A systematic framework to integrate preclinical data within PBPK models: from global sensitivity analysis to middle-out approaches

A thesis submitted to The University of Manchester
for the degree of Doctor of Philosophy
in the Faculty of Biology, Medicine and Health

2021

Estelle Yau

School of Health Sciences
Division of Pharmacy and Optometry

Contents

Word count: 48,031

List of Figures

List of Tables

List of Abbreviations

ADME: absorption, distribution, metabolism and elimination

AAG : α1-acid glycoprotein

AP: acidic phospholipids

AUC: area under the curve

BP: blood-to-plasma ratio

BW: body weight

CL: clearance

CL_p: plasma clearance

CL_b: blood clearance

CLint: intrinsic clearance

CL_R: renal clearance

CO: cardiac output

Css: steady-state plasma drug concentration

Cu_t: unbound concentration in tissues

Cu_p: unbound concentration of drug in blood plasma

CV: coefficient of variation

CYP: cytochrome P450

DDI: drug-drug interaction

EMA: European Medicines Agency

EW: extracellular water

FDA: Food and Drug Administration

FIH: First in human

FO: first-order maximum likelihood estimation method

FOCE-I: first-order conditional with interaction maximum likelihood estimation Method

fu_b: fraction unbound in blood

fu_p: fraction unbound in plasma

fu_t: fraction unbound in tissue

GSA: Global sensitivity analysis

IIV: inter-individual variability

IV or *iv:* intravenous

IVIVC: *in vitro-in vivo* correlations

IVIVE: *in vitro-in vivo* extrapolation

IW: intracellular water

Ka_{AP}: association constant of the drug compound with acidic phospholipids

Ka_{PR}: association constant of the drug compound with extracellular proteins

Kb: blood-to-tissue partition coefficient

Kp: plasma-to-tissue partition coefficient

Kpu: unbound plasma-to-tissue partition coefficient

LC-MS/MS: liquid chromatography with tandem mass spectrometry detection

LHS: Latin hypercube sampling

LogD: logarithm of the distribution coefficient

LogP: logarithm of the partition coefficient (octanol:water)

LSA: local sensitivity analysis

MCMC: Markov chain Monte Carlo

M&S: modelling and simulation

MRT: mean residence time

MW: molecular weight

NDA: New Drug Applications

NHP: Non-human primates

NL: neutral lipids

NLME: non-linear mixed effects

NP: neutral phospholipids

OATP: organic anion polypeptide transporter

ODE: ordinary differential equation

PBPK: Physiologically-based pharmacokinetic(s)

PD: pharmacodynamic(s)

P-gp: P-glycoprotein

PK: Pharmacokinetics

pKa: acid dissociation constant

PK/PD: pharmacokinetic/pharmacodynamics

PR: protein

PRCC: partial rank correlation coefficient

Q: blood flow

QSAR: quantitative structure-activity relationship

RBC: red blood cells

RoB: rest of the body

RMSE: Root mean squared error

R&R: Rodgers and Rowland

RSE: relative standard error

SAEM/IMP: Stochastic Approximation Estimation method followed by Method Monte-Carlo importance sampling

SF: scaling factor

UGT: UDP-Glucoronosyltransferases

V: volume

Vdss or Vss: Volume of distribution at steady state

Vuss: unbound volume of distribution

WBPBPK: Whole-body physiologically-based pharmacokinetic(s)

Abstract

THE UNIVERSITY OF MANCHESTER
Abstract of thesis submitted by Estelle Yau for the degree of Doctor of Philosophy, entitled: "A systematic framework to integrate preclinical data within PBPK models: from global sensitivity analysis to middle-out approaches".
Month and year of submission: June 2021

Whole-body physiologically-based pharmacokinetic (PBPK) models have many applications in academic and pharmaceutical research and drug development. However, optimization of parameters in such complex models by fitting models to observed data is a challenging and time-consuming process. The models are often unidentifiable/over-parameterized given the large number of parameters and availability of data which are mostly limited to plasma observations. It is common practice to fix some parameters and estimate others. However, the decision on which parameters to fix and which to estimate is subjective and therefore the final model and parameters may vary significantly between different investigators. This was a concern highlighted by regulatory agencies in their guidance documents for PBPK modelling. Hence, the overall aim of this thesis was to develop a systematic approach for integrating preclinical data within PBPK models to address this issue.

The first part of this thesis was focused on identifying key drug-dependent and physiological parameters that influence predictions of tissue-to-unbound plasma partition coefficients (Kpus) and thus drug distribution in PBPK models. The impact of these parameters was evaluated on the Kpus predicted by the Rodgers and Rowland model using sensitivity and uncertainty analyses. For most drug classes, LogP and fraction unbound in plasma (fu_p) were generally the most influential parameters for Kpu predictions. Uncertainty in tissue composition parameters especially acidic phospholipid concentrations and extracellular protein tissue:plasma ratios, could have a large impact on Kpu predictions for all classes. For parameter estimation involving PBPK models and dimensionality reduction purposes, less influential parameters might be assigned fixed values depending on the parameter space, while influential parameters could be subject to parameter estimation.

Secondly, several model reduction approaches were investigated to simplify PBPK model structure or dimensionality and thus facilitate the estimation process during PBPK model development. Tissues were clustered according to physiological information reducing the number of unknown parameters without changing the overall PBPK model structure. The investigated mechanistic models in conjunction with preclinical in vivo data were able to provide suitable estimates of Kpus using the nonlinear mixed effect method. To that end, diazepam was used as a case example. This analysis provided a basic framework for PBPK model development and estimation of distribution parameters and subsequent applications of PBPK modelling, especially translation of drug distribution from animals to humans.

Subsequently, the use of the investigated mechanistic models for interspecies extrapolation was evaluated. The models that could best fit data in rats and monkeys were applied for translation of drug distribution to humans. The performance of these best models was assessed for three compounds (diazepam, midazolam and basmisanil) and compared to the whole-body PBPK model with Kpu predictions from the Rodgers and Rowland model. Using the approach of simplified PBPK models with common scalars and the best models, PK profiles could be well described in preclinical species and plasma profiles were successfully predicted in human for diazepam and midazolam. This proof of concept was shown for lipophilic weakly basic compounds. For an exhaustive evaluation, the work and models proposed herein may be extended to different drug classes and more compounds. The PBPK modelling framework presented in this work for drug distribution and prediction of human PK could also be applied to translation within species e.g., from an adult to a paediatric population.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

Acknowledgements

First and foremost, I would like to acknowledge and declare my gratitude to my supervisors, Doctor Andrés Olivares-Morales, Doctor Michael Gertz, Doctor Kayode Ogungbenro and Professor Leon Aarons, for their time, expertise, guidance, constant support, and inspiration throughout this PhD project. I will always be grateful to you for allowing me to evolve as an independent researcher and giving me the opportunity to work with such brilliant and lovely people.

I acknowledge Roche pRED especially PLMS department for providing me with financial support, and I thank them for providing various learning, enriching, professional and personal development experiences. Doctor Thierry Lavé, Doctor Sherri Dudal and Doctor Benjamin Ribba have been especially supportive and helpful.

I would like to express my thanks to Neil Parrott and Dr Adam Darwich for all their insightful comments and scientific discussions on this project.

I would also like to thank my friends and colleagues at Roche from whom I have learnt a lot, and who made it a great place to work. Many thanks for all the scientific and non- scientific discussions, the fun times and help, which made all this period particularly enjoyable and completing this thesis would have been considerably more difficult otherwise.

Many thanks to my friends and colleagues in the Centre for Applied Pharmacokinetic Research (CAPKR) of the University of Manchester, for an amazing working environment during my research visits in Manchester. I enjoyed all the excellent discussions, the beers on Fridays, the dinners, the coffees, and laughter.

I would also like to take this opportunity to thank my friends from Basel for their kindness, support, and memorable moments that I really enjoyed. My thanks are extended to my friends from Paris for showing their support from a distance. I shall also thank my hiking buddies for all the exchanges, smiles and experiences spent outdoors. These adventures were always fulfilling, refreshing, and exhilarating breaks during a challenging period.

Much love to my parents, my brother and sister, and thank you for consistently caring for me, supporting me and encouraging me from abroad. Thank you to my past self for deciding one day to step into this PhD journey of thousands of miles...

Chapter 1: General introduction

1.1. Introduction

During the development of a new drug, it is important to characterise the absorption, distribution, metabolism, and excretion (ADME), to understand drug exposure especially after oral dosing and to link drug dose to efficacy and safety, and thus to predict a relevant dose in clinics. Absorption describes the rate and extent at which drugs enter the body. Once the drug is absorbed from the absorption site, it is distributed throughout the body, mostly via the blood circulation. The drug can be biotransformed into metabolites by organs or tissues (primarily the liver, the gastrointestinal tract, and kidneys). The drug can be transformed into more water-soluble metabolites to facilitate its excretion from the body via faeces or urine. Key enzymes to consider for drug metabolism are hydroxylation enzymes such as cytochrome P450 (CYPs) and conjugation enzymes such as UDP-glucuronosyltransferases (UGTs). The study of the ADME processes and the quantitative description of how these processes affect the concentration-time course of the drug in plasma, serum, or whole blood, tissue target and target organs is defined as pharmacokinetics (PK) [1].

Small molecule drug discovery and development is a long, challenging, and expensive process [2, 3]. It involves thousands of compounds being screened and tested in order to identify a novel drug that is useful and safe for the treatment of a disease or a condition. This process can be divided into different stages: discovery, pre-clinical, clinical Phases I to III and post-marketing. Each stage has its own aims and requirements and is designed to gather evidence and information about the PK, the safety and efficacy of the drug. Various PK data are generated from preclinical- to late-stage clinical development. Before the clinical phases, diverse animal, *in vitro* and *in silico* data are available. All the information available from those studies can be integrated through modelling and simulation (M&S) for supporting the design of human studies [4]. Thanks to the advancement in computer sciences in the last decades, M&S tools have become widely used in drug development and have helped to improve drug development and decision making [5, 6]. Various models can be used to predict PK parameters and subsequently to estimate human PK and clinical dose. These models vary from simple allometric scaling and *in vitro-in vivo* extrapolation (IVIVE) to more complex physiologically based pharmacokinetic (PBPK) modelling. The increased consideration for understanding the ADME properties has led to a reduction of attrition rate in drug research and development

due to PK reasons [7]. Additionally, confidence in the prediction of human PK allows assessment of potential drug-drug interactions (DDI) and mitigation of safety risks [8].

Regardless of the mechanism of drug action, the unbound concentration of drug attained at its site of action is the one driving therapeutic effect. Establishing and maintaining the optimal exposure help to determine the required dose and dosing interval for a compound. However, the unbound concentration of drug at these sites of action are generally unknown, especially if they are in tissues. The general assumption is to estimate the unbound concentration in tissues (Cu_t) based on the unbound concentration of drug in blood plasma (Cu_p), which are defined as follow (Eq. 1.1-Eq. 1.2):

$$Cu_t = C_{total,t} \times fu_t \quad \textbf{Eq. 1.1}$$

$$Cu_p = C_{total,p} \times fu_p \quad \textbf{Eq. 1.2}$$

where $C_{total,t}$ and $C_{total,p}$ are total concentration in the tissue and in plasma, respectively; fu_t and fu_p are the fraction of drug unbound in the tissue and in plasma, respectively.

At equilibrium, unbound plasma concentration and unbound tissue concentration would be equal for drugs distributed only by passive diffusion without efflux or influx by transporters, consequently (Eq. 1.3-Eq. 1.4):

$$C_{total,t} \times fu_t = C_{total,p} \times fu_p \quad \textbf{Eq. 1.3}$$

$$\frac{C_{total,t}}{C_{total,p}} = \frac{fu_p}{fu_t} = Kp \quad \textbf{Eq. 1.4}$$

The ratio of the total drug concentration in a tissue to plasma concentration at steady state is defined as the tissue-to-plasma partition coefficient (Kp). Kps are a measure of the steady-state distribution of the total drug concentration in tissue and will be described later in this chapter.

Together with clearance, the volume of distribution (Vd) are key PK parameters which determine the drug half-life and dosing regimen. The extent of drug distribution within the body is described by the volume of distribution [1]. Vd does not represent a physiological space, it is the apparent volume of fluid needed to contain the total amount of drug, assuming the drug is evenly distributed throughout the body, to produce the observed concentrations in plasma. Vd (Eq. 1.5) is the proportionality constant that relates the plasma concentration of drug (Cp) to its total amount in the body (A):

$$Vd = \frac{A}{Cp} \quad \textbf{Eq. 1.5}$$

As a measure of the extent of distribution into tissues, Vd can also be related to the protein binding and the physiological volumes of plasma (V_{plasma}) and tissues (V_{tissue}) and can be calculated as follows (Eq. 1.6) [9]:

$$Vd = V_{plasma} + \sum V_{tissue,i} * \frac{fu_p}{fu_t} \qquad \textbf{Eq. 1.6}$$

where fu_p and fu_t are the fraction unbound in plasma and body tissues, respectively, $V_{tissue,i}$ is the volume of the i^{th}-tissue.

The volume of distribution values can vary widely between individuals as it can be influenced by age, gender, disease state, and body composition [10, 11]. Consequently, this may affect the dose and dosing interval needed to maintain therapeutic concentrations e.g., when considering interactions in comorbidity and making translation between ethnicities, extrapolation to children. Factors explaining differences in drug distribution are factors that can influence the rate and/or extent of distribution of xenobiotics into tissues and which are detailed in the next section.

1.2. Factors influencing small molecule drug distribution

The distribution of a drug within the body refers to a generally reversible partitioning of drug from one compartment (systemic circulation) into another (extravascular tissues). This is mainly governed by blood flow rates and the ability of the drug to cross tissue and cellular membranes. Thus, tissue-to-plasma-levels of drug at steady state depend on compound properties and the physiological composition of different tissues. The rate of tissue distribution can be influenced by tissue blood perfusions, transporter-mediated uptake rate, and the membrane permeability (**Table 1.1**). The extent of tissue distribution is dependent upon tissue partitioning coefficients as well as binding to constituents within blood and tissues (**Table 1.1**). Therefore, both physiological and physicochemical factors can influence tissue distribution of a drug. The factors are summarized in **Table 1.1** and further detailed below.

Table 1.1. Drug- and physiology- specific factors affecting the rate and extent of drug distribution into tissues. Adapted from [12]

Determinants of tissue distribution	Physiological properties	Drug characteristics
Blood perfusion rate	Regional blood flow, organ size	
Membrane permeability	Tight junctions in capillary wall (for large molecules)	Molecular size, lipophilicity, pKa
Transporter-mediated distribution	Distribution of influx/ efflux transporters in various tissues	Functional groups
Binding to blood components	Albumin, α1-acid glycoprotein, lipoproteins, globulins, haemoglobin, red blood cells	Lipophilicity, acid/base/neutral character
Tissue binding	Extracellular albumin or lipoproteins, acidic phospholipid or lysosomal binding	Lipophilicity, acid/base/neutral character
Tissue partitioning	Tissue composition and volume, regional pH	Ionization and lipophilicity

Perfusion rate is defined as the volume of blood per unit time (blood flow) per unit tissue volume (or mass).

1.2.1. Tissue perfusion rate:

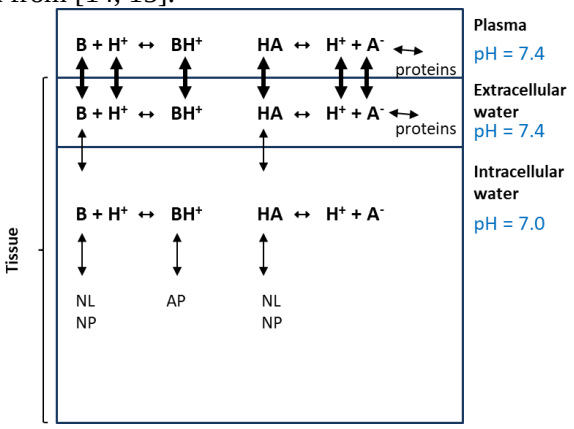
When drugs are highly lipophilic or the membrane is highly permeable, drug distribution becomes perfusion rate-limited. It is thus intrinsically dependent on the blood flow rate from the heart to the organs. Drug distribution thus occurs rapidly in highly perfused tissues such as heart, liver, kidney, spleen, or intestine, while it is generally slower in lowly-perfused tissues such as adipose tissues.

1.2.2. Membrane permeability

A small molecule drug can get across cell membranes by mainly two mechanisms: transcellular diffusion through the lipid membrane due to concentration gradients or active transport *via* carrier-mediated mechanisms [13]. Drug-related factors such as size, lipophilicity, ionisation, and functional groups (e.g., acidic or basic centres) influence the passive diffusion mechanism of molecules through the lipid membrane. When drugs are ionized at physiological pH, polar or water-soluble, drug distribution is permeability rate-limited, and the membrane and transporters are the rate-limiting step. Drugs can also be limited in their ability to penetrate into tissues depending on their ability to cross endothelial and cellular barriers, which can be highly selective (e.g., blood-brain barrier or placental barrier). Cell membranes are made up of lipid bilayers to which small

lipophilic drugs are highly permeable. Compared to acids and neutrals, basic compounds can enter more easily into cell membranes as lipid bilayers have negatively charged polar heads. Weak acids and bases contain at least one functional group that can reversibly disassociate or associate with a proton to form a negatively charged anion or a positively charged cation. Depending on the pH of the compartment media and the acid dissociation constant (pKa), the balance of the equilibrium can be shifted between ionised and unionised fractions (**Figure 1.1**). The extent of drug ionisation is determined by the pKa of the ionisable compound and the pH of the compartment fluid (tissues, plasma, and extracellular water). When a pH difference exists between physiological body compartments, then the extent of drug ionisation would influence its ability to distribute and cross cell membranes within the various body tissues. For large molecules, the rate of distribution into tissues is limited by their membrane permeability (tissue fenestration, and receptor mediated extravasation).

Figure 1.1: Binding of moderate to strong bases (B) to intracellular neutral lipids (NL), neutral phospholipids (NP) and acidic phospholipids (AP) and acids, weak bases and neutrals (A) to NL, NP and protein adapted from [14, 15].



Unionised fractions of moderate to strong bases (B) cross membranes and bind to intracellular NL, NP, while ionised fractions (BH^+) bind to AP. Unionised fractions (A) of acids, weak bases and neutrals cross membranes and bind to intracellular NL, NP, while ionised fractions (A^-) bind to extracellular proteins.

1.2.3. Transporters

Drugs distributed across cellular membranes *via* transporter proteins lead to accumulation in specific tissues or excretion from specific tissues. Because binding to drug transporters is a concentration-dependent and saturable process, it may introduce nonlinear distribution. Drug transporters are expressed at various levels in different

organs such as liver [16], kidney [17], intestine [18], and brain [19]. Efflux drug transporters, most importantly P-glycoprotein (P-gp), breast cancer resistance protein (BCRP), multidrug resistance protein (MRP), play an important role in drug transport across membranes and drug disposition [20]. Uptake transporters like organic anion transporting polypeptides (OATPs) and organic anion transporters (OATs) also play a significant role in drug absorption, distribution, and elimination [21]. Several *in vitro* and *in vivo* systems are commonly used to assess uptake and efflux transport kinetics [22]. Subsequent IVIVE of transporter-mediated processes use static and PBPK modelling approaches to predict transporter-mediated PK and DDIs.

1.2.4. Binding to blood components

Nonspecific drug binding to plasma proteins and binding of the drug to its target can also influence the extent of distribution as it is assumed that only the unbound drug material in blood distributes from systemic circulation into tissues [23]. Acidic drugs such as warfarin and acetylsalicylic acid mainly bind to albumin, the most abundant protein in plasma [24]. Neutral compounds can also bind to serum albumin [25]. Basic drugs such as lidocaine or propranolol usually bind to α1-acid glycoprotein (AAG), lipoproteins or globulins [26]. Additionally, various lipophilic compounds can bind to the membrane or the cellular components of the red blood cells (RBC) [27-29].

1.2.5. Binding to tissue components

Basic drugs generally more extensively distribute in tissues and thus have an apparent volume of distribution larger than the volume of acidic drugs. Many basic drugs are indeed ionized at physiological pH allowing them to bind to tissue constituents and also to accumulate in lysosomes which are able to trap lipophilic basic drugs, in particular cationic amphiphilic drugs (CADs) such as chlorpromazine, imipramine due to the low pH associated with them [30, 31]. Specific binding of drugs to tissues components (e.g., membrane phospholipids, DNA, proteins) can increase drug distribution by several fold in those tissues [32, 33]). And if the binding is irreversible, drugs tend to concentrate inside the tissue, leading to drug accumulation (e.g., tetracycline can distribute and accumulate in bone and teeth [34]).

1.2.6. Tissue partitioning

The unbound tissue partition coefficients reflect the extent of drug distribution into tissue and are dependent on both drug characteristics and tissue composition. Tissue compositions influence tissue partitioning and therefore the extent of drug distribution; hydrophilic drugs tend to distribute into water-rich tissues (e.g., muscle) or stay in the blood or interstitial spaces. On the other hand, lipophilic drugs may concentrate in fat-rich tissues (e.g., adipose, liver, brain, and kidney) which may increase the Vd and drug half-life, with a possible slow drug accumulation and delayed washout during and after chronic dosing. However, as plasma drug concentration decreases, tissues will release the accumulated drug back into the vascular space. Therefore, lipid contents of tissues (i.e., neutral lipids and phospholipids, acidic phospholipids) and physical tissue volumes appear as key physiological determinants of distribution.

1.3. Prediction of drug distribution

1.3.1. Volume of distribution at steady state (Vss)

Estimation of the volume of distribution at steady state (Vss) is a key parameter to predict the required dosing interval for a compound. Vss is the volume of distribution at equilibrium, i.e., when all the drug is distributed into the different tissues and is not influenced by the elimination (if elimination does not occur in any of the peripheral tissues).

Several methods can be used to estimate distribution volume based on experimental data [35]. Following an IV bolus dose, Vss can first be estimated by noncompartmental analysis using the following expression (Eq. 1.7):

$$Vss = \frac{Dose}{C_0} \qquad \textbf{Eq. 1.7}$$

Where C_0 is the initial concentration in plasma determined by plotting the log concentrations versus (linear) time and extrapolating back to time zero, before clearance began. This Eq. 1.7 assumes that the drug does not distribute into tissues and elimination half-life is very short, which may not be true. Assuming a mono-compartmental model, Vss can be estimated using the volume to link the elimination clearance (CL) to elimination half-life ($t_{1/2}$):

$$t_{1/2} = Vss \cdot \frac{0.693}{CL} \quad \textbf{Eq. 1.8}$$

However, the Vss in Eq. 1.8 is likely over-estimated due to the assumption of mono-exponential PK and no consideration of the contribution of multi-exponential elimination to the terminal half-life.

1.3.2. Prediction of human Vss

Various methods have been proposed to predict Vss (especially in human) based on preclinical data in early drug development (before any clinical data is obtained). Most common methods can be classified into empirical, semi-mechanistic and mechanistic methods. Although many studies have compared the Vss prediction accuracy of different methods [36-46], there is no clear agreement on which method is the most accurate as they reported varying results. This could be due to different sizes of datasets used and different types of compounds included in the comparison studies.

Empirical methods

The first empirical method is a single species allometric scaling where human Vss is simply derived from the Vss of preclinical species [38]. It is assumed that the unbound Vss (Vuss) normalised by body weight is the same between species (Eq. 1.9):

$$Vuss_{human} = Vuss_{animal} \quad \textbf{Eq. 1.9}$$

where Vuss=Vss/fu_p. This method is based on the assumptions of similarity in tissue composition across species, species-independence of drug partitioning into tissue lipids, and species difference in plasma protein binding.

Another empirical approach is allometric scaling where the physiological or pharmacokinetic parameter (here, Vss) is related to the power of body weight (Eq. 1.10) [47]:

$$Vss = a \cdot BW^b \quad \textbf{Eq. 1.10}$$

where BW is the body weight, a and b are the allometric coefficient and exponent, respectively. The coefficient and exponent are determined by fitting the allometric function to preclinical information (e.g., PK parameters) of various animal species of differing body weights. The fitted function is then used to extrapolate Vss from chosen

preclinical species to human. For multiple species, the exponents of volume of distribution for a drug generally tend to be close to 1, implying that Vss normalised by body weight is constant across species [48]. However, this is not observed for all drugs. One reason is that because of its empirical nature, allometry does not consider species differences in protein expression, enzymes, transporters, and plasma protein binding, amongst others. Alternatively, several studies have attempted to improve allometric predictions by correcting for plasma and/or tissue protein binding [45, 49].

Alternatively, Wajima et al. suggested to scale human Vss from two species: rat and dog using multivariate regression analysis (Eq. 1.11) [50]:

$$Log(Vss_{human}) = 0.07714 \cdot Log\,(Vss_{rat}) \cdot Log\big(Vss_{dog}\big) + 0.5147 \cdot Log\big(Vss_{dog}\big) + 0.586 \quad \textbf{Eq. 1.11}$$

Semi-mechanistic methods

The semi-mechanistic models also require preclinical *in vivo* data, but they consider species differences in physiology. Based on the relationship reported by Gillette et al [9], the Øie-Tozer method was one of the first physiological models for predicting human Vss [51]. In this model, it is assumed that the unbound drug distributes into extra- and intra-cellular space where it can interact with extracellular proteins and tissue components, as described in the following equation (Eq. 1.12):

$$Vss_{human} = V_{plasma} + \left(\mathrm{fu_{p,human}} \cdot V_e\right) + \left[\left(1 - \mathrm{fu_{p,human}}\right) \cdot R_{e/i} \cdot V_{plasma}\right] + V_r \cdot \frac{fu_{p,human}}{fu_{t,animal}} \quad \textbf{Eq. 1.12}$$

where $R_{e/i}$ is the ratio of total binding sites in extracellular fluids outside the plasma to that in plasma, Ve is the extracellular fluid volume, and Vr is the remaining fluid volume into which drug distributes. This approach assumes fu_t to be the same between species. Two different equations can be used to calculate fu_t and both need *in vivo* Vss, fu_p and some physiological data. A value of fu_t in preclinical species of less than or equal to zero would indicate no binding to tissue components. Recently, the Øie-Tozer model has been modified to consider other tissue components such as lysosomes that can influence the volume of distribution [52].

Wajima and Oie-Tozer approaches are commonly used in practice in the pharmaceutical industry to predict human Vss. In a study comparing 24 approaches using *in vitro* and *in vivo* data to predict Vss in human, Jones et al. reported that the Wajima model was the second most accurate model (94% of human Vss predicted within 3-fold of observed Vss) and the Øie-Tozer method being the first (78% of predicted Vss within 2-fold of observed Vss) [38]. However, these models generally require *in vivo* animal data which are not always available in drug discovery due to resource, time, and cost limitations as well as ethical issues (3Rs principles). Therefore, many efforts have been put into developing *in vitro* and *in silico* methods for predicting Vss early. Lombardo et al. rearranged the Øie-Tozer model to estimate fu_t based on physicochemical properties without requiring *in vivo* preclinical data for neutral and basic drugs (Eq. 1.13) [53, 54].

$$Log(fu_t) = 0.0080 - 0.2294 \cdot Log\ (D_{o:w}) - 0.9311 \cdot fi_{7.4} + 0.8885\ Log\ (fu_p) \quad \textbf{Eq. 1.13}$$

where $fi_{7.4}$ is the fraction of drug ionized at pH 7.4.

Additional quantitative structure-activity relationship (QSAR) models have been proposed to predict human Vss in early drug discovery when limited animal data are available [41, 46, 55-57]. These models are generally built using data mining (or machine learning) methods and predict Vss using physicochemical or molecular descriptors. Several studies suggested that these methodologies can predict Vss reasonably well within 2-fold difference of the actual value [55-57]. These *in silico* models are high-throughput approaches which can be used to assess compounds' PK even before their synthesis. One general limitation of these QSAR approaches could be the inability to predict the Vss of a compound in a new unknown chemical class, which was not included in the large dataset used for training and validation. Nevertheless, Lombardo et al. recently published a large dataset of human Vss values for 1352 structurally diverse compounds [58], presenting the opportunity to assess new QSAR models [41, 46, 59].

Mechanistic methods

Human Vss can also be estimated from tissue-to-plasma partition coefficients (Kp). Assuming tissue distribution is only driven by passive diffusion, the Vss can be related to physiological volumes of the body and a measure of the distribution at steady-state of the total drug concentration and tissues is reflected by Kp (Eq. 1.14) [60]:

$$Vss = V_{plasma} + Kp_{RBC} \cdot V_{RBC} + \sum_{i}^{n} Kp_i \cdot V_{tissue,i} \cdot (1 - E_i) \quad \textbf{Eq. 1.14}$$

where V_{plasma} and $V_{tissue,i}$ are the physiological volume of plasma and the tissue i, V_{RBC} is the volume occupied by red blood cells (V_{blood}-V_{plasma}). E is the extraction ratio of an eliminating tissue such as the liver. For non-eliminating tissues, E=0. Kp_{RBC} is the red blood cell-to-plasma ratio and can be determined from fu_p, blood:plasma ratio (BP) and the haematocrit (Hte) as in Eq. 1.15:

$$Kp_{RBC} = \frac{Hte - 1 + BP}{Hte} \quad \textbf{Eq. 1.15}$$

In the plasma, the drug is assumed to be in rapid equilibrium with plasma proteins and only the unbound and unionised fraction of drug can distribute into tissue. The tissue-to-plasma water partition coefficients (Kpu) reflect the steady state total concentration of drug in the tissues ($C_{total,t}$) to the unbound concentration in plasma ($C_{u,p}$) following a constant rate drug infusion (Eq. 1.16):

$$Kpu = \frac{C_{total,t}}{C_{u,p}} \quad \textbf{Eq. 1.16}$$

The unbound tissue partition coefficients (Kpu,u) illustrate the extent of steady-state distribution between the unbound fractions in tissues and in plasma (Eq. 1.17):

$$Kpu,u = Kp * \frac{fu_t}{fu_p} \quad \textbf{Eq. 1.17}$$

(Kpu,u=1 for passive diffusion, Kpu,u<1 for efflux and Kpu,u>1 for influx).

1.3.3. Prediction of human tissue partition coefficients

Kp values need to be known in each tissue in order to predict the human Vss for a compound. These Kp data can be determined experimentally from preclinical species, but it is however time consuming and costly. Consequently, many *in vitro* and *in silico* methods have been developed to estimate Kp values in order to predict Vss. The different methods for estimating Kp are described in greater details in the fourth section of this chapter. One of the methods is to use mechanistic tissue composition-based prediction models to predict tissue Kps. They have the advantage that they can be used in early drug discovery without requiring an *in vivo* study as they can rely on *in vitro* measured or

computed physicochemical properties of lipophilicity, ionisation, plasma protein binding, and blood-to-plasma concentration ratio.

Some of the tissue composition models use human physiological data [61] for prediction of human Kps and human Vss [36, 62]. However, experimental human tissue-to-plasma partition coefficients are rarely available and subsequently the reliability of the *in silico* partition coefficient estimates cannot easily be assessed. Nevertheless, a recent study showed that this approach is comparably accurate or superior to empirical allometric approaches based on the extrapolation of *in vivo* animal Vss data to predict Vss in human. It was also reported that prediction of human Vss is more accurate when the Vss in preclinical species is well predicted [36].

Alternatively, human Vss can be extrapolated from tissue-to-plasma partition coefficients in preclinical species (generally rodents) [44, 63]. Kpu values are commonly assumed to be directly translated between species by considering that tissue binding is conserved across species (Eq. 1.18) [1, 47, 49]:

$$Kpu_{human} = Kpu_{animal} \qquad \textbf{Eq. 1.18}$$

1.3.4. Prediction of PK profiles

The Kpu predictions can then be inputted into whole-body PBPK models. PBPK modelling provides a powerful tool for integrating preclinical data into human PK predictions of PK parameters and concentration-time profiles. The concept of PBPK modelling is based on an approach that integrates broad information on species physiology and a wide understanding of the mechanisms affecting the pharmacokinetic profile of a drug. PBPK model structure and how Kpus are integrated into the model are detailed in the section 5 of this chapter.

Other approaches for predicting human PK profiles based on preclinical data include Dedrick plots and Wajima's method [64-67]. The concept of the Dedrick plot or 'species-invariant time method' is based on allometic scaling and assumes that differences across species are based on body weight [64]. It assumes that after normalizing concentrations by body weights and transforming the chronological time to the equivalent time (kallynochrons and apolysichrons), the plasma concentration-time curves should be superimposable for all species [68]. A similar idea was later proposed by Wajima et al where the assumption is that after normalizing concentrations by steady-state plasma drug

concentration (Css=Dose/Vdss) and normalizing the time by mean residence time (MRT=Vss/CL), the plasma concentration–time curves should be superimposable across species [65]. These two empirical approaches can give good predictions [64, 65, 69] but inaccurate results are common especially if biliary clearance, renal secretion, or high metabolism exist and interspecies differences in these mechanisms are not well understood [70-73]. Application of correction factors can improve predictions in certain cases [68, 74, 75] but the assumption of linearity still has to be made and large interspecies differences in active processes are not addressed using these empirical approaches. Therefore, the use of PBPK modelling for predicting human PK profiles is preferred and several studies showed PBPK modelling performed better in comparison to empirical and allometric approaches [76-79].

1.4. Determination of tissue-plasma partition coefficients

Both physicochemical characteristics and tissue compositions are important determinants of the extent of drug distribution. Most of these factors are accounted for in a PBPK model by the tissue-to-plasma partition coefficients. The Kp values are key components for the characterization of the extent of drug distribution into different tissues in the body and reflect the degree of tissue distribution attributed to processes such as protein binding, lysosomal trapping, and lipid interaction. Kp values are defined as the ratio of total concentration of drug in the tissue to total concentration of drug in the plasma at equilibrium/steady-state. There are several methods employed to investigate drug distribution; the most general are *in vivo* animal and human PK studies. However, with the development of new *in vitro* and *in silico* predictive tools, it is now common to predict drug distribution in the early drug development stages.

1.4.1. *In vivo* methods

The most common way is to derive Kp values *in vivo* from blood and tissue concentrations after administration of the drug to the designed species (usually rat or mouse). One established approach is to calculate Kp values from steady-state plasma and tissue concentrations following a constant rate infusion of drug until steady state is obtained [80-82]. After sacrificing the animal at one time point, sampling is performed from the tissues and plasma. Kp is calculated as the ratio of concentration in the tissue to

the concentration in plasma at that time point. It is important to bear in mind that if the animal is sacrificed before the steady-state is reached in all tissues, the Kp values is likely to be under-estimated. To ensure that the drug has reached steady-state, the experiment is often repeated at several different time points and similar measured values would indicate that steady state has been reached. Alternatively, Kp can be calculated following an i.v bolus injection of drug and sampling of tissues and plasma at different time points [83-85]. Kps are then calculated as the ratio of areas under concentration time (AUC) profiles between tissues and plasma. For this approach, it is important to ensure that each tissue reaches one common terminal (elimination) phase after initial distribution (if tissues' terminal phases differ among one another, pseudo-steady state is not reached). Additionally, this method may not be suited for compounds that are rapidly eliminated from the body if the plasma and tissue AUC are not well estimated [86]. Additionally, the sampling site can also have an impact on measured drug concentrations with different concentrations being observed between arterial and venous sampling, and correction for eliminating tissues need to be consider [87-89]. It was also shown that venous and arterial concentrations are the same when lung distribution is negligible and distribution equilibrium is instantaneous.

All the *in vivo* techniques require at least one animal per sampling time. At each time point, the removed tissues are homogenized, and the total tissue concentrations can be determined by liquid chromatography/mass spectrometry [90] or by liquid scintillation counting [91, 92]. The extraction method should be adequately chosen depending on tissues to accurately measure the drug concentration within each tissue as some solvents may interact or alter the drug and some tissues need to be pre-treated [93, 94].

Other methods have also been used including imaging techniques such as quantitative whole-body autoradiography (QWBA) or positron emission tomography (PET) microdosing [91, 95-97]. In the QWBA, animals are dosed with radiolabelled compounds and then sacrificed in order to determine the tissue distribution of the radioactivity. No extraction step is necessary in this method and the radioactivity can be measured in the whole animal or in the individual tissues. However, one limitation may be its lack of specificity since by measuring the total radioactivity it does not distinguish parent compounds from metabolites. Another method is PET, a non-invasive method which makes it possible for use in human and where the drug is labelled with a positron-emitting radionuclide. Although the *in vivo* distribution of the labelled drug can be

detected externally and visualised as tomographic images, it is also difficult to distinguish radio-metabolites from the parent tracer. One of the main drawbacks of PET studies is its high cost and thus they are more carried out in drug development rather than in drug discovery.

1.4.2. *In vitro* methods

In order to reduce animal use in preclinical experiments, a number of *in vitro* methods have been developed to estimate Kp values. For volatile compounds, tissue:air partition coefficients can be measured using the vial equilibration method: the compound is allowed to equilibrate between the air and the blood (or tissue) at body temperature in a sealed vial after shaking [98, 99]. The sampling of compound in the vial headspace is performed by gas chromatography and tissue-to-plasma partition coefficients can be calculated simply as the ratio of tissue:air and plasma:air partition coefficients. The *in vitro* distribution or binding of drugs to tissues may be determined using tissue homogenates [84, 100], slices [101] and isolated tissue components [102]. For non-volatile compounds, tissue partitioning can be measured by equilibrium dialysis or ultrafiltration: the compound can distribute between blood (or plasma) and tissue homogenates through a membrane in a multi-chamber system [100, 103]. Tissue homogenates are the most extensively used although it has been argued that the compared to tissue slices homogenization may alter the cellular integrity of tissue (e.g., brain) without differentiating between intra- and extracellular drug distribution and thus tissue concentrations obtained may not be relevant [104]. This may lead to a measured *in vitro* Kp overestimating the actual *in vivo* distribution for drugs for which distribution is mainly restricted to the extracellular space *in vivo* since the total aqueous phase and intracellular binding sites have become available. Another method is to study *in situ* perfused organs (e.g., brain, liver) but their use is limited as it is low throughput [105, 106].

1.4.3. *In silico* methods

Various *in silico* models have been developed to predict tissue partition coefficient values for drugs (Table 1.2). Rather than requiring *in vivo* data, these mechanistic methodologies estimate the extent of tissue distribution from parameters already available in the literature or quickly and easily obtainable with simple experiments, such as

physicochemical and *in vitro* binding characteristics to lipids and proteins of the drug, and this allows the routine application of PBPK methods in early drug discovery. Lipophilicity (logP), ionisation (pKa), plasma and/or blood protein binding (fu_p or fu_b) are drug-parameters routinely measured during drug development. Several methods for the calculation of Kps are widely used and have been reviewed in the literature. More details can be found in the following publications [39, 107] and a summary can be found in Table 1.2.

