new pilot study would have provided no additional useful information and further formulation changes would have been required. Therefore, clinical trial simulations reduced tremendously the number of subjects exposed unnecessarily to Phase I studies. This allowed the pharmaceutical company to save time, money and provided a better study design and formulation design for the pivotal studies. It also yielded a greater understanding of their compound's absorption process. This helped the formulation scientists determine what changes were required in order to produce a drug that would be acceptable to the regulatory agencies. These change would have otherwise been made without knowledge of their impact on the BE potential of the drug.Lalonde presented an example with gabapentin where clinical trial simulations were used to avoid additional Phase III studies.[159] Although this provides an example where money was saved with clinical trial simulations, the clinical trial simulations were only performed after the regulatory agency had raised concerns with study designs. Therefore, the use of clinical trial simulations should be used throughout drug development rather than simply waiting for regulatory agencies to raise concerns.

For a supergeneric drug, the impact of formulation changes on equivalence conclusions are not as obvious to predict as for a simple generic drug. Although a supergeneric has the same active ingredient as a reference product, it has different delivery characteristics than the reference product. Therefore, formulation changes to the supergeneric product cannot be directly linked to the reference product's characteristics and non-active ingredients. The generic industry has yet to take advantage of population pharmacokinetics or clinical trial simulations. Even though an agency such as the Office of Generic Drugs (OGD) at the FDA prefers noncompartmental analysis as part of an abbreviated new drug submission, this should not stop the industry from using modeling and simulations to extract information whenever possible to minimize costs. Not all generic compounds are easy to market. Multiple reasons exist for which a typical bioequivalence study does not produce the desired results, and this is when researchers sometimes turn to compartmental analyses. Population compartmental analyses have been used for generic drug development when the compound was highly variable, could only be tested in patients or systemic and PD endpoints were hard to measure.[160] However, this article showed how clinical trial simulations can be used in a different equivalence context. Innovative, scientifically sound approaches to improve drug development should always be taken advantage of whether or not it is for the innovator or generic industry.

Our research was conducted on Phase I healthy subjects. No covariates or PK/PD modeling was required. For example, the genotype or phenotype for the CYP2C9 enzyme was not collected; therefore, it was not determined if the population in the Phase I study included any poor metabolizers. Therefore, the PK model did not contain genotype or

phenotype as a covariate. Although this could have helped explain the variability of the clearance, the observed half-life for the reference product only varied from 1.8 to 4.3 hours. Thus, even if a subject in a future study had a half-life that was two to three times longer than the longest observed half-life from this Phase I study, the AUCinf determined for this subject would still be included even if taking into consideration the new *a priori* criteria for rejecting individual AUCinf parameter. Therefore, the results of the modeling and simulations were able to provide valuable information even without this covariate information. However, it is expected that clinical trial simulations would be even more beneficial with covariates information or with the use of PK/PD modeling. PK/PD models are usually very variable and covariates are added to the model in an attempt to explain this variability. Many more questions arise from such data and careful planning for later clinical phases is crucial to avoid studies that fail to demonstrate efficacy or safety.

The clinical trial simulations performed in the last portion of this thesis demonstrated how clinical trial simulations can be used starting in Phase I and how it can reduce the number of inefficient studies. It allowed the extraction of meaningful data from a single pilot study, was instrumental in changing the planned development course of action and helped scientists have a better understanding of the compound being developed and bring appropriate changes to the formulation to increase the likelihood that the drug could be marketed. Clearly, clinical trial simulations were influential in the development of the drug being tested. The company has since changed their philosophy and after most pilot studies, clinical trial simulations are performed to determine the equivalence potential of their compound and to determine if the planned study designs for

the pivotal studies are adequate. In addition, steady-state concentrations for their supergeneric compounds are simulated to determine if they will meet equivalence criteria at steady-state.

CHAPTER IV: CONCLUSIONS

1. GENERAL CONCLUSIONS

The pharmaceutical industry cannot rely on previous business paradigms going forward. Medications for uncomplicated diseases have already been developed. Today, the focus of research and development is on diseases that are more complex, with mechanisms that are not as straightforward and which have drug targets that are harder to reach. These factors have contributed to making the current drug development process inefficient and have led to many studies that did not fulfill their objectives. Thus, this research aimed at providing some possible solutions to this issue and thereby improving drug development.

The first part of the research helped to better understand new tools available for compartmental analyses. The MLEM algorithm proved to be very stable, provided better results than other tools available and was associated with less shrinkage of the variance parameters estimated by the models. Accurate variance estimates are essential if simulations are to portray the expected variability in concentration-time profiles for the population. Simulations performed using a smaller population variability than the true variability will lead to decisions that do not account for some expected subjects' responses. The opposite is also true, in which simulations using a larger population variability than the true variability will lead to decisions that account for some subjects' responses that are unlikely to happen in the population. In both cases, an inaccurate quantification of the variability in the population will lead to false predictions and wrong conclusions.