Table 1.2: Summary of *in silico* Kp prediction models and their main inputs

Models	Approach	Main inputs
Arundel [108]	Correlation-based	Vss, logD
Bjorkman [109]	Correlation-based	Muscle Kp
Poulin [60, 110]	Tissue-composition-based	logP, $logK_{vo:w}$, fu_p
Berezhkovskiy [111]	Tissue-composition-based	logP, $logK_{vo:w}$, fu_p
Rodgers [14, 15]	Tissue-composition-based	logP, pKa, fu_p, B:P
Richter [112]	Correlation-based	muscle Kp
Lukacova [113]	Tissue-composition-based	logP, pKa, fu_p, B:P
Schmitt [114]	Tissue-composition-based	logP, logD, $logK_{vo:w}$, logMA, pKa, fu_p
Jansson [115]	Correlation-based	Vss, muscle Kp, logP, logD, $logK_{vo:w}$
Poulin and Theil [61]	Correlation-based	Muscle Kp, skin Kpu or *KpuRBC*
Peyret [116]	Tissue-composition-based	logP, $logK_{vo:w}$, pKa, fu_p, B:P, $K_{p:w}$, K_{prwp}
Poulin and Haddad [117]	Tissue-composition-based	logP, $logK_{vo:w}$, B:P, pKa, fu_p,
Yun and Edginton [118]	Correlation-based	Vss, logP, pKa, fu_p
Assmus [31]	Tissue-composition-based	logP, pKa, fu_p, B:P
Korzekwa and Nagar [52]	Tissue-composition-based	logP, B:P, pKa, fu_p, fu_m
Mayumi [119]	Correlation-based	Muscle Kp, logP, pKa, fu_p, B:P

Kp : Tissue-to-plasma partition coefficient ; Kpu :Tissue-to-unbound plasma partition coefficient; KpuRBC : Red blood cell to plasma partition coefficient data for unbound drugs ; B:P :Blood-to-plasma ratio ; $logK_{vo:w}$: Logarithmic value of vegetable oil–water partitioning adjusted for ionization at pH 7.4; LogP : Logarithmic value of N-octanol–water partition coefficient; logD: logP adjusted for ionisation at pH 7.4; logMA: Logarithmic value of the phosphatidylcholine-water partition coefficient at pH 7.4; $K_{p:w}$: protein:water partition coefficient ; K_{prwp} : plasma and interstitial fluid protein:water partition coefficient ; fu_p : Unbound fraction in plasma; fu_m : Unbound fraction in microsome; Vss : Volume of distribution at steady state

Overall, the equations of *in silico* models for predicting Kps are mostly based on physicochemical drug properties and physiological tissue composition information. However, they generally do not incorporate active transport (influx/efflux transporters) of drugs. Poulin et al. first suggested the development of a mechanistic model to estimate Kps by assuming that compounds dissolve into water content of tissues and partition into the lipids and phospholipids within tissue cells [60, 110]. Berezhkovskiy later modified the model by Poulin et al. and considered that only the unbound drug fraction in water bind to tissues [111]. Subsequently, Rodgers et al. developed two equations - one for strong bases with pKa>7 and one for neutrals, acids and weak bases) to take into account the impact of drug ionization on partitioning [14, 15]. To avoid the cut-off at a pKa of 7 in the Rodgers et al. model, Lukacova et al. suggested to transform the two equations into a single continuous equation [113]. A recent study found that the performances of Rodgers et al. and Lukacova model are very similar, with the exception for compounds having pKa around 7 [46]. Alternatively, Schmitt examined electrostatic interactions between charged molecules and acidic phospholipids and also considered distribution in cells and interstitial fluids in their unified algorithm [114]. On the other hand, Peyret et al. integrated previous mechanistic algorithms into a unified algorithm to predict tissue partition coefficients for drugs and environmental chemicals [116], whereas Poulin and Haddad developed a model for highly lipophilic compounds where Kp and Vss do not increase exponentially when LogP is above 6 [117]. Recently, Assmus et al. extended the Rodgers et al. model to include ion partitioning into acidic or basic intracellular compartments (lysosomes and mitochondria) which improved predictions especially for basic compounds [31]. Additionally, a recent study suggested the use of partitioning into microsomes (unsorted phospholipid membranes) to determine interactions with all phospholipids for the Kp prediction of both charged and uncharged compounds [52, 120].

An alternative to tissue composition-based models are correlation-based Kp prediction models which are empirical in nature and require more experimentally derived data. These correlation-based models use both physicochemical properties of a drug [115, 118, 119] and experimental data, such as muscle Kp [61, 109, 112, 115, 119, 121] or volume at steady-state Vss [108, 115, 118] and red blood cell partitioning (RBCu) data [61] as surrogate variables (Table 1.2).

Although many *in silico* models have been proposed, no general rules have emerged to determine which Kp prediction model is the best as one model cannot predict

the Kp of all model compounds accurately. A few studies have compared several Kp prediction methods [39, 62, 122]. Graham et al. performed a comparative study on the predictive performance of three mechanistic models [14, 15, 60, 111] and three empirical models [61, 108, 115] to predict the most accurate Kps for rat tissues [39]. They reported that the model proposed by Rodgers et. al. [14, 15] showed the most accurate Kp predictions with 77% of predicted values within threefold although transporter systems still need to be incorporated into these equations. More recently, Utsey et al. assessed the performance of five models for Kp predictions using a standardized tissue composition for humans [62]. However, their analysis found that no single Kp model consistently perform more accurately for Kp predictions than the other models. Nevertheless, the Rodgers et al. model generally showed good performance and is currently the most used model for predicting tissue Kp values in PBPK modelling. Therefore, the work of this thesis focussed on the Rodgers et al. model and further details of the model are provided below.

Rodgers et al. proposed two mechanistic equations: one for predicting Kpu for moderate-to-strong bases and group 1 zwitterions (at least one basic pKa>7) [14], and another for acids, very weak bases, neutrals and group 2 zwitterions (no basic pKa >7) [15]. The Rodgers et al. model integrated the electrostatic interactions between a drug and the tissue components. Acidic drugs and lipophilic neutrals bind to albumin and lipoproteins respectively, which are both present in the extracellular fluids in a tissue. Strong or moderate bases preferentially bind to acidic tissue phospholipids and to acidic components such as AAG present in plasma. The equations include drug partitioning into neutral lipids (NL) and phospholipids (PL), drug dissolution into the tissue intracellular (IW) and extracellular (EW) water, and also associations with extracellular proteins such as albumin or lipoprotein [14, 15, 94, 123] and are as follow:

Moderate to strong bases (Eq. 1.19):

$$Kpu = \left(\frac{X \cdot f_{IW}}{Y}\right) + f_{EW} + \left(\frac{P \cdot f_{NL} + (0.3P + 0.7) \cdot f_{IW}}{Y}\right) + \left(\frac{K_a \cdot [AP^-]_1}{Y}\right) \quad \textbf{Eq. 1.19}$$

Weak bases, acids and neutrals (Eq. 1.20):

$$Kpu = \frac{X \cdot f_{IW}}{Y} + f_{EW} + \left(\frac{P \cdot f_{NL} + (0.3P + 0.7) \cdot f_{NP}}{Y}\right) + \left[\left(\frac{1}{fu_p} - 1 - \left(\frac{P \cdot f_{NL_p} + (0.3P + 0.7) \cdot f_{NP_p}}{Y}\right)\right) \cdot \frac{[PR]_T}{[PR]_p}\right] \quad \textbf{Eq. 1.20}$$

where P refers to the partition coefficient of the unionized drug. f_{EW}, f_{IW}, f_{NL}, f_{NP}, and f_{AP}, represents fractions of extracellular water, intracellular water, neutral lipid, neutral phospholipids, and acid phospholipids respectively, with Σf=1, so that the fraction of residual tissue constituents f_{REM} =1 –f_{IW} –f_{EW} –f_{NL} –f_{NP}. P is the octanol:water partition coefficient for all tissues except adipose, for which P represents the vegetable oil:water partition coefficient. K_a is the association constant of strong basic compounds for acidic phospholipids. $[AP^-]$ is the tissue concentration of acidic phospholipids. $[PR]_T$ and $[PR]_p$ are the concentrations of albumin (for acids and weak bases) or lipoproteins (for neutrals) in the tissue and in plasma respectively. In Eq. 1.19 and Eq. 1.20, X and Y represent the degrees of ionization, as described by the Henderson-Hasselbalch equation, and are as follows:

Monoprotic bases:	$X = 1 + 10^{pKa - pH_{IW}}$	$Y = 1 + 10^{pKa - pH_p}$
Monoprotic acids:	$X = 1 + 10^{pH_{IW} - pKa}$	$Y = 1 + 10^{pH_p - pKa}$
Neutrals:	$X = 1$ (no ionization)	$Y = 1$ (no ionization)
Zwitterions:	$X = 1 + 10^{pKa_{base} - pH_{iw}} + 10^{pH_{iw} - pKa_{acid}}$	$Y = 1 + 10^{pKa_{base} - pH_p} + 10^{pH_p - pKa_{acid}}$

where pH_{IW} and pH_p refer to the pH of the intracellular water and plasma respectively.

K_a can be calculated from the red blood cell-to-plasma water partition coefficient *i.e.*, unbound drug concentration in blood cells (Kpu_{RBC}), based on the concentration of acidic phospholipids (AP) in red blood cells (Eq. 1.21) [15]:

$$K_a = \left[Kpu_{RBC} - \left(\frac{1 + 10^{pKa} - 10^{pH_{RBC}}}{1 + 10^{pKa - pH_p}} \cdot f_{IW,RBC}\right) - \left(\frac{P \cdot f_{NL,RBC} + (0.3P + 0.7) \cdot f_{NP,RBC}}{1 + 10^{pKa - pH_p}}\right)\right] \cdot \left(\frac{1 + 10^{pKa} - 10^{pH_p}}{[AP^-] \cdot 10^{pKa - pH_p}}\right) \quad \textbf{Eq. 1.21}$$

In Rodgers *et al*. [14], Kpu_{RBC} is estimated using the standard equation (Eq. 1.22)

$$Kpu_{RBC} = \frac{BP - (1 - Haematocrit)}{f_{up} \cdot Haematocrit} \quad \textbf{Eq. 1.22}$$

where BP is the blood: plasma-ratio and haematocrit value is 0.45.

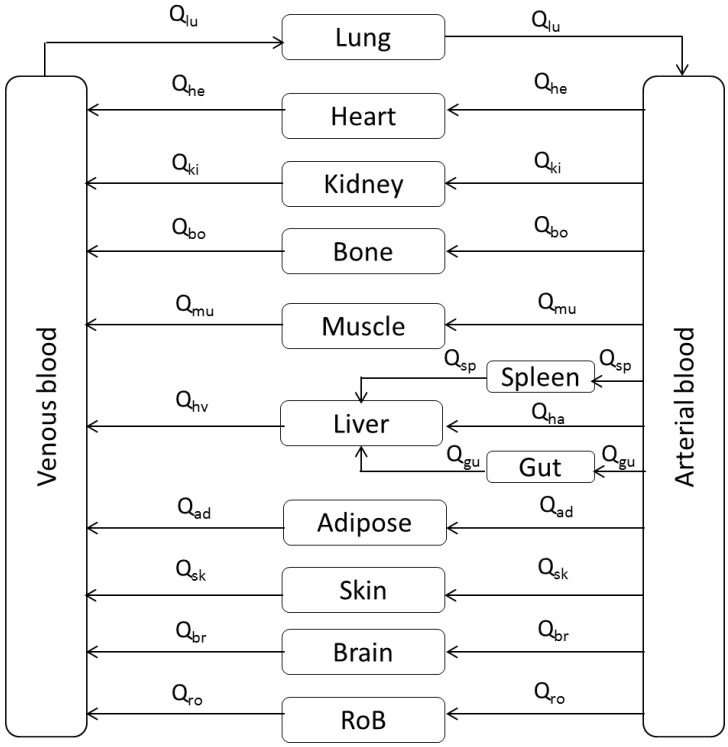
1.5. PBPK modelling

PBPK models are a mathematical representation of physiological and drug-specific processes that govern drug ADME. Such models allow simulations of the concentration-time profiles or pharmacokinetic (PK) profiles of a drug in various body compartments specified within the PBPK model which aids drug discovery and development and regulatory decision-making processes regarding optimal first in human (FIH) dose prediction and trial design, prediction of drug-drug interactions (DDIs), extrapolations of PK in special population and development of oral formulations, amongst others [78, 124-130]. The development and implementation of PBPK models have become more accessible with the availability of user-friendly PBPK software packages such as GastroPlus® (Simulations Plus, Lancaster, CA, USA), PK-Sim® (Bayer Technology Services, Leverkusen, Germany) and SimCYP® (Certara,Sheffield, UK), amongst others [76, 128, 131-134].

1.5.1. Model structure

The full (whole-body) PBPK modelling is particularly suited to integration of broad knowledge of different origins (e.g., *in silico*, *in vitro*, and *in vivo*) in an anatomically and physiologically relevant framework for each species ("bottom-up" approach). For small molecules, the structure of this model consists of compartments representing individual organs or tissues of the body, interlinked *via* the systemic blood circulation (**Figure 1.2**). For large molecules, PBPK models also include lymph flow and target-mediated drug disposition (TMDD). Each organ is defined as a homogenous compartment limited by perfusion rate or permeability rate [1, 135].

Figure 1.2: Schematic representation of a whole body PBPK model adapted from [127]



Q refers to blood flow: to the lungs (Q_{lu}), the heart (Q_{he}), the kidneys (Q_{ki}), the bones (Q_{bo}), the muscles (Q_{mu}), the spleen (Q_{sp}), the liver (Q_{ha}), the hepatic vein (Q_{hv}), the gut (Q_{gu}), the adipose (Q_{ad}), the skin (Q_{sk}), the brain (Q_{br}) and the rest of body (Q_{ro}).

Assuming tissues are perfusion-limited (well-stirred) compartments, each non-eliminating tissue can be described by the following equations (Eq. 1.23-Eq. 1.24):

Lung
$$V_{lung} \cdot \frac{dC_{lung}}{dt} = Q_{lung} \cdot \left(\mathrm{C_{venous}} - \frac{C_{lung}}{Kb_{lung}} \right) \quad \textbf{Eq. 1.23}$$

Other tissues
$$V_i \cdot \frac{dC_i}{dt} = Q_i \cdot \left(\mathrm{C_{arterial}} - \frac{C_i}{Kb_i} \right) \quad \textbf{Eq. 1.24}$$

where C_{lung}, C_i , C_{venous} and $C_{arterial}$ are the total drug blood concentrations in the lung, the i^{th}-tissue, the influent venous blood and arterial blood, respectively. V_{lung}, Q_{lung}, V_i and Q_i are the volumes and blood flows for the lung and each i^{th}-tissue, respectively. Kb_{lung} and Kb_i are the tissue-to-blood partition coefficient in the lung and in the i^{th} tissue and represent the tissue to venous blood concentration ratio at steady state. Kb values are determined from Kpu, the tissue-to-unbound plasma partition coefficients (Eq. 1.25).

$$K\mathrm{b} = Kp\mathrm{u} \cdot \frac{fu_p}{BP} \quad \textbf{Eq. 1.25}$$

where fu_p and BP are the fraction unbound in plasma and the blood-to-plasma ratio, respectively.

When renal elimination (CL_R) is part of the venous compartment, the rate equations for the arterial blood and venous blood compartments are defined as follows (Eq. 1.26- Eq. 1.27):

$$V_{arterial} \cdot \frac{dC_{arterial}}{dt} = Q_{lung} \cdot \left(\frac{C_{lung}}{Kb_{lung}} - \mathrm{C_{arterial}} \right) \quad \textbf{Eq. 1.26}$$

$$V_{venous} \cdot \frac{dC_{venous}}{dt} = \sum Q_i \cdot \frac{C_i}{Kb_i} - Q_{lung} \cdot C_{venous} - \mathrm{CL_R} \cdot C_{venous} \quad \textbf{Eq. 1.27}$$

where $\sum Q_i \cdot \frac{C_i}{Kb_i}$ includes all the ith tissues except the stomach, gut, pancreas, and spleen; $V_{arterial}$ and V_{venous} are the volume of arterial and venous blood, respectively.

For the liver, the rate equation is defined as (Eq. 1.28):

$$V_{liver} \cdot \frac{dC_{liver}}{dt} = Q_{ha} \cdot \mathrm{C_{arterial}} + \sum Q_{splan,i} \cdot \frac{C_{splan,i}}{Kb_{splan,i}} - Q_{liver} \cdot \frac{C_{liver}}{Kb_{liver}} - CLint_{liver} \cdot fu_b \cdot \frac{C_{liver}}{Kb_{liver}} \quad \textbf{Eq. 1.28}$$

where the $Q_{splan,i}$, $C_{splan,i}$ and $Kb_{splan,i}$ are the concentration, the blood flow, the volume and the blood partition coefficient of the i^{ith} splanchnic organs (stomach, gut, pancreas and spleen); Q_{ha} is the blood flow from the hepatic artery; C_{liver}, Q_{liver}, V_{liver}, Kb_{liver} are the concentration, the blood flow, the volume and the blood partition coefficient of the liver; CLint is the hepatic intrinsic clearance and fu_b is the fraction unbound in blood.

The PBPK model parameters can be generally classified into system-specific and drug-specific parameters. System-specific parameters characterise physiological and species-related properties (e.g., blood flow rate, organ volume, tissue composition, enzyme or transporter abundance [incl. genetic polymorphisms], plasma protein concentrations, haematocrit, etc.). Most of the values for these physiological parameters for humans and other species (commonly mouse, rats, and non-human primates (NHP)) are available in the literature [136-138]. However, these values usually refer to a "reference individual" while each of the biological parameters (physiological, anatomical, enzymatic, and transporters) may significantly vary between individuals and within populations.

In contrast, drug-specific parameters are the compound's physicochemical, biochemical and biopharmaceutical properties and other parameters that are relevant to the ADME processes of the drug (e.g., binding to blood, tissue-to-plasma distribution

coefficient, permeability, metabolic rate, etc.) and which generally differ between drugs. These parameters are generally incorporated into PBPK models by extrapolating them from *in vitro* systems to *in vivo* with the use of physiologically based scaling factors in the so-called *in vitro– in vivo* extrapolation (IVIVE) approach [139]. However, uncertainty in model parameters may exist due to the uncertainty associated with the *in vitro* measurements, the prediction errors of inadequate scaling factors or even the inability to capture the processes *in vitro*.

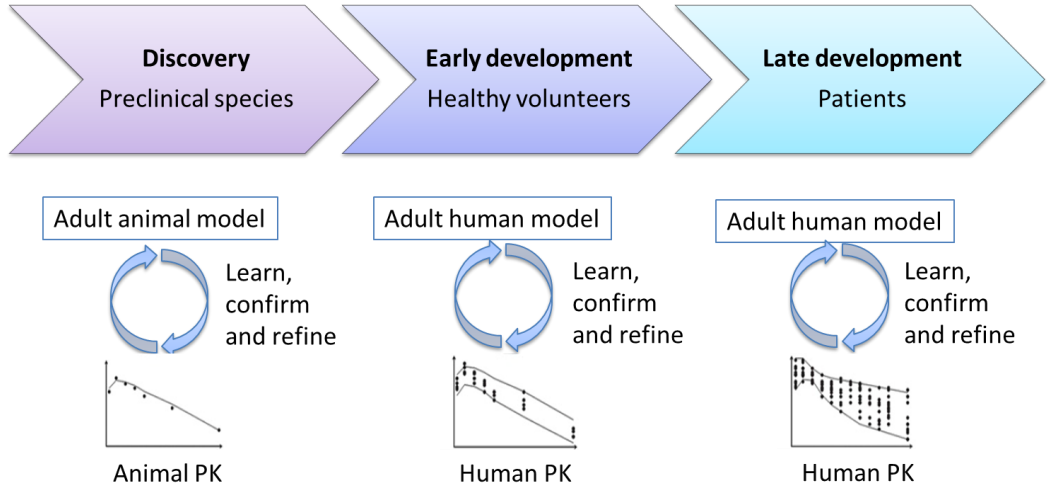
1.5.2. Refinement of PBPK models and parameter optimisation

The large number of biochemical and physiological parameters and the associated variability and uncertainty may be limiting the prediction accuracy of such models [140, 141]. The lack of proper input parameters and misspecifications of the model structure can result in poor predictions from PBPK models. There should be a balance between model complexity and available information; it is possible to have a very detailed model, but the limiting factor is the lack of all relevant parameters to populate it, therefore the model is structurally adequate but lack of understanding of parameters limits its use. On the other hand, with simpler models such as compartmental models the system is reduced into a one- or two- compartment model which lack mechanism and extrapolative power. Recent regulatory guidance regarding PBPK modelling [142, 143] indicated the need for addressing the uncertainty of model predictions and required a more systematic approach for reporting variability and uncertainty in parameter estimates. Indeed, many *in silico* methods, *in vitro* experiments, *in vivo* and clinical studies often report only point estimates of physiological and drug parameters without quantification of variability or uncertainty [144]. Variability refers to diversity of data attributed to environmental or genetic factors. Variability of the data cannot be reduced, but it can be better characterised. On the other hand, uncertainty is a variation that is due to errors in the experiment, measurements, modelling and assumptions of the studied system. Uncertainty can be reduced through optimization of the experiment with more or better data.

Simulation and prediction using a PBPK modelling approach can begin at a very early stage of drug development when the plasma concentration data are not yet available in humans or even in animals. When new data are generated during the development of a new drug, they could be used to revise the uncertainty and improve the predictive

performance of the model. Comparison of simulated data to *in vivo* data is needed to verify the model assumptions and information about parameters, to refine the PBPK model and to assess its ability for adequate prediction. If a mismatch exists between the model predictions and the observed data, additional studies may be performed to enable a more mechanistic understanding of the reasons for such a difference (Figure 1.3) [145]. If such differences cannot be explained from a mechanistic viewpoint or due to uncertainties in the initial parameters, it is a common practice to adjust some of these parameters and fit the model to experimental data in order to improve the PBPK model performance and optimise uncertain model parameters ("middle-out" approach). The initial parameters values of the PBPK model (e.g., those based on simple *in silico* descriptors or *in vitro* experiments) can be updated based on the PK data in the preclinical species; and the new values can be further updated when PK data in humans are available.

Figure 1.3: PBPK modelling strategy in drug discovery and development adapted from [145]



1.5.3. Parameter identifiability, uncertainty, and sensitivity analysis

In the case of parameter estimation, it is important to test if the model is said to be structurally identifiable or globally identifiable in that there is sufficient information contained in the data available and there is a unique correspondence between parameter values and the observed data [146-149]. As a result, all the parameters to be estimated for a given model are uniquely identifiable given perfect data. In contrast, if a limited set of parameter combinations can lead to the same input/output relationship the model is said to be locally identifiable. Another issue is numerical identifiability, which refers to an

increase uncertainty in the parameters of the structurally identifiable model given the observed data.

In early drug development, PK studies are carried out on a small number of animals exposed to a few administered dose levels and observation periods, range typically from several hours to several weeks. As a result, observed data is from a small number of repeated measurements on the same animal or time-course data from animals sacrificed at different sampling points. Additionally, for a PBPK model to be identifiable, data from the tissues should be available, but such data are generally not available or limited to samples at termination of the study. To overcome these issues, generation of additional data by sampling from additional tissue compartments can be proposed, as well as improvement of the collected data by considering an optimal design approach to optimise sample collection [150].

Because PBPK models are parametrised by a large number of physiological and biochemical parameters, simultaneous estimation of these parameters is rarely performed, and the fitting process does not usually follow a systematic and formal approach. In order to stabilise parameter estimation, many parameters are fixed to values that are generally not completely certain and then only a few parameters are therefore selected and estimated through mathematical computations, which could result in biased estimates. In addition, the decision on which parameters to fix and which to estimate is generally subjective and therefore the final model and model parameters may vary significantly between different modellers [144, 151].

Before trying to estimate an unknown parameter, it is recommended to perform a sensitivity analysis in order to determine the influence of the unknown model parameters on the output [152, 153]. The extent to which a change in the unknown model parameter affects quantitatively or qualitatively the chosen output depends on how sensitive it is to the parameter. If improvements in the measurement of the sensitive parameter values are not possible, the quantitative impact of the error in the input needs to be at least incorporated into the PBPK predictions. If the PBPK model is too complex with various compartments and parameters, the system may be over-parametrised and turn out to be structurally and numerically unidentifiable from the available data. A comprehensive review of sensitivity analysis can be found in [152, 154] and examples of application in PBPK modelling can be found in [155-160].

1.5.4. Lumping of PBPK models

To reduce the high-dimensionality of PBPK models for parameter estimation, model order reduction techniques can be used [161]. Proper lumping of tissue compartments is one of the techniques used to derive reduced PBPK models with a fewer number of parameters [162-165] that retains the kinetic behaviour of the more complex models. A minimal PBPK model further reduces the number of compartments to just one or a few compartments with comparable blood flow rates (Jones et al. 2013; Sager et al. 2015). Proper lumping involves lumping together kinetically similar tissue compartments, each state of the original model can be included in only one of the states of the reduced model, and therefore parameters can be directly related between original and lumped models [166]. In the context of PBPK models, a Bayesian lumping method can be of particular interest; it uses prior knowledge on the original model to construct a prior distribution for the reduced model parameters and accounts for parameter uncertainty during the lumping process [165, 167, 168]. Furthermore, physiological parameters that may be highly correlated with each other should be mechanistically considered when estimating parameters in a PBPK model, otherwise this may result in biased, imprecise or biologically implausible parameter estimates [169]. The recommendation would be to use a physiologically plausible value extracted from the literature for one of them, or to reparametrize the model in terms of a composite variable.

1.5.5. Parameter estimation approaches

Once the different previously described limitations and methodological issues related to the fitting of PBPK model are addressed, population modelling can be used to estimate the PBPK model parameters even when only sparse data are available. In addition to establishing correlations between parameters and covariates in order to explain parameter variability, population PK analysis also considers intra- and/or inter-individual variability in observed concentrations when estimating the parameters [170-172].

In order to fit PBPK models to observed data, global optimisation methods have been proposed such as Monte Carlo optimisation and the simplex method [173]. However, these approaches may lead to unrealistic parameter values for well-established

physiological parameters (e.g., flows and volumes) due to parameter identifiability issues. Therefore, using constraints for these bounded parameters is recommended. New techniques such as genetic algorithms have been developed to optimise many parameters simultaneously in complex models based on a natural selection process that mimics biological evolution [174]. Nonetheless, the most frequently used approach is to fix most of the model parameters to values from literature or experiment, and to only estimate a small number of unknown/uncertain parameters. These parameters are usually optimised against PK data by a trial-and-error visual calibration or by more systematic methods which can quantify uncertainties and variations in the parameters through their interval estimates. Many methods have been proposed to fit population PK parameters to observed data.

Nonlinear mixed-effects (NLME) models are more frequently used for the analysis of PK data and repeated measures over time. Maximum likelihood estimation (MLE) approaches constitute a large category of methods commonly used in NLME model analyses. The marginal likelihood density function is proportional to the probability of observing the data when the model and its parameters are assumed to be correct [175]. The principle of maximum likelihood is to estimate the model parameter values which maximise the likelihood of observed data given a specific set of parameters or effectively minimise the negative logarithm of the marginal likelihood density. However, for the models that are nonlinear with respect to the parameters, the integral in the likelihood function does not have a closed form expression and is difficult to compute exactly [176]. Different algorithms have been developed to apply the MLE approach. Classical estimation methods such as the first-order (FO), the first-order conditional estimation with interaction (FOCE) and Laplace, approximate the likelihood [176]. These linearization methods generally perform well for simple and low dimension models but may fail to converge and estimated parameters may be very approximate when model are complex. Other methods have also been proposed such as important sampling EM (IMP), stochastic approximation expectation maximization (SAEM), and Markov chain Monte Carlo Bayesian (BAYES). These expectation maximization (EM)-based estimation methods are more precise in computing the likelihood as they can use a variable step size to control the error of the approximation. The EM iterative algorithm was first developed for problems with missing data [177] and each iteration alternates two steps : the E-step involves computing the conditional expectation of the complete data likelihood given the

observed data, and the M-step only involves likelihood maximization of the complete data. This two-step algorithm was applied to nonlinear mixed-effects models considering the random effects as the missing data [178]. The most commonly used software for NLME modelling are Nonmem (ICON Development Solutions, Hanover, MD, USA) [179] and Monolix (Lixoft SAS, Antony, France) [180].

However, caution should be taken when using these approaches for extrapolating from the estimated parameters and for refining PBPK models because these parameter estimates are conditional on the values that have been assumed for the fixed parameters [173, 181]. Additionally, the fitted estimates present a certain level of uncertainty coming from physiological and drug-specific data or the model assumptions. Consequently, any extrapolations, conclusions or predictions performed based on these parameters may not be very reliable especially if only single values are reported for the estimated parameters. Thus, using these methods is challenging for the refinement of PBPK models because of their multidimensional nature limiting the estimation of the model parameters simultaneously. Therefore, the fit of PBPK models to data and estimation of relevant model parameters will be investigated in this thesis.

1.6. Thesis aims and objectives

The overall aim of this thesis was to develop a systematic framework for integrating observed preclinical PK data in PBPK model development; for refining model predictions in animal and to improve the prediction of human drug disposition during drug discovery and development programme. In order to refine predictions, the intention is to estimate model parameters when observed PK data are available. The current project proposes a formal framework for parameter estimation within PBPK modelling, with the specific focus on drug distribution. It critically examines the sensitivity of model predictions to key model parameters and highlights the importance of qualitative and quantitative data for drug-dependant and physiological parameters. It also provides different PBPK models that could be used for model fitting, parameter estimation and translation. In addition, several case examples illustrate the application of the systematic framework to estimate parameters in PBPK models and to perform translation, especially to drive predictions in human and support the design of first in human studies. This thesis has three result chapters, each with its own aim and objectives.

In the first result chapter, Chapter 2, the aim was to identify key drug-dependent and physiological parameters that influence predictions of Kpu and thus drug distribution in PBPK models. This study investigated the variability and uncertainty of drug and physiological values, and their impact on model predictions. The sensitivity analysis work also aimed at potentially reducing PBPK model dimensionality by excluding non-influential parameters. This published analysis is presented in Chapter 2 entitled "Global sensitivity analysis of the Rodgers and Rowland model for prediction of tissue: plasma partitioning coefficients: Assessment of the key physiological and physicochemical factors that determine small-molecule tissue distribution".

Subsequently, different model reduction approaches were investigated in order to simplify WBPBPK model structure or dimensionality. The aim was to obtain mechanistic models that have a reduced number of model parameters to estimate and thus facilitate the estimation process during PBPK model development. This study evaluated the ability of these investigated models to fit PK data and estimate meaningful Kpu parameters using the nonlinear mixed effect method. The performance of the investigated models for estimating Kpu parameters and fitting PK data was assessed using diazepam as an example. This analysis is presented in Chapter 3 entitled "Investigation of simplified physiologically based pharmacokinetic (PBPK) models in rat and human".

Finally, the investigated models were used to fit PK data in preclinical species (rat, monkey) for three compounds (diazepam, midazolam, basmisanil). The best models were selected according to criteria which then allow the use of these models for translation to human. The performance of these models for prediction of human drug disposition was verified against observed clinical data. The work presented here therefore provides a framework for prediction of human Kpus and volume of distribution and concentration time profiles from preclinical data. This study is presented in Chapter 4 entitled "Prediction of human drug disposition from preclinical data using a 'middle-out approach' to physiologically based pharmacokinetic (PBPK) modelling".

1.7. List of manuscripts and author contribution statement

In accordance with the University of Manchester guidance on the alternative thesis format for the Doctor of Philosophy degree, the following section describes the contribution of the candidate and co-authors to the manuscripts presented in this thesis that are either published or in preparation.

Chapter 2: Global sensitivity analysis of the Rodgers and Rowland model for prediction of tissue: plasma partitioning coefficients: Assessment of the key physiological and physicochemical factors that determine small-molecule tissue distribution

AAPS J. 2020, 22(2):41

E. Yau: Conducted the research, literature search, design, perform the analysis of the data, interpretation of the results and writing of the manuscript and figures.

A. Olivares-Morales: Supervision and design of the research, input in data analysis, and contribute to the writing and discussion of the manuscript.

M. Gertz: Supervision and design of the research, input in data analysis, and contribute to the writing and discussion of the manuscript.

N. Parrott: Provide advice on the approach and contribute to the discussion and writing of the manuscript.

A.S. Darwich: Supervision and design of the research, input in data analysis, and contribute to the writing and discussion of the manuscript.

L. Aarons: Supervision and design of the research, input in data analysis, and contribute to the writing and discussion of the manuscript.

K. Ogungbenro: Supervision and design of the research, input in data analysis, and contribute to the writing and discussion of the manuscript.

Chapter 3: Investigation of simplified physiologically based pharmacokinetic (PBPK) models in rat and human

[In preparation]

E. Yau: Conducted the research, literature search, design, perform analysis of the data and modelling work, interpretation of the results and writing of the manuscript and figures.

A. Olivares-Morales: Supervision and design of the research, input in data analysis, and contribute to the writing and discussion of the manuscript.

K. Ogungbenro: Supervision and design of the research, input in data analysis, and contribute to the writing and discussion of the manuscript.

L. Aarons: Supervision and design of the research, input in data analysis, and contribute to the writing and discussion of the manuscript.

M. Gertz: Supervision and design of the research, input in data analysis, and contribute to the writing and discussion of the manuscript.

Chapter 4: Prediction of human drug disposition from preclinical data using a 'middle-out approach' to physiologically based pharmacokinetic (PBPK) modelling

[In preparation]

E. Yau: Conducted the research, literature search, collated *in vitro* and clinical data from literature and from Roche-internal data, design, perform analysis of the data and modelling work, interpretation of the results and writing of the manuscript and figures.

M. Gertz: Supervision and design of the research, input in data analysis, and contribute to the writing and discussion of the manuscript.

K. Ogungbenro: Supervision and design of the research, input in data analysis, and contribute to the writing and discussion of the manuscript.

L. Aarons: Supervision and design of the research, input in data analysis, and contribute to the writing and discussion of the manuscript.

A. Olivares-Morales: Supervision and design of the research, input in data analysis, and contribute to the writing and discussion of the manuscript.

1.8. References

[1] Rowland M, Tozer TN. Clinical Pharmacokinetics: Concepts and Applications. 3rd ed: Lippincott Williams & Wilkins; 1995.

[2] Wouters OJ, McKee M, Luyten J. Estimated Research and Development Investment Needed to Bring a New Medicine to Market, 2009-2018. JAMA. 2020;323(9):844-53.

[3] DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: New estimates of R&D costs. J Health Econ. 2016;47:20-33.

[4] Peck CC, Barr WH, Benet LZ, Collins J, Desjardins RE, Furst DE, et al. Opportunities for integration of pharmacokinetics, pharmacodynamics, and toxicokinetics in rational drug development. J Clin Pharmacol. 1994;34(2):111-9.

[5] Lesko LJ, Rowland M, Peck CC, Blaschke TF, Breimer D, de Jong HJ, et al. Optimizing the science of drug development: opportunities for better candidate selection and accelerated evaluation in humans. Eur J Pharm Sci. 2000;10(4):iv-xiv.

[6] EFPIA MID3 Workgroup, Marshall SF, Burghaus R, Cosson V, Cheung SY, Chenel M, et al. Good Practices in Model-Informed Drug Discovery and Development: Practice, Application, and Documentation. CPT Pharmacometrics Syst Pharmacol. 2016;5(3):93-122.

[7] Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? Nat Rev Drug Discov. 2004;3(8):711-5.

[8] Fowler S, Morcos PN, Cleary Y, Martin-Facklam M, Parrott N, Gertz M, et al. Progress in Prediction and Interpretation of Clinically Relevant Metabolic Drug-Drug Interactions: a Minireview Illustrating Recent Developments and Current Opportunities. Curr Pharmacol Rep. 2017;3(1):36-49.

[9] Gillette JR. Factors affecting drug metabolism. Ann N Y Acad Sci. 1971;179:43-66.

[10] Mangoni AA, Jackson SH. Age-related changes in pharmacokinetics and pharmacodynamics: basic principles and practical applications. Br J Clin Pharmacol. 2004;57(1):6-14.

[11] Ueno K, Sato H. Sex-related differences in pharmacokinetics and pharmacodynamics of anti-hypertensive drugs. Hypertens Res. 2012;35(3):245-50.

[12] Peters SA. Physiological Model for Distribution. Physiologically-Based Pharmacokinetic (PBPK) Modeling and Simulations: John Wiley & Sons, Inc.; 2012. p. 89-117.

[13] Seddon AM, Casey D, Law RV, Gee A, Templer RH, Ces O. Drug interactions with lipid membranes. Chem Soc Rev. 2009;38(9):2509-19.

[14] Rodgers T, Leahy D, Rowland M. Physiologically based pharmacokinetic modeling 1: predicting the tissue distribution of moderate-to-strong bases. J Pharm Sci. 2005;94(6):1259-76.

[15] Rodgers T, Rowland M. Physiologically based pharmacokinetic modelling 2: predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. J Pharm Sci. 2006;95(6):1238-57.

[16] Jetter A, Kullak-Ublick GA. Drugs and hepatic transporters: A review. Pharmacol Res. 2020;154:104234.

[17] Ivanyuk A, Livio F, Biollaz J, Buclin T. Renal Drug Transporters and Drug Interactions. Clin Pharmacokinet. 2017;56(8):825-92.

[18] Muller J, Keiser M, Drozdzik M, Oswald S. Expression, regulation and function of intestinal drug transporters: an update. Biol Chem. 2017;398(2):175-92.

[19] Liu L, Liu X. Contributions of Drug Transporters to Blood-Brain Barriers. Adv Exp Med Biol. 2019;1141:407-66.

[20] International Transporter C, Giacomini KM, Huang SM, Tweedie DJ, Benet LZ, Brouwer KL, et al. Membrane transporters in drug development. Nat Rev Drug Discov. 2010;9(3):215-36.

[21] Roth M, Obaidat A, Hagenbuch B. OATPs, OATs and OCTs: the organic anion and cation transporters of the SLCO and SLC22A gene superfamilies. Br J Pharmacol. 2012;165(5):1260-87.

[22] Zamek-Gliszczynski MJ, Lee CA, Poirier A, Bentz J, Chu X, Ellens H, et al. ITC recommendations for transporter kinetic parameter estimation and translational modeling of transport-mediated PK and DDIs in humans. Clin Pharmacol Ther. 2013;94(1):64-79.

[23] van Waterschoot RAB, Parrott NJ, Olivares-Morales A, Lave T, Rowland M, Smith DA. Impact of target interactions on small-molecule drug disposition: an overlooked area. Nat Rev Drug Discov. 2018;17(4):299.

[24] Rowland M, Riegelman S. Determination of acetylsalicylic acid and salicylic acid in plasma. J Pharm Sci. 1967;56(6):717-20.

[25] Endo S, Goss K-U. Serum Albumin Binding of Structurally Diverse Neutral Organic Compounds: Data and Models. Chemical Research in Toxicology. 2011;24(12):2293-301.

[26] Smith SA, Waters NJ. Pharmacokinetic and Pharmacodynamic Considerations for Drugs Binding to Alpha-1-Acid Glycoprotein. Pharm Res. 2018;36(2):30.

[27] Bickel MH. Binding of chlorpromazine and imipramine to red cells, albumin, lipoproteins and other blood components. J Pharm Pharmacol. 1975;27(10):733-8.

[28] Bayne WF, Chu LC, Theeuwes F. Acetazolamide binding to two carbonic anhydrase isoenzymes in human erythrocytes. J Pharm Sci. 1979;68(7):912-3.

[29] Ehrnebo M, Agurell S, Boreus LO, Gordon E, Lonroth U. Pentazocine binding to blood cells and plasma proteins. Clin Pharmacol Ther. 1974;16(3):424-9.

[30] Yokogawa K, Ishizaki J, Ohkuma S, Miyamoto K. Influence of lipophilicity and lysosomal accumulation on tissue distribution kinetics of basic drugs: a physiologically based pharmacokinetic model. Methods Find Exp Clin Pharmacol. 2002;24(2):81-93.

[31] Assmus F, Houston JB, Galetin A. Incorporation of lysosomal sequestration in the mechanistic model for prediction of tissue distribution of basic drugs. Eur J Pharm Sci. 2017;109:419-30.