Having the right tools to perform modeling and simulations is important; however, it is also crucial to understand how certain subjects could influence conclusions and set appropriate guidelines regarding these subjects to ensure that the correct conclusions are obtained. Otherwise, incorrectly performed clinical trial simulations will have the opposite effect and hinder drug development. Therefore, the second portion of the research focused on which subject's PK parameters might unduly influence study conclusions due specifically to a short sampling scheme. Results demonstrated that maintaining a subject whose PK parameters are based on an unreliable profile will probably be more detrimental than beneficial in the analyses. With clinical trial simulations, numerous study designs are tested in order to answer many different questions. Consequently, with multiple simulations it is likely that subjects with unreliable profiles are simulated, so this aspect of my research project is especially important because it examines the influence of such subjects and proposes how to deal with them.

Lastly, the research demonstrated how clinical trial simulations could be used to improve the development process with a drug under development. Information was extracted from one study in order to answer many questions which would only have been possible by performing additional studies. They were also helpful in explaining how the drug could be modified to help achieve the profiles that were sought after by the researchers. Ultimately, clinical trial simulations allowed a better understanding of the compound being developed. This permitted the scientists to focus on ways to improve the compound while removing guesswork from the development process. More importantly, clinical trial simulations were able to improve our understanding while

reducing the number of subjects exposed unnecessarily to the compound, decrease the number of studies actually dosed and save time and money. The results were also instrumental in changing the way a specific pharmaceutical company is developing their drugs. Clinical trial simulations are now included in their drug development process with the goal of better understanding their drugs and reducing unnecessary studies.

Research presented in this thesis provided concrete ways to improve the drug development process by demonstrating the strengths and weaknesses of some newly available tools for compartmental analyses, setting standards stipulating which estimated PK parameters should be excluded from certain PK analyses and illustrating how clinical trial simulations are useful to answer many questions in the drug development process.

REFERENCES

1. DiMasi JA, Hansen RW, Grabowski HG. The price of innovation: new estimates of drug development costs. J Health Econ. 2003 Mar;22(2):151-85.
2. Adams CP, Brantner VV. Estimating the cost of new drug development: is it really 802 million dollars? Health Aff. 2006 Mar-Apr;25(2):420-8.
3. Gilbert J, Henske P, Singh A. Rebuilding Big Pharma's Business Model. In Vivo, the Business & Medicine Report, 2003 Nov; 21(10).
4. Rawlins MD. Cutting the cost of drug development? Nat Rev Drug Discov. 2004 Apr;3(4):360-4.
5. Challenge and Opportunity on the Critical Path to Medical Products. U.S. Department of Health and Human Services Food and Drug Administration. March 2004. Available at http://www.fda.gov/downloads/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm113411.pdf
6. DiMasi JA, Feldman L, Seckler A, Wilson A. Trends in risks associated with new drug development: success rates for investigational drugs. Clin Pharmacol Ther. 2010 Mar;87(3):272-7.
7. Smith CG, O'Donnell JT (2006). The process of new drug discovery and development. (2nd ed.). New York: Informa Healthcare USA, Inc.
8. Schmid EF, Smith DA. Is declining innovation in the pharmaceutical industry a myth? Drug Discov Today. 2005 Aug 1;10(15):1031-9.
9. Kaitin KI, DiMasi JA. Measuring the pace of new drug development in the user fee era. Drug Information Journal. 2000; 34:673–680.
10. Heilman RD. Drug development history, "overview," and what are GCPs? Qual Assur. 1995 Mar;4(1):75-9.
11. Gordian M, Singh N, Zemmel R, Elias T. Why products fail in Phase III. In Vivo, the Business & Medicine Report, 2006 Apr.
12. Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov. 2004 Aug;3(8):711-5.
13. Ette EI, Williams PJ (2007). Pharmacometrics the science of quantitative pharmacology. New Jersey: John Wiley & Sons, Inc.
14. Atkinson AJ, Abernethy DR, Daniels CE, Dedrick RL, Markey SP (2007). Principal of clinical pharmacology. 2nd ed. Elsevier; Burlington MA.
15. Burton ME, Shaw LM, Schentag JJ, Evans WE (2006). Applied pharmacokinetics & pharmacodynamics – principle of therapeutic drug monitoring. 4th ed. Lippincott Williams & Wilkins; Philadelphia PA.
16. Holford NH, Sheiner LB. Understanding the dose-effect relationship: clinical application of pharmacokinetic-pharmacodynamic models. Clin Pharmacokinet. 1981 Nov-Dec;6(6):429-53.
17. Fuseau E, Sheiner LB. Simultaneous modeling of pharmacokinetics and pharmacodynamics with a nonparametric pharmacodynamic model. Clin Pharmacol Ther. 1984 Jun;35(6):733-41.
18. Verotta D, Sheiner LB (1993). Pharmacokinetic/pharmacodynamic models and methods. In Integration of pharmacokinetics, pharmacodynamics, and

toxicokinetics in rational drug development. Yacobi A et al editors. Plenum Press; New York, NY.
19. Derendorf H, Meibohm B. Modeling of pharmacokinetic/pharmacodynamic (PK/PD) relationships: concepts and perspectives. Pharm Res. 1999 Feb;16(2):176-85.
20. Dunne A, King P. Estimation of noncompartmental parameters: A technical note. J Pharmacokinet Biopharm. 1989; 17(1):131-137.
21. Nüesch EA. Noncompartmental approach in pharmacokinetics using moments. Drug Metab Rev. 1984;15(1-2):103-31.
22. Ette EI, Williams PJ. Population pharmacokinetics II: estimation methods. Ann Pharmacother. 2004 Nov;38(11):1907-15.
23. Beal SL, Sheiner LB. Methodology of population pharmacokinetics (1985). In: Drug fate of metabolism Vol 5. Garrett ER, Hirtz JL eds. Marcel Dekker Inc, New York, NY.
24. Gillespie WR. Noncompartmental versus compartmental modelling in clinical pharmacokinetics. Clin Pharmacokinet. 1991 Apr;20(4):253-62.
25. Gibaldi M, Perrier D. Pharmacokinetics. Second edition revised and expanded. Informa Healthcare; New York, 2007.
26. Rowland M, Tozer TN. Clinical pharmacokinetics concepts and applications. 3rd ed. Lippincott Williams & Wilkins, Philadelphia, 1995.
27. Yamaoka K, Nakagawa T, Uno T. Statistical moments in pharmacokinetics. J Pharmacokinet Biopharm. 1978 Dec;6(6):547-58.
28. Cutler DJ. Theory of the mean absorption time, an adjunct to conventional bioavailability studies. J Pharm Pharmacol. 1978 Aug;30(8):476-8.
29. Riegelman S, Collier P. The application of statistical moment theory to the evaluation of in vivo dissolution time and absorption time. J Pharmacokinet Biopharm. 1980 Oct;8(5):509-34.
30. Purves RD. Numerical estimation of the noncompartmental pharmacokinetic parameters variance and coefficient of variation of residence times. J Pharm Sci. 1994 Feb;83(2):202-5.
31. Yeh KC, Kwan KC. A comparison of numerical integrating algorithms by trapezoidal, Lagrange, and spline approximation. J Pharmacokinet Biopharm. 1978 Feb;6(1):79-98.
32. Veng-Pedersen P. Mean time parameters in pharmacokinetics. Definition, computation and clinical implications (Part I). Clin Pharmacokinet. 1989 Nov;17(5):345-66.
33. Veng-Pedersen P. Mean time parameters in pharmacokinetics. Definition, computation and clinical implications (Part II). Clin Pharmacokinet. 1989 Dec;17(6):424-40.
34. Shargel L, Wu-Pong S, YU ABC. Applied biopharmaceutics & Pharmacokinetics. 5th ed. McGraw-Hill Medical Publishing Division; New York, 2005.
35. Benet LZ (1993). The role of pharmacokinetics in the drug development process. In Integration of pharmacokinetics, pharmacodynamics, and toxicokinetics in rational drug development. Yacobi A et al editors. Plenum Press; New York, NY.
36. Guidance for Industry. Bioavailability and bioequivalence studies for orally administered drug products — general considerations. U.S. Department of Health