[32] Terasaki T, Iga T, Sugiyama Y, Sawada Y, Hanano M. Nuclear binding as a determinant of tissue distribution of adriamycin, daunomycin, adriamycinol, daunorubicinol and actinomycin D. Journal of Pharmacobio-Dynamics. 1984;7(5):269-77.

[33] Gustafson DL, Rastatter JC, Colombo T, Long ME. Doxorubicin pharmacokinetics: Macromolecule binding, metabolism, and excretion in the context of a physiologic model. J Pharm Sci. 2002;91(6):1488-501.

[34] Skinner HC, Nalbandian J. Tetracyclines and mineralized tissues: review and perspectives. Yale J Biol Med. 1975;48(5):377-97.

[35] Atkinson AJ. Chapter 3 - Compartmental Analysis of Drug Distribution. In: Atkinson AJ, Huang S-M, Lertora JJL, Markey SP, editors. Principles of Clinical Pharmacology (Third Edition): Academic Press; 2012. p. 27-40.

[36] Chan R, De Bruyn T, Wright M, Broccatelli F. Comparing Mechanistic and Preclinical Predictions of Volume of Distribution on a Large Set of Drugs. Pharmaceutical Research. 2018;35(4):87.

[37] Lombardo F, Waters NJ, Argikar UA, Dennehy MK, Zhan J, Gunduz M, et al. Comprehensive assessment of human pharmacokinetic prediction based on in vivo animal pharmacokinetic data, part 1: volume of distribution at steady state. J Clin Pharmacol. 2013;53(2):167-77.

[38] Jones RD, Jones HM, Rowland M, Gibson CR, Yates JW, Chien JY, et al. PhRMA CPCDC initiative on predictive models of human pharmacokinetics, part 2: comparative assessment of prediction methods of human volume of distribution. J Pharm Sci. 2011;100(10):4074-89.

[39] Graham H, Walker M, Jones O, Yates J, Galetin A, Aarons L. Comparison of in-vivo and in-silico methods used for prediction of tissue: plasma partition coefficients in rat. J Pharm Pharmacol. 2012;64(3):383-96.

[40] Petersson C, Papasouliotis O, Lecomte M, Badolo L, Dolgos H. Prediction of volume of distribution in humans: analysis of eight methods and their application in drug discovery. Xenobiotica. 2019:1-10.

[41] Murad N, Pasikanti KK, Madej BD, Minnich A, McComas JM, Crouch S, et al. Predicting Volume of Distribution in Humans: Performance of In Silico Methods for a Large Set of Structurally Diverse Clinical Compounds. Drug Metab Dispos. 2021;49(2):169-78.

[42] Berry LM, Li C, Zhao Z. Species differences in distribution and prediction of human V(ss) from preclinical data. Drug Metab Dispos. 2011;39(11):2103-16.

[43] Zou P, Zheng N, Yang Y, Yu LX, Sun D. Prediction of volume of distribution at steady state in humans: comparison of different approaches. Expert Opin Drug Metab Toxicol. 2012;8(7):855-72.

[44] Nigade PB, Gundu J, Pai KS, Nemmani KVS, Talwar R. Prediction of volume of distribution in preclinical species and humans: application of simplified physiologically based algorithms. Xenobiotica. 2019;49(5):528-39.

[45] Obach RS, Baxter JG, Liston TE, Silber BM, Jones BC, MacIntyre F, et al. The prediction of human pharmacokinetic parameters from preclinical and in vitro metabolism data. J Pharmacol Exp Ther. 1997;283(1):46-58.

[46] Mathew S, Tess D, Burchett W, Chang G, Woody N, Keefer C, et al. Evaluation of Prediction Accuracy for Volume of Distribution in Rat and Human Using In Vitro, In Vivo, PBPK and QSAR Methods. J Pharm Sci. 2021;110(4):1799-823.

[47] Boxenbaum H. Interspecies scaling, allometry, physiological time, and the ground plan of pharmacokinetics. J Pharmacokinet Biopharm. 1982;10(2):201-27.

[48] Ward KW, Smith BR. A comprehensive quantitative and qualitative evaluation of extrapolation of intravenous pharmacokinetic parameters from rat, dog, and monkey to humans. II. Volume of distribution and mean residence time. Drug Metab Dispos. 2004;32(6):612-9.

[49] Sawada Y, Hanano M, Sugiyama Y, Harashima H, Iga T. Prediction of the volumes of distribution of basic drugs in humans based on data from animals. J Pharmacokinet Biopharm. 1984;12(6):587-96.

[50] Wajima T, Fukumura K, Yano Y, Oguma T. Prediction of human pharmacokinetics from animal data and molecular structural parameters using multivariate regression analysis: volume of distribution at steady state. J Pharm Pharmacol. 2003;55(7):939-49.

[51] Oie S, Tozer TN. Effect of altered plasma protein binding on apparent volume of distribution. J Pharm Sci. 1979;68(9):1203-5.

[52] Korzekwa K, Nagar S. Drug Distribution Part 2. Predicting Volume of Distribution from Plasma Protein Binding and Membrane Partitioning. Pharm Res. 2017;34(3):544-51.

[53] Lombardo F, Obach RS, Shalaeva MY, Gao F. Prediction of human volume of distribution values for neutral and basic drugs. 2. Extended data set and leave-class-out statistics. J Med Chem. 2004;47(5):1242-50.

[54] Lombardo F, Obach RS, Shalaeva MY, Gao F. Prediction of volume of distribution values in humans for neutral and basic drugs using physicochemical measurements and plasma protein binding data. J Med Chem. 2002;45(13):2867-76.

[55] Freitas AA, Limbu K, Ghafourian T. Predicting volume of distribution with decision tree-based regression methods using predicted tissue:plasma partition coefficients. J Cheminform. 2015;7:6.

[56] Lombardo F, Obach RS, Dicapua FM, Bakken GA, Lu J, Potter DM, et al. A hybrid mixture discriminant analysis-random forest computational model for the prediction of volume of distribution of drugs in human. J Med Chem. 2006;49(7):2262-7.

[57] Ghafourian T, Barzegar-Jalali M, Hakimiha N, Cronin MT. Quantitative structure-pharmacokinetic relationship modelling: apparent volume of distribution. J Pharm Pharmacol. 2004;56(3):339-50.

[58] Lombardo F, Berellini G, Obach RS. Trend Analysis of a Database of Intravenous Pharmacokinetic Parameters in Humans for 1352 Drug Compounds. Drug Metab Dispos. 2018;46(11):1466-77.

[59] Lombardo F, Bentzien J, Berellini G, Muegge I. In Silico Models of Human PK Parameters. Prediction of Volume of Distribution Using an Extensive Data Set and a Reduced Number of Parameters. J Pharm Sci. 2021;110(1):500-9.

[60] Poulin P, Theil FP. Prediction of pharmacokinetics prior to in vivo studies. 1. Mechanism-based prediction of volume of distribution. J Pharm Sci. 2002;91(1):129-56.

[61] Poulin P, Theil FP. Development of a novel method for predicting human volume of distribution at steady-state of basic drugs and comparative assessment with existing methods. J Pharm Sci. 2009;98(12):4941-61.

[62] Utsey K, Gastonguay MS, Russell S, Freling R, Riggs MM, Elmokadem A. Quantification of the Impact of Partition Coefficient Prediction Methods on Physiologically Based Pharmacokinetic Model Output Using a Standardized Tissue Composition. Drug Metab Dispos. 2020;48(10):903-16.

[63] Poulin P, Jones RD, Jones HM, Gibson CR, Rowland M, Chien JY, et al. PHRMA CPCDC initiative on predictive models of human pharmacokinetics, part 5: prediction of plasma concentration-time profiles in human by using the physiologically-based pharmacokinetic modeling approach. J Pharm Sci. 2011;100(10):4127-57.

[64] Dedrick R, Bischoff KB, Zaharko DS. Interspecies correlation of plasma concentration history of methotrexate (NSC-740). Cancer Chemother Rep. 1970;54(2):95-101.

[65] Wajima T, Yano Y, Fukumura K, Oguma T. Prediction of human pharmacokinetic profile in animal scale up based on normalizing time course profiles. J Pharm Sci. 2004;93(7):1890-900.

[66] Mahmood I, Yuan R. A comparative study of allometric scaling with plasma concentrations predicted by species-invariant time methods. Biopharm Drug Dispos. 1999;20(3):137-44.

[67] Vuppugalla R, Marathe P, He H, Jones RD, Yates JW, Jones HM, et al. PhRMA CPCDC initiative on predictive models of human pharmacokinetics, part 4: prediction of plasma concentration-time profiles in human from in vivo preclinical data by using the Wajima approach. J Pharm Sci. 2011;100(10):4111-26.

[68] Boxenbaum H. Interspecies pharmacokinetic scaling and the evolutionary-comparative paradigm. Drug Metab Rev. 1984;15(5-6):1071-121.

[69] Mahmood I, Balian JD. Interspecies scaling: predicting pharmacokinetic parameters of antiepileptic drugs in humans from animals with special emphasis on clearance. J Pharm Sci. 1996;85(4):411-4.

[70] Mahmood I. Interspecies scaling of biliary excreted drugs: a comparison of several methods. J Pharm Sci. 2005;94(4):883-92.

[71] Mahmood I. Interspecies scaling of renally secreted drugs. Life Sci. 1998;63(26):2365-71.

[72] Mahmood I, Balian JD. Interspecies scaling: predicting clearance of drugs in humans. Three different approaches. Xenobiotica. 1996;26(9):887-95.

[73] Lave T, Portmann R, Schenker G, Gianni A, Guenzi A, Girometta MA, et al. Interspecies pharmacokinetic comparisons and allometric scaling of napsagatran, a low molecular weight thrombin inhibitor. J Pharm Pharmacol. 1999;51(1):85-91.

[74] Lave T, Coassolo P, Reigner B. Prediction of hepatic metabolic clearance based on interspecies allometric scaling techniques and in vitro-in vivo correlations. Clin Pharmacokinet. 1999;36(3):211-31.

[75] Mahmood I, Sahajwalla C. Interspecies scaling of biliary excreted drugs. J Pharm Sci. 2002;91(8):1908-14.

[76] Jones HM, Parrott N, Jorga K, Lave T. A novel strategy for physiologically based predictions of human pharmacokinetics. Clin Pharmacokinet. 2006;45(5):511-42.

[77] De Buck SS, Sinha VK, Fenu LA, Nijsen MJ, Mackie CE, Gilissen RA. Prediction of human pharmacokinetics using physiologically based modeling: a retrospective analysis of 26 clinically tested drugs. Drug Metab Dispos. 2007;35(10):1766-80.

[78] Jones HM, Gardner IB, Collard WT, Stanley PJ, Oxley P, Hosea NA, et al. Simulation of human intravenous and oral pharmacokinetics of 21 diverse compounds using physiologically based pharmacokinetic modelling. Clin Pharmacokinet. 2011;50(5):331-47.

[79] Zhang T, Heimbach T, Lin W, Zhang J, He H. Prospective Predictions of Human Pharmacokinetics for Eighteen Compounds. J Pharm Sci. 2015;104(9):2795-806.

[80] Benowitz N, Forsyth FP, Melmon KL, Rowland M. Lidocaine disposition kinetics in monkey and man. I. Prediction by a perfusion model. Clin Pharmacol Ther. 1974;16(1):87-98.

[81] Harrison LI, Gibaldi M. Physiologically based pharmacokinetic model for digoxin distribution and elimination in the rat. J Pharm Sci. 1977;66(8):1138-42.

[82] Bernareggi A, Rowland M. Physiologic modeling of cyclosporin kinetics in rat and man. J Pharmacokinet Biopharm. 1991;19(1):21-50.

[83] Harris PA, Gross JF. Preliminary pharmacokinetic model for adriamycin (NSC-123127). Cancer Chemother Rep. 1975;59(4):819-25.

[84] Igari Y, Sugiyama Y, Awazu S, Hanano M. Comparative physiologically based pharmacokinetics of hexobarbital, phenobarbital and thiopental in the rat. J Pharmacokinet Biopharm. 1982;10(1):53-75.

[85] Granero L, Chesa-Jiménez J, Monserrat V, Almela M, Gimeno M-J, Torres-Molina F, et al. Physiological pharmacokinetic model for ceftazidime disposition in the rat and its application to prediction of plasma concentrations in humans. European Journal of Pharmaceutical Sciences. 1993;1(1):3-11.

[86] Blakey GE, Nestorov IA, Arundel PA, Aarons LJ, Rowland M. Quantitative structure-pharmacokinetics relationships: I. Development of a whole-body physiologically based model to characterize changes in pharmacokinetics across a homologous series of barbiturates in the rat. J Pharmacokinet Biopharm. 1997;25(3):277-312.

[87] Musther H, Gill KL, Chetty M, Rostami-Hodjegan A, Rowland M, Jamei M. Are Physiologically Based Pharmacokinetic Models Reporting the Right C(max)? Central Venous Versus Peripheral Sampling Site. AAPS J. 2015;17(5):1268-79.

[88] Levitt DG. Physiologically based pharmacokinetic modeling of arterial - antecubital vein concentration difference. BMC Clin Pharmacol. 2004;4:2.

[89] Huang W, Isoherranen N. Sampling Site Has a Critical Impact on Physiologically Based Pharmacokinetic Modeling. J Pharmacol Exp Ther. 2020;372(1):30-45.

[90] Prasain JK, Peng N, Moore R, Arabshahi A, Barnes S, Wyss JM. Tissue distribution of puerarin and its conjugated metabolites in rats assessed by liquid chromatography-tandem mass spectrometry. Phytomedicine. 2009;16(1):65-71.

[91] Chay SH, Pohland RC. Comparison of quantitative whole-body autoradiographic and tissue dissection techniques in the evaluation of the tissue distribution of [14C]daptomycin in rats. J Pharm Sci. 1994;83(9):1294-9.

[92] Li Q, Xie L, Zhang J, Weina PJ. The distribution pattern of intravenous [(14)C] artesunate in rat tissues by quantitative whole-body autoradiography and tissue dissection techniques. J Pharm Biomed Anal. 2008;48(3):876-84.

[93] Thomas N, Tirand L, Chatelut E, Plenat F, Frochot C, Dodeller M, et al. Tissue distribution and pharmacokinetics of an ATWLPPR-conjugated chlorin-type photosensitizer targeting neuropilin-1 in glioma-bearing nude mice. Photochem Photobiol Sci. 2008;7(4):433-41.

[94] Rodgers T, Leahy D, Rowland M. Tissue distribution of basic drugs: accounting for enantiomeric, compound and regional differences amongst beta-blocking drugs in rat. J Pharm Sci. 2005;94(6):1237-48.

[95] Jacob S, Ahmed AE. Effect of route of administration on the disposition of acrylonitrile: quantitative whole-body autoradiographic study in rats. Pharmacol Res. 2003;48(5):479-88.

[96] Cornford EM, Bocash WD, Braun LD, Crane PD, Oldendorf WH, MacInnis AJ. Rapid distribution of tryptophol (3-indole ethanol) to the brain and other tissues. J Clin Invest. 1979;63(6):1241-8.

[97] Bergstrom M, Grahnen A, Langstrom B. Positron emission tomography microdosing: a new concept with application in tracer and early clinical drug development. Eur J Clin Pharmacol. 2003;59(5-6):357-66.

[98] Kaneko T, Wang PY, Sato A. Partition coefficients of some acetate esters and alcohols in water, blood, olive oil, and rat tissues. Occup Environ Med. 1994;51(1):68-72.

[99] Fiserova-Bergerova V, Diaz ML. Determination and prediction of tissue-gas partition coefficients. Int Arch Occup Environ Health. 1986;58(1):75-87.

[100] Bickel MH, Raaflaub RM, Hellmuller M, Stauffer EJ. Characterization of drug distribution and binding competition by two-chamber and multi-chamber distribution dialysis. J Pharm Sci. 1987;76(1):68-74.

[101] Ludden TM, Schanker LS, Lanman RC. Binding of organic compounds to rat liver and lung. Drug Metab Dispos. 1976;4(1):8-16.

[102] Bickel MH, Steele JW. Binding of basic and acidic drugs to rat tissue subcellular fractions. Chem Biol Interact. 1974;8(3):151-62.

[103] Murphy JE, Janszen DB, Gargas ML. An in vitro method for determination of tissue partition coefficients of non-volatile chemicals such as 2,3,7,8-tetrachlorodibenzo-p-dioxin and estradiol. J Appl Toxicol. 1995;15(2):147-52.

[104] Friden M, Gupta A, Antonsson M, Bredberg U, Hammarlund-Udenaes M. In vitro methods for estimating unbound drug concentrations in the brain interstitial and intracellular fluids. Drug Metab Dispos. 2007;35(9):1711-9.

[105] Abbott NJ, Dolman DE, Patabendige AK. Assays to predict drug permeation across the blood-brain barrier, and distribution to brain. Curr Drug Metab. 2008;9(9):901-10.

[106] Kulkarni P, Korzekwa K, Nagar S. Intracellular Unbound Atorvastatin Concentrations in the Presence of Metabolism and Transport. J Pharmacol Exp Ther. 2016;359(1):26-36.

[107] Poulin P. Drug Distribution to Human Tissues: Prediction and Examination of the Basic Assumption in In Vivo Pharmacokinetics-Pharmacodynamics (PK/PD) Research. J Pharm Sci. 2015;104(6):2110-8.

[108] Arundel PA. A multi-compartmental model generally applicable to physiologically-based pharmacokinetics. 3rd IFAC Modelling and Control in Biomedical Systems; Warwick, UK1997. p. 5.

[109] Bjorkman S. Prediction of the volume of distribution of a drug: which tissue-plasma partition coefficients are needed? J Pharm Pharmacol. 2002;54(9):1237-45.

[110] Poulin P, Schoenlein K, Theil FP. Prediction of adipose tissue: Plasma partition coefficients for structurally unrelated drugs. Journal of Pharmaceutical Sciences. 2001;90(4):436-47.

[111] Berezhkovskiy LM. Volume of distribution at steady state for a linear pharmacokinetic system with peripheral elimination. Journal of Pharmaceutical Sciences. 2004;93(6):1628-40.

[112] Richter WF, Starke V, Whitby B. The distribution pattern of radioactivity across different tissues in quantitative whole-body autoradiography (QWBA) studies. Eur J Pharm Sci. 2006;28(1-2):155-65.

[113] Lukacova V, Parrott N, Lavé T, Fraczkiewicz G, Bolger MB. General Approach to Calculation of Tissue:Plasma Partition Coefficients for Physiologically Based Pharmacokinetic (PBPK) Modeling. AAPS; November 17-19, 2008; Atlanta, GA2008.

[114] Schmitt W. General approach for the calculation of tissue to plasma partition coefficients. Toxicol In Vitro. 2008;22(2):457-67.

[115] Jansson R, Bredberg U, Ashton M. Prediction of drug tissue to plasma concentration ratios using a measured volume of distribution in combination with lipophilicity. J Pharm Sci. 2008;97(6):2324-39.

[116] Peyret T, Poulin P, Krishnan K. A unified algorithm for predicting partition coefficients for PBPK modeling of drugs and environmental chemicals. Toxicol Appl Pharmacol. 2010;249(3):197-207.

[117] Poulin P, Haddad S. Advancing prediction of tissue distribution and volume of distribution of highly lipophilic compounds from a simplified tissue-composition-based model as a mechanistic animal alternative method. J Pharm Sci. 2012;101(6):2250-61.

[118] Yun YE, Edginton AN. Correlation-based prediction of tissue-to-plasma partition coefficients using readily available input parameters. Xenobiotica. 2013;43(10):839-52.

[119] Mayumi K, Tachibana M, Yoshida M, Ohnishi S, Kanazu T, Hasegawa H. The Novel In Vitro Method to Calculate Tissue-to-Plasma Partition Coefficient in Humans for Predicting Pharmacokinetic Profiles by Physiologically-Based Pharmacokinetic Model With High Predictability. J Pharm Sci. 2020;109(7):2345-55.

[120] Holt K, Ye M, Nagar S, Korzekwa KR. Prediction of tissue - plasma partition coefficients using microsomal partitioning: Incorporation into physiologically-based pharmacokinetic models and steady state volume of distribution predictions. Drug Metab Dispos. 2019.

[121] Nigade PB, Gundu J, Sreedhara Pai K, Nemmani KVS. Prediction of Tissue-to-Plasma Ratios of Basic Compounds in Mice. Eur J Drug Metab Pharmacokinet. 2017;42(5):835-47.

[122] De Buck SS, Sinha VK, Fenu LA, Gilissen RA, Mackie CE, Nijsen MJ. The prediction of drug metabolism, tissue distribution, and bioavailability of 50 structurally diverse compounds in rat using mechanism-based absorption, distribution, and metabolism prediction tools. Drug Metab Dispos. 2007;35(4):649-59.

[123] Rodgers T, Rowland M. Mechanistic approaches to volume of distribution predictions: understanding the processes. Pharm Res. 2007;24(5):918-33.

[124] Parrott N, Lave T. Applications of physiologically based absorption models in drug discovery and development. Mol Pharm. 2008;5(5):760-75.

[125] Bouzom F, Walther B. Pharmacokinetic predictions in children by using the physiologically based pharmacokinetic modelling. Fundam Clin Pharmacol. 2008;22(6):579-87.

[126] Bois FY. Physiologically based modelling and prediction of drug interactions. Basic Clin Pharmacol Toxicol. 2010;106(3):154-61.

[127] Rowland M, Peck C, Tucker G. Physiologically-based pharmacokinetics in drug development and regulatory science. Annu Rev Pharmacol Toxicol. 2011;51:45-73.

[128] Zhao P, Zhang L, Grillo JA, Liu Q, Bullock JM, Moon YJ, et al. Applications of physiologically based pharmacokinetic (PBPK) modeling and simulation during regulatory review. Clin Pharmacol Ther. 2011;89(2):259-67.

[129] Jones HM, Chen Y, Gibson C, Heimbach T, Parrott N, Peters SA, et al. Physiologically based pharmacokinetic modeling in drug discovery and development: a pharmaceutical industry perspective. Clin Pharmacol Ther. 2015;97(3):247-62.

[130] Min JS, Bae SK. Prediction of drug-drug interaction potential using physiologically based pharmacokinetic modeling. Arch Pharm Res. 2017.

[131] Chen Y, Jin JY, Mukadam S, Malhi V, Kenny JR. Application of IVIVE and PBPK modeling in prospective prediction of clinical pharmacokinetics: strategy and approach during the drug discovery phase with four case studies. Biopharm Drug Dispos. 2012;33(2):85-98.

[132] Huang SM, Abernethy DR, Wang Y, Zhao P, Zineh I. The utility of modeling and simulation in drug development and regulatory review. J Pharm Sci. 2013;102(9):2912-23.

[133] Jones HM, Dickins M, Youdim K, Gosset JR, Attkins NJ, Hay TL, et al. Application of PBPK modelling in drug discovery and development at Pfizer. Xenobiotica. 2012;42(1):94-106.

[134] Parrott N, Jones H, Paquereau N, Lave T. Application of full physiological models for pharmaceutical drug candidate selection and extrapolation of pharmacokinetics to man. Basic Clin Pharmacol Toxicol. 2005;96(3):193-9.

[135] Kwon Y, Morris ME. Membrane transport in hepatic clearance of drugs. I: Extended hepatic clearance models incorporating concentration-dependent transport and elimination processes. Pharm Res. 1997;14(6):774-9.

[136] Davies B, Morris T. Physiological parameters in laboratory animals and humans. Pharm Res. 1993;10(7):1093-5.

[137] Brown RP, Delp MD, Lindstedt SL, Rhomberg LR, Beliles RP. Physiological parameter values for physiologically based pharmacokinetic models. Toxicol Ind Health. 1997;13(4):407-84.

[138] Valentin J. Basic anatomical and physiological data for use in radiological protection: reference values. ICRP Publication 89. Ann ICRP. 2002;32(3-4):1-277.

[139] Rostami-Hodjegan A. Physiologically based pharmacokinetics joined with in vitro-in vivo extrapolation of ADME: a marriage under the arch of systems pharmacology. Clin Pharmacol Ther. 2012;92(1):50-61.

[140] Parrott N, Paquereau N, Coassolo P, Lave T. An evaluation of the utility of physiologically based models of pharmacokinetics in early drug discovery. J Pharm Sci. 2005;94(10):2327-43.

[141] Paixao P, Gouveia LF, Morais JA. Prediction of the human oral bioavailability by using in vitro and in silico drug related parameters in a physiologically based absorption model. Int J Pharm. 2012;429(1-2):84-98.

[142] European Medicines Agency, Guideline on the qualification and reporting of physiologically based pharmacokinetic (PBPK) modelling and simulation 2018 [Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-reporting-physiologically-based-pharmacokinetic-pbpk-modelling-simulation_en.pdf; accessed May 2019]

[143] U.S. Food and Drug Administration, Physiologically Based Pharmacokinetic Analyses — Format and Content : Guidance for Industry 2018 [Available from: https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM531207.pdf; accessed May 2019]

[144] Tsamandouras N, Rostami-Hodjegan A, Aarons L. Combining the 'bottom up' and 'top down' approaches in pharmacokinetic modelling: fitting PBPK models to observed clinical data. Br J Clin Pharmacol. 2013;79(1):48-55.

[145] Jones H, Rowland-Yeo K. Basic concepts in physiologically based pharmacokinetic modeling in drug discovery and development. CPT Pharmacometrics Syst Pharmacol. 2013;2:e63.

[146] Bellman R, Åström KJ. On structural identifiability. Mathematical Biosciences. 1970;7(3):329-39.

[147] Slob W, Janssen PH, van den Hof JM. Structural identifiability of PBPK models: practical consequences for modeling strategies and study designs. Crit Rev Toxicol. 1997;27(3):261-72.

[148] Yates JW. Structural identifiability of physiologically based pharmacokinetic models. J Pharmacokinet Pharmacodyn. 2006;33(4):421-39.

[149] Yates JW, Jones RD, Walker M, Cheung SY. Structural identifiability and indistinguishability of compartmental models. Expert Opin Drug Metab Toxicol. 2009;5(3):295-302.

[150] Gueorguieva I, Aarons L, Ogungbenro K, Jorga KM, Rodgers T, Rowland M. Optimal design for multivariate response pharmacokinetic models. J Pharmacokinet Pharmacodyn. 2006;33(2):97-124.

[151] Margolskee A, Darwich AS, Pepin X, Aarons L, Galetin A, Rostami-Hodjegan A, et al. IMI - Oral biopharmaceutics tools project - Evaluation of bottom-up PBPK

prediction success part 2: An introduction to the simulation exercise and overview of results. Eur J Pharm Sci. 2017;96:610-25.

[152] Saltelli A, Chan K, Scott EM. Sensitivity analysis. Chichester ; New York: Wiley; 2000. xv, 475 p. p.

[153] McNally K, Cotton R, Loizou GD. A Workflow for Global Sensitivity Analysis of PBPK Models. Front Pharmacol. 2011;2:31.

[154] Marino S, Hogue IB, Ray CJ, Kirschner DE. A methodology for performing global uncertainty and sensitivity analysis in systems biology. J Theor Biol. 2008;254(1):178-96.

[155] Clewell HJ, 3rd, Lee TS, Carpenter RL. Sensitivity of physiologically based pharmacokinetic models to variation in model parameters: methylene chloride. Risk Anal. 1994;14(4):521-31.

[156] Liu D, Li L, Jamei M. Application of Global Sensitivity Analysis Methods to Determine the most Influential Parameters of a Minimal PBPK Model of Quinidine. The PAGE 2017 meeting, Budapest, Hungary2017.

[157] Hsieh NH, Reisfeld B, Bois FY, Chiu WA. Applying a Global Sensitivity Analysis Workflow to Improve the Computational Efficiencies in Physiologically-Based Pharmacokinetic Modeling. Front Pharmacol. 2018;9:588.

[158] Liu D, Li L, Rostami-Hodjegan A, Bois FY, Jamei M. Considerations and Caveats when Applying Global Sensitivity Analysis Methods to Physiologically Based Pharmacokinetic Models. AAPS J. 2020;22(5):93.

[159] Melillo N, Aarons L, Magni P, Darwich AS. Variance based global sensitivity analysis of physiologically based pharmacokinetic absorption models for BCS I-IV drugs. J Pharmacokinet Pharmacodyn. 2019;46(1):27-42.

[160] Melillo N, Darwich AS, Magni P, Rostami-Hodjegan A. Accounting for inter-correlation between enzyme abundance: a simulation study to assess implications on global sensitivity analysis within physiologically-based pharmacokinetics. J Pharmacokinet Pharmacodyn. 2019.

[161] Okino MS, Mavrovouniotis ML. Simplification of Mathematical Models of Chemical Reaction Systems. Chem Rev. 1998;98(2):391-408.

[162] Nestorov IA, Aarons LJ, Arundel PA, Rowland M. Lumping of whole-body physiologically based pharmacokinetic models. J Pharmacokinet Biopharm. 1998;26(1):21-46.

[163] Gueorguieva I, Nestorov IA, Rowland M. Reducing whole body physiologically based pharmacokinetic models using global sensitivity analysis: diazepam case study. J Pharmacokinet Pharmacodyn. 2006;33(1):1-27.

[164] Pilari S, Huisinga W. Lumping of physiologically-based pharmacokinetic models and a mechanistic derivation of classical compartmental models. J Pharmacokinet Pharmacodyn. 2010;37(4):365-405.

[165] Wendling T, Tsamandouras N, Dumitras S, Pigeolet E, Ogungbenro K, Aarons L. Reduction of a Whole-Body Physiologically Based Pharmacokinetic Model to Stabilise the Bayesian Analysis of Clinical Data. AAPS J. 2016;18(1):196-209.

[166] Dokoumetzidis A. Lumping of compartments. PAGE 212012.

[167] Dokoumetzidis A, Aarons L. A method for robust model order reduction in pharmacokinetics. J Pharmacokinet Pharmacodyn. 2009;36(6):613-28.

[168] Dokoumetzidis A, Aarons L. Proper lumping in systems biology models. IET Syst Biol. 2009;3(1):40-51.

[169] Jamei M, Dickinson GL, Rostami-Hodjegan A. A framework for assessing inter-individual variability in pharmacokinetics using virtual human populations and integrating general knowledge of physical chemistry, biology, anatomy, physiology and genetics: A tale of 'bottom-up' vs 'top-down' recognition of covariates. Drug Metab Pharmacokinet. 2009;24(1):53-75.

[170] Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetics parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data. J Pharmacokinet Biopharm. 1980;8(6):553-71.

[171] Sheiner LB, Beal SL. Evaluation of methods for estimating population pharmacokinetic parameters. II. Biexponential model and experimental pharmacokinetic data. J Pharmacokinet Biopharm. 1981;9(5):635-51.

[172] Aarons L. Population pharmacokinetics: theory and practice. Br J Clin Pharmacol. 1991;32(6):669-70.

[173] Woodruff TJ, Bois FY. Optimization issues in physiological toxicokinetic modeling: a case study with benzene. Toxicol Lett. 1993;69(2):181-96.

[174] Goldberg DE. Genetic Algorithms in Search, Optimization and Machine Learning: Addison-Wesley Longman Publishing Co., Inc.; 1989. 372 p.

[175] Edwards AWF. Likelihood: Cambridge University Press; 1984.

[176] Wang Y. Derivation of various NONMEM estimation methods. J Pharmacokinet Pharmacodyn. 2007;34(5):575-93.

[177] 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.

[178] Mentre F, Gomeni R. A two-step iterative algorithm for estimation in nonlinear mixed-effect models with an evaluation in population pharmacokinetics. J Biopharm Stat. 1995;5(2):141-58.

[179] Bauer RJ. NONMEM Tutorial Part I: Description of Commands and Options, with Simple Examples of Population Analysis. CPT Pharmacometrics Syst Pharmacol. 2019.

[180] Monolix version 2020 Antony, France: Lixoft SAS; 2020 [accessed

[181] Jonsson F, Bois FY, Johanson G. Assessing the reliability of PBPK models using data from methyl chloride-exposed, non-conjugating human subjects. Arch Toxicol. 2001;75(4):189-99.

Chapter 2: Global sensitivity analysis of the Rodgers and Rowland model for prediction of tissue: plasma partitioning coefficients: Assessment of the key physiological and physicochemical factors that determine small-molecule tissue distribution

Estelle Yau, Andrés Olivares-Morales, Michael Gertz, Neil Parrott, Adam S. Darwich, Leon Aarons and Kayode Ogungbenro

AAPS J. 2020, 22(2):41

2.1. Abstract

In physiologically based pharmacokinetic (PBPK) modelling, the large number of input parameters, limited amount of available data and the structural model complexity generally hinder simultaneous estimation of uncertain and/or unknown parameters. These parameters are generally subject to estimation. However, the approaches taken for parameter estimation vary widely. Global sensitivity analyses are proposed as a method to systematically determine the most influential parameters that can be subject to estimation. Herein, a global sensitivity analysis was conducted to identify the key drug and physiological parameters influencing drug disposition in PBPK models and to potentially reduce the PBPK model dimensionality. The impact of these parameters was evaluated on the tissue-to-unbound plasma partition coefficients (Kpus) predicted by the Rodgers and Rowland model using Latin hypercube sampling combined to partial rank correlation coefficients (PRCC). For most drug classes, PRCC showed that LogP and fraction unbound in plasma (fu_p) were generally the most influential parameters for Kpu predictions. For strong bases, blood:plasma partitioning was one of the most influential parameters. Uncertainty in tissue composition parameters had a large impact on Kpu and Vss predictions for all classes. Among tissue composition parameters, changes in Kpu outputs were especially attributed to changes in tissue acidic phospholipid concentrations and extracellular protein tissue:plasma ratio values. In conclusion, this work demonstrates that for parameter estimation involving PBPK models and dimensionality reduction purposes, less influential parameters might be assigned fixed values depending on the parameter space, while influential parameters could be subject to parameters estimation.

2.2. Introduction

Physiologically based pharmacokinetic (PBPK) models are complex mechanistic models which are used during all stages of drug development for various analyses (e.g., *in vitro-in vivo* extrapolation (IVIVE), interspecies extrapolation and predictions of drug exposure [1-3]). Key parameters in PBPK models are tissue:unbound plasma partition coefficients (Kpus). They describe the extent of drug distribution in each organ/tissue and are defined as the ratio of the tissue drug concentration to the unbound plasma concentration at steady state. In drug discovery and early development stages, Kpus can also be used to predict the volume of distribution at steady state (Vss), a key pharmacokinetic parameter describing the overall drug distribution within the body relevant for selecting the first dose in human and dosing frequency. As *in vivo* Kpu measurements require excessive resources in terms of animals' numbers and bioanalytics for the large number of compounds considered during drug discovery, these experiments are not usually performed at this stage. Instead, several approaches have been proposed to predict Kpu from *in vitro* and *in silico* data [4-7].

The models proposed by Rodgers and Rowland [5, 8, 9] are one of the most commonly used method. In an analysis comparing models to predict tissue:plasma partition coefficients (Kp), Graham *et al.* reported that the Rodgers and Rowland (R&R) model was the most accurate for rat Kp predictions with 77% within three-fold of experimental values and the second most accurate for rat Vss prediction with 80% within three-fold [10]. Similarly, for human Vss, the PhRMA consortium reported that the R&R model had the best prediction accuracy with 78% of compounds within three-fold compared to five other mechanistic methods [11]. The main assumptions of the R&R model are that drugs partition into neutral lipids and neutral phospholipids of tissue cells, and also partition within intra- and extra- cellular tissue water. Additionally, the electrostatic interactions that form between basic drugs and tissue acidic phospholipids (AP) are incorporated for compounds with pKa≥7, while acids interact with extracellular proteins (PR) and weakly basic compounds bind predominately to albumin and neutral drugs to lipoproteins [5, 8]. The R&R model has many input parameters: drug specific input parameters are experimentally measured or calculated *in silico* from empirical models; and tissue composition parameters are set values reported in the literature for an individual or average subject of the target species. And although sometimes overlooked, variability and uncertainty do exist regarding the true value of each input parameter for a particular chemical/bioanalytical assay or a particular group of animals or human population:

uncertainty in the measured values, lack of information, inter/within species variability, etc.

During PBPK model development, the Kpus predicted by the R&R model, or similar models, are usually considered as point estimates to simulate drug distribution into tissues. Simulations are then typically validated against observed *in vivo* plasma pharmacokinetic data from pre-clinical species (namely mice, rat, dog and non-human primates) and humans. If a mismatch exists additional studies may be performed to enable a mechanistic understanding of the reasons for the difference. However, if differences cannot be explained then it is common practice to adjust parameter estimates to fit the experimental data. The complexity of the PBPK model structure, the large number of input parameters and the limited data available generally hinder estimation of uncertain or unknown model parameters. Heterogeneous and subjective approaches for parameter estimation using PBPK models exist in the literature [12-17]. For instance, modellers generally fit a few specific Kpus as pharmacokinetic data become available while fixing others subjectively. This may lead to inaccurate parameter estimates or underestimation of uncertainty, and overall poor extrapolation. Bayesian methods could be a powerful approach for aiding parameter estimation of PBPK models [18]. The current work is a first step to reduce PBPK model dimensionality by assessing the sensitivities and excluding non-influential parameters (Kpus) and model states (tissue compartments) as recommended by recent regulatory guidelines [19, 20]. The comprehensive sensitivity analysis (SA) was also conducted to identify key parameters responsible for variability/uncertainty in predicted drug distribution (Kpus and Vss).

2.3. Methods

2.3.1. Rodgers and Rowland model

The current work is focussed exclusively on the R&R equations as previous work had shown that it gave the highest degree of prediction accuracy for Kpu values [21]. More details of these equations can be found in the original articles [5, 8, 9]. The R&R model includes one equation for the prediction of Kpu for moderate-to-strong bases with one pKa≥7 (Eq. 2.1) and a second equation for other drug classes (Eq. 2.2):

$$Kpu = \left(\frac{X \cdot f_{IW}}{Y}\right) + f_{EW} + \left(\frac{P \cdot f_{NL} + (0.3P + 0.7) \cdot f_{NP}}{Y}\right) + \left(\frac{Ka_{AP} \cdot [AP^{-}]_{T} \cdot (X - 1)}{Y}\right) \quad \textbf{Eq. 2.1}$$

$$Kpu = \left(\frac{X \cdot f_{IW}}{Y}\right) + f_{EW} + \left(\frac{P \cdot f_{NL} + (0.3P + 0.7) \cdot f_{NP}}{Y}\right) + (Ka_{PR} \cdot [PR]_T) \qquad \textbf{Eq. 2.2}$$

where P is the octanol:water partition coefficient for all tissues except adipose (vegetable oil:water). The vegetable oil:water partition coefficient, $LogP_{vo:w}$, is calculated as: $LogP_{vo:w} = 1.115 \cdot logP_{o:w} - 1.35$.

f is the fractional tissue volume and subscripts EW and IW refer to extra- and intra-cellular tissue water, NL and NP refer to neutral lipids and neutral phospholipids, respectively ; $[AP^-]_T$ and $[PR]_T$ are the tissue concentrations of acidic phospholipids (AP) and extracellular albumin (for acids and weak bases) or lipoprotein (for neutrals), respectively; Ka_{AP} and Ka_{PR} are the association constants of the drug compound with AP and either extracellular albumin or lipoprotein, respectively; and X and Y are terms accounting for the drug ionisation in intracellular water and in plasma defined as follows:

For monoprotic bases: $= 1 + 10^{pKa - pH_{IW}}$, $Y = 1 + 10^{pKa - pH_p}$

For monoprotic acids: $= 1 + 10^{pH_{IW} - pKa}$, $Y = 1 + 10^{pH_p - pKa}$

For neutrals: $X = Y = 1$ *(no ionization).*

pH_{IW} : pH of the intracellular water (7), pH_p : pH of plasma (7.4) [5]

Ka_{AP} represents an overall affinity constant for various AP. The model makes the assumption that Ka_{AP} in red blood cells (RBCs) is representative of the Ka_{AP} in all tissues. Furthermore, the model assumes that Ka_{PR} determined from the plasma data is representative of the Ka_{PR} in all tissues, where Ka_{PR} represents the affinity constant to extracellular binding proteins. The affinity constants to bind to Ka_{AP} and Ka_{PR} were determined from Kpu_{RBC} or fu_p data/information using Eq. 2.3 and Eq. 2.4, respectively.

$$Ka_{AP} = \left[Kpu_{RBC} - \left(\frac{1 + 10^{pKa - pH_{RBC}}}{Y} \cdot f_{IW,RBC}\right) - \left(\frac{P \cdot f_{NL,RBC} + (0.3P + 0.7) \cdot f_{NP,RBC}}{Y}\right)\right] \cdot \left(\frac{Y}{[AP^-]_{RBC} \cdot 10^{pKa - pH_{RBC}}}\right) \qquad \textbf{Eq. 2.3}$$

$$Ka_{PR} = \left[\frac{1}{fu_p} - 1 - \left(\frac{P \cdot f_{NL_p} + (0.3P + 0.7) \cdot f_{NP_p}}{Y}\right)\right] \cdot \frac{1}{[PR]_p} \qquad \textbf{Eq. 2.4}$$

where subscripts RBC and P indicate red blood cells and plasma, respectively; pH_{RBC} is the intracellular pH of red blood cells (7.22) [5]; fu_p is the unbound fraction of drug in plasma; and Kpu_{RBC} is the red blood cell:plasma water concentration ratio.

Kpu_{RBC} can be determined *in vitro* or calculated from fu_p, blood:plasma ratio (BP) and the haematocrit using Eq. 2.5 [22, 23]:

$$Kpu_{RBC} = \frac{Haematocrit - 1 + BP}{fu_p \cdot Haematocrit} \qquad \textbf{Eq. 2.5}$$

Using these equations, negative values can be obtained for Ka_{AP} and Ka_{PR}, in such a case the affinity constants are set to zero [9].

By decomposing the R&R model (Eq. 2.1 and Eq. 2.2), three terms can actually be distinguished and each can dominate the Kpu outputs under certain circumstances:

- Term 1 ($\left(\frac{X \cdot f_{IW}}{Y}\right) + f_{EW}$), related to fractional tissue water volumes (f_{IW}, f_{EW}) and is only pKa-dependent: it has the greatest relevance if terms 2 and 3 are negligible (*e.g.*, high fu_p - low LogP compound). Under such conditions, the distribution space is the total water.
- Term 2 ($\frac{P \cdot f_{NL} + (0.3P + 0.7) \cdot f_{NP}}{Y}$), related to tissue lipid partitioning (neutral lipids (f_{NL}), and neutral phospholipids (f_{NP})) and is LogP- and pKa-dependent: it might be the most relevant term when Ka_{AP} or Ka_{PR} are zero.
- Term 3 ($\frac{Ka_{AP} \cdot [AP^-]_T \cdot (X-1)}{Y}$ or $Ka_{PR} \cdot [PR]_T$), related to interactions with tissue AP ($Ka_{AP} \cdot [AP^-]_T$) or nonspecific protein binding ($Ka_{PR} \cdot [PR]_T$), is dependent on LogP, pKa, fu_p (and BP for strong bases). It has relevance if partitioning into RBCs lipids or into plasma lipids is greater than binding to RBCs or to plasma proteins (*e.g.*, low fu_p relative to LogP), *i.e.* : $\left(\frac{1+10^{pKa-pH_{RBC}}}{Y} \cdot f_{IW,RBC}\right) + \left(\frac{P \cdot f_{NL,RBC} + (0.3P+0.7) \cdot f_{NP,RBC}}{Y}\right) \geq \frac{Haematocrit - 1 + BP}{fu_p \cdot Haematocrit}$ from equations Eq. 2.3 and Eq. 2.5, or $1 + \left(\frac{P \cdot f_{NL_p} + (0.3P+0.7) \cdot f_{NP_p}}{Y}\right) \geq \frac{1}{fu_p}$ from equation Eq. 2.4. Ka_{AP} and Ka_{PR} in term 3, are deconvolved from expressions for partitioning into RBCs or plasma lipids ($f_{NL,RBC}$ and $f_{NP,RBC}$, or $f_{NL,p}$ and $f_{NP,p}$) and interactions with RBCs acid phospholipids or plasma proteins [5, 8]. Therefore, Ka_{AP} and Ka_{PR} might be negligible or even become zero when partitioning into RBCs or plasma lipids determine fu_p and BP regardless of LogP: Kpu_{RBC} and subsequently $Ka_{AP} = 0$ if BP = 1-Haematocrit; and $Ka_{PR} = 0$ if $fu_p = 1$.

Finally, the plasma Vss can be calculated using the predicted Kpu values as follows (Eq. 2.6):

$$Vss = \sum\left((Kpu_i \cdot fu_p) \cdot V_{tissue,i} \cdot (1 - E_i)\right) + Kpu_{RBC} \cdot fu_p \cdot V_{RBC} + V_{plasma} \quad \textbf{Eq. 2.6}$$

where E_i and $V_{tissue,i}$ represents respectively the extraction ratio and the volume of i^{th} tissue (values used are given in Table S1). V_{RBC} is the volume of red blood cells. V_{plasma} is the plasma volume, and V_{plasma} = 3.15L for a reference man of 70 kg [9]. The tissue volume V_{tissue} was calculated as a fraction of body weight (BW) corrected by the tissue density (kg/L): $V_{tissue} = f_{BW} \times BW/density$.

2.3.2. Global sensitivity analysis on drug parameters: PRCC

Global sensitivity analyses (GSA) use a set of samples representative of the parameter space of inputs to explore the design space which are simulated according to their distribution functions and possible correlations [24]. Monte Carlo sampling of input parameters generates output variable distributions to be used in assessing model uncertainty [25]. Based on Monte Carlo simulations, scatterplots of the tissue Kpus predicted from the R&R model and each drug parameter were generated in order to identify visually the relationships between the inputs and outputs (tissue Kpus) which were all monotonic and mostly nonlinear (**Figure A1.2**-**Figure A1.5**).

Partial rank correlation coefficient (PRCC) is a method for GSA based on rank-transformed linear regression analysis. It is a powerful method to evaluate the statistical input-output relationships after eliminating the linear influence of other input variables and when there is a nonlinear monotonic trend between input parameters and output variables as it requires a non-parametric test of ranked data [25]. It has been demonstrated that for nonlinear non-monotonic trends, PRCC does not perform as well as variance-based methods, such as extended Fourier amplitude sensitivity test (eFAST). Nevertheless, when applied to monotonic trends, combining Latin hypercube sampling (LHS) [26, 27] with PRCC this methods is robust, reliable and less computationally costly ([28]). Additionally, PRCC can also consider interactions between parameters [29]. Calculated PRCC is a standardised similar sensitivity measure between -1 and 1 that can be compared among different parameters, with a value of |PRCC| close to 1 indicating the parameter has a strong impact on the model output. Sensitivity of Kpu predictions to the physicochemical input parameters, *i.e.* LogP, pKa, fu_p and BP, was investigated for hypothetical but relevant neutral, acid, weak basic and moderate-to-strong basic compounds in R v.3.4.2 with Rstudio v.1.0.153 [30]. Zwitterions and multiple charged compounds were not explicitly considered in this work as in the R&R model they are

expected to behave like strong bases or acids/weak bases depending primarily on their most basic/acidic pKa value. For each drug class, a total of 10,000 compounds with different properties were generated by LHS with the R package "lhs" [31]. Each simulated set of compounds per drug class was uniformly sampled by selecting values for each physicochemical drug input parameter from the ranges specified in Table 2.1. The ranges represent realistic intervals for the drug-dependent parameters [32, 33]. For each simulated compound, tissue Kpu values were then calculated using the Eq. 2.1 and Eq. 2.2. The input parameter (drug physicochemical properties) and output parameter (tissue Kpu) values were transformed into their ranks and PRCCs were calculated following the procedure described previously [28]. The significance of a non-zero PRCC value was tested using a two-sided Student's t test [28] as the number of tests performed is large, a Bonferroni multiple test correction was used [34]. Details of the PRCC analysis are provided in Appendix A1.1.

Table 2.1: Constrained bounds of drug parameters

Input parameters	Neutral	Acid	Weak Base	Strong Base
pKa	-	from 2 to 8	from 3 to 7	from 7 to 11
LogP	from -3 to 6	from -3 to 6	from -3 to 6	from -3 to 6
fu_p	from 0.001 to 1	from 0.001 to 1	from 0.001 to 1	from 0.001 to 1
BP	from 0.55 to 2.4	from 0.55 to 2.4	from 0.55 to 2.4	from 0.55 to 2.4

2.3.3. Relationship between LogP and fu_p

Model input variables are typically assumed to be independent for practical reasons, as non-independent inputs samples are more complex to generate and may need a very large sample size to compute accurate sensitivity measures. However, the assumption of independence among input variables may not be appropriate given the nature of the relationship between, for instance, lipophilicity and plasma protein binding [35-37]. Consequently, several degrees of dependency between LogP and fu_p were considered when sampling the LogP and fu_p values:

1) Independence of LogP and fu_p: LogP and fu_p was each sampled independently from its defined uniform distribution with the LHS method.
2) A linear relationship between LogP and fu_p while investigating different correlation coefficient ρ=-0.3, -0.5 and -0.9 following Iman and Conover's procedure [38].

3) A nonlinear empirical relationship between LogP/LogD and fu_p (ρ =-0.8 for neutral basic drugs, ρ =0.5 for acids) with added noise [35, 39].

The sample size of the GSA for the different degrees of dependency evaluated was chosen to be the same (N=10,000). Further details regarding the setup of the dependency between LogP and fu_p are provided in Appendix A1.2.

2.3.4. Local sensitivity analysis of drug parameters with Vss output

Preliminary analyses based on Monte Carlo sampling of input parameters revealed that the relationship between the drug input parameters and Kp and Vss was not strictly monotonic (data not shown), a key prerequisite for the use of PRCC. Consequently, this type of analysis is inappropriate for investigating the influence of drug input parameters on Vss output. Alternatively, local sensitivity analyses (SA) were performed investigating how small changes in one parameter value at a time affect the model output. The sensitivity coefficient (S_{ij}) for a definite independent variable X was calculated (Eq. 2.7) based on the partial differentiation of each output of interest with respect to each model input parameter [40] and normalised by both the output and input parameter to remove the influence of units [25]. S_{ij} quantifies the relative change of the model output at a relative change of the input parameter (Eq. 2.7). It should be noted that each input parameter is perturbed to a small extent while holding all other parameters fixed.

$$S_{ij} = \frac{\partial Y_i}{\partial X_j} \times \frac{X_j}{Y_i} \qquad \textbf{Eq. 2.7}$$

where Y_i is the model output i, X_j is the input parameter j.

As drug input parameters influence Kpu and subsequently Vss predictions, a local SA of the R&R model was carried out with the most and least influential drug input parameters found from the GSA for each drug class to assess their impact on Vss. Indeed, some tissues have a small physical volume and therefore may become negligible contributors to Vss sensitivity despite having sensitive Kpu values. On the other hand, other tissues can have a large volume and not sensitive Kpu values but be very influential on Vss. Additionally, tissues with a high drug extraction ratio (close to 1) may have a low influence on Vss. The sensitivity coefficients depend on the specific set of parameter values used and the ranges of the input variable were selected to be the same as the ones used in the GSA (Table 2.1). This analysis was not meant to be exhaustive and to investigate all possible scenarios. As such, cases of a high lipophilicity-high protein bound compound (LogP=3, fu_p=0.01) and a low lipophilicity-low protein bound

compound (LogP=-0.3, fu_p=0.9) were investigated to illustrate the impact on drug types frequently encountered during drug discovery and development that can influence significantly drug distribution.

2.3.5. Incorporation of uncertainties in physiological parameters

Tissue composition data originates from a very limited number of studies. This section investigated how uncertainties in tissue composition data may propagate into Kpu and Vss predictions. The term uncertainty includes both biological variability (different sources/animals used to generate the data) and analytical uncertainty. Physiological input parameters for PBPK models especially tissue composition data are generally fixed to average values obtained from the literature (Tables S2 and S3) [5, 8, 41, 42]. However, most studies that focused on collection of human and rat tissue composition data (*i.e.*, interstitial, intracellular, and vascular volumes, albumin and lipoprotein concentrations) reported uncertainties in those measurements [43-48]. The reported coefficients of variation (CV) varied widely from 2 to 66% depending on the fractional tissue volume but measurements were often very sparse, with only one individual measurement reported in certain cases. Measurements could also be inaccurate due to experimental limitations. Tissue composition data uncertainty can also be due to data arising from subjects of differing ages, races/strains, weights and sex. Additionally, all physiological parameters are associated with inherent biological variability in a population (animal or human). Uncertainties are generally not explicitly incorporated in Kpu predictions. Therefore, parameter uncertainty and variability in tissue compositions values could have an influence on the accuracy of Kpu and therefore the accuracy of Vss predictions. The sensitivity of physiological parameters on drug distribution was explored by incorporating 30% uncertainty on the following set of input parameters in Eq. 2.1 and Eq. 2.2 (f_{NL}, $f_{NP,}$ f_{IW}, f_{EW}, $[AP^-]_T$, $[AP^-]_{RBC}$, $[PR]_T/[PR]_p$). Three scenarios were investigated which included 30% uncertainty on different terms of the equation: (i) all tissue fractions (f_{NL}, $f_{NP,}$ f_{IW}, f_{EW}), acidic phospholipids ($[AP^-]_T$, $[AP^-]_{RBC}$) and protein ratios ($[PR]_T/[PR]_p$); (ii) all tissue fractions only and (iii) acidic phospholipids and protein ratios only. This analysis served to identify the most influential physiological parameter(s). Distributions of fractional tissue volumes had mean values matching the typical average values used for tissue composition-based models [42] and a CV of 30 % were generated.

The tissue fractions were assumed to follow logistic-normal distributions to constrain the values between 0 and 1 [49]. For this, the j normalised fractional tissue volumes per tissue were assumed to follow a (j+1)-dimensional logistic-normal distribution which is derived after the transformation of a standard (j+1)-multivariate normal distribution with mean vector M and variance-covariance matrix Σ. Examples of how to generate samples from a logistic-normal distribution were previously reported [49, 50]. It was assumed that $F = [f_1, f_2, ..., f_j]^T \sim N_j(M, \Sigma)$, where F is a j-dimensional vector that follows a standard multivariate normal distribution with mean vector M defined as a null-vector of length j, and variance-covariance matrix Σ defined as a j-diagonal matrix of 0.0862 ($=0.30^2$ in order to have 30 % CV). For plasma, the two normalised fractional parameters (f_{NP} and f_{NL}) were assumed to follow a two-dimensional logistic normal distribution, whereas for adipose, bone, brain, gut, heart, kidney, liver, lung muscle, skin, and spleen, four normalised fractional parameters (f_{NP}, f_{NL}, f_{EW} and f_{IW}) were assumed to follow a four-dimensional logistic normal distribution. For blood cells, the three normalised fractional parameters (f_{NP}, f_{NL}, and f_{IW}) were assumed to follow a three-dimensional logistic normal distribution. Here, M and Σ parameters were fixed to generate population distributions of fractional tissue volumes that have means matching the average physiological parameter values [42] and a CV of 30% in the logistic domain. On the other hand, $[AP^-]_T$, $[AP^-]_{RBC}$, $[PR]_T/[PR]_p$ were sampled from a normal distribution ($N(\mu, \sigma)$) where μ is the average value given and σ is the variance of the associated normal in order to generate distributions of these physiological parameters with mean matching the average values and a CV of 30%. Additional details of the incorporation of uncertainties are provided in Appendix A1.3.

Finally, Kpu values were estimated based on R&R equations (Eq. 2.1 and Eq. 2.2) for a hypothetical compound of each class (neutral, acid, weak base, strong base) under the three different correlation assumptions outlined above. Additionally, the analysis was done according to four case scenarios of LogP and fu_p which corresponded to: a hydrophilic, a lipophilic, a highly bound and a lowly bound compound. These cases represented examples frequently encountered during drug development (Table 2.2). Simulations of fractional tissue volumes, $[AP-]_T$, $[AP^-]_{RBC}$ and $[PR]_T/[PR]_p$ ratios (N=1000 each) and calculation of Kpu and Vss values for each set of simulated tissue composition values (Eq. 2.1, Eq. 2.2, and Eq. 2.6) were implemented in R. We defined the effect of uncertainties in input on output as very influential if the CV% on the output was greater than 10% for a 30 CV% of the input.

Table 2.2: Simulation case scenarios investigated using sampled fractional tissue volumes and extracellular proteins levels

Drug characteristics	CASE A	CASE B	CASE C	CASE D
Neutral BP:1	LogP: -0.3 fu_p: from 0.001 to 1	LogP:3 fu_p: from 0.001 to 1	LogP: from -0.3 to 3 fu_p:0.01	LogP: from -0.3 to 3 fu_p:0.9
Acid pKa:3 BP:0.55				
Weak Base pKa:6.5 BP:1				
Strong Base pKa:9 BP:1				

2.4. Results

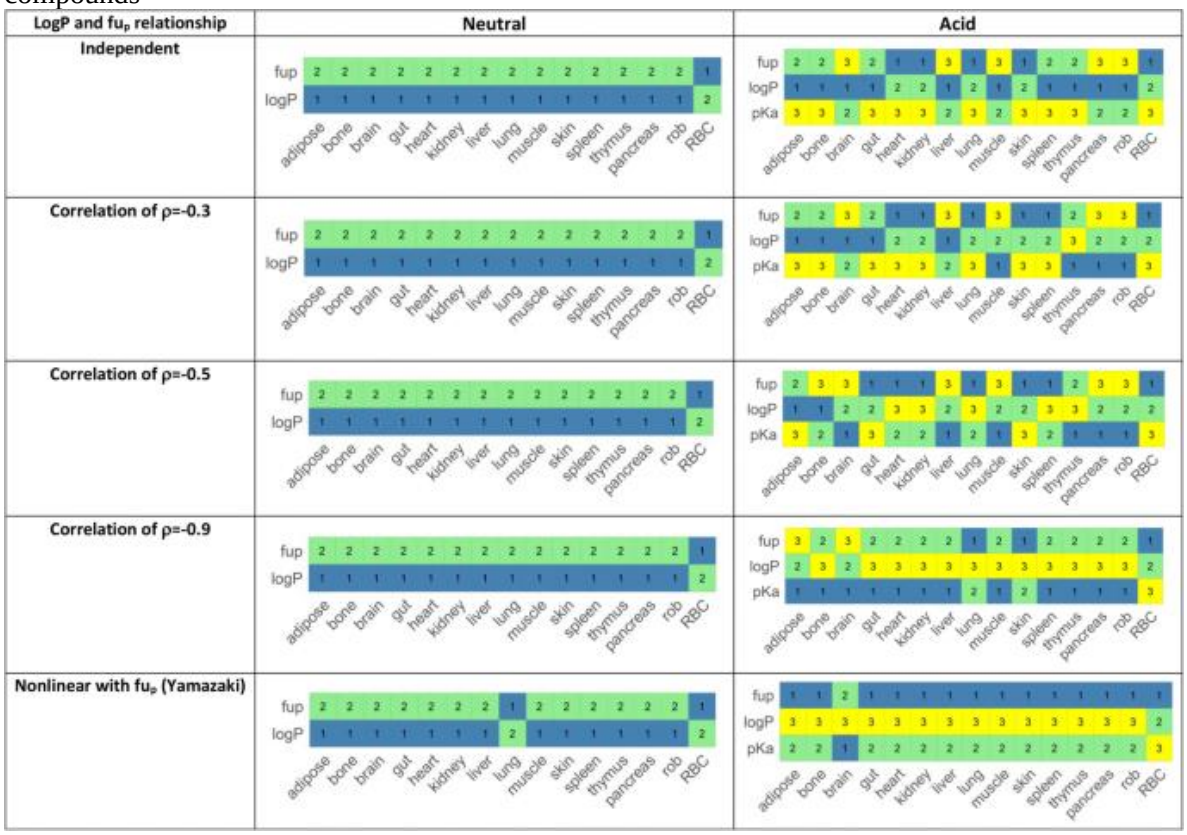
2.4.1. Global sensitivity analysis of drug-specific parameters: PRCC

The drug-specific input parameters used for the Kpu predictions were pKa, LogP, fu_p and BP, depending on drug class (Eq. 2.1 and Eq. 2.2). The results of GSA combining LHS and PRCC with the different relationships between LogP and fu_p investigated are summarised in Table 2.3. The interpretation of the results depended on the sampled input space and on the correlation between LogP and fu_p parameters. However, all parameters were influential and showed statistically significant PRCC values (p-value<0.001 after Bonferroni correction) except for pKa for adipose and skin Kpus for weak bases (**Figure A1.7**).

When all input parameters were sampled independently, the PRCCs assessment showed that LogP played a major role in the Kpu predictions for all drug classes, indicating generally the highest sensitivity of all input parameters (Figure 2.1). The second most influential parameter overall was fu_p. For acidic drugs, it was even the most influential parameter for a few of the tissue Kpus (heart, kidney, lung and skin) which represented tissues with a high fractional volume of extracellular water and albumin ratio (Figure 2.1). For strong bases, fu_p was actually the most influential input parameter for most Kpus except for the tissues that displayed the smallest tissue concentration of AP, namely adipose, bone and brain, where LogP was the most influential input parameter. For strong bases, BP was found to have a strong impact on the tissue Kpu outputs (Figure 2.1). In general, among the investigated parameters, pKa tended to be the least influential parameter with absolute PRCC values between 0.18 and 0.72 for acids and strong bases

during independent sampling of LogP and fu_p. In general, LogP, pKa and BP were generally positively correlated with the Kpu outputs across all classes except for the strong bases where pKa was negatively correlated with the outputs (**Figure A1.7**). On the other hand, fu_p was strongly negatively correlated with the outputs across all drug classes (**Figure A1.6** and **Figure A1.7**).

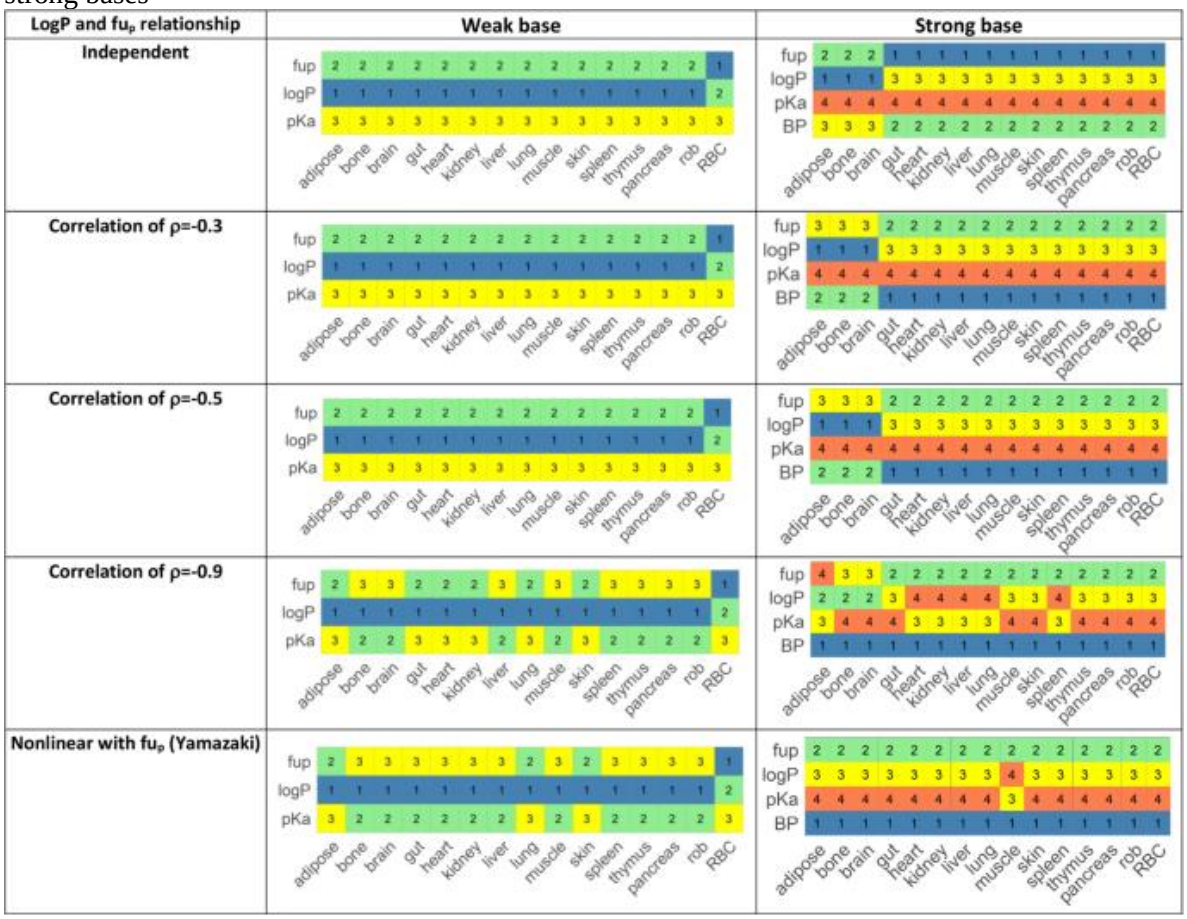
Figure 2.1: Parameter ranking determined by the PRCC of tissue Kpus for each drug class with different relationships between LogP and fraction unbound in plasma (fu_p) for neutral and acidic compounds



rob: rest of body; RBC: Red blood cell
Parameter sensitivity ranking from most to less: blue (1), green (2), yellow (3)

When considering a low to moderate correlation of ρ= -0.3 or -0.5 between LogP and fu_p, the results were similar to the ones obtained where LogP and fu_p were independent, where the sensitivity of Kpu to LogP was the highest followed by fu_p and pKa (Figure 2.1-Figure 2.2). However, for acids, fu_p became a more influential parameter for some of the tissue Kpus although the difference between the PRCC values of LogP and fu_p of these tissues was actually slight (PRCC < 0.1) (**Figure A1.6**). For strong bases, BP became a more influential parameter than fu_p for a majority of the tissue Kpus (maximal PRCC of |0.79|, Figure 2.1).

Figure 2.2 : Parameter ranking determined by the PRCC of tissue Kpus for weak and strong bases with different relationship between LogP and fraction unbound in plasma (fu_p) for weak and strong bases



rob: rest of body; RBC: Red blood cell
Parameter sensitivity ranking from most to less: blue (1), green (2), yellow (3), red (4)

When considering a strong correlation of ρ= -0.9, a minimal change was observed for neutral and weak basic compounds where LogP remained the most sensitive parameter (Figure 2.1-Figure 2.2). For acidic compounds, LogP now had the smallest PRCCs (between 0.49 and 0.69, Figure 2.1). For strong bases, BP was now the most sensitive parameter (PRCC between 0.74 and 0.84), followed by fu_p (PRCC between -0.56 and -0.72), LogP (PRCC between 0.02 and 0.34) and pKa (PRCC between -0.04 and -0.31) (Figure 2.2). When considering a nonlinear relationship where fu_p was considered dependent on LogP, sensitivity ranking was similar to when considering a strong correlation of -0.9 between LogP and fu_p. The exception was for acids where fu_p was the most influential parameter (PRCC between -0.76 and -0.92, Figure 2.1). It appeared that a nonlinear relationship where fu_p depended on LogP, the distribution of fu_p was mostly concentrated around its lower range (more than 60% of simulated compounds had an

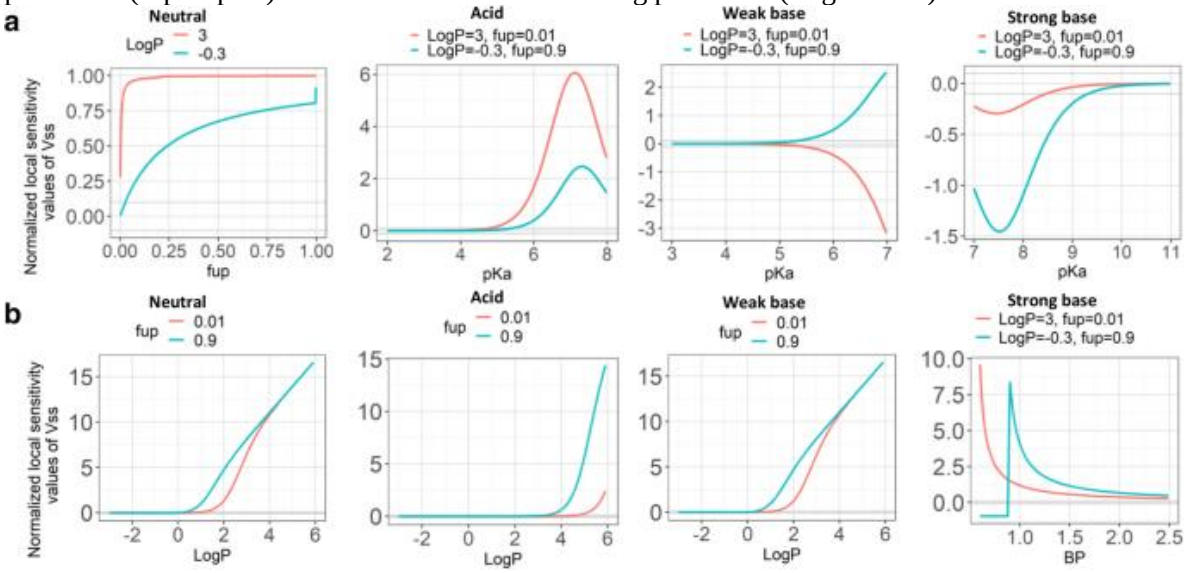
fu_p<0.1) while with a correlation of -0.9 fu_p stayed uniformly distributed within the whole defined range.

2.4.2. Local sensitivity analysis with Vss output

Drug input parameters influenced Kpu and subsequently Vss predictions. **Figure 2.3**A and **Figure 2.3**B illustrated the influence of these least and most influential parameters, respectively, on a normalized score of the Vss. For this assessment, two prototypical compounds, one high lipophilicity-high protein bound (LogP = 3, fup = 0.01) and one low lipophilicity-low protein bound (LogP = − 0.3, fup = 0.9) compound, were investigated. Although pKa for acids, strong and weak bases and fu_p for neutrals were identified as the least relevant, these parameters still had an important influence on the Vss predictions, especially between pKa of 5 and 9 where acids, strong and weak bases can be ionised at physiological pH and when plasma protein binding was high (fu_p <0.1) as it has an inverse influence on tissue Kpu (**Figure 2.3**). For neutrals, an increase in sensitivity of Vss was observed when fu_p approached 1 ; when considering a lipophilic compound with LogP=3, the most apparent change was seen between fu_p 0.001 and 0.1 while for a hydrophilic compound with LogP=-0.3, it was between 0.9 and 1. In this latter case of high fu_p, Ka_{PR} was set to zero as negative values were obtained [9], term 2 became zero in Eq. 2.2, and term 3 (function of LogP and pKa) was dominant in Kpu predictions resulting in a sudden change in Vss and a sharp profile. For acidic compounds, a positive change in the Vss was observed from pKa values ranging from 2 to 7.2 approximately, and negative change for pKa values between 7.2 and 8 for compounds with high and low lipophilicity although the normalized sensitivity coefficient at pKa 7.2 was around 6 for lipophilic acidic compound and only 2.2 for hydrophilic acidic compound. No change in Vss was observed for weak bases with pKa values between 3 and 5 where compounds are mainly unionised (78% with < 5% ionisation in plasma) and a positive change was observed from a pKa of 5 to 6.9 for an unbound hydrophilic weakly basic compound, while a negative change was observed for a bound lipophilic weak basic compound. In this latter case, lipid partitioning was greater than nonspecific protein binding (*e.g.*, low fu_p relative to LogP), the terms 1 and 3 became negligible in Eq. 2.2 and the term 2 (dependent on a mixture of LogP, pKa, fu_p) was predominant and decreased as pKa increased. For strong basic compounds, negative change in the Vss was observed from pKa 7 to 7.5, a positive change was observed from pKa 7.5 to 10, and no change from pKa 10 to 11 (fully ionized from pKa > 9.5); the normalized sensitivity coefficient at pKa

7.5 was -1.5 for free hydrophilic strongly basic compounds and only -0.3 for bound lipophilic strongly basic compound. In this latter case, the affinity constant Ka_{AP} varied greatly as Kpu_{RBC} was very high (Eq. 2.1) causing a higher change in Vss.

Figure 2.3. Normalised local SA values (Sij) of Vss with respect to '**a'** the least influential drug parameter (fup or pKa) or '**b'** the most influential drug parameter (LogP or BP).



fup: fraction of unbound drug; BP: blood-to-plasma ratio
Red line: Compound with high lipophilicity/low binding; Blue line: Compound with low lipophilicity/high binding. fup was varied in increments of 0.001 from 0.001 to 1; pKa values were varied in increments of 0.01; LogP was varied in increments of 0.1; BP was varied in increments of 0.01 from 0.6 to 2.5 for an acid (pKa = 3), a weak base (pKa = 6) and a strong base (pKa = 8)

Compared to **Figure 2.3**A, the normalized local sensitivity values of Vss to the most influential drug parameters axis in **Figure 2.3**B were indeed larger corroborating that the parameters LogP and BP were considerably more influential. For a strongly basic compound with pKa=8, LogP=-0.3 and fu_p=0.9 (high fu_p), Ka_{AP} was set to zero as negative values were obtained when BP<fu_p; consequently term 2 became null in Eq. 2.1, and term 3 (function of LogP and pKa which were fixed) was dominant in Kpu predictions resulting in no change in Vss and then a sharp profile when BP>fu_p.

2.4.3. Incorporation of uncertainties in tissue composition parameters

The impact of uncertainties in tissue composition input data on Vss predictions is illustrated in Figure 2.4 and the individual results of all tissue Kpu outputs can be found in **Figure A1.8**. For all compound classes, the effect of tissue composition data on model output (Kpu and Vss) was influential as the CV varied up to 43% in Kpu outputs and up to 32% in Vss outputs when considering a CV of 30% in tissue composition. In order to identify which parameter was primarily responsible for the observed changes and sensitivities, three scenarios were investigated.

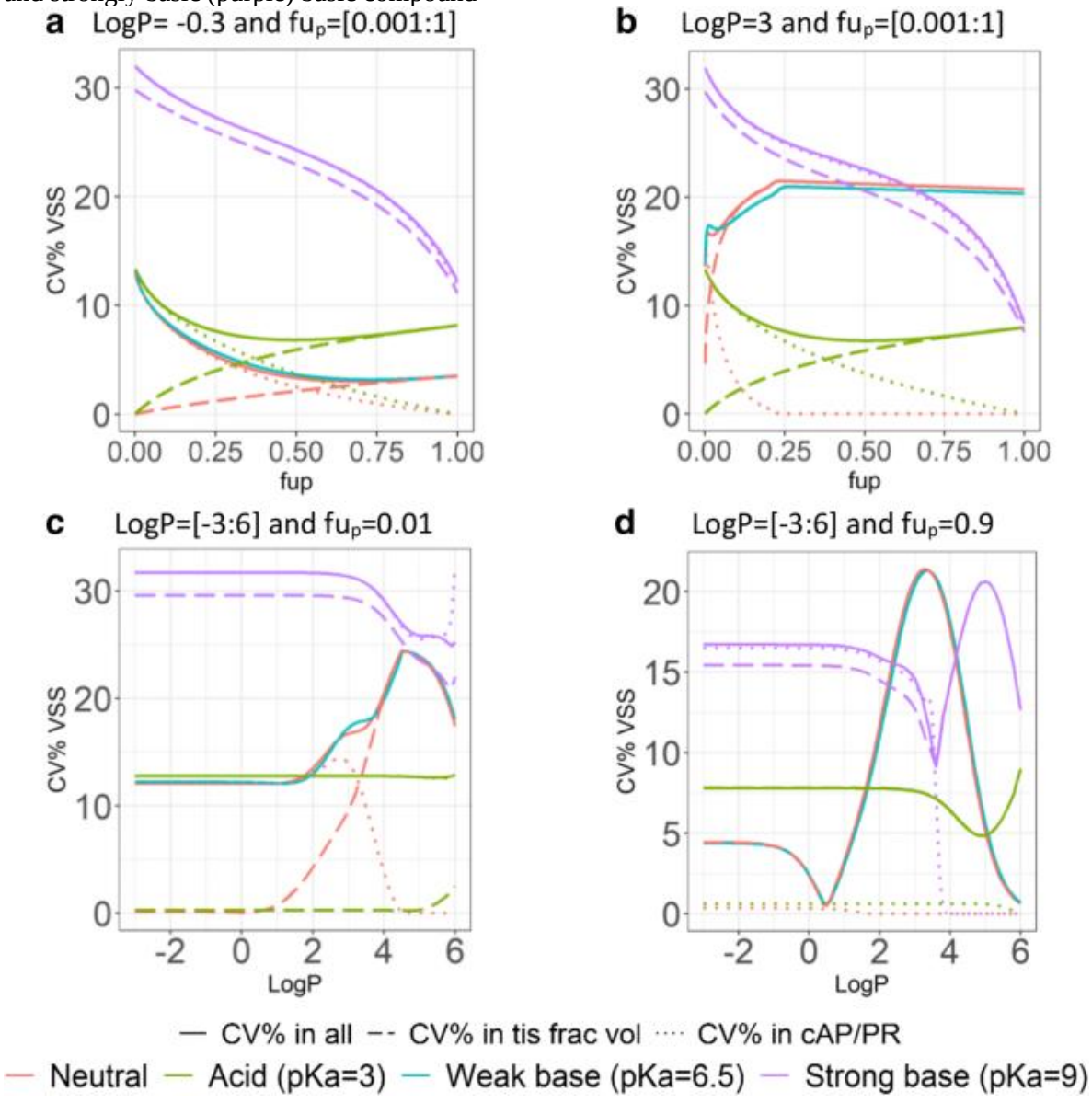
Firstly, uncertainties of 30% in all tissue composition parameters resulted in the highest output variability for highly bound strong bases (fu_p<0.01) with output CVs between 25 and 45% and between 25 and 32%, respectively, for Kpu and Vss (solid lines, Figure 2.4C). The CV values in output parameters greater than 30% resulted from the variability of AP being included at the level of the tissue $[AP^-]_T$ and of the red blood cell values $[AP^-]_{RBC}$ in Eq. 2.1 and Eq. 2.3 leading to a ratio of $[AP^-]_T/[AP^-]_{RBC}$ with CV%>30 for Kpu outputs (*e.g.*, 41%CV for Kpu lung, 48%CV for Kpu skin). For unbound strong bases in plasma (fu_p>0.9, Figure 2.4D) and other compound classes, the influence of uncertainties in tissue compositions on Kpu and Vss outputs was smaller but still influential with CVs between 10 and 30%, and 9 and 21% for Kpu and Vss, respectively (solid lines, Figure 2.4D).