and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER). March 2003.
37. Chiou WL. Critical evaluation of the potential error in pharmacokinetic studies of using the linear trapezoidal rule method for the calculation of the area under the plasma level--time curve. J Pharmacokinet Biopharm. 1978 Dec;6(6):539-46.
38. Purves RD. Bias and variance of extrapolated tails for area-under-the-curve (AUC) and area-under-the-moment-curve (AUMC). J Pharmacokinet Biopharm. 1992 Oct;20(5):501-10.
39. Vandenhende F, Comblain M, Delsemme MH, Dewe W, Hoffman WP, Boulanger B. Construction of an optimal destructive sampling design for noncompartmental AUC estimation. J Pharmacokinet Biopharm. 1999 Apr;27(2):191-212.
40. D'Argenio DZ, Katz D. Sampling strategies for noncompartmental estimation of mean residence time. J Pharmacokinet Biopharm. 1983 Aug;11(4):435-46.
41. Guidance for Industry. Conduct and Analysis of Bioavailability and Bioequivalence Studies - Part A: Oral Dosage Formulations Used for Systemic Effects. Canada Health Products and Food Branch Guidance Document. 1992.
42. Levy G. Pharmacokinetics of salicylate elimination in man. J Pharm Sci. 1965 Jul;54(7):959-67.
43. Nagashima R, Levy G, O'Reilly RA. Comparative pharmacokinetics of coumarin anticoagulants. IV. Application of a three-compartmental model to the analysis of the dose-dependent kinetics of bishydroxycoumarin elimination. J Pharm Sci. 1968 Nov;57(11):1888-95.
44. Levy G. Pharmacokinetics of LSD effect. Clin Pharmacol Ther. 1969 Jan-Feb;10(1):134-5.
45. Levy G. Pharmacokinetics of succinylcholine in newborns. Anesthesiology. 1970 Jun;32(6):551-2.
46. Levy G. Relationship between elimination rate of drugs and rate of decline of their pharmacologic effects. J Pharm Sci. 1964 Mar;53:342-3.
47. Levy G, Gibaldi M. Pharmacokinetics of drug action. Annu Rev Pharmacol. 1972;12:85-98.
48. Sheiner LB, Rosenberg B, Melmon KL. Modelling of individual pharmacokinetics for computer-aided drug dosage. Comput Biomed Res. 1972 Oct;5(5):411-59.
49. Peck CC, Sheiner LB, Martin CM, Combs DT, Melmon KL. Computer-assisted digoxin therapy. N Engl J Med. 1973 Aug 30;289(9):441-6.
50. Galeazzi RL, Benet LZ, Sheiner LB. Relationship between the pharmacokinetics and pharmacodynamics of procainamide. Clin Pharmacol Ther. 1976 Sep;20(3):278-89.
51. Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetics parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data. J Pharmacokinet Biopharm. 1980 Dec;8(6):553-71.
52. Sheiner LB, Stanski DR, Vozeh S, Miller RD, Ham J. Simultaneous modeling of pharmacokinetics and pharmacodynamics: application to d-tubocurarine. Clin Pharmacol Ther. 1979 Mar;25(3):358-71.
53. Sheiner BL, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters. II. Biexponential model and experimental pharmacokinetic data. J Pharmacokinet Biopharm. 1981 Oct;9(5):635-51.

54. Martin E, Tozer TN, Sheiner LB, Riegelman S. The clinical pharmacokinetics of phenytoin. J Pharmacokinet Biopharm. 1977 Dec;5(6):579-96.
55. Beal SL. Population pharmacokinetic data and parameter estimation based on their first two statistical moments. Drug Metab Rev. 1984;15(1-2):173-93.
56. Bonate PL (2006). Pharmacokinetic-Pharmacodynamic Modeling and Simulation. New York: Springer Science+Business Media, Inc.
57. Sheiner LB, Rosenberg B, Marathe VV. Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. J Pharmacokinet Biopharm. 1977 Oct;5(5):445-79.
58. Sheiner LB, Beal SL. Pharmacokinetic parameter estimates from several least squares procedures: superiority of extended least squares. J Pharmacokinet Biopharm. 1985 Apr;13(2):185-201.
59. Krewski D, Burnett RT, Ross W (1991). Statistical Foundations of pharmacokinetic modeling. In: New trends in pharmacokinetics. Rescigno A, Thakur AK editors. Plenum Press; New York, NY.
60. Rodman JH, D'Argenio DZ, Peck CC (2006). Analysis of Pharmacokinetic Data for Individualizing Drug Dosage Regimens. In ME Burton, LM Shaw, JJ Schentag and WE Evans (Eds.), Applied Pharmacokinetics & Pharmacodynamics. (pp. 40-59). Baltimore, MD: Lippincott Williams & Wilkins.
61. Sheiner LB. Analysis of pharmacokinetic data using parametric models. II. Point estimates of an individual's parameters. J Pharmacokinet Biopharm. 1985 Oct;13(5):515-40.
62. Jelliffe RW, Maire P, Sattler F, Gomis P, Tahani B. Adaptive control of drug dosage regimens: basic foundations, relevant issues, and clinical examples. Int J Biomed Comput. 1994 Jun;36(1-2):1-23.
63. Racine-Poon A, Wakefield J. Statistical methods for population pharmacokinetic modelling. Stat Methods Med Res. 1998 Mar;7(1):63-84.
64. Sheiner LB, Beal SL. Bayesian individualization of pharmacokinetics: simple implementation and comparison with non-Bayesian methods. J Pharm Sci. 1982 Dec;71(12):1344-8.
65. Sheiner LB. The population approach to pharmacokinetic data analysis: rationale and standard data analysis methods. Drug Metab Rev. 1984;15(1-2):153-71.
66. D'Argenio, D.Z., A. Schumitzky and X. Wang. ADAPT 5 User's Guide: Pharmacokinetic/Pharmacodynamic Systems Analysis Software. Biomedical Simulations Resource, Los Angeles, 2009.
67. D'Argenio, D. and A. Schumitzky, ADAPT-II Users Manual. 1997, Biomedical Simulations Resource, University of Southern California: Los Angeles.
68. Sheiner LB, Rosenberg B, Melmon KL. Modeling of individual pharmacokinetics for computer-aided drug dosage. Comp Biomed Res 1972;5:441-59.
69. Computer-assisted digoxin therapy. Peck CC, Sheiner LB, Martin CM, Combs DT, Melmon KL. N Engl J Med. 1973 Aug 30;289(9):441-6.
70. Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters. III. Monoexponential model: routine clinical pharmacokinetic data. J Pharmacokinet Biopharm. 1983 Jun;11(3):303-19.
71. Sheiner LB, Grasela TH. Experience with NONMEM: analysis of routine phenytoin clinical pharmacokinetic data. Drug Metab Rev. 1984;15(1-2):293-303.