Secondly, uncertainties of 30% in fractional tissue volumes only resulted in limited output variability with CV of 0 to 30 % for both Kpu and Vss parameters (dashed lines, Figure 2.4). For the tissue volumes to be influential, the specific and nonspecific binding of drug needed to be negligible (*i.e.*, high fu_p and low LogP). As specific and nonspecific binding increased, the relevance of the fractional volume terms diminished (Eq. 2.1 and Eq. 2.2) since they represented only a small proportion of the total tissue volumes. It should be noted that the adipose Kpu stood out in its behaviour compared to the other tissues, probably because of the high fractional volume of neutral lipids compared to other tissues (**Figure A1.8**A and B).When varying LogP for strong bases at an fu_p of 0.9, adding uncertainties in fractional tissue volumes had an impact on tissue Kpus with CVs between 5 and 28% (dashed lines, Figure 2.4) but resulting in a CV of 16% in Vss predictions possibly due to the errors cancelling each other out (Figure 2.4D).

Finally, uncertainties of 30% in only $[AP]_T$, $[AP]_{RBC}$ and extracellular $[PR]_T/[PR]_p$ resulted in similar output variability of 30% in all tissue composition parameters,

especially when fu_p was low (fu_p<0.1) (solid and dotted lines were blended, Figure 2.4B), suggesting that uncertainties in tissue AP and extracellular PR levels data dominated the changes in Kpu and Vss outputs in this fu_p parameter space (an average 43%CV in tissue Kpu and a CV close to 30% in Vss output). Indeed, the term related to AP and Ka_{AP} in Eq. 2.1 was set to zero and plasma binding was exclusively driven by nonspecific binding which explained the observed sharp profiles in **Figure A1.8B**.

Figure 2.4: Effect of inputting CV30% simultaneously or individually on fractional tissue volumes and/or all tissue acidic phospholipids (cAP)/ extracellular protein ratios (PR) when varying fup or LogP on Vss for a hypothetical neutral (red), acidic (green), weakly basic (blue) and strongly basic (purple) basic compound



2.5. Discussion

In this paper, global and local SA were conducted to identify the key drug and physiological parameters of tissue distribution based on the mechanistic Kpu equations of R&R [5, 8]. These equations were selected as they represent an accurate prediction tool for Kpu values [10] and mechanistically integrate many of the underlying distribution mechanisms along with physiological information. However, a similar analysis is of course applicable to other Kpu mechanistic equations for Kpu predictions as well [4, 7, 51-54]. The decomposition of equations Eq. 2.1 and Eq. 2.2 into three terms in the Methods section allowed a better understanding of the equations and how uncertainty propagates into Kpu and Vss when varying drug and tissue composition parameters. However, several key distribution processes are not considered in the R&R model such as tissues being divided into interstitial and intracellular spaces with differing pH values [7], lysosomal trapping [55], microsomal partitioning [56] and active transport across membranes (*e.g.*, for poorly permeable molecules) and could contribute to the Vss misprediction.

The GSA combining LHS and PRCC was performed on input parameter ranges covering a wide parameter space (Table 2.1) [28] and with several degrees of dependency between LogP and fu_p. Contradicting reports exist in the literature regarding the relationship [35, 57-61] or lack of relationship [62-64] between fu_p and LogP (or LogD). Degrees of correlation were selected based on a few nonlinear negative relationships reported in the literature with correlation values ranging from -0.91 to -0.36 depending on the investigated dataset of compounds and classes and also the measured or calculated LogP/LogD term considered [35, 57, 65, 66]. Ultimately, the correlation can slightly change the sensitivity ranking of the input parameters especially for acids and strong bases and additional information on correlations between LogP and fu_p within the specific chemotype could be useful as input for the parameter estimation process for these classes.

In the current analysis, different correlations between fup and LogP were investigated and overall, two patterns could be distinguished: (1) a case with no-to-moderate extent of correlation and (2) a case with high correlation (Table 2.3). When all input parameters were assumed to be independent, LogP was generally the most influential parameter for neutral drugs and weak bases which were predominantly unionised at physiological pH; while fu_p was the most influential input parameter for most Kpus for strong bases. Given its considerable importance, any errors in computational or experimental LogP determination can significantly influence Kpu predictions (Figure 2.1).

A recent study showed that among different LogP methods investigated, the LogP prediction of the best model against a compound dataset representing pharmaceutical space was within one log unit approximately 70% of the time [67]. Introducing a correlation between LogP and fu_p constrained the sampling space and made it less likely that term 2 of Eq. 2.1 and Eq. 2.2 became zero; therefore, term 3 became the only relevant mechanism of tissue partitioning. This attenuated the importance of LogP especially for acids and strong bases. A correlated sample may be a more plausible combination of drug input parameters and possibly a better representation of the behaviour of drug compounds when correlations are known [68, 69]. However, when LogP was uniformly sampled and fu_p was calculated using Yamazaki *et al.*'s relationship [35], the main limitation was that many of the simulated compounds had a very small fu_p value (<0.05). This may be an artefact caused by the large experimental errors of the drugs as the data were compiled from many different sources [70]. Free fraction can be accurately measured up to 0.1% (0.001) for highly bound compounds and becomes more uncertain below this value [71]. As fu_p was identified as a highly influential parameter, uncertainty in its determination is likely going to contribute significantly to the variability of Kpu and Vss outputs.

Table 2.3: Summary of drug parameter sensitivity and ranking based on the performed GSA

	Independently (or low correlation) sampled LogP and fu_p		**Correlation of -0.9 or nonlinear relationship between LogP and fu_p**	
	High sensitivity parameter (overall PRCC >0.5)	**Low sensitivity parameter (overall PRCC <0.5)**	**High sensitivity parameter (overall PRCC >0.5)**	**Low sensitivity parameter (overall PRCC <0.5)**
Neutrals	LogP>fu_p		LogP>fu_p	
Acids	LogP>fu_p>pKa		pKa~fu_p >LogP	
Weak bases	LogP>fu_p	pKa	LogP	pKa~fu_p
Strong bases	fu_p> BP~LogP	pKa	BP>fu_p	LogP>pKa

>: greater sensitivity ranking; ~: similar sensitivity ranking

PRCC analyses only gave ranking of parameter relevance for Kpu values but did not evaluate the contribution of each input to the output uncertainty. Other GSA methods including screening methods, variances could be applied although they are more computationally intensive and the interpretation is difficult in the presence of statistical dependence between inputs [72, 73]. In contrast to GSA, the local SA assesses the impact of single parametric perturbations on the model output (Kpus, Vss). Although the GSA ranking differed from the local SA ranking in certain parameter spaces, this should not be considered inconsistent as they arise from different design and purposes. The local SA on

Vss predictions showed that even the least influential drug parameters (based on PRCC) on Kpu were found to have a relevant influence on the Vss predictions. Particularly when pKa was around physiological pH where acid, weak and strong bases are unionised and when plasma protein binding was high (fu_p <0.1) [9]. On the other hand, the high sensitivity of Kpu and subsequently Vss to BP illustrates how important it is to measure this value for strong basic compounds as BP serves as surrogate for drug interaction with AP in the body (Eq. 2.3). Therefore, when BP is frequently assumed to be one for strong bases due to unavailable measurements, it can actually lead to errors in AP drug affinity calculations and subsequent Kpu and Vss predictions.

Due to the lack of tissue composition data in human especially regarding concentrations of AP, albumin and lipoprotein ratios, rat data are used instead [9, 42]. However, prediction success of human Vss might be affected by the assumption that rat and human tissue compositions are the same. Our uncertainty analysis illustrated the influence of uncertainties in tissue composition data (*i.e.*, tissue fractions, phospholipids and protein ratios) on Kpu and Vss predictions when a CV of 30% was considered for the tissue composition values. The choice of CV 30% was a realistic average value as a 5 or 15 CV % had been reported for several fractional tissue volumes and up to 60% for other fractional volumes in rat and human [43-46, 74]. However, this may inadequately represent the variability in tissue composition in the general population as a large interspecies variation can exist in these measured parameters. For example, a fractional AP content of 0.0004 was found in rat brain, whereas a content of 0.02 was found in human brain [75]. Moreover, lipid composition of neutral lipids and acidic phospholipids was shown to differ between species, especially compositions of fatty acids [74, 76], which may lead to variable interactions with drugs from one species to another. Alternatively, modelling and simulations in conjunction with imaging techniques (e.g., MRI, PET, PET-CT scan) can be used to help characterise tissue distribution in the body and different tissues [77-79].

Uncertainties in tissue composition are likely going to have a considerable impact on the success of Vss predictions for all classes and especially for strong bases with low fu_p mostly due to the uncertainty in data on tissue specific acidic phospholipid and protein levels. Further examination by separation of fractional tissue volumes from [AP^-] or extracellular PR revealed that fractional tissue volumes had less impact than AP and extracellular PR exception for a compound with high LogP and fu_p (Figure 2.4). Overall, this uncertainty analysis indicates that additional research and a better characterisation of

AP and extracellular PR (albumin and lipoprotein) levels in tissues and plasma would improve the confidence in Kpu and Vss prediction accuracies across species. However, it should be noted that this assessment is based on the R&R model, which assumes that ionised bases interact predominately with AP and uncharged with neutral phospholipids and neutral lipids. This assumption has been questioned recently [80, 81]. In addition to model refinements, additional data particularly AP and extracellular PR in the different tissues will help to reduce uncertainty and obtain more reliable Kpu predictions in the future.

Finally, for the parameter estimation process, less influential parameters for Kpu predictions in each drug class might be assigned fixed values depending on the sensitivity of the parameter space, while influential parameters could be fitted using priors and uncertainty associated with experimental methods and data. In this work, we found that the sensitivity ranking depends on the degree of dependence between LogP and fu_p for acids and strong bases, therefore this needs to be taken into account when fixing certain parameters.

2.6. Conclusions

Based on the GSA using a wide range of drug input parameters, the most influential parameters on Kpu predictions in the R&R model were generally LogP and fup for the drug-specific parameters.

Uncertainties in tissue composition have a considerable influence on Kpu and Vss predictions for all classes and especially for strong bases with low fu_p mostly due to the uncertainty in data on tissue specific acidic phospholipid and protein levels.

In the context of parameter estimation for PBPK models and dimensionality reduction, less influential parameters for Kpu predictions in each drug class might be assigned fixed values depending on the sensitivity of the parameter space, while influential parameters could be fitted, for instance using a Bayesian approach, where priors and uncertainty associated with experimental methods and data are accounted for.

2.7. References

[1] Jones HM, Chen Y, Gibson C, Heimbach T, Parrott N, Peters SA, et al. Physiologically based pharmacokinetic modeling in drug discovery and development: a pharmaceutical industry perspective. Clin Pharmacol Ther. 2015;97(3):247-62.

[2] Jamei M. Recent Advances in Development and Application of Physiologically-Based Pharmacokinetic (PBPK) Models: a Transition from Academic Curiosity to Regulatory Acceptance. Curr Pharmacol Rep. 2016;2:161-9.

[3] Luzon E, Blake K, Cole S, Nordmark A, Versantvoort C, Berglund EG. Physiologically based pharmacokinetic modeling in regulatory decision-making at the European Medicines Agency. Clin Pharmacol Ther. 2016.

[4] Poulin P, Theil FP. A Priori Prediction of Tissue:Plasma Partition Coefficients of Drugs to Facilitate the Use of Physiologically-Based Pharmacokinetic Models in Drug Discovery. Journal of Pharmaceutical Sciences. 2000;89(1):16-35.

[5] Rodgers T, Leahy D, Rowland M. Physiologically based pharmacokinetic modeling 1: predicting the tissue distribution of moderate-to-strong bases. J Pharm Sci. 2005;94(6):1259-76.

[6] Berezhkovskiy LM. Volume of distribution at steady state for a linear pharmacokinetic system with peripheral elimination. Journal of Pharmaceutical Sciences. 2004;93(6):1628-40.

[7] Schmitt W. General approach for the calculation of tissue to plasma partition coefficients. Toxicol In Vitro. 2008;22(2):457-67.

[8] Rodgers T, Rowland M. Physiologically based pharmacokinetic modelling 2: predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. J Pharm Sci. 2006;95(6):1238-57.

[9] Rodgers T, Rowland M. Mechanistic approaches to volume of distribution predictions: understanding the processes. Pharm Res. 2007;24(5):918-33.

[10] Graham H, Walker M, Jones O, Yates J, Galetin A, Aarons L. Comparison of in-vivo and in-silico methods used for prediction of tissue: plasma partition coefficients in rat. J Pharm Pharmacol. 2012;64(3):383-96.

[11] Jones RD, Jones HM, Rowland M, Gibson CR, Yates JW, Chien JY, et al. PhRMA CPCDC initiative on predictive models of human pharmacokinetics, part 2: comparative assessment of prediction methods of human volume of distribution. J Pharm Sci. 2011;100(10):4074-89.

[12] Woodruff TJ, Bois FY. Optimization issues in physiological toxicokinetic modeling: a case study with benzene. Toxicol Lett. 1993;69(2):181-96.

[13] Yang J, Jamei M, Heydari A, Yeo KR, de la Torre R, Farre M, et al. Implications of mechanism-based inhibition of CYP2D6 for the pharmacokinetics and toxicity of MDMA. J Psychopharmacol. 2006;20(6):842-9.

[14] Peters SA. Identification of intestinal loss of a drug through physiologically based pharmacokinetic simulation of plasma concentration-time profiles. Clinical Pharmacokinetics. 2008;47(4):245-59.

[15] Xia B, Heimbach T, Gollen R, Nanavati C, He H. A simplified PBPK modeling approach for prediction of pharmacokinetics of four primarily renally excreted and CYP3A metabolized compounds during pregnancy. AAPS J. 2013;15(4):1012-24.

[16] Ke A, Barter Z, Rowland-Yeo K, Almond L. Towards a Best Practice Approach in PBPK Modeling: Case Example of Developing a Unified Efavirenz Model Accounting for Induction of CYPs 3A4 and 2B6. CPT Pharmacometrics Syst Pharmacol. 2016;5(7):367-76.

[17] Budha NR, Ji T, Musib L, Eppler S, Dresser M, Chen Y, et al. Evaluation of Cytochrome P450 3A4-Mediated Drug-Drug Interaction Potential for Cobimetinib Using Physiologically Based Pharmacokinetic Modeling and Simulation. Clin Pharmacokinet. 2016;55(11):1435-45.

[18] Tsamandouras N, Rostami-Hodjegan A, Aarons L. Combining the 'bottom up' and 'top down' approaches in pharmacokinetic modelling: fitting PBPK models to observed clinical data. Br J Clin Pharmacol. 2013;79(1):48-55.

[19] European Medicines Agency, Guideline on the qualification and reporting of physiologically based pharmacokinetic (PBPK) modelling and simulation 2018 [Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-reporting-physiologically-based-pharmacokinetic-pbpk-modelling-simulation_en.pdf; accessed May 2019]

[20] U.S. Food and Drug Administration, Physiologically Based Pharmacokinetic Analyses — Format and Content : Guidance for Industry 2018 [Available from: https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM531207.pdf; accessed May 2019]

[21] Graham H. Predicting drug distribution in rat and human: University of Manchester; 2012.

[22] Davies B, Morris T. Physiological parameters in laboratory animals and humans. Pharm Res. 1993;10(7):1093-5.

[23] Rowland M, Tozer TN. Clinical Pharmacokinetics: Concepts and Applications. 3rd ed: Lippincott Williams & Wilkins; 1995.

[24] Kamboj S, Cheng JJ, Yu C. Deterministic vs. probabilistic analyses to identify sensitive parameters in dose assessment using RESRAD. Health Phys. 2005;88(5 Suppl):S104-9.

[25] Hamby DM. A review of techniques for parameter sensitivity analysis of environmental models. Environ Monit Assess. 1994;32(2):135-54.

[26] McKay MD, Beckman RJ, Conover WJ. Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. Technometrics. 1979;21(2):239-45.

[27] Blower SM, Dowlatabadi H. Sensitivity and Uncertainty Analysis of Complex-Models of Disease Transmission - an Hiv Model, as an Example. Int Stat Rev. 1994;62(2):229-43.

[28] Marino S, Hogue IB, Ray CJ, Kirschner DE. A methodology for performing global uncertainty and sensitivity analysis in systems biology. J Theor Biol. 2008;254(1):178-96.

[29] Saltelli A, Chan K, Scott EM. Sensitivity analysis. Chichester ; New York: Wiley; 2000. xv, 475 p. p.

[30] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria2018.

[31] Carnell R. lhs: Latin Hypercube Samples. R package version 1.0.1 ed: Comprehensive R Archive Network (CRAN); 2019.

[32] Margolskee A, Darwich AS, Pepin X, Pathak SM, Bolger MB, Aarons L, et al. IMI - oral biopharmaceutics tools project - evaluation of bottom-up PBPK prediction success part 1: Characterisation of the OrBiTo database of compounds. Eur J Pharm Sci. 2017;96:598-609.

[33] Poulin P, Jones HM, Jones RD, Yates JW, Gibson CR, Chien JY, et al. PhRMA CPCDC initiative on predictive models of human pharmacokinetics, part 1: goals, properties of the PhRMA dataset, and comparison with literature datasets. J Pharm Sci. 2011;100(10):4050-73.

[34] Abdi H. The Bonferonni and Šidák Corrections for Multiple Comparisons. Encyclopedia of measurement and statistics. 2007;3.

[35] Yamazaki K, Kanaoka M. Computational prediction of the plasma protein-binding percent of diverse pharmaceutical compounds. J Pharm Sci. 2004;93(6):1480-94.

[36] Laznicek M, Laznickova A. The effect of lipophilicity on the protein binding and blood cell uptake of some acidic drugs. J Pharm Biomed Anal. 1995;13(7):823-8.

[37] Ghafourian T, Amin Z. QSAR models for the prediction of plasma protein binding. Bioimpacts. 2013;3(1):21-7.

[38] Iman RL, Conover WJ. A Distribution-Free Approach to Inducing Rank Correlation among Input Variables. Commun Stat B-Simul. 1982;11(3):311-34.

[39] Saltelli A. Global sensitivity analysis : the primer. Chichester, England ; Hoboken, NJ: John Wiley; 2008. x, 292 p. p.

[40] Saltelli A, Ratto M, Tarantola S, Campolongo F, Commission E, Ispra JRC. Sensitivity analysis practices: Strategies for model-based inference. Reliab Eng Syst Safe. 2006;91(10-11):1109-25.

[41] Poulin P, Theil FP. Development of a novel method for predicting human volume of distribution at steady-state of basic drugs and comparative assessment with existing methods. J Pharm Sci. 2009;98(12):4941-61.

[42] Poulin P, Jones RD, Jones HM, Gibson CR, Rowland M, Chien JY, et al. PHRMA CPCDC initiative on predictive models of human pharmacokinetics, part 5: prediction of plasma concentration-time profiles in human by using the physiologically-based pharmacokinetic modeling approach. J Pharm Sci. 2011;100(10):4127-57.

[43] Simon G, Rouser G. Species variations in phospholipid class distribution of organs. II. Heart and skeletal muscle. Lipids. 1969;4(6):607-14.

[44] Rouser G, Simon G, Kritchevsky G. Species variations in phospholipid class distribution of organs. I. Kidney, liver and spleen. Lipids. 1969;4(6):599-606.

[45] Hof H, Simon RG. Phospholipid content of human and guinea pig muscle: post-mortem changes and variations with muscle composition. Lipids. 1970;5(5):485-7.

[46] Diagne A, Fauvel J, Record M, Chap H, Douste-Blazy L. Studies on ether phospholipids. II. Comparative composition of various tissues from human, rat and guinea pig. Biochim Biophys Acta. 1984;793(2):221-31.

[47] Miller SP, Zirzow GC, Doppelt SH, Brady RO, Barton NW. Analysis of the lipids of normal and Gaucher bone marrow. J Lab Clin Med. 1996;127(4):353-8.

[48] Ellmerer M, Schaupp L, Brunner GA, Sendlhofer G, Wutte A, Wach P, et al. Measurement of interstitial albumin in human skeletal muscle and adipose tissue by open-flow microperfusion. Am J Physiol Endocrinol Metab. 2000;278(2):E352-6.

[49] Aitchison J, Shen SM. Logistic-Normal Distributions - Some Properties and Uses. Biometrika. 1980;67(2):261-72.

[50] Tsamandouras N, Wendling T, Rostami-Hodjegan A, Galetin A, Aarons L. Incorporation of stochastic variability in mechanistic population pharmacokinetic models: handling the physiological constraints using normal transformations. J Pharmacokinet Pharmacodyn. 2015;42(4):349-73.

[51] Berezhkovskiy LM. Determination of volume of distribution at steady state with complete consideration of the kinetics of protein and tissue binding in linear pharmacokinetics. J Pharm Sci. 2004;93(2):364-74.

[52] Peyret T, Poulin P, Krishnan K. A unified algorithm for predicting partition coefficients for PBPK modeling of drugs and environmental chemicals. Toxicol Appl Pharmacol. 2010;249(3):197-207.

[53] Poulin P, Haddad S. Advancing prediction of tissue distribution and volume of distribution of highly lipophilic compounds from a simplified tissue-composition-based model as a mechanistic animal alternative method. J Pharm Sci. 2012;101(6):2250-61.

[54] Poulin P, Theil FP. Prediction of pharmacokinetics prior to in vivo studies. 1. Mechanism-based prediction of volume of distribution. J Pharm Sci. 2002;91(1):129-56.

[55] Assmus F, Houston JB, Galetin A. Incorporation of lysosomal sequestration in the mechanistic model for prediction of tissue distribution of basic drugs. Eur J Pharm Sci. 2017;109:419-30.

[56] Holt K, Ye M, Nagar S, Korzekwa KR. Prediction of tissue - plasma partition coefficients using microsomal partitioning: Incorporation into physiologically-based pharmacokinetic models and steady state volume of distribution predictions. Drug Metab Dispos. 2019.

[57] Davis AM, Webborn PJ, Salt DW. Robust assessment of statistical significance in the use of unbound/intrinsic pharmacokinetic parameters in quantitative structure-pharmacokinetic relationships with lipophilicity. Drug Metab Dispos. 2000;28(2):103-6.

[58] Laruelle M, Slifstein M, Huang Y. Relationships between radiotracer properties and image quality in molecular imaging of the brain with positron emission tomography. Mol Imaging Biol. 2003;5(6):363-75.

[59] Laznicek M, Kvetina J, Mazak J, Krch V. Plasma protein binding-lipophilicity relationships: interspecies comparison of some organic acids. J Pharm Pharmacol. 1987;39(2):79-83.

[60] Lobell M, Sivarajah V. In silico prediction of aqueous solubility, human plasma protein binding and volume of distribution of compounds from calculated pKa and AlogP98 values. Mol Divers. 2003;7(1):69-87.

[61] van de Waterbeemd H, Smith DA, Jones BC. Lipophilicity in PK design: methyl, ethyl, futile. J Comput Aided Mol Des. 2001;15(3):273-86.

[62] Kratochwil NA, Huber W, Muller F, Kansy M, Gerber PR. Predicting plasma protein binding of drugs: a new approach. Biochem Pharmacol. 2002;64(9):1355-74.

[63] Liu JZ, Yang L, Li Y, Pan DH, Hopfinger AJ. Constructing plasma protein binding model based on a combination of cluster analysis and 4D-fingerprint molecular similarity analyses. Bioorgan Med Chem. 2006;14(3):611-21.

[64] Saiakhov RD, Stefan LR, Klopman G. Multiple computer-automated structure evaluation model of the plasma protein binding affinity of diverse drugs. Perspect Drug Discov. 2000;19(1):133-55.

[65] Watanabe R, Esaki T, Kawashima H, Natsume-Kitatani Y, Nagao C, Ohashi R, et al. Predicting Fraction Unbound in Human Plasma from Chemical Structure: Improved Accuracy in the Low Value Ranges. Mol Pharm. 2018;15(11):5302-11.

[66] Gleeson MP. Plasma protein binding affinity and its relationship to molecular structure: an in-silico analysis. J Med Chem. 2007;50(1):101-12.

[67] Plante J, Werner S. JPlogP: an improved logP predictor trained using predicted data. J Cheminformatics. 2018;10.

[68] Doki K, Darwich AS, Achour B, Tornio A, Backman JT, Rostami-Hodjegan A. Implications of intercorrelation between hepatic CYP3A4-CYP2C8 enzymes for the evaluation of drug-drug interactions: a case study with repaglinide. Br J Clin Pharmacol. 2018;84(5):972-86.

[69] Melillo N, Darwich AS, Magni P, Rostami-Hodjegan A. Accounting for inter-correlation between enzyme abundance: a simulation study to assess implications on global sensitivity analysis within physiologically-based pharmacokinetics. J Pharmacokinet Pharmacodyn. 2019.

[70] Vozeh S, Schmidlin O, Taeschner W. Pharmacokinetic drug data. Clin Pharmacokinet. 1988;15(4):254-82.

[71] Kalvass JC, Phipps C, Jenkins GJ, Stuart P, Zhang X, Heinle L, et al. Mathematical and Experimental Validation of Flux Dialysis Method: An Improved Approach to Measure Unbound Fraction for Compounds with High Protein Binding and Other Challenging Properties. Drug Metab Dispos. 2018;46(4):458-69.

[72] Da Veiga S, Wahl F, Gamboa F. Local Polynomial Estimation for Sensitivity Analysis on Models With Correlated Inputs. Technometrics. 2009;51(4):452-63.

[73] Saltelli A, Tarantola S. On the relative importance of input factors in mathematical models: Safety assessment for nuclear waste disposal. Journal of the American Statistical Association. 2002;97(459):702-9.

[74] Gray GM, Yardley HJ. Lipid compositions of cells isolated from pig, human, and rat epidermis. J Lipid Res. 1975;16(6):434-40.

[75] Ruark CD, Hack CE, Robinson PJ, Mahle DA, Gearhart JM. Predicting passive and active tissue:plasma partition coefficients: interindividual and interspecies variability. J Pharm Sci. 2014;103(7):2189-98.

[76] Kuksis A. Fatty Acid Composition of Glycerolipids of Animal Tissues. In: Kuksis A, editor. Fatty Acids and Glycerides. Boston, MA: Springer US; 1978. p. 381-442.

[77] Ulloa JL, Stahl S, Yates J, Woodhouse N, Kenna JG, Jones HB, et al. Assessment of gadoxetate DCE-MRI as a biomarker of hepatobiliary transporter inhibition. NMR Biomed. 2013;26(10):1258-70.

[78] Bergstrom M, Grahnen A, Langstrom B. Positron emission tomography microdosing: a new concept with application in tracer and early clinical drug development. Eur J Clin Pharmacol. 2003;59(5-6):357-66.

[79] Chang YJ, Chang CH, Chang TJ, Yu CY, Chen LC, Jan ML, et al. Biodistribution, pharmacokinetics and microSPECT/CT imaging of 188Re-bMEDA-liposome in a C26 murine colon carcinoma solid tumor animal model. Anticancer Res. 2007;27(4B):2217-25.

[80] Holt K, Nagar S, Korzekwa K. Methods to Predict Volume of Distribution. Current Pharmacology Reports. 2019;5(5):391-9.

[81] Korzekwa K, Nagar S. On the Nature of Physiologically-Based Pharmacokinetic Models -A Priori or A Posteriori? Mechanistic or Empirical? Pharm Res. 2017;34(3):529-34.

Chapter 3: Investigation of simplified physiologically based pharmacokinetic (PBPK) models in rat and human

3.1. Abstract

3.1.1. Background

Whole-body physiologically based pharmacokinetic (PBPK) models have many applications in academic and pharmaceutical research and drug development. It is often necessary to adjust these models with emerging animal or clinical data, improving model parameters and making the model more predictive for future applications. This provides an opportunity as well as a challenge given the large number of parameters in such models. The aim of this work was to propose new mechanistic models of drug distribution that reduce the complexity of the PBPK model structure and/or the number of parameters for optimization. These models are then evaluated for the ability to estimate physiologically relevant values for unbound tissue to plasma partition coefficients (Kpu) and to simulate observed pharmacokinetic (PK) data.

3.1.2. Methods

Two approaches are being proposed for this purpose. Firstly, using established kinetic lumping methods based on tissue time constants and secondly, using clustering analysis on Kpus to identify tissues sharing common Kpu values or Kpu scalars (i.e., scaling factors) based on similarities of tissue composition parameters while maintaining the PBPK model structure. PBPK models derived from these approaches were assessed using rat and human PK data of diazepam. Model performance was compared based on physiological plausibility, visual and numerical predictive checks.

3.1.3. Results

Several models using either of the two approaches were found to have highly similar abilities to describe IV data compared to empirical models (reasonable fits, good precision). These included a model with 3 kinetically lumped compartments and several PBPK models with 3 or 4 common Kpus or scalars (different tissue groupings possible). Although the clustering analyses produced minor differences in grouping of tissues depending on the clustering method, kidney and liver were generally grouped together whereas bone, brain, muscle, pancreas were found more similar. Clustering into 4 tissue groups appeared more physiologically relevant in terms of tissue composition, with adipose as a separate group due to its particular composition. For diazepam, these models described the rat concentration-time profiles well and allowed estimation of physiologically relevant partitioning coefficients. Additionally, these mechanistic models produced Kpu estimates that were comparable to experimental Kpu values for diazepam.

While the kinetically lumped model is useful for compounds with low tissue to plasma partitioning and when data are available in the species of interest, the PBPK models with common scalars show more promise for data-driven PK studies for a wider set of drugs and for interspecies translation.

3.1.4. Conclusions

A new method of combining tissues (*a priori* lumping), based on their constituents has been proposed and assessed. For diazepam, the best model (generally, the model with scalars using k-means clustering) captured the plasma PK profile well and the predictions of tissue concentrations were consistent with available measurements in rat, suggesting the potential use of these models for inferring tissue distribution based on plasma concentration-time data alone. These methods enable PBPK translation from preclinical species to human (see Chapter 4).

3.2. Introduction

Pharmacokinetic (PK) models describe the drug concentration within the body as a function of time. They cover models with various degree of complexity, from simple empirical models such as non-compartmental and parsimonious compartmental models, to semi-mechanistic and more complex whole-body physiologically based pharmacokinetic (PBPK) models [1]. Empirical models are frequently built based only on the observed PK data without the need for prior knowledge of the system. These models can describe available data well but may have limited utility due to a lack of physiological mechanisms and consequently lack extrapolative power. On the other hand, whole-body PBPK models are particularly suited to integrate knowledge of different origins (e.g., *in silico*, *in vitro*, and *in vivo*) in an anatomically and physiologically relevant framework. PBPK modelling provides a powerful tool for integrating preclinical data into human PK predictions [2, 3]. Although PBPK models become more and more informed as knowledge about the drug and the system increases, numerous and often untested assumptions are inherent to these models. When observed PK data become available certain PBPK model parameters can be optimized by combining the bottom-up and top-down approaches [4]. Yet, the high dimensionality and the complexity of PBPK models, as well as the limited amount of data available (e.g., plasma or blood observations only and relatively sparse sampling) may limit the simultaneous estimation of the large number of parameters due to computational and numerical issues, as well as kinetic processes being missed as they unfold too rapidly. In order to stabilise PBPK parameter estimation, many parameters are fixed to typical values and then only a few selected parameters are estimated through mathematical computations, which could result in biased estimates. In addition, the decision on which parameters to fix and which to estimate is often subjective and therefore the final model and model parameters may vary significantly between different modellers [4, 5]. Indeed, each modeller visualises a model from a different perspective, each one of these models may be valid to describe the data, and equally none of them may be correct. Consequently, PBPK models are not often used for data-driven PK analysis compared to empirical compartmental models.

Several lumping approaches have been proposed to reduce the dimensionality and complexity of whole body PBPK models by aggregating model states (here, tissue compartments) or parameters. Formal or proper lumping is a method of model reduction, where each state of the original model contributes to only one state of the reduced system. Nestorov et al. were the first to incorporate lumping principles in PBPK modelling,

consisting on lumping tissues with identical structural positions (serial or parallel connection) and similar kinetic properties (time-constants), and applied them for barbiturates in rats [6]. The reduced model with lumped compartments retains the kinetic behaviour of the original model, and the associated parameters maintain their meaning. An algorithm for proper lumping was proposed by Dokoumetzidis and Aarons, which is automatic and can be applied to PBPK models [7, 8]. Similarly, Pilari and Huisinga proposed an algorithm for lumping tissues with similar kinetics (normalized concentration-time profiles) to obtain simplified perfusion and permeability rate limited PBPK models, which was applied to 25 small molecules [9]. Other examples of lumping compartments to simplify PBPK models exist in the literature [10-12]. Additional model reduction approaches, such as global sensitivity analysis and balanced truncation, have been suggested for reducing PBPK models [13, 14]. However, kinetic lumping approaches have to typically be tailor-made for each new compound as they are generally only valid locally for a specific set of parameter values [15].

The limitation of drug-dependence in model reduction might be overcome by using a general simplified PBPK model. A few generalized minimal PBPK models have been proposed as simplification of whole-body PBPK for estimation of physiologically relevant PK parameters without tissue concentration data available and for interspecies extrapolation but showed some limitations and did not include mixed effects [16-18]. Using lumping principles, Arundel proposed a multi-compartmental approach with 2-blood compartments and 6-tissue groups which are each characterised by its time constant defined as the disappearance rate from a tissue [17]. It was found that for each lumped tissue group, the product of tissue time constant and volume of distribution at steady state (Vss) was relatively constant except for adipose tissue for a range of 10 structurally diverse compounds. Therefore, based on the Vss obtained from plasma concentrations, tissue time constants and consequently tissue partition coefficients (Kp) can be estimated. This approach, however, cannot be applied for the adipose Kp. Cao and Jusko proposed minimal-PBPK models with two- or three tissue compartments and applied them for capturing the blood (or plasma) profiles of 27 drugs from four different therapeutic classes [16]. In these hybrid models with properties in between whole-body PBPK and compartmental PK models, venous and arterial blood compartments are separated, and tissues are lumped separately. Physiological restrictions are integrated on blood and tissue volumes and on fractions of cardiac output. However, the lumped tissues remain

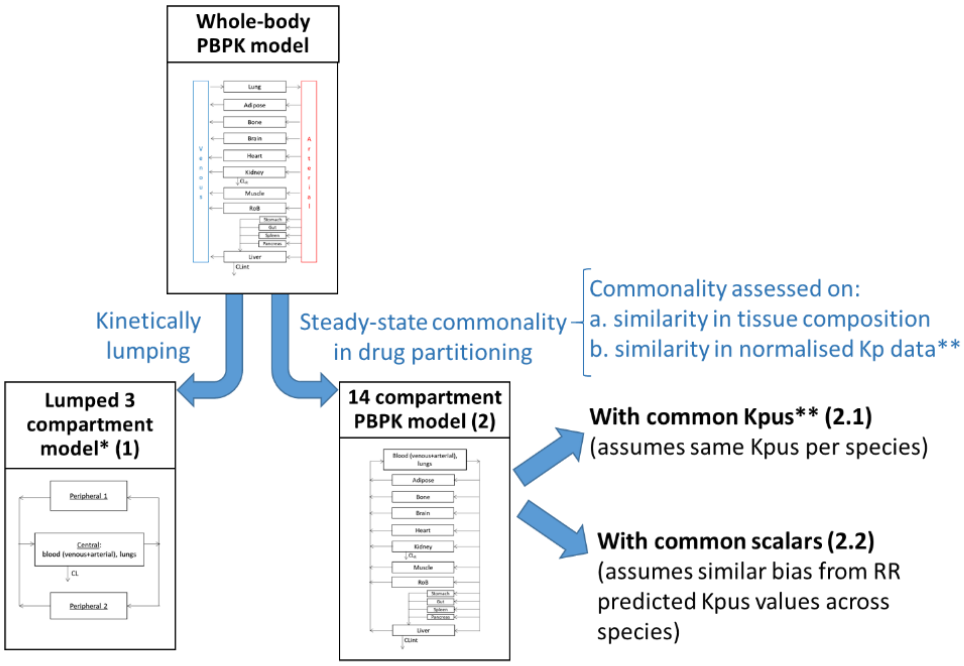
empirical tissue compartments and independent Kps cannot always be identified for each tissue.

The purpose of this work was to investigate generalised approaches for the simplification of PBPK models that can be easily used for PK studies, translation of PK properties from animals to humans and even population PK analysis. The proposed models should be mechanistic allowing separation of system- and drug-specific parameters and the model parameters can be estimated without tissue concentration data. This approach could thus allow the optimisation of a PBPK model when Vss is not well predicted from a bottom-up approach with Rodgers and Rowland (R&R) model. An evaluation of these simplified PBPK models will be shown for diazepam. These models could be applied for the prediction of human PK from preclinical data. The approach with these models retains physiological interpretability and can be performed a priori, e.g., in the absence of data in human.

3.3. Methods

In this section, PBPK models are classified into 3 categories, as shown in Figure 3.1: (1) Whole-body PBPK model, (2) kinetically lumped model with 3 compartments, (3) kinetically lumped model with 14 compartment and common physiology. (2) and (3) are simplified PBPK models proposed in order to limit the number of parameters for estimation and yet maintaining physiological aspects of the whole-body PBPK model. They are two clearly distinct approaches: kinetically lumping reduces the complexity of the model from a mathematical point of view; whereas in contrast, steady-state commonality in drug partitioning, reduces the number of parameters needed in the complex model.

Figure 3.1: Approaches investigated for simplifying a whole-PBPK model



*different lumped model per species unless assuming they have the same model structure
**applicable to rat and to human only under the assumption that Kpus are the same between species

3.3.1. Whole-body PBPK model

A typical whole-body PBPK model for small molecules used for animals to human PK predictions is shown in **Figure 3.1** [6, 19, 20]. In this model, compartments represent individual organs or tissues of the body, connected via the systemic blood circulation. Commonly, there are two blood compartments (arterial and venous) and 14 tissue compartments (lungs, heart, kidneys, bone, muscle, brain, adipose, skin, spleen, pancreas, liver, stomach, gut, bones and rest of body), which makes 16 states in total.

Assuming tissues are perfusion-limited (well-stirred) compartments, each non-eliminating tissue can be described by the following equations (Eq. 3.1-Eq. 3.2):

Lung $$V_{lung} \cdot \frac{dC_{lung}}{dt} = Q_{lung} \cdot \left(\mathrm{C_{venous}} - \frac{C_{lung}}{Kb_{lung}} \right) \quad \textbf{Eq. 3.1}$$

Other tissues $$V_i \cdot \frac{dC_i}{dt} = Q_i \cdot \left(\mathrm{C_{arterial}} - \frac{C_i}{Kb_i} \right) \quad \textbf{Eq. 3.2}$$

where C_{lung}, C_i , C_{venous} and $C_{arterial}$ are the total drug blood concentrations (mg/L) in the lung, the i^{th}-tissue, the influent venous blood and arterial blood, respectively. V_{lung}, Q_{lung}, V_i and Q_i are the volumes (L) and blood flows (L/min) for the lung and each i^{th}-tissue, respectively. Kb_{lung} and Kb_i are the tissue-to-blood partition coefficient in the lung and in the i^{th} tissue, and represent the tissue to venous blood concentration ratio at steady state.

When renal elimination (CL_R) is part of the venous compartment, the rate equations for the arterial blood and venous blood compartments are defined as follows (Eq. 3.3-Eq. 3.4):

$$V_{arterial} \cdot \frac{dC_{arterial}}{dt} = Q_{lung} \cdot \left(\frac{C_{lung}}{Kb_{lung}} - \mathrm{C_{arterial}} \right) \quad \textbf{Eq. 3.3}$$

$$V_{venous} \cdot \frac{dC_{venous}}{dt} = \sum Q_i \cdot \frac{C_i}{Kb_i} - Q_{lung} \cdot C_{venous} - \mathrm{CL_R} \cdot C_{venous} \quad \textbf{Eq. 3.4}$$

where $\sum Q_i \cdot \frac{C_i}{Kb_i}$ includes all the i^{th} tissues except the stomach, gut, pancreas and spleen; $V_{arterial}$ and V_{venous} are the volume of arterial and venous blood, respectively; CL_R is calculated as the fraction excreted (fe) of the total blood clearance (L/min) and values of fe are reported in **Table 3.2**. Plasma concentrations can be derived by dividing C_{venous} by the blood-to-plasma ratio, BP.

For the liver, the rate equation is defined as (Eq. 3.5):

$$V_{liver} \cdot \frac{dC_{liver}}{dt} = Q_{ha} \cdot \mathrm{C_{arterial}} + \sum Q_{splan,i} \cdot \frac{C_{splan,i}}{Kb_{splan,i}} - Q_{liver} \cdot \frac{C_{liver}}{Kb_{liver}} - CLint_{liver} \cdot fu_b \cdot \frac{C_{liver}}{Kb_{liver}} \quad \textbf{Eq. 3.5}$$

where the $Q_{splan,i}$, $C_{splan,i}$ and $Kb_{splan,i}$ are the concentration, the blood flow, the volume and the blood partition coefficient of the i^{th} splanchnic organs (stomach, gut, pancreas and spleen); Q_{ha} is the blood flow from the hepatic artery; C_{liver}, Q_{liver}, V_{liver}, Kb_{liver} are the concentration, the blood flow, the volume and the blood partition coefficient of the liver; CLint is the hepatic intrinsic clearance (L/min) and fu_b is the fraction unbound in blood. In addition, assuming the liver is represented by a well-stirred model, the blood hepatic clearance (CL_H) can be related to its intrinsic clearance and extraction ratio (ER_H) [21]:

$$CL_H = \frac{Q_{liver} \cdot fu_b \cdot CLint_{liver}}{Q_{liver} + fu_b \cdot CLint_{liver}} = Q_{liver} \cdot ER_H \quad \textbf{Eq. 3.6}$$

Kb values are a function of drug and species-specific parameters. Kb is determined from Kpu, the tissue-to-unbound plasma partition coefficient (Eq. 3.7). Kpu values are key components for the characterization of the rate and extent of drug distribution into different tissues in the body and reflect the degree of tissue distribution attributed to processes such as protein binding, lipid interaction, lysosomal trapping, etc. [22-27].

$$K\mathrm{b} = Kp\mathrm{u} \cdot \frac{fu_p}{BP} \quad \textbf{Eq. 3.7}$$

where fu_p and BP are the fraction unbound in plasma and the blood-to-plasma ratio, respectively.

The Rodgers and Rowland (R&R) model was selected as it was demonstrated to be the most accurate in a study comparing multiple tissue:plasma partition coefficient prediction methods [28]. Rodgers et al. proposed two mechanistic equations for predicting Kpus: one for moderate-to-strong bases and group 1 zwitterions (at least one basic pKa>7) [23], and another for acids, very weak bases, neutrals and group 2 zwitterions (no basic pKa >7) [24].

Tissue blood flows and volumes are species-specific parameters. Blood flows and tissue volumes were calculated for a standard 70-kg man and for a standard 250-g-rat (Table A2.1). The volume of distribution at steady state based on whole blood (L) can be calculated as (Eq. 3.8):

$$Vss, b = V_{venous} + V_{arterial} + \sum Kb_i \cdot V_i \cdot (1 - ER_i) \quad \textbf{Eq. 3.8}$$

where V_i is the volume of the i^{th}-tissue, and Kb_i and ER_i are respectively its tissue-to-blood partitioning coefficient and extraction ratio.

3.3.2. Lumped PBPK model with 3 compartments

For an extensive description of lumping principles adopted in this section, the reader is referred to a previous publication [6]. The main rule is that only tissues with identical model specifications (i.e., connected in parallel or in series) and with similar time constants, can be grouped together. Time constants (T) are defined as follows [6] (Eq. 3.9-Eq. 3.11):

Eliminating tissue (i.e., liver)
$$T_i = \frac{V_i \cdot Kb_i}{Q_i + fu_b \cdot CL_{int,i}} \quad \textbf{Eq. 3.9}$$

Non-eliminating tissue
$$T_i = \frac{V_i \cdot Kb_i}{Q_i} \quad \textbf{Eq. 3.10}$$

Arterial and venous blood
$$T_i = \frac{V_i}{Q_i} \quad \textbf{Eq. 3.11}$$

To lump serial tissues (splanchnic and liver), they should have low time constants, i.e., equilibrate very rapidly with each other (rapid equilibration condition). For example, the lungs, venous and arterial compartments could be lumped as one ‘central’

compartment since they are connected serially and have similar low time constants (Table **3.1**). This assumption holds in many cases except if lungs depict a deep compartment for a particular compound (i.e., high Kb_{lung} value) or if sampling in very early time events are of interest [6]. Therefore, this transformation illustrated in Eq. 3.12 was applied for all the models proposed in this work.

Table 3.1: Tissue time constants for a reference man (70kg) and rat (250g)[29-33]

Tissues	Human			Rat		
	Time constant * (min)	Rank **	Compar-tment	Time constant * (min)	Rank **	Compar-tment
Adipose	$53.5*Kb_{adipose}$	16	3	$3.40*Kb_{adipose}$		2
Bone	$17.8*Kb_{bone}$	14	2	$0.53*Kb_{bone}$	7	1
Brain	$1.92*Kb_{brain}$	10	2	$0.83*Kb_{brain}$	11	2
Gut	$1.92*Kb_{gut}$	9	2	$0.59*Kb_{gut}$	9	1
Heart	$1.35*Kb_{heart}$	7	2	$0.20*Kb_{heart}$	5	2
Kidney	$0.296*Kb_{kidney}/(1.11+fub*Clint_{kidney})$	3	2	$1.76*Kb_{kidney}/(11.72+fub*Clint_{kidney})$	3	2
Liver	$1.67*Kb_{liver}/(1.49+fub*Clint_{liver})$	6	2	$8.47*Kb_{liver}/(12.55+fub*Clint_{liver})$	10	2
Lung	$0.088*Kb_{lung}$	1	1	$0.014*Kb_{lung}$	1	1
Muscle	$27.1*Kb_{muscle}$	15	2	$4.21*Kb_{muscle}$	14	2
Pancreas	$1.61*Kb_{pancreas}$	8	2	$0.51*Kb_{pancreas}$	6	2
Skin	$8.55*Kb_{skin}$	13	2	$9.49*Kb_{skin}$	16	3
Spleen	$1.00*Kb_{spleen}$	5	2	$0.58*Kb_{spleen}$	8	2
Stomach	$2.42*Kb_{stomach}$	11	2	$1.02*Kb_{stomach}$	12	2
RoB	$3.99*Kb_{RoB}$	12	2	$3.87*Kb_{RoB}$	13	2
Arterial	0.228	2	1	0.079	2	1
Venous	0.683	4	1	0.157	4	1

** The term time constant is defined in Eq. 3.9-Eq. 3.11*
*** Ranking of time constant from lowest to highest value and assuming Kb=1 and Clint=0*

To lump parallel tissues, they should have similar time constants (similarity conditions). As a rule of thumb, larger tissues with time constants not differing by more than 50-60% from each other may be considered as kinetically equivalent and can be lumped together, while lumping of smaller tissues can be extended to an order of magnitude [6]. A three-compartment model was chosen for the structure of the lumped model in order to be simply adapted and reduce computation times in PK modelling software. This model specification implies a rapidly equilibrating (central), a moderately equilibrating (peripheral 1), and a slowly equilibrating (peripheral 2) compartment. Tissues were attributed to one of the latter two compartments (rapidly-to-moderately or slowly equilibrating) depending on the percentage difference and magnitude of the tissue time constants (Table **3.1**). A closed form solution can be derived (see Appendix A2.2) facilitating and expanding the use of this model with computationally intensive method such as Bayesian estimation methods. This model was implemented with the ADVAN11

subroutine and TRANS1 parameterization in NONMEM 7.3 (ICON Development Solutions, Hanover, MD, USA) [34].

In the situation where elimination is considered part of the central compartment (arterial, venous blood and lungs), the rate equation is defined as (Eq. 3.12):

$$V_{central}\frac{dC_{central}}{dt} = \frac{Q_{p1}\cdot \mathrm{C_{p1}}}{Kb_{p1}} + \frac{Q_{p2}\cdot \mathrm{C_{p2}}}{Kb_{p2}} - Q_{central}\cdot\frac{C_{central}}{Kb_{central}} - CL_{blood}\cdot\frac{C_{central}}{Kb_{central}} \quad \textbf{Eq. 3.12}$$

Where $C_{central}$, C_{p1}, C_{p2} are the venous blood total drug concentration in the central, the peripheral 1 and peripheral 2 compartments, respectively. The blood flow $Q_{central}$ and the volume $V_{central}$ of the central compartment are defined as $Q_{central}$=cardiac output and $V_{central}= V_{arterial}+V_{venous}+ V_{lung}$; $Kb_{central}$ is defined as the blood tissue-to-plasma partitioning coefficient of the central compartment and corresponding to $(V_{arterial}+V_{venous}+V_{lung}\ Kb_{lung})/(V_{lung}+V_{arterial}+V_{venous})$; Q_{p1} and Kb_{p1} are the blood flow and blood tissue-to-plasma partitioning coefficient of all tissues in the lumped peripheral 1 compartment; Kb_{p1} is weighted by the following volume $(\Sigma V_i - V_{liver} - V_{kidney}) + V_{liver}\ (1-CL_H/Q_{liver}) + V_{kidney}\ (1-CL_R/Q_{renal}))/\Sigma V_i$ where i includes tissue lumped in the peripheral 1 compartment; Q_{p2} and Kb_{p2} are the blood flow and blood tissue-to-plasma partitioning coefficient of all the tissues in the lumped peripheral 2 compartment; CL_{blood} is the total blood clearance and corresponds here to the sum of the hepatic and renal clearance (CL_H and CL_R, respectively). It should be noted that the clearance is from the central compartment although it does not contain an eliminating organ, however this renders the lumped model similar to a model with 3 compartments where elimination is central. Additionally, individual concentrations profiles (venous, arterial and lung) can be derived as follows (Eq. 3.13-Eq. 3.14):

$$V_{arterial}\cdot\frac{dC_{arterial}}{dt} = V_{venous}\cdot\frac{dC_{venous}}{dt} = \frac{dC_{central}}{dt}\cdot\frac{1}{V_{arterial}+V_{venous}+\left(V_{lung}\cdot Kb_{central}\right)} \quad \textbf{Eq. 3.13}$$

$$V_{lung}\cdot\frac{dC_{lung}}{dt} = \frac{dC_{central}}{dt}\cdot\frac{1}{\frac{(V_{arterial}+V_{venous})}{Kb_{central}}+V_{lung}} \quad \textbf{Eq. 3.14}$$

And for each of the lumped peripheral compartments 1 and 2, the rate equation is defined as (Eq. 3.15-Eq. 3.16):

$$V_{p1} \cdot \frac{dC_{p1}}{dt} = Q_{p1} \cdot \left(\frac{C_{central}}{Kb_{central}} - \frac{C_{p1}}{Kb_{p1}} \right) \quad \textbf{Eq. 3.15}$$

$$V_{p2} \cdot \frac{dC_{p2}}{dt} = Q_{p2} \cdot \left(\frac{C_{central}}{Kb_{central}} - \frac{C_{p2}}{Kb_{p2}} \right) \quad \textbf{Eq. 3.16}$$

Where $Q_{p1} = \Sigma Q_i$ and $V_{p1} = \Sigma V_i$. Q_i and V_i are the blood flow and volume of the i^{th}-tissue lumped in the peripheral 1 compartment; $Q_{p2} = \Sigma Q_j$ and $V_{p2} = \Sigma V_j$. Q_j and V_j are the blood flow and volume of the j^{th}-tissue lumped in the peripheral 2 compartment.

The volume of distribution at steady state based on whole blood (Vss,b) can be calculated as (Eq. 3.17):

$$Vss, b = V_{central} + V_{p1} + V_{p2} \quad \textbf{Eq. 3.17}$$

3.3.3. PBPK model with common Kpus or common scalars

Compared to the previous lumped-PBPK model, the 14 compartmental PBPK model with common Kpus or scalars only makes the kinetic assumption that the lung, venous and arterial blood equilibrate quasi instantly which is valid in many cases unless the Kb_{lung} value of the particular drug is high or its very early time events are used [6]. The term 'PBPK model' will be used from now on to refer to the 14-compartmental PBPK model (model 2 in Figure 3.1). Derivation of this model into differential equations is provided in Appendix A2.2. This PBPK model is primarily based on steady state similarity in drug partitioning into tissues while individual tissue blood-flows and volumes are preserved. In the PBPK model with common Kpus (model 2.1 in Figure 3.1), similar tissues are assumed to share a similar value of Kpu. While in the PBPK model with common scalars (model 2.2 in Figure 3.1), tissues are assumed to have their Kpu values (e.g., predicted from R&R model) and share a similar tissue-scaling factor (referred to scalar) which is a drug-specific scaling factor (Eq. 3.18):

$$Kpu_i = Kpu_{predRR,i} \cdot SF \quad \textbf{Eq. 3.18}$$

Where Kpu_i is the true value of Kpu for tissue i whereas $Kpu_{predRR,i}$ is the Kpu predicted using the R&R model and SF is the scaling factor for the tissue i.

A global sensitivity analysis showed that tissues could behave similarly in terms of Kpu across different compounds with different properties, suggesting grouping of

correlated tissues [35]. For identifying groups of tissues that share commonality in steady-state drug partitioning, clustering analyses were performed. A clustering analysis aims to group observations of a given dataset into mutually exclusive groups (clusters). Therefore, observations within the same group are as similar as possible while observations from different clusters are as dissimilar as possible. Many clustering algorithms have been proposed and fully described in the literature [36, 37]. Among them, the k-means clustering and hierarchical clustering are the most widely used. Given the utility and limitations of each method (detailed in Appendix A2.3), both clustering methods were ultimately explored in the current analysis. Briefly, both methods cluster elements by minimizing the distance between elements and a centroid measure for the former [37-39], and by analysing the dissimilarity (or similarity) between each observation for the latter [40]. The tissue similarities were assessed using two different datasets (Figure 3.1): (a) tissue composition data (rat and human) or (b) normalised *in vivo* Kp values (rat data only). More details about the clustering analysis can be found in Appendix A2.3.

Clustering based on tissue composition data

Species tissue composition data were first standardized (z-score standardization) to identify clusters of observations with the same overall profiles regardless of their magnitude. Based on previous work [35], acid phospholipids and extracellular protein levels were found to be more influential on Kpu outputs predicted using the R&R model and consequently should be assigned a higher weight, so that it would influence the cluster formation more than other tissue component variables. To include this effect, each standardized variable was multiplied by a higher weight for acid phospholipids and tissue proteins (i.e., albumin and lipoproteins) than for other variables (a weighting factor of 2 and 1 was respectively considered after investigating several scenarios of weighting in tissue composition). The k-means and hierarchical clustering methods were applied to classify the tissues clusters using the R package ‘stats’ [41]. For the k-means clustering, the tissue Kpu were preferentially grouped into 3 or 4 clusters and the algorithm was executed 50 times with different initial centres for better stability and to avoid local solutions, and 10 iterations were typically sufficient for convergence. The hierarchical clustering was applied with Euclidean distance as the distance metric and Ward’s method as the linkage method, and 3 or 4 clusters were determined subsequently.

Clustering based on normalised rat steady-state Kp data

Similarly, a hierarchical clustering analysis was performed using available experimental Kp data for identifying groups of tissues that share similar Kp outcomes because of physiological similarities in tissues or as a consequence of multiple processes leading to similar results. However, this approach using experimental Kp data is not using the R&R model by including all mechanisms of tissue partitioning and binding (known and unknown) and not only the known mechanisms included in the R&R model. *In vivo* Kp values have been previously determined experimentally in rat and collected for various compounds [23, 24, 42, 43]. The dataset of rat Kps collected from the literature was composed of 107 compounds (71 strong bases, 9 weak bases, 21 acids and 6 neutrals), 48% of the Kp values are missing (Table A2.6). The missing Kp data in the different tissues ranged from 17% (muscle) to 85% (pancreas). Much research has been performed on the issue of missing data, but it is not in the scope of this work. To address the problem of clustering data with missing values, the method of multiple imputation combined with clustering was used [44, 45]. 100 imputed datasets were created, and each data set was then analysed by hierarchical clustering as described in the previous section (with 3 or 4 clusters). After adjusting for the labelling of cluster assignments, the tissues were assigned to clusters based on the most frequent cluster assignment per tissue among the 100 datasets. Multiple imputation was performed using the R package 'mice' [46]. Predictive mean matching is the default modelling method in the mice package and was used for imputation [47]. The choice of m=100 iterations was adopted [45]. Imputed Kp values were normalised by Vss to avoid compound bias in the analysis which was tested, and no clear trend existed between normalised Kps and compound specific properties (pKa, BP, fup, LogP). Alternatively, hierarchical clustering could have been used directly as it has the ability to cluster even when missing values are present by still obtaining a full distance matrix. However, this can lead to some information bias as one distance between two observations is calculated based on many variables while another distance between two other observations may be based on only a few variables. Additionally, it was found that the classification by hierarchical clustering with missing values was not accurate with 48% of values missing for this type of multivariate data even when full information is given [44] and consequently, this method was not used.

3.3.4. Diazepam data

The models were fitted to diazepam PK data in rat and human. Individual and average PK profiles following intravenous administration were available in humans (N=35 profiles) and in rat (N=6 profiles). Concentration-time data in plasma, blood or serum were digitized using WebPlotDigitizer (version 4.2, https://automeris.io/WebPlotDigitizer). The majority of the PK profiles represent mean profiles except for a few studies which reported individual profiles (details can be found in **Table A2.4** and **Table A2.5**). Physicochemical and *in vitro* PK data for diazepam are summarized in Table 3.2.

Table 3.2: Physicochemical properties and *in vitro* PK data of diazepam

LogP	pKa	Human			Rat		
		fup	BP	fe	fup	BP	fe
2.82[a]	3.4[a]	0.009[e] (0.014-0.032)[d]	0.559[e] (0.65)[d]	0.0005[b]	0.1[e] (0.03-0.15)[d]	0.836[e] (1-1.19)[d]	0.009[c]

pKa: acid dissociation constant; LogP: n-octanol/water partition coefficient; fup: fraction unbound in plasma; BP: blood-plasma ratio; fe: urinary excretion
[a] *From* [30]; [b] *From* [31]*;* [c] *From* [32]; [d] *From literature [48-53]*
[e] *fup and BP were measured internally (details can be found in Appendix A2.4).*

3.3.5. Data analysis

The suitability of the proposed models was explored: (1) to be able to fit data, (2) to produce meaningful Kpu values (which can be compared in rat as experimental data are available) and (3) to provide a suitable model structure for translation as a potential application.

The investigated models were fitted to the human and rat PK data and their model parameter values were estimated using the first order conditional estimation with interaction (FOCE-I) method as implemented in NONMEM v7.3. The FOCE-I method allows for interaction between inter-individual variability (IIV) and residual variability, and it is fast to converge for simple structured models [54, 55]. The following general model was fitted to the data (Eq. 3.19):

$$Y_{ij} = f(\theta_i, t_{ij}) \cdot (1 + \varepsilon_{ij}) \qquad \textbf{Eq. 3.19}$$

where Y_{ij} is the observed data for study i at time t_{ij}, f is the structural model, which is identical for all the individual studies, θ_i is the vector of p individual PK parameters for the individual i, and ε_{ij} is the residual error. An exponential model was used to account for IIV. Inter-study variability (ISV) is confounded with IIV due to the pooling of studies and may over-estimate IIV if ISV is high [56]. For the k^{th} parameter : $\theta_{ik} = \mu \cdot \exp(\eta_{ik})$

where μ is a vector of fixed effects parameters representing the typical population parameter values and η_i is a vector of random effects for individual study i. The random effects are assumed to follow a multinormal distribution with mean 0 and variance Ω , N(0, Ω). Ω is the p × p variance-covariance matrix of the inter- individual variability. The residual variability was modelled using a proportional error model and ε_{ij} is assumed to follow a normal distribution with mean 0 and variance σ^2, N(0, σ^2).

Models were evaluated using standard goodness-of-fit (GOF) diagnostics, numerical evaluation and decrease in objective function value (OFV, approximated by −2*Log Likelihood) [57]. Model comparison was made using the Bayesian information criterion (BIC) due to changes in the structural model. Biological plausibility of parameter estimates was also evaluated and checked with consistency to existing data, in particular rat Kpu values of diazepam available from the literature (Table 3.3). Following an analysis based on the large dataset of rat Kp where distributions of Kpu prediction success using R&R model were characterized (not shown), it was found that 95% R&R Kpu predictions would be within approximately 21-fold error. As a result, a 25-fold under/over predictions can be considered unlikely given the available data.

Table 3.3: Rat Kpu values of diazepam

Tissue	*In vivo* Kpu*	Kpu predicted by RR
Lung	29.93	28.17
Splanchnic	22.67	53.77
Stomach	31.93	53.77
Pancreas	NA	35.88
Liver	57.27	38.32
Bone	NA	60.69
Brain	13.53	51.06
Heart	35.87	29.17
Kidney	31.07	32.14
Skin	19.67	52.5
Muscle	24.47	23.81
Adipose	140	55.99
RoB	24.47	23.81

**Kpu values were calculated after adjusting the Kps by the fup and BP reported in the study [58]*
Rob: rest of body

A common study design was simulated using the R package 'RxODE' [59] allowing the comparison of the performance of the different models since reported studies

have different designs. Observations in human and in rat were simulated from a reference model (structure, parameter estimates, and associated uncertainty) which was the compartmental model (two compartments in human and three compartments in rat) fitted to describe the diazepam data. For each species, 1,000 drug concentration-time courses were simulated for a single IV dose of diazepam and infusion rate to achieve a steady state plasma concentration determined by the drug elimination rate. A 16.1h infusion of 10 mg in man and a 4.3h infusion of 1mg in rat were chosen to best illustrate the different kinetic phases. The predictions using the different investigated models (structure, parameter estimates, and associated uncertainty but the clearance value was fixed to the one from the fitted empirical model) were superimposed onto the simulated observations to obtain a visual display of the investigated models' ability to describe the data. The median Vss,b predicted from the investigated models (Eq. 3.8 and Eq. 3.17) were compared against the median Vss,b simulated 'observed' from the fitted empirical model, and median concentration profiles were compared by calculating the root-mean-square error (RMSE, Eq. 3.20) to assess the precision of the predictions, with lower RMSE value representing greater precision of the model:

$$RMSE = \sqrt{\frac{\sum(PRED_t - OBS_t)^2}{n}} \qquad \textbf{Eq. 3.20}$$

Where OBS_t is the median 'observed' concentration value at time t and $PRED_t$ is the median simulated concentration at time t, and n is the sample size.

3.4. Results

3.4.1. Lumped PBPK model with 3 compartments

The first approach for simplifying PBPK models was mainly to lump tissue compartments with similar kinetics. The time constants of the tissue compartments and their ranking (with Kb=1 and CLint=0) are shown in Table **3.1**. In rat, rest of body, muscle and adipose were considered large tissues (>10% body volume) whereas only muscle and adipose were in human. In general, the tissue time constant is low when the volume is large or when the blood flow is small. The tissues were classified into three classes, according to the values of their time constants:

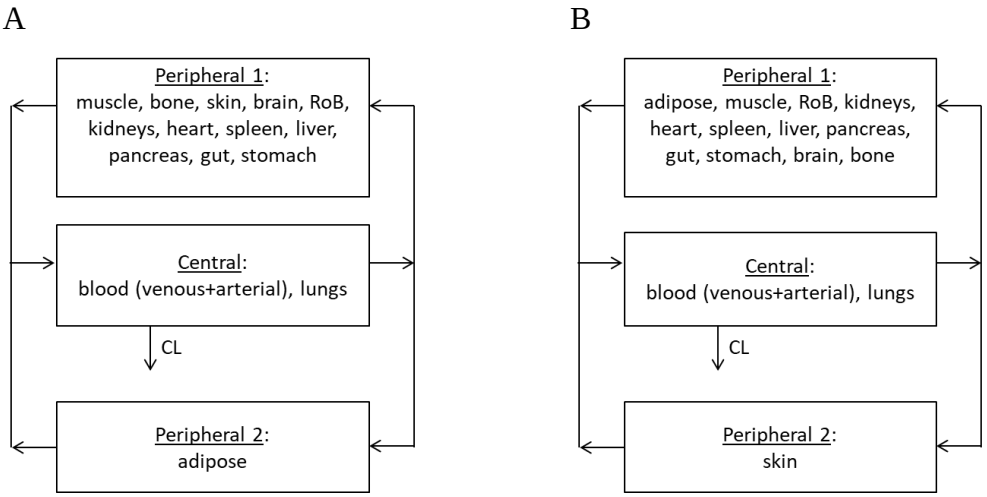
(i) the quasi-instantaneously or very rapidly equilibrating tissues with a time constant virtually equal to zero – lung in human and rat and arterial and venous blood (arterial blood and lung will behave similarly for IV injections);

(ii) the rapidly-moderately equilibrating tissues with very small to medium time constants – kidney, splanchnic organs, liver, brain, rest of body, skin, bone and muscle in man, or kidney, heart, splanchnic organs, bone, liver, brain, adipose, rest of body and muscle in rat;

(iii) the slowly equilibrating tissues with large time constants –adipose in human or skin in rat. This could be due to the physiological difference in the volume of adipose in rat which is considerably smaller than in human.

The lumped PBPK models in human and rat with one central and two peripheral compartments are illustrated in Figure 3.2.

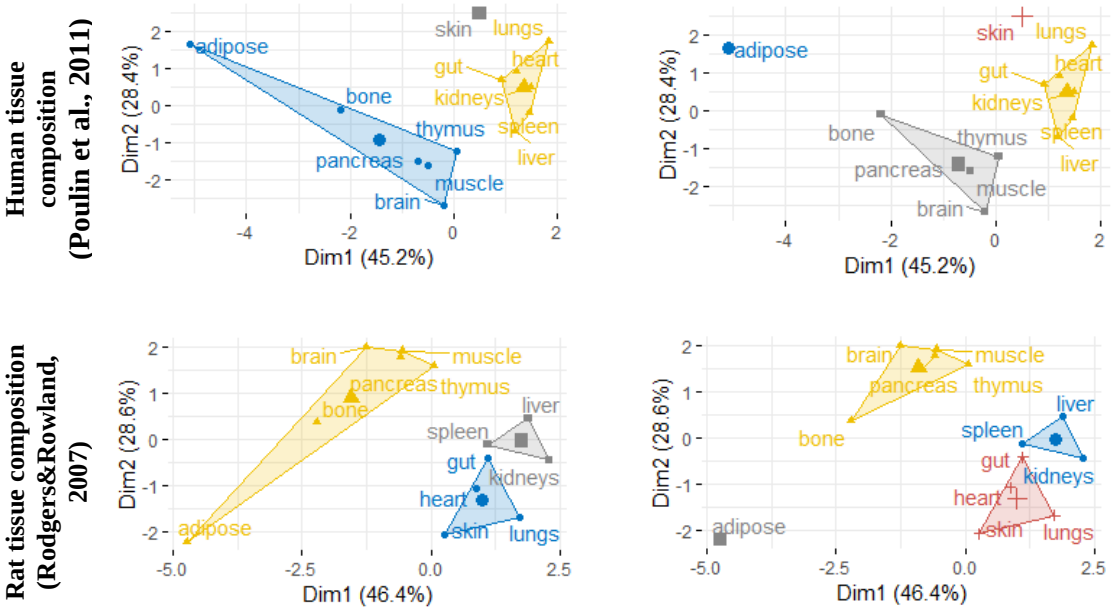
Figure 3.2 : Schematic representation of the lumped 3compartment-model in human (A) and in rat (B)



3.4.2. PBPK model with common Kpu values or common scalars

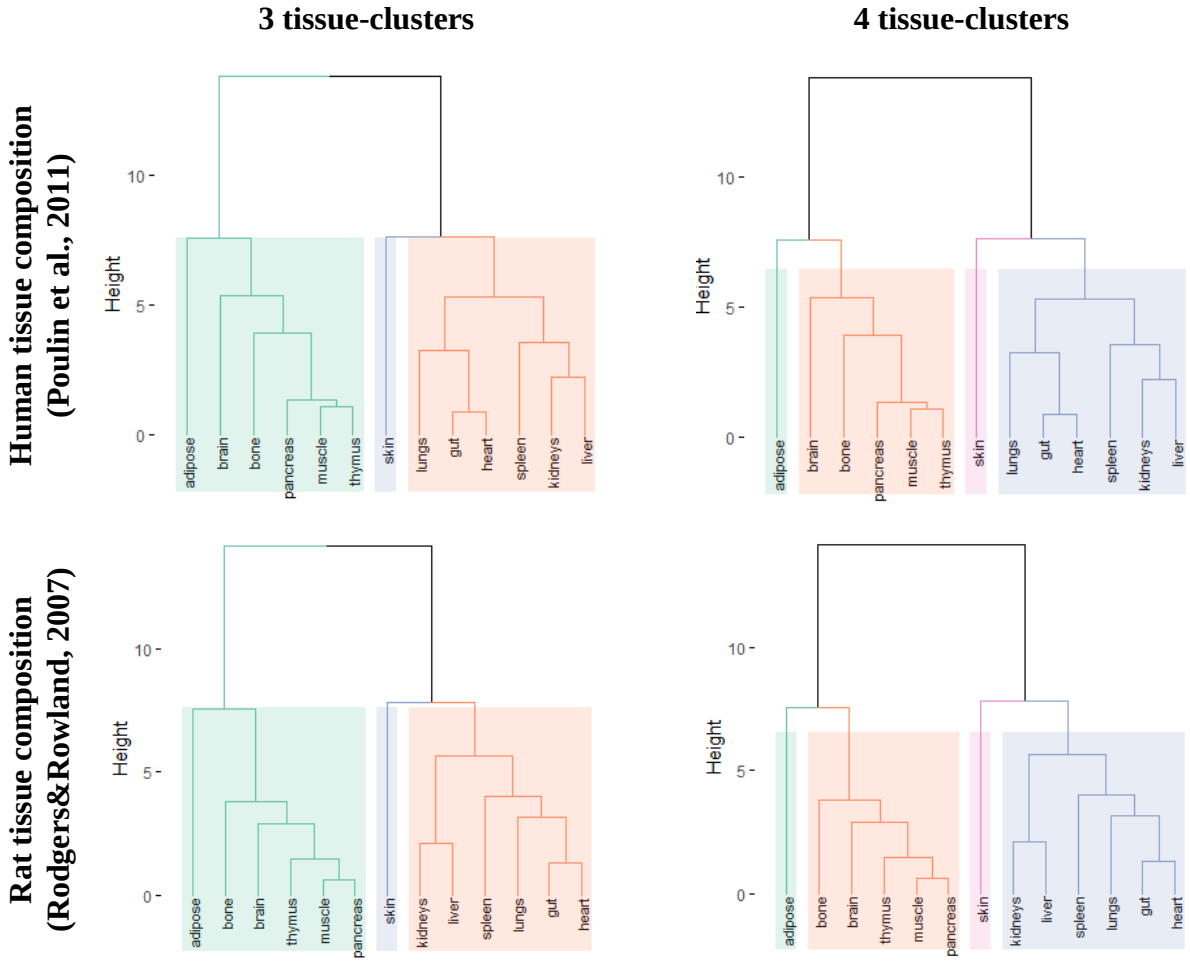
The second approach for simplifying PBPK models was to have a 14-compartmental model with tissue compartments sharing common Kpu values or common Kpu scalars (Figure 3.1). In order to identify tissue with common Kpu values or scalars, a clustering analysis was performed using different clustering methods on (a) tissue composition data in human and rat and (b) imputed *in vivo* rat Kp data. Results of this clustering analysis in man and in rat are summarised in **Figure 3.3**-**Figure 3.5**.

Figure 3.3: Principal component plots for k-means clustering with 3 and 4 groups in man and rat



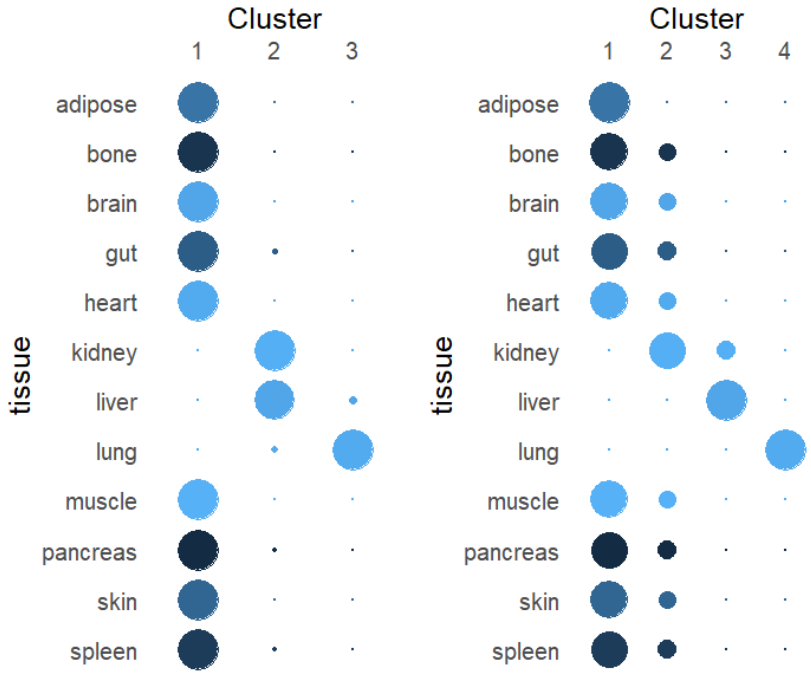
The dots represent a two-dimensional matrix of tissues generated by principal component analysis of the tissue composition data. Horizontal axis, factor score of the first component extracted from the composition of each tissue; vertical axis, factor score of the second component. Distinct clusters are shown in different colours.

Figure 3.4: Dendrograms obtained from hierarchically clustering of the human and rat tissue composition data with Euclidian distance and Ward's method, cut into three or four distinct clusters



Distinct clusters are shown in different colours.

Figure 3.5: Frequencies of cluster assignment and missing values per tissue for 3 and 4 clusters after multiple imputation (n=100) and normalisation of rat steady state Kpus combined with hierarchical clustering



The proportion of missing value per tissue is represented by the intensity of the colour (darker colour relating to more missing data), whereas the radius of the circles represents the frequency of cluster assignment (bigger size relating to more frequent cluster assignment).

Using hierarchical clustering, the same 3 or 4 tissue groups were found for man and rat, whereas slightly different tissue groups were found for rat using k-means clustering (**Figure 3.3** and **Figure 3.4**). For human, skin tissue composition (very high fraction of extracellular water, high albumin and low lipoprotein concentrations) is different from other tissue groups (**Table A2.2**). The group 2 (kidney, gut, lung, heart, liver, spleen) includes tissues with high acid phospholipid and lipoprotein contents, whereas group 1 (adipose, bone, brain, pancreas, muscle) contains tissues with low tissue water, low acid phospholipid and lipoprotein contents but very high neutral lipid and phospholipids. Adipose has a very particular composition (low water, very high neutral lipids) compared to the other tissues of group 1, and thus having 4 groups may be more appropriate (**Table A2.2**). Likewise in the rat, the adipose tissue composition stands out compared to other tissues of group 1 and thus it is more relevant to again consider 4 groups with adipose as a separate tissue group (**Table A2.3**). On the other hand, when using k-means clustering, skin in rat was not considered to be different to the tissues of group 3 (gut, heart, lung, stomach) which are also tissues with higher fractional extracellular water, higher albumin than other tissues (**Table A2.3**).

Alternatively, for clustering based on rat Kp values, tissue grouping was based on frequencies of tissue assignments to a cluster as some tissues have the probability of arising from different clusters (**Figure 3.5**). The tissue groups obtained based on the dataset of Kps were surprisingly different from the previous clustering analysis based on tissue composition data. Here, all the tissues were grouped together except the lung, which was separate from the other tissues, as well as the two eliminating tissues (kidney and liver), which could be further differentiated into two different tissue groups on their own. This tissue grouping seemed to be similar to the range of tissue blood flows.

In general, the results showed that clustering into 4 tissue groups should be favoured as it appeared more physiologically relevant in terms of tissue composition, in particular adipose having a different tissue composition.

3.4.3. Estimation of Kpu values for diazepam using the simplified PBPK models

After defining and identifying all the models and their structures, these models were investigated further for parameter estimation using diazepam data in human and rat. A summary of the various models is presented in Table 3.4.

Initial exploration of structural models revealed that a linear two-compartment model best described the concentration-time profiles of diazepam in human while a linear three-compartment model best described the rat data of diazepam. A clearance value of 3.71 L/h and a total volume of distribution of 145 L (Ca. 2 L/kg) were estimated in man. In rat, the clearance and total volumes were estimated to be 0.920 L/h and 0.914 L (Ca. 4 L/kg). These values were considered as the reference values in the current study and were overall in good agreement with values reported in published studies (**Table A2.4** and **Table A2.5**).