72. Beal, S.L., Sheiner, L.B. & Boeckmann, A. NONMEM Users Guide (1989-2006) Version VI. Icon Development Solutions, Ellicott city, MD, 2008.
73. Ludden TM. Population pharmacokinetics. J Clin Pharmacol. 1988 Dec;28(12):1059-63.
74. Sheiner LB, Grasela TH. An introduction to mixed effect modeling: Concepts. Definitions and justification. J Pharmacokinet Biopharm. 1991;19(3):11s-24s.
75. White DB, Walawander CA, Tung Y, Grasela TH. An evaluation of point and interval estimates in population pharmacokinetics using NONMEM analysis. J Pharmacokinet Biopharm. 1991 Feb;19(1):87-112.
76. Steimer JL, Mallet A, Golmard JL, Boisvieux JF. Alternative approaches to estimation of population pharmacokinetic parameters: comparison with the nonlinear mixed-effect model. Drug Metab Rev. 1984;15(1-2):265-92.
77. Vermes A, Math t RA, van der Sijs IH, Dankert J, Guchelaar HJ. Population pharmacokinetics of flucytosine: comparison and validation of three models using STS, NPEM, and NONMEM. Ther Drug Monit. 2000 Dec;22(6):676-87.
78. Ette EI, Sun H, Ludden TM. Balanced designs in longitudinal population pharmacokinetic studies. J Clin Pharmacol. 1998 May;38(5):417-23.
79. Ette EI, Kelman AW, Howie CA, Whiting B. Analysis of animal pharmacokinetic data: performance of the one point per animal design. J Pharmacokinet Biopharm. 1995 Dec;23(6):551-66.
80. White DB, Walawander CA, Tung Y, Grasela TH. An evaluation of point and interval estimates in population pharmacokinetics using NONMEM analysis. J Pharmacokinet Biopharm. 1991 Feb;19(1):87-112.
81. Grasela TH Jr, Antal EJ, Townsend RJ, Smith RB. An evaluation of population pharmacokinetics in therapeutic trials. Part I. Comparison of methodologies. Clin Pharmacol Ther. 1986 Jun;39(6):605-12.
82. Forrest, A., J. Hawtoff, and M. J. Egorin. Evaluation of a new program for population PK/PD analysis: applied to simulated phase I data. Clin Pharmacol Ther 1991;49(2):153.
83. Forrest A, Ballow CH, Nix DE, Birmingham MC, Schentag JJ. Development of a population pharmacokinetic model and optimal sampling strategies for intravenous ciprofloxacin. Antimicrob Agents Chemother. 1993 May;37(5):1065-72.
84. Auclair B, Sirois G, Ngoc AH, Ducharme MP. Novel pharmacokinetic modelling of transdermal nitroglycerin. Pharm Res. 1998 Apr;15(4):614-9.
85. Auclair B, Ducharme MP. Piperacillin and tazobactam exhibit linear pharmacokinetics after multiple standard clinical doses. Antimicrob Agents Chemother. 1999 Jun;43(6):1465-8.
86. Drusano GL, Liu W, Perkins R, Madu A, Madu C, Mayers M, Miller MH. Determination of robust ocular pharmacokinetic parameters in serum and vitreous humor of albino rabbits following systemic administration of ciprofloxacin from sparse data sets by using IT2S, a population pharmacokinetic modeling program. Antimicrob Agents Chemother. 1995 Aug;39(8):1683-7.
87. Williams ML, Wainer IW, Granvil CP, Gehrcke B, Bernstein ML, Ducharme MP. Pharmacokinetics of (R)- and (S)-cyclophosphamide and their dechloroethylated metabolites in cancer patients. Chirality. 1999;11(4):301-8.