Table 3.4: Summary of models and tissue grouping

Model	Model description	Tissue grouping (by lumping or clustering)
1	Lumped 3 compartment model	*Kpu1: blood, lungs, kidneys, heart, spleen, liver, pancreas, gut, stomach, bone, brain* *Kpu2: adipose, muscle, rest of body* *Kpu3: skin*
2A	14-compartment PBPK model with 3 common Kpus (H)	*Kpu1: adipose, bone, brain, muscle, pancreas, muscle, rest of body* *Kpu2: lung, gut, stomach, kidney, heart, spleen, liver* *Kpu3: skin*
2B	14-compartment PBPK model with 4 common Kpus (H)	*Kpu1:bone, brain, muscle, pancreas, muscle, rest of body* *Kpu2: lung, gut, stomach, kidney, heart, spleen, liver* *Kpu3: skin* *Kpu4: adipose*
2C	14-compartment PBPK model with 3 common Kpus (Km)	*Kpu1: adipose, bone, brain, muscle, pancreas, rest of body* *Kpu2: kidney, spleen, liver* *Kpu3: skin, lung, gut, stomach, heart*
2D	14-compartment PBPK model with 4 common Kpus (Km)	*Kpu1: bone, brain, muscle, pancreas, rest of body* *Kpu2: kidney, spleen, liver* *Kpu3: skin, lung, gut, stomach, heart* *Kpu4: adipose*
2E	14-compartment PBPK model with 3 common Kpus (ss)	*Kpu1: adipose, bone, brain, muscle, pancreas, muscle, rest of body, skin, gut, stomach, heart, spleen* *Kpu2: kidney, liver* *Kpu3: lung*
2F	14-compartment PBPK model with 4 common Kpus (ss)	*Kpu1: adipose, bone, brain, muscle, pancreas, muscle, rest of body, skin, gut, stomach, heart, spleen* *Kpu2: kidney* *Kpu3: liver* *Kpu4: lung*
3A	14-compartment PBPK model with 3 scalars (H)	*SF1: adipose, bone, brain, muscle, pancreas, muscle, rest of body* *SF2: lung, gut, stomach, kidney, heart, spleen, liver* *SF3: skin*
3B	14-compartment PBPK model with 4 scalars (H)	*SF1: bone, brain, muscle, pancreas, muscle, rest of body* *SF2: lung, gut, stomach, kidney, heart, spleen, liver* *SF3: skin* *SF4: adipose*
3C	14-compartment PBPK model with 3 scalars (Km)	*SF1: adipose, bone, brain, muscle, pancreas, rest of body* *SF2: kidney, spleen, liver* *SF3: skin, lung, gut, stomach, heart*
3D	14-compartment PBPK model with 4 scalars (Km)	*SF1: bone, brain, muscle, pancreas, rest of body* *SF2: kidney, spleen, liver* *SF3: skin, lung, gut, stomach, heart* *SF4: adipose*
3E	14-compartment PBPK model with 3 scalars (ss)	*SF1: adipose, bone, brain, muscle, pancreas, muscle, rest of body, skin, gut, stomach, heart, spleen* *SF2: kidney, liver* *SF3: lung*
3F	14-compartment PBPK model with 4 scalars (ss)	*SF1: adipose, bone, brain, muscle, pancreas, muscle, rest of body, skin, gut, stomach, heart, spleen* *SF2: kidney* *SF3: liver* *SF4: lung*

H: hierarchical clustering on rat tissue composition data; Km: k-means clustering on rat tissue composition data; ss: clustering on in vivo rat Kps data; SF: scaling factor

Estimation in human

In total five different PBPK model variants (models 1, 2A, 2B, 3A and 3B) were fitted to the concentration-time data of diazepam in humans. Table 3.5 shows the final estimates of the parameters and the corresponding precision together with the BIC values and lists the number of parameters in the models differentiating between fixed and random (IIV and residual) effects:

Table 3.5: Parameter estimates of the different investigated mechanistic models for diazepam in man and comparison of median Vss,b and RMSE of simulated concentration profiles

Model	1	WBPBPK models			
		2A	2B	3A	3B
Number of parameters (structural/IIV/residual)	4/1/1	4/1/1	5/1/1	4/1/1	5/1/1
Median plasma concentration RMSE	2.56	2.11	1.92	6.60	2.02
BIC	-2632.71	-2662.18	-2597.66	-2344.21	-2596.42
CLb (L/h)	3.63 (4%)	3.67 (4%)	3.67 (4%)	3.71 (4%)	3.56 (5%)
IIV_{CLb}	33.1% (17%)	34.3% (18%)	32.1% (19%)	26% (16%)	35.2% (20%)
Kpu1 or SF1	1153 (5%)	29.1 (5%)	32.4 (6%)	3.35 (7%)	0.206 (42%)
Kpu2 or SF2	24.3 (10%)	72.2 (11%)	89.1 (10%)	1.66 (41%)	5.70 (11%)
Kpu3 or SF3	483 (2%)	3429 (2%)	323.8 (5%)	0.26 (10%)	6.42 (14%)
Kpu4 or SF4			483 (8%)		8.67 (6%)
Residual error	39.3% (8%)	38.7% (8%)	38.6% (8%)	53.9% (15%)	38.6% (8%)
Vss,b median (L)	154.02	158.45	159.45	114.58	159.46

WBPBPK: whole-body physiologically-based pharmacokinetic; IIV: intra-individual variability; CLb: blood clearance; SF, scaling factor; Vss,b: volume of distribution in blood at steady-state
Parameter estimates are listed together with the coefficient of variation [CV (%)] in parentheses.
Abbreviations for model are defined in Table 3.4.
The RMSE is calculated relative to the median concentration simulated from the reference model (2 compartment) during the simulation time interval (0-240 h).

As data came from different individuals (and studies), IIV in clearance was estimated. The values estimated for clearance were similar to the one from the empirical model (3.71L/h) and to that reported previously in literature (**Table A2.4**). Additionnaly, the value estimated for IIV on clearance was similar between the models (Table 3.5). IIV on Kpu parameters were not estimated for better comparison between the different models.

All kinetically lumped and simplified PBPK models could be fitted to diazepam data in man. They could also all recapture an estimated Vss,b close to the Vss,b observed (146.4 L) in man within 20-25% error. Estimated Kpu values from the investigated models were generally in the same range except for the model 1 for which very high

values were obtained for Kpu1 (1153) and may not be physiologically plausible to represent the lung Kpu (Table 3.5). Similarly, Kpu3 in the model 2A has a very high value (3429) and may not be physiologically plausible to represent the skin Kpu (Table 3.5). The value estimated for adipose Kpu is very close for the models 1, 2B and 3B (respectively 483, 483, 485 after multiplying adipose Kpu predicted by the estimated scalar). Additionally, the concentration-time profiles and the Vss,b of all investigated models, except model 3A, were similar to the ones of the empirical two-compartment model (low RMSE, Table 3.5). Figure 3.6 shows the model-predicted population profiles following a single IV infusion of diazepam for each of the simplified PBPK models investigated. All models except model 3A could adequately describe the median of the observed data (simulations using the empirical model), whereas the 10^{th} and the 90^{th} percentiles were not always well captured (Figure 3.6). The shaded area incorporates inter-individual variability on clearance and other unexplained variability (proportional residual error).

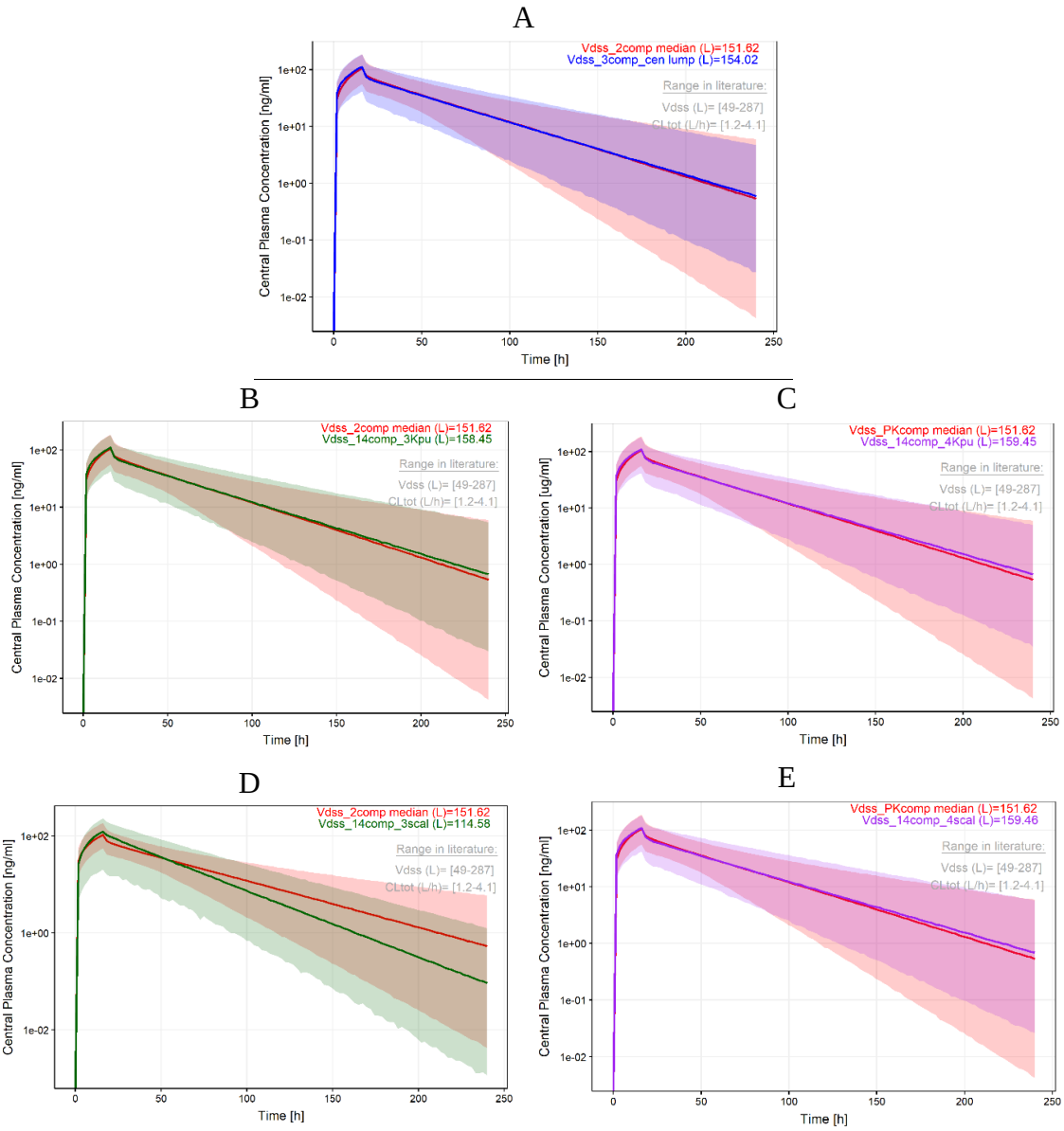
The model 2A had the lowest BIC, followed by the models 1 and 2B. The lowest residual error was obtained with the models 2A, 2B and 3B (Table 3.5). Based on the BIC, plausibility of estimated values and the model performance, the models 2B and 3B seemed to be the best models to describe diazepam data in human (Table 3.6). The model 1 performed better in terms of numerical predictive checks and could also be considered although the value of Kpu1 may be high. On the contrary, due to poor physiological plausibility of skin Kpu and poor model performance (BIC), the models 2A and 3A would not be selected here.

Table 3.6: Comparison of model performance for estimating diazepam data in human according to evaluation criteria

Models	**Criteria**				
	BIC ranking	**Goodness of fits and convergence plots**	**Precision of estimates (CV≤0.4)**	**Plausibility of Kpus**	**Vss within 20% (YYY), 25% (YY), 30% (Y)**
1	2	OK	Yes	?	YYY
2A	1	OK	Yes	?	YYY
2B	3	OK	Yes	Yes	YYY
3A	5	OK	Yes	Yes	YY
3B	4	OK	Yes	Yes	YYY

Abbreviations for model are defined in Table 3.4.
BIC: Bayesian information criterion ranking from lowest to highest value
Plausibility of Kpus: (?) one Kpu may not be plausible

Figure 3.6: Simulated time profiles of diazepam following an infusion dose (10 mg during 16.1h) in human for the different investigated mechanistic models vs the reference model



The solid red line represents the median concentrations and the semi-transparent red field represents a simulation based 90% confidence interval for the median using the reference model (empirical two-compartmental model). The other solid line represents the median concentrations and the other semi-transparent field represents a simulation based 90% confidence interval for the median using the model 1 (A), the models 2A and 2B (B, C) and the models 3A and 3B (D,E). Abbreviations for model are defined in Table 3.4.

Estimation in rat

Compared to the human, the rat PK (plasma concentration) data for diazepam were more limited which rendered the analysis less stable and precise. However, rat Kpu data were available from the literature (Table 3.3) allowing a direct comparison of the models' estimates to the *in vivo* data. Table 3.7 shows the final estimates of the parameters and the corresponding precision together with the BIC values, and lists the number of parameters in the models differentiating between fixed and random (IIV and residual) effects.

Similar to the analysis of human data, rat data (average) originated from different studies and an IIV (confounded with ISV) in clearance was estimated. The clearance estimates are similar between the mechanistic models (Table 3.7) and within the range reported previously (**Table A2.5**) but slightly different to the one from the empirical model (0.92L/h). Because of sparse data, IIV on Kpu parameters were not estimated in order to compare the different models. Additionally, a high correlation (>0.95) between estimates of distribution parameters (Kpu2 and Kpu4 in model 2D; Kpu2 and Kpu3 in model 3D; Kpu1 and Kpu2 in models 2E and 2F) was sometimes observed but this may be an artefact due to the small size of the dataset.

Table 3.7: Parameter estimates of the different investigated mechanistic models for diazepam in rat and comparison of median Vss and RMSE of simulated concentration profiles

Model	1	WBPBPK models											
		2A	2B	2C	2D	2E	2F	3A	3B	3C	3D	3E	3F
Number of param (typical/IIV/residual)	4/1/1	4/1/1	5/1/1	4/1/1	5/1/1	4/1/1	5/1/1	4/1/1	5/1/1	4/1/1	5/1/1	4/1/1	5/1/1
Median plasm con-centration RMSE	7.61	9.04	8.94	10.35	6.58	19.26	19.55	48.13	9.93	7.32	6.06	6.63	6.33
BIC	-177.743	-175.898	-133.547	-178.184	-136.008	-141.567	-99.266	-170.424	-133.886	-176.836	-136.747	-175.255	-133.346
CLb (L/h)	1.14 (47)	1.18 (35)	1.16 (37)	1.14 (33)	1.14 (33)	1.22 (21)	1.22 (23)	1.21 (25)	1.19 (31)	1.14 (32)	1.13 (35)	1.14 (33)	1.13 (34)
IIV_{CLb}	15.1% (33)	15% (31)	15.2% (32)	14.9% (32)	14.9% (33)	16.6% (23)	16.5% (24)	15.1% (32)	15.1% (31)	15.2% (32)	14.2% (33)	14.8% (33)	14.8% (33)
Kpu1 or SF1*	73.0 (54)	28.8 (4)	30.0 (4)	27.9 (4)	29.4 (4)	47.9 (2)	47.9 (2)	1.97 (40)	3.25 (10)	3.39 (8)	3.13 (12)	1.80 (21)	1.73 (24)
Kpu2 or SF2*	30.6 (0.3)	334 (7)	334 (7)	428 (4)	446 (2)	351 (1)	351 (1)	29.1 (5)	21.5 (7)	19.5 (9)	23.3 (11)	43.4 (5)	45.6 (5)
Kpu3 or SF3*	90.0 (3)	90.0 (3)	90.0 (3)	89.1 (3)	89.1 (3)	518 (3)	6374	1.80 (32)	2.07 (20)	0.109	2.04 (20)	0.0265	49021
Kpu4 or SF4*			21.8 (2)		18.2 (5)		518 (3)		0.204 (6)		0.391		0.0174
Residual error	20.2% (9)	20.4% (6)	20.4% (6)	20.0% (4)	20.0% (4)	29.8% (8)	29.8% (8)	21.9% (5)	21.9% (5)	20.4% (6)	19.8% (6)	20.6% (5)	20.6% (5)
Vss median (L)	1.13	1.44	1.43	1.53	1.19	1.34	1.40	1.87	1.68	1.10	1.23	1.12	8.56

WBPBPK: whole-body physiologically-based pharmacokinetic; IIV: intra-individual variability; CLb: blood clearance; SF, scaling factor; Vss,b: volume of distribution in blood at steady-state
Parameter estimates are listed together with the coefficient of variation [CV (%)] in parentheses.
Abbreviations for model are defined in Table 3.4.
The RMSE is calculated relative to the median concentration simulated from the reference model (2 compartment) during the simulation time interval (0-240 h).

The kinetically lumped model (model 1) performed well (the second lowest BIC and the estimated Vss,b within 25% error of the Vss,b observed) whereas the performance of simplified PBPK models is more heterogeneous (Table 3.8). The Vss,b in rat estimated from model 3C and 3E were within 25% error of the Vss,b observed in rat (0.91 L) while none of the Vss,b estimated by models 2A-2F showed a similar performance. Moreover, the models 2E and 2F had the highest residual error compared to the other models (Table 3.7). Contrary to the models 3B, 3D, 3F, the models 3C and 3E showed better performance with Vss,b estimated within 25% error of the Vss,b observed in rat. (Table 3.7 and Table 3.8). Additionally, compared to experimental rat Kpu values reported in the literature (Table 3.3), the Kpu values estimated with the mechanistic models were generally in the same range or order of magnitude. Interestingly, very high values of Kpu3 were estimated in the models 2F and 3F whereas a very small value of Kpu3 or Kpu4 were estimated respectively in the models 3E and 3F (Table 3.7). In the former case, it would be unexpected to have such high values of kidney and liver Kpus (6374 or more than 700000 after multiplying by the estimated scalar) whereas in the latter case, lung Kpu is unlikely to be so low (0.50 or 0.33 after multiplying the predicted Kpu by the estimated scalar, respectively). It was found that 95% of Kpus predictions using R&R model would be successfully predicted within around 21-fold error according to the analysis of a dataset of compiled rat experimental tissue partitioning coefficients from the literature (not shown). The concentration-time profiles and the volumes of distribution of models 1, 2D, 3C, 3D and 3E were the closest to the ones of the empirical two-compartment model (Table 3.7).

Figure 3.7 shows the model-predicted population profiles following a single IV infusion of diazepam in rat for each of the mechanistic models investigated. All models except the models 3A, 2E and 2F could adequately describe the median of the observed data (simulations using the empirical model as surrogate for observed data), whereas the 10^{th} and the 90^{th} percentiles were not always well captured (Figure 3.7). The shaded area incorporates inter-study variability on clearance and other unexplained variability (proportional residual error).

Due to poor physiological plausibility of several Kpus and poor model performance, the models 3A, 2E, 2F, 3E and 3F would not be selected here. Other models such as models 2A, 2B, 2C,2D, 3C and 3D performed well and can be selected for further use. Based on the BIC, plausibility of estimated values, the model performance, three reduced models were selected as primary candidates (Table 3.8): the models 3C and 2D.

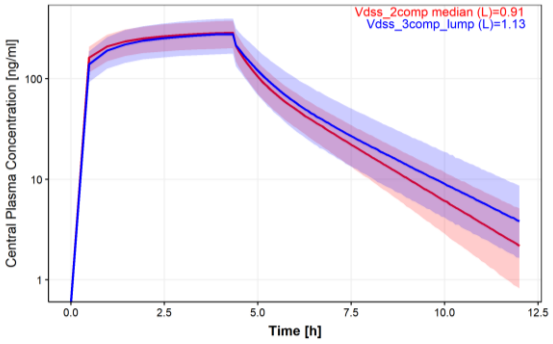
Table 3.8: Comparison of model performance for estimating diazepam data in rat according to evaluation criteria

Models	Criteria				
	BIC ranking	**Goodness of fits and convergence plots**	**Precision of estimates (CV≤0.4)**	**Plausibility of Kpus**	**Vss within 20% (YYY), 25% (YY), 30% (Y)**
1	2	Yes	No	Yes	YY
2A	4	Yes	Yes	Yes	No
2B	11	Yes	Yes	Yes	No
2C	1	Yes	Yes	Yes	No
2D	9	Yes	Yes	Yes	Y
2E	7	Yes	Yes	Yes	No
2F	13	Yes	Yes	?	No
3A	6	Yes	Yes	?	No
3B	10	Yes	Yes	Yes	No
3C	3	Yes	Yes	Yes	YY
3D	8	Yes	Yes	Yes	No
3E	5	Yes	No	?	YY
3F	12	Yes	Yes	?	No

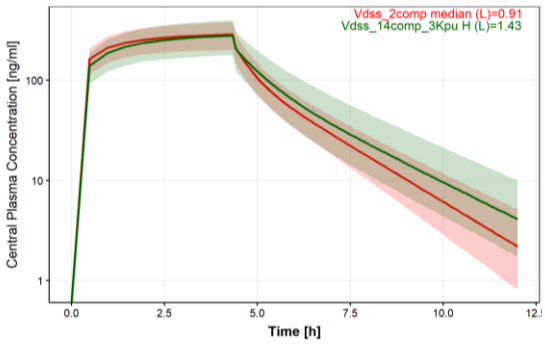
Abbreviations for model are defined in Table 3.4.
BIC: Bayesian information criterion ranking from lowest to highest value
Plausibility of Kpus: (?) one Kpu may not be plausible

Figure 3.7: Simulated time profiles of diazepam following an infusion dose (1 mg during 4.33h) in rat for the different investigated mechanistic models vs the reference model

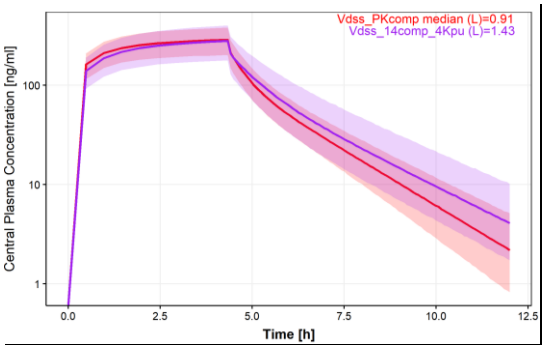
A



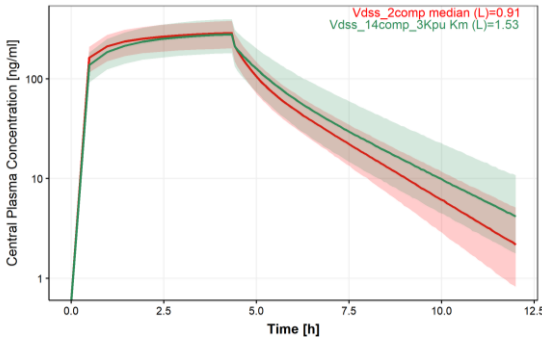
B



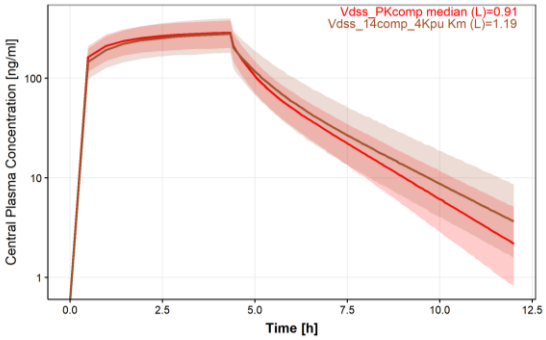
C



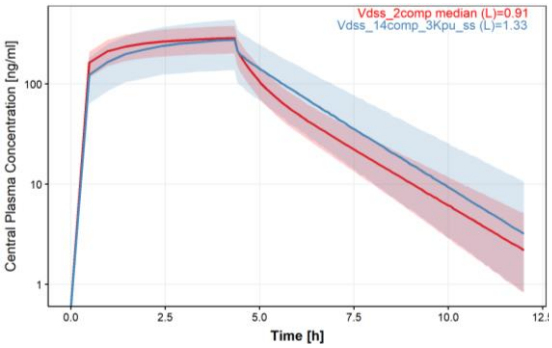
D



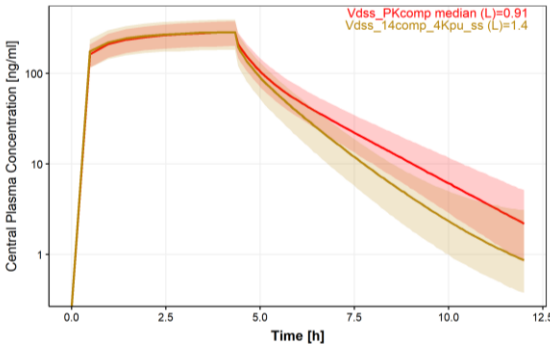
E



F

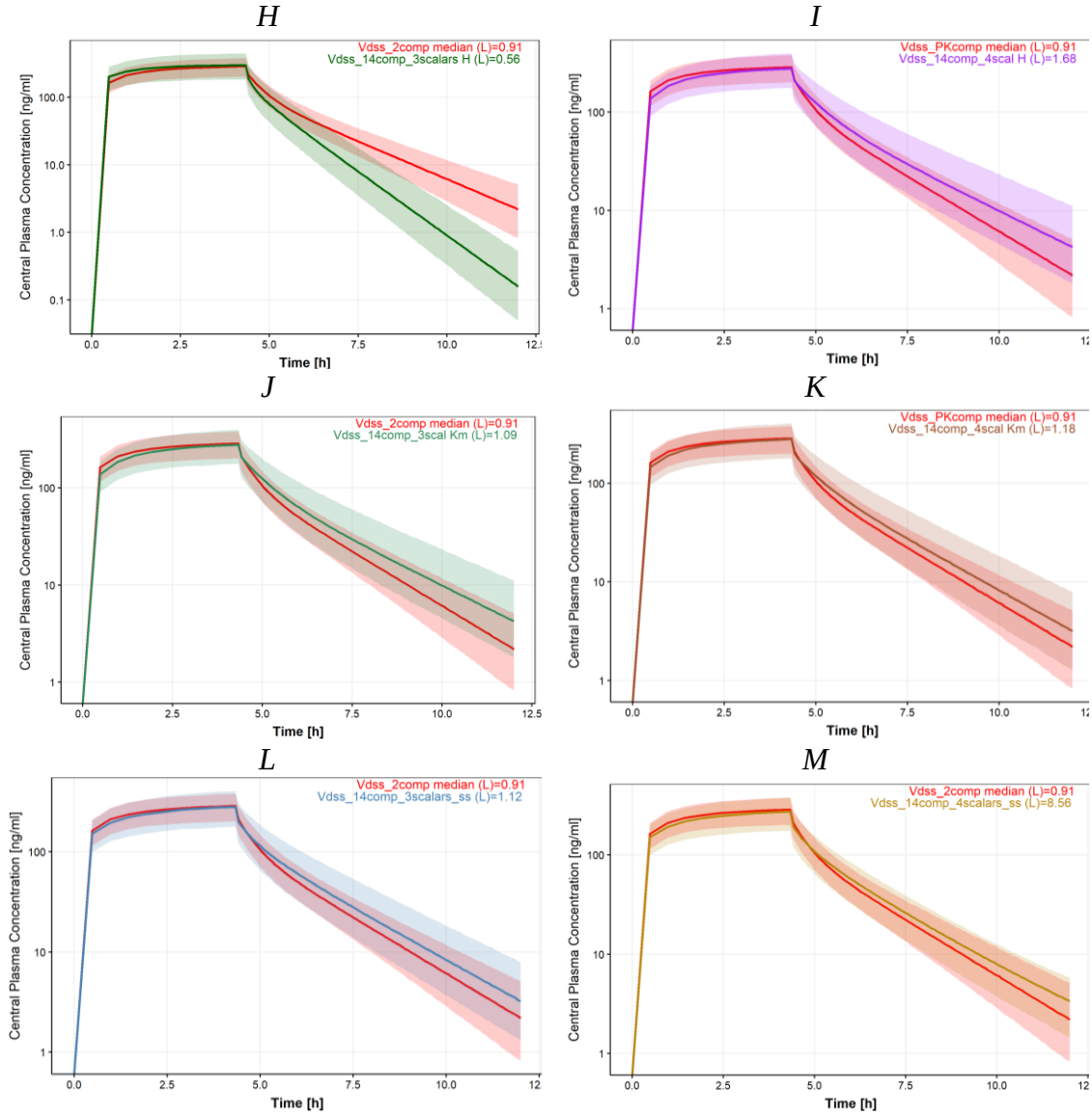


G



(to continue)

Figure 3.7: Simulated time profiles of diazepam following an infusion dose (1 mg during 4.33h) in rat for the different investigated mechanistic models vs the reference model (continued)



The solid red line represents the median concentrations and the semi-transparent red field represents a simulation based 90% confidence interval for the median using the reference model (empirical 3-compartmental model). The other solid line represents the median concentrations and the other semi-transparent field represents a simulation based 90% confidence interval for the median using the models 1 (A), 2A, 2B, 3A, 3B (B, C, H, I respectively), 2C, 2D, 3C, 3D (D, E, J, K respectively), and the models 2E, 2F, 3E and 3F (F, G, L, M respectively).
*(NB: as the hepatic clearance (CLH) in rat was greater that the hepatic blood flow (Q_HV), it was assumed that CLH=0.99*Q_HV for simulations)*
Abbreviations for model are defined in Table 3.4.

3.5. Discussion

In this work, two mechanistic approaches for simplifying whole-body PBPK models were presented and are summarised in Table 3.9. Both approaches were sufficiently flexible as the investigated models generally showed a good ability to fit data and to estimate model parameters.

Table 3.9: Summary and characteristics of the approaches investigated for simplifying PBPK models

Models and assumptions		Advantages	Disadvantages
Lumped 3 compartment model: • QSS approximation for blood and lung • Lumping of compartments based on similar tissue kinetics • Drug distributes homogenously among the lumped tissue at the same distribution rate		• Easy to implement as closed form solution • Possible to use computer-intensive methods for estimation • Possible to derive individual lung profiles	• May be difficult to extrapolate due to different lumped models between species • Assumptions on tissue kinetics (blood-flows and volumes) • Limited to compounds with relatively low tissue-to-plasma partitioning • Assumption of low clearance compounds
14 compartment PBPK model: • QSS approximation for blood and lung • Grouping of tissue based on steady states similarities in drug partitioning or physiological similarities	**With common Kpus:** Assumption of same Kpus between species	• No assumption on tissue kinetics (except for lung) • Better extrapolation potentially due to stronger similarity of tissue composition compared to tissue kinetics between species • Possible to derive individual tissue profiles • Not limited by any drug properties	• Long running time due to model high dimensionality
	With common scalars: Assumption of similar bias from predicted Kpus (by R&R model) across species		

QSS: quasi-steady state

The first proposed method followed Nestorov's lumping principles for PBPK models [6] where lumping of tissues is mainly based on similar or low time constants. This is consistent with other minimal PBPK models, which reduce the number of compartments to just one or a few compartments with comparable blood flow rates [60, 61]. A lumped reduced PBPK models can be easily used during development as it is simple to implement and numerically stable while investigating the kinetics of a few tissue and plasma compartments. The number of compartments considered was in good agreement with previous work by Pilari and Huisinga where they presented lumped models that comprised generally of three to four compartments for bases and acids which are moderately to highly unbound, while a single compartment model was sufficient for acids with high plasma protein binding [9]. However, lung was often lumped with other tissues, which may be inconsistent as the lung blood flow is in the opposite direction compared to other tissue compartments. Arundel suggested a model with more compartments (six tissue compartments) but estimating six Kpu parameters may be challenging [17]. In the current approach based on ranking of tissue distribution rate, the drug's tissue time constant was considered to only depend on volume and blood flow and the effect of drug-specific partition coefficient was considered minimal (Kb close to 1 and Clint close to 0). However, this assumption seems invalid in many cases e.g., diazepam. In the manner of previous studies [6, 62], a general "quasi-steady state" (QSS) approximation was postulated for blood and lung lumping due to the time scale differences between lung distribution relative to any observations. This assumption holds true in the artery, vein and lungs in general except for highly lipophilic compounds. Being in the central compartment, the lung Kpu might represent additional processes more rapid than considered blood flows (e.g., extravasation from the site of injection or during the transit from the site of injection to the site of observation). However, scaling of this first approach across species may not work well as some tissues (e.g. muscle, skin, adipose) have different flows and volumes (Table **3.1**), and consequently different tissue kinetics between human and animals resulting in slightly different lumped tissue compartments (Figure 3.2). Moreover, rat has low information for adipose tissue as it is a much leaner species than human and human has a smaller proportion of skin compared to other tissues than rat. It could be possible to extrapolate to human with this first approach by assuming a same model structure in rats and humans, and thus use the lumped human model to fit parameters in rat before extrapolating to human. Additionally, the kinetically lumped model can be expressed in analytical form that could be of great utility when using time-

intensive estimation methods, although it would be reserved for compounds with relatively low Kp values (satisfying the kinetic assumptions).

The second approach followed the assumption that a drug has similar steady state binding/partitioning behaviour in tissues with similar compositions, but different kinetic behaviours given their different volumes and flow rates of tissues. This assumption is consistent with the results of a previous work where it was observed several tissues had similar Kp behaviour while adipose standing out possibly due to different tissue composition [35]. Additionally, a previous study observed correlation between muscle Kpu and other tissues Kpu and proposed to have muscle Kpu representing lean tissues [63]. To formalize this idea, a clustering analysis was performed to identify tissues with common Kpus or scalars thus reducing the number of unknown parameters while keeping the complexity of PBPK model structure and making minimal assumption about tissue kinetics. Like in the first approach, it was assumed that the kinetics of venous, arterial blood and lungs are quasi-instantaneous (QSS approximation) and these tissues can be lumped into one central compartment [6]. Two clustering methods were applied to formally group tissues using two different types of information: human or rat tissue composition data or rat normalised Kps data, which both contain uncertainty and variability. The tissue composition data are a hybrid combination of measured, calculated, approximated individual or mean values [23-25, 64]. Moreover, rat data could originate from different strains such as Sprague-Dawley or Wistar. For human, actual data from humans are very limited and often surrogate information is taken from another species (generally rat or monkey). On the other hand, rat Kps data are generally derived from *in vivo* studies following a constant rate infusion of drug until steady state is supposedly reached. Alternatively, Kps data can be calculated based on plasma and tissue areas under concentration time (AUC) profiles following an iv bolus. Additionally, these *in vivo* Kp methods require at least one animal per time point. Therefore, the different tissue grouping from the clustering using rat experimental Kps compared to the clustering analysis based on tissue composition data could be because of the size of the data set, the availability of data (high percentage of missing data), the quality of reported experimental Kps, as well as the varying source of studies and methods used for measuring Kps. And when using the tissue grouping from the clustering of rat Kps for a PBPK model in human, the assumption is that Kpus are the same between species which may not be always true. Indeed, Kpu values may be different across species especially for relatively lipophilic compounds and for ionised bases if lipid levels highly differ between species [65].

In this work, the mechanistic models using either a kinetic lumping approach or an approach with common Kpus or scalars were found equivalent at describing IV data compared to empirical models. Data in human showed better bioanalytical resolution and slower kinetics, which may reveal kinetic phases that are hidden in rat. It should be noted the quality of the plasma PK data used for fitting was not ideal as the data were mostly arithmetic mean data which may misrepresent the terminal phase in addition to coming from multiple studies following different protocols in different populations. Thus, the recommendation would be to perform this analysis on individual data or on geometric mean data if data are aggregated.

For the example of diazepam, the proposed mechanistic models produced estimates of Kpu that were comparable to those obtained in distribution studies from the IV data alone. Compared to previous works where diazepam data was fitted in man [48, 49, 58], the models investigated here have the advantage of not requiring a highly-dimensional PBPK model structure and the use of prior information or full Bayesian estimation methods for estimating the multiple parameters. The current work is an example of middle-out approach linking PBPK modelling (bottom-up) and population PK modelling (top-down) and where PBPK models can be fitted to observed preclinical or clinical data. Contrary to empirical compartment models that are frequently used in all stages of drug development, the investigated PBPK models retain important physiological features giving estimates with physiological meaning and thus supporting extrapolation purposes while requiring a similar number of free parameters.

The value of PBPK modelling is indeed its ability to be used for extrapolation to another population or species and experimental conditions. The aspect of the interspecies extrapolation from preclinical species (rat and monkey) to human will be investigated in a future study (Chapter 4) for diazepam and additional compounds. Among the various investigated PBPK models, the PBPK models with scalars show more promise for a wider set of drugs and for inter-species translation. Although the PBPK models with 3 scalars are more parsimonious and should be considered first when data are limited, the PBPK models with 4 scalars are more physiologically relevant with regards to the tissue composition and conclusions of the clustering analysis and the plausibility of Kpu values estimated. However, the PBPK model with 4 scalars using hierarchical clustering and the PBPK models with 3 or 4 scalars using rat Kps did not find plausible estimates for all rat Kpus for diazepam and therefore these models may not be appropriate. Additionally, the PBPK model with commonalities based on the clustering using normalized rat Kps need

to assume that Kpus are the same between species for it to be extrapolated to human, which may not be true as mentioned above. For a comprehensive comparison of the investigated mechanistic models, evaluation of different compounds would be needed.

3.6. Conclusion

The main goal of the current study was to determine simplified PBPK models that can be easily used for fitting purposes. The two approaches for simplifying PBPK models presented in this work allow more mechanistic models to be explored instead of only empirical models for fitting PK data. These models using either a kinetic lumping approach, a widely established method, or a novel approach defining common Kpus or scalars. Both types of models were overall found equivalent at describing IV data compared to empirical models but are likely to have different behaviour for PK translation (see Chapter 4). These mechanistic models generally produced Kpu estimates that were physiologically plausible. Although the current study is based on diazepam, the findings provide insights that the kinetically lumped model has great utility for compounds with relatively low tissue-to-plasma partitioning (satisfying the kinetic assumptions), and the PBPK models with scalars show more promise for a wider set of drugs and for the purpose of extrapolation. The PBPK models with 4 scalars where adipose is separated could be the most relevant for a simplified PBPK model as adipose has quite a different tissue composition compared to other tissues. Further research with additional compounds should be conducted to assess the different models.

3.7. Acknowledgement

We thank Isabelle Walter for assistance with *in vitro* measurements of diazepam. We also thank Neil Parrott for scientific discussion and reviewing this manuscript.

3.8. References

[1] Aarons L. Physiologically based pharmacokinetic modelling: a sound mechanistic basis is needed. Br J Clin Pharmacol. 2005;60(6):581-3.

[2] Reddy MB, Iii HJC, Lave T, Andersen ME. Physiologically Based Pharmacokinetic Modeling: A Tool for Understanding ADMET Properties and Extrapolating to Human. In: Gowder S, editor. New Insights into Toxicity and Drug Testing. Rijeka: InTech; 2013. p. Ch. 09.

[3] Poulin P, Jones RD, Jones HM, Gibson CR, Rowland M, Chien JY, et al. PHRMA CPCDC initiative on predictive models of human pharmacokinetics, part 5: prediction of plasma concentration-time profiles in human by using the physiologically-based pharmacokinetic modeling approach. J Pharm Sci. 2011;100(10):4127-57.

[4] Tsamandouras N, Rostami-Hodjegan A, Aarons L. Combining the 'bottom up' and 'top down' approaches in pharmacokinetic modelling: fitting PBPK models to observed clinical data. Br J Clin Pharmacol. 2013;79(1):48-55.

[5] Margolskee A, Darwich AS, Pepin X, Aarons L, Galetin A, Rostami-Hodjegan A, et al. IMI - Oral biopharmaceutics tools project - Evaluation of bottom-up PBPK prediction success part 2: An introduction to the simulation exercise and overview of results. Eur J Pharm Sci. 2017;96:610-25.

[6] Nestorov IA, Aarons LJ, Arundel PA, Rowland M. Lumping of whole-body physiologically based pharmacokinetic models. J Pharmacokinet Biopharm. 1998;26(1):21-46.

[7] Dokoumetzidis A, Aarons L. Proper lumping in systems biology models. IET Syst Biol. 2009;3(1):40-51.

[8] Dokoumetzidis A, Aarons L. A method for robust model order reduction in pharmacokinetics. J Pharmacokinet Pharmacodyn. 2009;36(6):613-28.

[9] Pilari S, Huisinga W. Lumping of physiologically-based pharmacokinetic models and a mechanistic derivation of classical compartmental models. J Pharmacokinet Pharmacodyn. 2010;37(4):365-405.

[10] Brochot C, Toth J, Bois FY. Lumping in pharmacokinetics. J Pharmacokinet Pharmacodyn. 2005;32(5-6):719-36.

[11] Bjorkman S. Reduction and lumping of physiologically based pharmacokinetic models: prediction of the disposition of fentanyl and pethidine in humans by successively simplified models. J Pharmacokinet Pharmacodyn. 2003;30(4):285-307.

[12] Fronton L, Pilari S, Huisinga W. Monoclonal antibody disposition: a simplified PBPK model and its implications for the derivation and interpretation of classical compartment models. J Pharmacokinet Pharmacodyn. 2014;41(2):87-107.

[13] Gueorguieva I, Nestorov IA, Rowland M. Reducing whole body physiologically based pharmacokinetic models using global sensitivity analysis: diazepam case study. J Pharmacokinet Pharmacodyn. 2006;33(1):1-27.

[14] Snowden TJ, van der Graaf PH, Tindall MJ. Model reduction in mathematical pharmacology : Integration, reduction and linking of PBPK and systems biology models. J Pharmacokinet Pharmacodyn. 2018.

[15] Dokoumetzidis A. Lumping of compartments. PAGE 212012.

[16] Cao Y, Jusko WJ. Applications of minimal physiologically-based pharmacokinetic models. J Pharmacokinet Pharmacodyn. 2012;39(6):711-23.

[17] Arundel PA. A multi-compartmental model generally applicable to physiologically-based pharmacokinetics. 3rd IFAC Modelling and Control in Biomedical Systems; Warwick, UK1997. p. 5.