88. Collins DG, Forrest A. IT2S User's guide. Buffalo: State University of New York at Buffalo, 1995.
89. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. J. Roy. Statist. Soc. Ser. B. 1977; 39(1):1-38.
90. Schumitzky A. EM Algorithms and two stage methods in pharmacokinetic population analysis. In: Advanced methods of pharmacokinetic and pharmacodynamic system analyses. Vol. II. D'Argenio, DZ editor. Plenum Press; New York:1995. p 145-160.
91. Aarons L. The estimation of population pharmacokinetic parameters using an EM algorithm. Comput Methods Programs Biomed. 1993 Sep;41(1):9-16.
92. Wang X, Schumitzky A, D'Argenio DZ. Nonlinear Random Effects Mixture Models: Maximum Likelihood Estimation via the EM Algorithm. Comput Stat Data Anal. 2007 Aug 15;51(12):6614-6623.
93. Wang J, Weiss M, D'Argenio DZ. A note on population analysis of dissolution-absorption models using the inverse Gaussian function. J Clin Pharmacol. 2008 Jun;48(6):719-25.
94. Wang X, Schumitzky A, D'Argenio DZ. Population Pharmacokinetic/Pharmacodyanamic Mixture Models via Maximum a Posteriori Estimation. Comput Stat Data Anal. 2009 Oct 1;53(12):3907-3915.
95. Sheiner LB, Beal SL. Some suggestions for measuring predictive performance. J Pharmacokinet Biopharm. 1981 Aug;9(4):503-12.
96. Soy D, Beal SL, Sheiner LB. Population one-compartment pharmacokinetic analysis with missing dosage data. Clin Pharmacol Ther. 2004 Nov;76(5):441-51.
97. Wade JR, Beal SL, Sambol NC. Interaction between structural, statistical, and covariate models in population pharmacokinetic analysis. J Pharmacokinet Biopharm. 1994 Apr;22(2):165-77.
98. Mandema JW, Verotta D, Sheiner LB. Building population pharmacokinetic--pharmacodynamic models. I. Models for covariate effects. J Pharmacokinet Biopharm. 1992 Oct;20(5):511-28.
99. Ludden TM, Beal SL, Sheiner LB. Comparison of the Akaike Information Criterion, the Schwarz criterion and the F test as guides to model selection. J Pharmacokinet Biopharm. 1994 Oct;22(5):431-45.
100. Yamaoka K, Nakagawa T, Uno T. Application of Akaike's information criterion (AIC) in the evaluation of linear pharmacokinetic equations. J Pharmacokinet Biopharm. 1978 Apr;6(2):165-75.
101. Rodman JH, Jelliffe RW, Kolb E, Tuey DB, de Guzman MF, Wagers PW, Haywood LJ. Clinical studies with computer-assisted initial lidocaine therapy. Arch Intern Med. 1984 Apr;144(4):703-9.
102. Jelliffe RW. Computer-controlled administration of cardiovascular drugs. Prog Cardiovasc Dis. 1983 Jul-Aug;26(1):1-14.
103. Jelliffe RW, *et al*. Model-based, goal-oriented, individualised drug therapy. Linkage of population modelling, new 'multiple model' dosage design, bayesian feedback and individualised target goals. Clin Pharmacokinet. 1998 Jan;34(1):57-77.
104. Sheiner LB, Grasela TH. An introduction to mixed effect modeling: concepts, definitions, and justifications. J Pharmacokinet Biopharm 1991 Jun; 19(3):11S-24S

105. Gillespie WR. Noncompartmental versus compartmental modelling in clinical pharmacokinetics. Clin Pharmacokinet. 1991 Apr;20(4):253-62.
106. Rodman JH, Sunderland M, Kavanagh RL, Ochs J, Yalowich J, Evans WE, Rivera GK. Pharmacokinetics of continuous infusion of methotrexate and teniposide in pediatric cancer patients. Cancer Res. 1990 Jul 15;50(14):4267-71.
107. Levy G, Gibaldi M, Jusko WJ. Multicompartment pharmacokinetic models and pharmacologic effects. J Pharm Sci. 1969 Apr;58(4):422-4.
108. Slattery JT, Levy G. Pharmacokinetic model of acetaminophen elimination. Am J Hosp Pharm. 1979 Apr;36(4):440.
109. Katz D, D'Argenio DZ. Implementation and evaluation of control strategies for individualizing dosage regimens, with application to the aminoglycoside antibiotics. J Pharmacokinet Biopharm. 1986 Oct;14(5):523-37.
110. D'Argenio DZ, Katz D. Sampling strategies for noncompartmental estimation of mean residence time. J Pharmacokinet Biopharm. 1983 Aug;11(4):435-46.
111. Grevel J, Whiting B, Kelman AW, Taylor WB, Bateman DN. Population analysis of the pharmacokinetic variability of high-dose metoclopramide in cancer patients. Clin Pharmacokinet. 1988 Jan;14(1):52-63.
112. Benet LZ. The role of pharmacokinetics in the drug development process. In: Integration of Pharmacokinetics, pharmacodynamics, and toxicokinetics in rational drug development. Yacobi A, Skelly JP, Shah VP and Benet LZ editors. Plenum Press; New York:1993. p 115-123.
113. Sheiner LB, Ludden TM. Population pharmacokinetics/dynamics. Annu Rev Pharmacol Toxicol. 1992;32:185-209.
114. Beach CL, Farringer JA, Peck CC, Crawford MH, Ludden TM, Clementi WA. Clinical assessment of a two-compartment Bayesian forecasting method for lidocaine. Ther Drug Monit. 1988;10(1):74-9.
115. Privitera MD, Homan RW, Ludden TM, Peck CC, Vasko MR. Clinical utility of a Bayesian dosing program for phenytoin. Ther Drug Monit. 1989;11(3):285-94.
116. Ludden TM, Gillespie WR, Bachman WJ. Commentary on "Physiologically based pharmacokinetic modeling as a tool for drug development". J Pharmacokinet Biopharm. 1995 Apr;23(2):231-5.
117. Peck CC, *et al*. Opportunities for integration of pharmacokinetics, pharmacodynamics, and toxicokinetics in rational drug development. J Clin Pharmacol. 1994 Feb;34(2):111-9.
118. Peck CC. Drug development: improving the process. Food Drug Law J. 1997;52(2):163-7.
119. Guidance for Industry. Population Pharmacokinetics. U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) Center for Biologics Evaluation and Research (CBER). February 1999.
120. Cost B1 Steering Committee on the population approach (Geneva 1997). The population approach: measuring and managing variability in response, concentration and dose. European communities, printed in Belgium.
121. Aarons L, Balant LP, Mentré F, Morselli PL, Rowland M, Steimer JL, Vozeh S. Population approaches in drug development. Report on an expert meeting to discuss

population pharmacokinetic/pharmacodynamic software. Eur J Clin Pharmacol. 1994;46(5):389-91.
122. Lesko LJ, Rowland M, Peck CC, Blaschke TF. Optimizing the Science of Drug Development: Opportunities for Better Candidate Selection and Accelerated Evaluation in Humans. Pharm Res. 2000 Nov;17(11):1335-44.
123. Aarons L, Balant LP, Mentre F, Morselli PL, Rowland M, Steimer JL, Vozeh S. Practical experience and issues in designing and performing population pharmacokinetic/pharmacodynamic studies. Eur J Clin Pharmacol. 1996;49(4):251-4.
124. Steimer J, Vozeh S, Racine-Poon A, Holford N, O'Neil R (1994). The population approach: rationale, methods, applications in clinical pharmacology and drug development. In: Pharmacokinetics of drugs, handbook of experimental pharmacology. Edited by Welling P, Balant L. Springer-Verlag, Berlin. P. 405-451.
125. Sheiner LB. Learning versus confirming in clinical drug development. Clin Pharmacol Ther. 1997 Mar;61(3):275-91.
126. Zhang L, Sinha V, Forgue ST, Callies S, Ni L, Peck R, Allerheiligen SR. Model-based drug development: the road to quantitative pharmacology. J Pharmacokinet Pharmacodyn. 2006 Jun;33(3):369-93.
127. Lockwood PA, Cook JA, Ewy WE, Mandema JW. The use of clinical trial simulation to support dose selection: application to development of a new treatment for chronic neuropathic pain. Pharm Res. 2003 Nov;20(11):1752-9.
128. Lalonde RL et al. Model-based drug development. Clin Pharmacol Ther. 2007 Jul;82(1):21-32.
129. Lockwood P, Ewy W, Hermann D, Holford N. Application of clinical trial simulation to compare proof-of-concept study designs for drugs with a slow onset of effect; an example in Alzheimer's disease. Pharm Res. 2006 Sep;23(9):2050-9.
130. Miller R et al. How modeling and simulation have enhanced decision making in new drug development. J Pharmacokinet Pharmacodyn. 2005 Apr;32(2):185-97.
131. Dickson M, Gagnon JP. Key factors in the rising cost of new drug discovery and development. Nat Rev Drug Discov. 2004 May;3(5):417-29.
132. Booth B, Zemmel R. Prospects for productivity. Nat Rev Drug Discov. 2004 May;3(5):451-6.
133. Prentis RA, Lis Y, Walker SR. Pharmaceutical innovation by the seven UK-owned pharmaceutical companies (1964-1985). Br J Clin Pharmacol. 1988 Mar;25(3):387-96.
134. Vozeh S, Steimer JL, Rowland M, Morselli P, Mentre F, Balant LP, Aarons L The use of population pharmacokinetics in drug development. Clin Pharmacokinet. 1996 Feb;30(2):81-93.
135. Mahmood I. Interspecies scaling for the prediction of drug clearance in children: application of maximum lifespan potential and an empirical correction factor. Clin Pharmacokinet. 2010 Jul 1;49(7):479-92.
136. Wang W, Prueksaritanont T. Prediction of human clearance of therapeutic proteins: simple allometric scaling method revisited. Biopharm Drug Dispos. 2010 May;31(4):253-63.
137. Lin JH. Species similarities and differences in pharmacokinetics. Drug Metab Dispos. 1995 Oct;23(10):1008-21.

138. Guidance for Industry. Estimating the Maximum Safe Starting Dose in Initial Clinical Trials for Therapeutics in Adult Healthy Volunteers. U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER). July 2005.
139. Splawinski J, Kuzniar J, Kurianowicz R, Wanczura P. Bioequivalence of two preparations of ticlopidine evaluated using a pharmacodynamic end point. Int J Clin Pharmacol Ther. 2005 Sep;43(9):452-6.
140. Geoffroy P, Lalonde RL, Ahrens R, Clarke W, Hill MR, Vaughan LM, Grossman J. Clinical comparability of albuterol delivered by the breath-actuated inhaler (Spiros) and albuterol by MDI in patients with asthma. Ann Allergy Asthma Immunol. 1999 Apr;82(4):377-82.
141. Di Marco M, *et al*. Pharmacodynamic equivalence of two orlistat capsule formulations in healthy volunteers under fed conditions. Int J Clin Pharmacol Ther. 2008 Jun;46(6):319-26.
142. Krivoshiev S, *et al*. Therapeutic equivalence of epoetin zeta and alfa, administered subcutaneously, for maintenance treatment of renal anemia. Adv Ther. 2010 Feb;27(2):105-17.
143. Schuirmann DJ. A comparison of the two one-sided tests procedure and the power approach for assessing the equivalence of average bioavailability. J Pharmacokinet Biopharm. 1987 Dec;15(6):657-80.
144. Karlsson MO, Sheiner LB. The importance of modeling interoccasion variability in population pharmacokinetic analyses. J Pharmacokinet Biopharm. 1993 Dec;21(6):735-50.
145. Yu DK, Hutcheson SJ, Wei G, Bhargava VO, Weir SJ. A comparison of population and standard two-stage pharmacokinetic analyses of vigabatrin data. Biopharm Drug Dispos. 1994 Aug;15(6):473-84.
146. Savic RM, Karlsson MO. Importance of shrinkage in empirical bayes estimates for diagnostics: problems and solutions. AAPS J. 2009 Sep;11(3):558-69.
147. Hunt CA, Givens GH, Guzy S. Bootstrapping for pharmacokinetic models: visualization of predictive and parameter uncertainty. Pharm Res. 1998 May;15(5):690-7.
148. Efron B. Better Bootstrap Confidence Intervals. J Amer Statl Ass, Vol. 82(397) (Mar., 1987), pp. 171-185.
149. DiCiccio TJ, Efron B. Bootstrap Confidence Intervals. Statistical Science, Vol. 11(3), Aug., 1996, pp. 189-212
150. Mick R, Ratain MJ. Bootstrap validation of pharmacodynamic models defined via stepwise linear regression. Clin Pharmacol Ther. 1994 Aug;56(2):217-22.
151. Viberg A, Lannergård A, Larsson A, Cars O, Karlsson MO, Sandström M. A population pharmacokinetic model for cefuroxime using cystatin C as a marker of renal function. Br J Clin Pharmacol. 2006 Sep;62(3):297-303.
152. Yoo HD, Kim MS, Cho HY, Lee YB. Population pharmacokinetic analysis of glimepiride with CYP2C9 genetic polymorphism in healthy Korean subjects. Eur J Clin Pharmacol. 2011 Apr 8.
153. Purves RD. Bias and variance of extrapolated tails for area-under-the-curve (AUC) and area-under-the-moment-curve (AUMC). J Pharmacokinet Biopharm. 1992 Oct;20(5):501-10.

154. Holford N, Ma SC, Ploeger BA. Clinical trial simulation: a review. Clin Pharmacol Ther. 2010 Aug;88(2):166-82.
155. De Ridder F. Predicting the outcome of phase III trials using phase II data: a case study of clinical trial simulation in late stage drug development. Basic Clin Pharmacol Toxicol. 2005 Mar;96(3):235-41.
156. Klueglich M, Ring A, Scheuerer S, Trommeshauser D, Schuijt C, Liepold B, Berndl G. Ibuprofen extrudate, a novel, rapidly dissolving ibuprofen formulation: relative bioavailability compared to ibuprofen lysinate and regular ibuprofen, and food effect on all formulations. J Clin Pharmacol. 2005 Sep;45(9):1055-61.
157. Application review. Center for drug evaluation and research. Apllication number 74-937. http://www.accessdata.fda.gov/drugsatfda_docs/anda/98/74-937_Ibuprofen_bioeqr.pdf.
158. Lockwood PA, Cook JA, Ewy WE, Mandema JW. The use of clinical trial simulation to support dose selection: application to development of a new treatment for chronic neuropathic pain. Pharm Res. 2003 Nov;20(11):1752-9.
159. Lalonde RL et al. Model-based drug development. Clin Pharmacol Ther. 2007 Jul;82(1):21-32.
160. Colucci P, Marier JF, Ducharme MP. 'Chapter 8 – Population Pharmacokinetic Approaches for Assessing Bioequivalence' in the book 'Population PK approaches for assessing bioequivalence. In; Kanfer I, Shargel, eds. Generic Drug Product Development – Bioequivalence'. Informa Healthcare; Ney York, New York, 2008.

APPENDICE I – Permissions