[18] Luttringer O, Theil FP, Poulin P, Schmitt-Hoffmann AH, Guentert TW, Lave T. Physiologically based pharmacokinetic (PBPK) modeling of disposition of epiroprim in humans. J Pharm Sci. 2003;92(10):1990-2007.

[19] Bernareggi A, Rowland M. Physiologic modeling of cyclosporin kinetics in rat and man. J Pharmacokinet Biopharm. 1991;19(1):21-50.

[20] Rowland M, Peck C, Tucker G. Physiologically-based pharmacokinetics in drug development and regulatory science. Annu Rev Pharmacol Toxicol. 2011;51:45-73.

[21] Pang KS, Rowland M. Hepatic clearance of drugs. I. Theoretical considerations of a "well-stirred" model and a "parallel tube" model. Influence of hepatic blood flow, plasma and blood cell binding, and the hepatocellular enzymatic activity on hepatic drug clearance. J Pharmacokinet Biopharm. 1977;5(6):625-53.

[22] Assmus F, Houston JB, Galetin A. Incorporation of lysosomal sequestration in the mechanistic model for prediction of tissue distribution of basic drugs. Eur J Pharm Sci. 2017;109:419-30.

[23] Rodgers T, Leahy D, Rowland M. Physiologically based pharmacokinetic modeling 1: predicting the tissue distribution of moderate-to-strong bases. J Pharm Sci. 2005;94(6):1259-76.

[24] Rodgers T, Rowland M. Physiologically based pharmacokinetic modelling 2: predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. J Pharm Sci. 2006;95(6):1238-57.

[25] Poulin P, Theil FP. Prediction of pharmacokinetics prior to in vivo studies. 1. Mechanism-based prediction of volume of distribution. J Pharm Sci. 2002;91(1):129-56.

[26] Berezhkovskiy LM. Volume of distribution at steady state for a linear pharmacokinetic system with peripheral elimination. Journal of Pharmaceutical Sciences. 2004;93(6):1628-40.

[27] Schmitt W. General approach for the calculation of tissue to plasma partition coefficients. Toxicol In Vitro. 2008;22(2):457-67.

[28] Graham H, Walker M, Jones O, Yates J, Galetin A, Aarons L. Comparison of in-vivo and in-silico methods used for prediction of tissue: plasma partition coefficients in rat. J Pharm Pharmacol. 2012;64(3):383-96.

[29] Valentin J. Basic anatomical and physiological data for use in radiological protection: reference values. ICRP Publication 89. Ann ICRP. 2002;32(3-4):1-277.

[30] Williams LR, Leggett RW. Reference values for resting blood flow to organs of man. Clin Phys Physiol Meas. 1989;10(3):187-217.

[31] Nestorov I. Modelling and simulation of variability and uncertainty in toxicokinetics and pharmacokinetics. Toxicol Lett. 2001;120(1-3):411-20.

[32] Brown RP, Delp MD, Lindstedt SL, Rhomberg LR, Beliles RP. Physiological parameter values for physiologically based pharmacokinetic models. Toxicol Ind Health. 1997;13(4):407-84.

[33] Kuwahira I, Moue Y, Ohta Y, Mori H, Gonzalez NC. Distribution of pulmonary blood flow in conscious resting rats. Respir Physiol. 1994;97(3):309-21.

[34] Bauer RJ. NONMEM Tutorial Part I: Description of Commands and Options, with Simple Examples of Population Analysis. CPT Pharmacometrics Syst Pharmacol. 2019.

[35] Yau E, Olivares-Morales A, Gertz M, Parrott N, Darwich AS, Aarons L, et al. Global Sensitivity Analysis of the Rodgers and Rowland Model for Prediction of Tissue: Plasma Partitioning Coefficients: Assessment of the Key Physiological and Physicochemical Factors That Determine Small-Molecule Tissue Distribution. AAPS J. 2020;22(2):41.

[36] Jain AK. Data clustering: 50 years beyond K-means. Pattern Recogn Lett. 2010;31(8):651-66.

[37] Xu D, Tian Y. A Comprehensive Survey of Clustering Algorithms. Annals of Data Science. 2015;2(2):165-93.

[38] MacQueen J, editor Some methods for classification and analysis of multivariate observations. Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics; 1967 1967; Berkeley, Calif.: University of California Press.

[39] Hartigan JA, Wong MA. Algorithm AS 136: A K-Means Clustering Algorithm. Journal of the Royal Statistical Society Series C (Applied Statistics). 1979;28(1):100-8.

[40] Johnson SC. Hierarchical clustering schemes. Psychometrika. 1967;32(3):241-54.

[41] Murtagh F, Legendre P. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? Journal of Classification. 2014;31(3):274-95.

[42] Poulin P, Theil FP. A Priori Prediction of Tissue:Plasma Partition Coefficients of Drugs to Facilitate the Use of Physiologically-Based Pharmacokinetic Models in Drug Discovery. Journal of Pharmaceutical Sciences. 2000;89(1):16-35.

[43] Jansson R, Bredberg U, Ashton M. Prediction of drug tissue to plasma concentration ratios using a measured volume of distribution in combination with lipophilicity. J Pharm Sci. 2008;97(6):2324-39.

[44] Wilson SE. Methods for clustering data with missing values: Universiteit Leiden; 2015.

[45] Basagana X, Barrera-Gomez J, Benet M, Anto JM, Garcia-Aymerich J. A framework for multiple imputation in cluster analysis. Am J Epidemiol. 2013;177(7):718-25.

[46] van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software; Vol 1, Issue 3 (2011). 2011.

[47] Little RJA. Missing-Data Adjustments in Large Surveys. Journal of Business & Economic Statistics. 1988;6(3):287-96.

[48] Tsiros P, Bois FY, Dokoumetzidis A, Tsiliki G, Sarimveis H. Population pharmacokinetic reanalysis of a Diazepam PBPK model: a comparison of Stan and GNU MCSim. J Pharmacokinet Pharmacodyn. 2019.

[49] Langdon G, Gueorguieva I, Aarons L, Karlsson M. Linking preclinical and clinical whole-body physiologically based pharmacokinetic models with prior distributions in NONMEM. Eur J Clin Pharmacol. 2007;63(5):485-98.

[50] Greenblatt DJ, Allen MD, Harmatz JS, Shader RI. Diazepam disposition determinants. Clin Pharmacol Ther. 1980;27(3):301-12.

[51] Gueorguieva I, Nestorov IA, Murby S, Gisbert S, Collins B, Dickens K, et al. Development of a whole body physiologically based model to characterise the pharmacokinetics of benzodiazepines. 1: Estimation of rat tissue-plasma partition ratios. J Pharmacokinet Pharmacodyn. 2004;31(4):269-98.

[52] Igari Y, Sugiyama Y, Sawada Y, Iga T, Hanano M. Prediction of diazepam disposition in the rat and man by a physiologically based pharmacokinetic model. J Pharmacokinet Biopharm. 1983;11(6):577-93.

[53] Mandelli M, Tognoni G, Garattini S. Clinical pharmacokinetics of diazepam. Clin Pharmacokinet. 1978;3(1):72-91.

[54] Johansson AM, Ueckert S, Plan EL, Hooker AC, Karlsson MO. Evaluation of bias, precision, robustness and runtime for estimation methods in NONMEM 7. J Pharmacokinet Pharmacodyn. 2014;41(3):223-38.

[55] Bauer RJ, Guzy S, Ng C. A survey of population analysis methods and software for complex pharmacokinetic and pharmacodynamic models with examples. AAPS J. 2007;9(1):E60-83.

[56] Laporte-Simitsidis S, Girard P, Mismetti P, Chabaud S, Decousus H, Boissel JP. Inter-study variability in population pharmacokinetic meta-analysis: when and how to estimate it? J Pharm Sci. 2000;89(2):155-67.

[57] Nguyen TH, Mouksassi MS, Holford N, Al-Huniti N, Freedman I, Hooker AC, et al. Model Evaluation of Continuous Data Pharmacometric Models: Metrics and Graphics. CPT Pharmacometrics Syst Pharmacol. 2017;6(2):87-109.

[58] Gueorguieva I, Aarons L, Rowland M. Diazepam pharamacokinetics from preclinical to phase I using a Bayesian population physiologically based pharmacokinetic model with informative prior distributions in WinBUGS. J Pharmacokinet Pharmacodyn. 2006;33(5):571-94.

[59] Wang W, Hallow KM, James DA. A Tutorial on RxODE: Simulating Differential Equation Pharmacometric Models in R. CPT Pharmacometrics Syst Pharmacol. 2016;5(1):3-10.

[60] Jones HM, Mayawala K, Poulin P. Dose selection based on physiologically based pharmacokinetic (PBPK) approaches. AAPS J. 2013;15(2):377-87.

[61] Sager JE, Yu J, Ragueneau-Majlessi I, Isoherranen N. Physiologically Based Pharmacokinetic (PBPK) Modeling and Simulation Approaches: A Systematic Review

of Published Models, Applications, and Model Verification. Drug Metab Dispos. 2015;43(11):1823-37.

[62] Yates JW, Arundel PA. On the volume of distribution at steady state and its relationship with two-compartmental models. J Pharm Sci. 2008;97(1):111-22.

[63] Bjorkman S. Prediction of the volume of distribution of a drug: which tissue-plasma partition coefficients are needed? J Pharm Pharmacol. 2002;54(9):1237-45.

[64] Ruark CD, Hack CE, Robinson PJ, Mahle DA, Gearhart JM. Predicting passive and active tissue:plasma partition coefficients: interindividual and interspecies variability. J Pharm Sci. 2014;103(7):2189-98.

[65] Rodgers T, Rowland M. Mechanistic approaches to volume of distribution predictions: understanding the processes. Pharm Res. 2007;24(5):918-33.

Chapter 4: Prediction of human drug disposition from preclinical data using a ‘middle-out approach’ to physiologically based pharmacokinetic (PBPK) modelling

4.1. Abstract

4.1.1. Background

In Chapter 3, several simplified physiologically based pharmacokinetic (PBPK) models, using estimated tissue-to-unbound plasma partition coefficients (Kpu), were investigated by fitting to *in vivo* pharmacokinetic (PK) data: a kinetically lumped 3 compartment model, 14 compartment models with 3 or 4 common Kpus and 14 compartment models with 3 or 4 common scalars. In the 14 compartment models, commonality of the 3 or 4 tissue groups was based on the similarity in tissue composition following a clustering analysis using K-means or hierarchical methods. The current work evaluated the performance of several PBPK models for the prediction of human concentration-time profiles and volumes of distribution at steady state (Vss) after fitting the models to preclinical PK data. After optimisation with preclinical data, the performance of these models for translation of distribution kinetics from rat and monkey to human was compared with the traditional whole-body PBPK (WBPBPK) modelling strategy ('bottom-up'), in order to determine the best approaches for the prediction of human drug disposition.

4.1.2. Methods

The analysis was performed for 3 lipophilic bases (diazepam, midazolam and basmisanil) for which intravenous PK data were available in rat, monkey and human. *In vivo* and *in vitro* preclinical data were used in fitting the simplified PBPK models in preclinical species by analysing the blood or plasma PK profile while fixing clearance. The best models and the three or four estimated parameters were then used to predict the human PK profile and Vss.

4.1.3. Results

The models with scalars using K-means were generally the best for fitting the data in the preclinical species and the estimated parameters gave plausible Kpu values. Predictions of plasma concentrations for diazepam and midazolam using the best models and parameters obtained from fitting rat or monkey data were consistent with observed clinical data. The best model for fitting and for translation were generally the models with 3 or 4 common Kpu scalars. For diazepam, the Vss,b in human could be predicted within 1.1 and 1.5-fold error after optimisation in rats and within 3.1- and 2.5-fold error after optimisation in monkeys. These predictions are better compared to the Vss,b estimated from the traditional WBPBPK modelling approach (Vss,b=41L, 3.7-fold error). For

midazolam, the Vss,b in human could be predicted within 1.1- and 2.2- -fold error after optimisation in rats and within 3.1-fold and 2.5-fold error after optimisation in monkeys. The predictions are better or similar compared to the Vss,b estimated from the traditional WBPBPK modelling approach (Vss,b=41L, 3.7-fold error). For basmisanil, the poor quality of preclinical data available could have affected the model performance for fitting in the preclinical species and subsequently extrapolation to human.

4.1.4. Conclusions

Overall, this work provides a rational strategy to predict human drug distribution using preclinical PK data within the PBPK modelling strategy.

4.2. Introduction

During drug discovery and development, characterizing the pharmacokinetic (PK) properties (absorption, distribution, metabolism, and excretion) of a drug in the early stage is crucial to avoid failures due to poor PK properties. Prediction of human PK helps to design the optimal phase 1 studies and allows the selection of the starting dose in humans that is safe and effective in order to maintain a rapid dose escalation, saving time and cost [1]. Different strategies for predicting human PK profiles based on preclinical data have been proposed, such as Dedrick plots, the steady-state plasma drug concentration (Css)–mean residence time (MRT) method and physiologically-based pharmacokinetic (PBPK) modelling [2-5]. Among them, the whole-body physiologically-based pharmacokinetic (WBPBPK) modelling has the advantage to provide a biological and mechanistic understanding for inter-species scaling and intra-species scaling as well as a better understanding of the drug behaviour [6-8]. Several studies showed better accuracy of the PBPK modelling approach for predicting PK compared to empirical and allometric approaches [9-12].

PBPK model development is an iterative 'predict, learn, confirm' process [13-15]. As a bottom-up approach, WBPBPK modelling and simulation are based on physiological input data (tissue volumes, blood flows, composition, etc.) and *in vitro* input data (plasma protein binding, microsomal or hepatocyte intrinsic clearance, cell membrane permeability, etc.). The predictions from WBPBPK models are then compared with observed PK data. However, mismatches often occur between model predictions and observations; and thus, model parameters need to be adjusted. In this case, the PBPK model is built to fit the observed data following a 'middle-out' approach which allows available *in vivo* data to estimate unknown or uncertain model parameters [16-18]. However, parameter estimation with WBPBPK models is challenging due to the large number of parameters and the paucity of observed data usually available (especially preclinical in vivo data). When data are associated with high uncertainty and/or variability, this would be mirrored in the estimated parameters. It is thus a common practice to fix some parameters and estimate others. The decision on which parameters to fix and which to estimate is subjective and therefore the final model and parameters may vary significantly between different modellers [16, 19]. Consequently, this approach does not usually follow a systematic and formal approach. The previous work (Chapter 3) showed the possibility of using simplified PBPK models for fitting PK data. In these models,

tissue distribution was described by mechanistic equations that calculate tissue-to-unbound plasma partition coefficients (Kpus) [20, 21]. Kinetically lumped models based on tissue time constants [22] were investigated as well as PBPK models with common tissue Kpus or common tissue scaling factors. The commonality in tissue partitioning was considered when observing similar behaviour of tissue Kpus across compounds with different properties [23]. More details of the model structures and lumping approaches can be found in this chapter Methods section. While the kinetically lumped model may not extrapolate well across species due to interspecies differences in tissue blood flows and volumes, the PBPK models with a full structure may perform better due to stronger similarity of tissue composition compared to tissue kinetics between species. By assuming identical model structure across species and species independence of Kpus or scaling factors, the refined model and estimated parameters using preclinical data can be applied to predict human PK. For prediction of clinical PK from pre-clinical data, the use of empirical scaling factors derived by model fitting to animal PK has to assume that these scaling factors are species independent.

The aim of this work was to develop and evaluate a strategy for the extrapolation of non-clinical PK to humans using the aforementioned simplified PBPK models and subsequently improving the quality of predictions used to support candidate drug selection and internal decision making during preclinical development. Herein the current paper presents a proof-of-concept work using diazepam, midazolam and basmisanil as model drugs for which IV PK data in rats, monkeys and humans were available for the fitting and evaluation of the translation to man.

4.3. Methods

The proposed models were applied to three compounds: diazepam, basmisanil, midazolam. Criteria for compound selection were the availability of intravenous plasma concentrations profiles after intravenous administration in rats, monkeys and humans and the availability of relevant *in vitro* information for Kpu-predictions (sensitivity to LogP, fup and B:P).

4.3.1. Experimental data

Physicochemical and in vitro data

Physicochemical properties and in vitro data for the 3 weak basic compounds are listed in Table 4.1. Molecular weight, pKa, and LogP data were extracted from the literature. fup and BP were measured in-house (details of experiments can be found in Appendix A3.1-A3.2).

Table 4.1. Physicochemical properties, in vitro PK data for the 3 compounds

Drug	Molecular weight [g/mol]	pKa	LogP	Rat		Monkey		Human	
				fup	BP	fup	BP	fup	BP
Diazepam	284.7[a]	3.4[a]	2.82[a]	0.1	0.836	0.084	0.606	0.0087	0.559
Midazolam	325.8[a]	5.88[a]	3.13[a]	0.054	0.742	0.0603	0.594	0.055	0.57
Basmisanil	445.5[a]	2.07[b]	1.86[b]	0.181	0.99	0.065	0.59	0.055	0.59

pKa: acid dissociation constant; LogP: n-octanol/water partition coefficient; fup: fraction unbound in plasma; BP: blood-plasma ratio
[a] From [24]; [b] From [25]; fup and BP are from internal data

In vivo pharmacokinetic data

For diazepam and midazolam, the literature was searched for PK studies where plasma (or blood or serum) concentrations were reported following an intravenous administration in rats (Wistar, Sprague Dawley and Holtzman strains), Cynomolgus monkeys, and humans (Table 4.2). These studies provided a mixture of average profiles and individual concentration time profiles which were digitized using WebPlotDigitizer (version 4.2, https://automeris.io/WebPlotDigitizer). For basmisanil, data were collected from Roche-internal PK studies (Appendix A3.2) and these were historical data subject to retrospective analysis. Summaries of the pharmacokinetic data available for each compound are shown in Table 4.2.

Table 4.2. Summary of the animal and clinical IV studies available

Drug	Species	Study	Number of PK profiles	IV dose (length of infusion)	Reported total clearance (s.d.) (plasma or blood)	Volume of distribution (s.d)	fup and BP (s.d)	fe (%)	Reference
Diazepam	Rat	1	1 average of 10 male Wistar	4mg/kg (bolus);	1.1 (0.2) L/h (plasma)	1.3 (0.2) L	fup: 0.036 (0.009); BP:1.06 (0.04)		[26]
		2	1 average of 4 male Wistar	1.2mg/kg (bolus)	1.2 (0.03) L/h (plasma)	1.4 (0.1) L	Fup:0.14 (0.003); BP:1.037 (0.007)		[27]
		3	1 average of 24 male Sprague Dawley	1mg (5min infusion)	N/A (intrinsic clearance: 400 ml/min)	N/A	Fup:0.15; BP:1		[28]
		4	2 average of 5 middle-aged male Wistar and of 5 old male Wistar	5mg/kg (bolus)	1.1 L/h for middle-aged rat, 3.3 L/h for old rats (plasma)	1.2 L for middle-aged rat, 13 L for old rats	N/A		[29]
		5	1 average of 3 male Wistar	5mg/kg (bolus)	3.9 (0.5) L/h (blood)	1.1 (0.2) L for	fup:0.137 (0.011); BP:0.38 (0.07)		[30]
	Monkey	1	2 average of 6 young cynomolgus and 6 aged cynomolgus	0.04mg/kg (bolus)	6.6 (0.6) L/h for young monkeys; 4.7 (0.8) L/h for aged monkeys (plasma)	4.1 (1.7) L/ for young monkeys; 5.8 (2.1) for aged monkey L	Fup:0.0843 (0.0111) for young monkeys; 0.0780 (0.0103) for aged monkeys; BP: 0.662	0.05	[31]
	Human	1	1 average of 6 male and 5 female	0.15mg/kg (1min)	2.1 (0.3) L/h (plasma)	87 (2) L	N/A		[32]
		2	4 average : 1 young male, 1 elderly male, 1 young female, 1 elderly female	0.15mg/kg (0.375min)	1.3 to 1.7 L/h (plasma)	80 to 161 L	fup: 0.009 to 0.027		[33]
		3	1 average of 6 males	10mg (2min)	1.2 (0.4) L/h (plasma)	63 (19) L	N/A		[34]
		4	1 individual (healthy male)	0.1mg/kg (2min)	1.6 (0.5) L/h (plasma)	66 (13) L	fup: 0.026 (0.01); BP: 0.58 (0.15)		[30]
		5	1 average of 10 young males	0.1mg/kg (2min)	1.6 (0.2) L/h (plasma)	79 (20) L	fup: 0.032 (0.008); BP:0.58 (0.11)		[35]
		6	4 individuals (healthy males)	10mg (bolus)	3.3 to 4.0 L/h (blood)	105 to 174 L	N/A		[36]
		7	23 individuals (healthy males/females)	0.15mg/kg (0.375min)	1.8 L/h (plasma)	N/A	fup: 0.015 ; BP:0.65		[33, 37, 38]
Midazolam	Rat	1	1 average of 10 male Wistar	5mg/kg (bolus)	1.5 (0.04) L/h (plasma)	1.0 (0.2) L			[39]
		2	2 average of 3 male Sprague Dawley and of 3 Sprague Dawley	2.5mg/kg (bolus)	N/A	N/A			[40]
		3	1 average of 6 male Holtzman	0.55mg/kg (bolus)	0.78 (0.19) L/h (plasma)	0.95 (0.36) L			[41]
		4	1 average of 8 male Wistar	10mg/kg (5min)	0.9 (0.03) L/h (plasma)	0.385 (0.025) L			[42]
		5	1 average of 10 male Wistar	0.1mg/kg (bolus)	0.63 (0.10) L/h (plasma)	0.33 (0.16) L			[43]
		6	1 average of 7male Wistar	10mg/kg (15min)	1.16 (0.04) L/h (plasma)	0.49 (0.018) L	fup:0.031 (0.001)		[44]
	Monkey	1	1 average of 4 male cynomolgus	0.3mg/kg (bolus)	2.5 (0.32) L/h (plasma)	3.2 (0.16) L			[45]
		2	1 average of 3 male cynomolgus	1mg/kg (bolus)	4.6 (0.63) L/h (plasma)	8.2 (1.9) L			[46]
		3	1 average of 3 male cynomolgus	1mg/kg (15min infusion)	6.3 (0.77) L/h (plasma)	N/A			[46]
	H n	1	1 average of 6 male	0.075mg/kg (bolus)	17 (2.6) L/h (plasma)	48 (11) L			[47]

Table 4.2. Summary of the animal and clinical IV studies available (continued)

Drug	Species	Study	Number of PK profiles	IV dose (length of infusion)	Reported total clearance (s.d.) (plasma or blood)	Volume of distribution (s.d)	fup and BP (s.d)	fe (%)	Reference
		2	1 individual (male)	5mg (bolus) + 4mg (120min infusion)] + [3mg (bolus) + 5.4mg (120min infusion)] + [2.2mg (bolus) +6.8mg (120min infusion)	17 (2.6) L/h (plasma)	48 (11) L			[47]
		3	1 average of 8 male	0.15mg/kg (bolus)	21 (5.2) L/h (plasma)	50 L			[48]
		4	6 individuals (healthy males)	0.15mg/kg (bolus)	13-28 L/h (plasma)	39-68 L			[48]
		5	1 average of 8 males/females	0.1mg/kg (15min infusion)	31 (0.18) L/h (plasma)	60 (7) L			[42]
		6	1 average of 7 males	7.5mg (bolus)	27 (1.9) (plasma)	112 (7) L			[49]
		7	1 average of 6 males/females	1mg (bolus)	33 (6.8) L/h (plasma)	83 (41) L			[50]
		8	2 individuals (females)	5mg (bolus)	26; 17 L/h (serum)	78 ; 45 L			[51]
		9	1 average of 9 males/females	2mg (bolus)	33 (12) L/h (plasma)	136 (57) L			[52]
		10	2 average of 7 females and of 7 males	0.15mg/kg (2min infusion)	20-33 L/h (plasma)	70-218 L			[53]
		11	1 average of 7 females	0.15mg/kg (bolus)	31 (4.5) L/h (plasma)	79 (18) L			[54]
		12	1 average of 11 males/females	0.1mg/kg (1min infusion)	46 (5.0)L/h (plasma)	162 (12) L			[32]
Basmisanil	Rat	1	2 individuals (male Wistar)	4.3 mg/kg (bolus)	0.77 L/h	0.69 L		0.2	Internal source
	Monkey	1	3 individuals (male cynomolgus)	1 mg/kg (bolus)	2.0 (0.47) L/h	10 (6.8) L			Internal source
	Human	1	6 individuals (healthy males)	0.001 mg/kg (15min)	4.4 (1.2) L/h	49 (5.8) L		0.4	Internal source

4.3.2. Fitting and extrapolation approach of the simplified PBPK models

Model structures

The approaches for simplifying PBPK models and the investigated model structures have been described in detail in the previous chapter (Chapter 3). These models showed a good ability to fit data in rats and in humans for diazepam. Figure 4.1 shows the lumped 3-compartment model with one central and two peripheral compartments and the 14-compartment model with common Kpus or common scalars. In the first model, tissue compartments are lumped together based on similar tissue time constants [22] In the latter model, it is assumed that certain tissues either share the same Kpu or share a common scalar (bias from the Kpu predicted from the R&R model). The tissue

commonality was determined based on clustering analysis (using hierarchical or k-means method) of either rat tissue composition data or of a dataset of imputed normalized *in vivo* rat Kps. In the kinetically lumped approach, clearance was assumed to be from the plasma/blood compartment whereas in the 14-compartment model clearance can be considered to be from eliminating organs such as the liver and kidney. In the current work, these models are applied for translation between species. Grouping of tissues that are common in the 14-compartment models is detailed in Table 4.3.

Figure 4.1: Structure of the investigated models: lumped 3-compartment model (left) and 14 compartment model with common Kpus or common scalars (right**)**

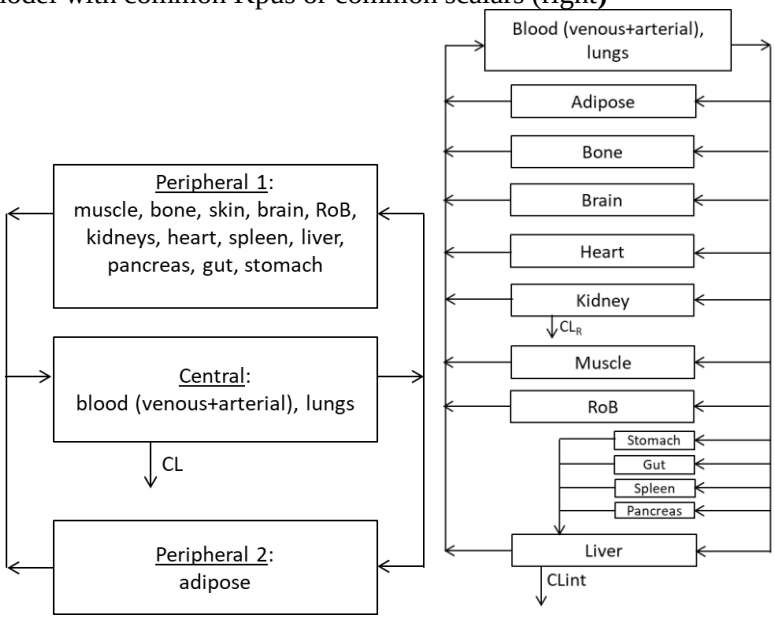


Table 4.3: Models and details of tissue grouping

Model	Model description	Tissue grouping (by lumping or clustering)
1	Lumped 3 compartment model	*Kpu1: blood, lungs, kidneys, heart, spleen, liver, pancreas, gut, stomach, bone, brain* *Kpu2: adipose, muscle, rest of body* *Kpu3: skin*
2A	14-compartment PBPK model with 3 common Kpus (H)	*Kpu1: adipose, bone, brain, muscle, pancreas, muscle, rest of body* *Kpu2: lung, gut, stomach, kidney, heart, spleen, liver* *Kpu3: skin*
2B	14-compartment PBPK model with 4 common Kpus (H)	*Kpu1:bone, brain, muscle, pancreas, muscle, rest of body* *Kpu2: lung, gut, stomach, kidney, heart, spleen, liver* *Kpu3: skin* *Kpu4: adipose*
2C	14-compartment PBPK model with 3 common Kpus (Km)	*Kpu1: adipose, bone, brain, muscle, pancreas, rest of body* *Kpu2: kidney, spleen, liver* *Kpu3: skin, lung, gut, stomach, heart*
2D	14-compartment PBPK model with 4 common Kpus (Km)	*Kpu1: bone, brain, muscle, pancreas, rest of body* *Kpu2: kidney, spleen, liver* *Kpu3: skin, lung, gut, stomach, heart* *Kpu4: adipose*
2E	14-compartment PBPK model with 3 common Kpus (ss)	*Kpu1: adipose, bone, brain, muscle, pancreas, muscle, rest of body, skin, gut, stomach, heart, spleen* *Kpu2: kidney, liver* *Kpu3: lung*
2F	14-compartment PBPK model with 4 common Kpus (ss)	*Kpu1: adipose, bone, brain, muscle, pancreas, muscle, rest of body, skin, gut, stomach, heart, spleen* *Kpu2: kidney* *Kpu3: liver* *Kpu4: lung*
3A	14-compartment PBPK model with 3 scalars (H)	*SF1: adipose, bone, brain, muscle, pancreas, muscle, rest of body* *SF2: lung, gut, stomach, kidney, heart, spleen, liver* *SF3: skin*
3B	14-compartment PBPK model with 4 scalars (H)	*SF1: bone, brain, muscle, pancreas, muscle, rest of body* *SF2: lung, gut, stomach, kidney, heart, spleen, liver* *SF3: skin* *SF4: adipose*
3C	14-compartment PBPK model with 3 scalars (Km)	*SF1: adipose, bone, brain, muscle, pancreas, rest of body* *SF2: kidney, spleen, liver* *SF3: skin, lung, gut, stomach, heart*
3D	14-compartment PBPK model with 4 scalars (Km)	*SF1: bone, brain, muscle, pancreas, rest of body* *SF2: kidney, spleen, liver* *SF3: skin, lung, gut, stomach, heart* *SF4: adipose*
3E	14-compartment PBPK model with 3 scalars (ss)	*SF1: adipose, bone, brain, muscle, pancreas, muscle, rest of body, skin, gut, stomach, heart, spleen* *SF2: kidney, liver* *SF3: lung*
3F	14-compartment PBPK model with 4 scalars (ss)	*SF1: adipose, bone, brain, muscle, pancreas, muscle, rest of body, skin, gut, stomach, heart, spleen* *SF2: kidney* *SF3: liver* *SF4: lung*

H: hierarchical clustering on rat tissue composition data; Km: k-means clustering on rat tissue composition data; ss: clustering on in vivo rat Kps data; SF: scaling factor

Optimization of the distribution with preclinical data and model evaluation

The performance of the developed models in preclinical species was assessed by comparing the simulated concentration-time profiles with experimental data of intravenous studies from the literature (diazepam, midazolam) and from internal studies (basmisanil) (Table 4.2).

For diazepam and midazolam, as the data were coming from various studies with different designs (Table 4.2), a model-based analysis was first performed: rat and monkey data were fitted using compartmental analysis in order to estimate the Vss and CL of the observed data using the first order estimation with interaction (FO-I) and the first order conditional estimation with interaction (FOCE-I) methods in NONMEM 7.3 (ICON Development Solutions, Hanover, MD, USA) [55]. For basmisanil, the Vss and CL of the observed data were obtained by noncompartmental analysis of the PK studies. These 'observed' Vss and 'observed' CL were considered as the reference values in this study. The error model was defined as a proportional error. For all compounds, the different simplified PBPK models were then fitted to the rat and monkey PK data while fixing the clearance to the value estimated from the compartmental analysis (diazepam and midazolam) or the value obtained by noncompartmental analysis reported in the study (basmisanil) in order for the model performance to be comparable between the different investigated models. Estimating clearance would otherwise influence the estimation of Kpu parameters. FO-I, FOCE-I and stochastic approximation expectation maximization (SAEM) methods were used for the fitting, SAEM was used when either FO-I or FOCE-I failed to converge due to model complexity and data sparsity. A study comparing performance of FOCE and EM methods reported that EM methods such as SAEM were more robust for more complex models and sparse data [56]. Estimated parameters were mu-referenced (log-transform of a Gaussian random vector to meet with constraints of positivity) to improve the computational efficiency [57]. The AUTO option was first switched on (AUTO=1) allowing NONMEM to determine the best options for Monte Carlo using the SAEM method, and optimized options were tested and used if significant improvement was observed (decrease of BIC, convergence, better precision of estimates, etc.). Initial values for lumped Kpu and common Kpu parameters were calculated as a weighted sum of all the tissue Kpus (predicted by the Rodgers and Rowland model) that are lumped or clustered together. Similarly, by expecting Kpu values predicted using the R&R model to be true, initial values of the common scalars were chosen to be 1. After

fitting the simplified models to the preclinical data, several criteria were used to evaluate the model performance and determine its suitability for extrapolation to human. These criteria included: Bayesian information criterion (BIC), convergence, goodness-of-fit plots, precision of estimates, physiological plausibility of Kpu estimates, and closeness of estimated Vss value to the 'observed' Vss. For each compound, models which match all the above criteria are considered the 'best' models. Biological plausibility could be assessed based on the analysis of a dataset of compiled rat experimental tissue partitioning coefficients from the literature where it was found that 95% of these Kpu predictions would be within around 21-fold error. There is a higher probability of Kpu predictions being much closer than 21-fold error and obtaining predicted values beyond are less likely to occur (further details can be found in Appendix A3.3).

The reader is referred to Chapter 3 for extensive details about the models. For each model, blood Vss (Vss,b) in the preclinical species was calculated as follows:

For the lumped model:

$$Vss,b = V_{central} + V_{p1} + V_{p2} \quad \textbf{Eq. 4.1}$$

where $V_{central}$= $V_{arterial}$+V_{venous}+ $Kb_{central}$*V_{lung} with $Kb_{central}$ being the blood tissue-to-plasma partitioning coefficient of the central compartment. V_{p1}= (ΣV_i-V_{liver}-V_{kidney}) + V_{liver} (1-CL_H/Q_{liver}) + V_{kidney} (1-CL_R/Q_{renal}). Q_i and V_i are the blood flow and volume of the i^{th}-tissue lumped in the peripheral 1 compartment. V_{p2}= ΣV_i. Q_i and V_i are the blood flow and volume of the i^{th}-tissue lumped in the peripheral 2 compartment model.

For the 14-compartment PBPK model with common Kpus or Kpu scalars:

$$Vss,b = V_{venous} + V_{arterial} + \sum Kb_i \cdot V_i \cdot (1 - ER_i) \quad \textbf{Eq. 4.2}$$

where V_i is the volume of the i^{th}-tissue, and Kb_i and ER_i are respectively its tissue-to-blood partitioning coefficient and extraction ratio. Kb can be determined from Kpu:

$$Kb = Kpu \cdot \frac{fup}{BP} \quad \textbf{Eq. 4.3}$$

Where fup and BP are the fraction unbound in plasma and the blood-to-plasma ratio.

Intravenous prediction in human of simplified model

The best model for fitting was then used for extrapolation to human. For diazepam and midazolam, as the clinical observations were from various PK studies with different

doses and lengths of infusion, it was not feasible to use a single visual predictive check (VPC) plot to compare predicted *vs* observed data. Therefore, in order to compare performance of the different models and to remove differences in study design, a common study design was simulated using the R package 'RxODE' [58]. Observations in human were simulated from a reference model (structure, parameter estimates, and associated uncertainty) which was the compartmental model (two or three compartments) fitted to the diazepam or midazolam data. 1,000 drug concentration-time courses were simulated for a single IV dose and infusion rate to achieve a steady state plasma concentration determined by the drug elimination rate: a 16.1h infusion of 10 mg for diazepam and a 0.01h infusion of 5 mg for midazolam were chosen to illustrate the different kinetic phases. For basmisanil, CL was fixed to the value calculated from non-compartmental analysis in the clinical study (7.26 L/h) and the simulations were performed similarly to the design in the reported clinical study (Table 4.2) and predicted concentrations were compared to observed data. The extrapolation for each approach and model was then performed as follows:

In the kinetically lumped model and the 14-compartment PBPK model with common Kpus: the assumption is that tissue compositions are similar and Kpus are the same across species. While adjusting for the differences in species physiology and species fup and BP, Kp human were estimated by assuming Kpu_{human} is equal to $Kpu_{species}$ where species correspond to rat or monkey:

$$Kp_{human} = Kpu_{species} \cdot fup_{human} \quad \textbf{Eq. 4.4}$$

In the 14-compartment PBPK model with common scalars: the assumption is that the same bias from R&R predicted Kpu (KpuRR) exist across species and no assumptions on interspecies similarity of tissue composition or Kpus are made. While adjusting for the differences in species physiology and species fup and BP, Kp_{human} were estimated by assuming Kpu_{human} is equal to $KpuRR_{human}$* $SF_{speceis}$ where $SF_{species}$ correspond to the scalar estimated in rat or monkey and KpuRR is the tissue to plasma unbound partition coefficient calculated using the equations developed by Rodgers and Rowland [20, 21]:

$$Kp_{human} = \left(KpuRR_{human} \cdot SF_{species}\right) * fup_{human} \quad \textbf{Eq. 4.5}$$

Interspecies extrapolation was conducted using the species-specific tissue blood flow rates and tissue volumes and tissue composition when using the models with common scalars (**Table A3.1**-**Table A3.6**).

Assessment of human prediction accuracy /performance assessment of extrapolation

For the studied compounds, simulation of intravenous concentration-time profiles in human using the best models were compared against the simulation using the traditional WBPBPK modelling approach with 16 compartments and tissue Kpus calculated from the Rodgers and Rowland equations [20, 21].

For each best model, the Vss in human was calculated using equations Eq. 4.1 and Eq. 4.2 with human physiological parameters. For the traditional WBPBPK modelling approach, the Vss,b in human was calculated using equation Eq. 4.2. These Vss,b values were assessed against the 'observed' Vss (simulated from the fitted empirical model) for diazepam and midazolam or the Vss,b reported in clinical trials (noncompartmental analysis) for basmisanil. When Vss in plasma (Vss,p) was reported in the literature study, plasma Vss was converted to blood Vss ($\mathrm{Vss,b} = \mathrm{Vss,p} \cdot \mathrm{BP}$).

Additionally, simulated concentration profiles from each best model ('middle-out' approach) and from the traditional WBPBPK model with Kpus predicted from the R&R model ('bottom-up' approach) were compared against the median 'observed' concentration profile by calculating the root-mean-square error (RMSE, Eq. 4.6) to assess the accuracy of the predictions, with lower RMSE value representing greater precision of the model:

$$RMSE = \sqrt{\frac{\sum(PRED_t - OBS_t)^2}{n}} \qquad \textbf{Eq. 4.6}$$

Where OBS_t is the median 'observed' concentration value at time t and $PRED_t$ is the median simulated concentration at time t, and n is the number of time points evaluated.

4.5. Results

The three studied compounds were weak bases and among them, diazepam (DZP) was selected as the first compound for optimization due to the number of PK studies in humans (n=36), monkeys (n=2) and rats (n=5). Based on the preliminary results of the fitting of the models to rat data of DZP, it was decided to reduce the number of models to be further tested for the fitting of DZP model to monkey data. The number of models tested was also reduced for the fitting of midazolam and basmisanil. The best simplified models fitted in rat and monkey for all 3 compounds are summarised in **Table 4.4**.

Table 4.4: Summary of the best simplified models fitted in preclinical species for each of the 3 compounds

	Diazepam	Midazolam	Basmisanil
Rat	3C, 3D	3D	None
Monkey	3C, 3D	3C, 3D	None

Abbreviations for model are defined in Table 4.3.

Results shown in this work are based on using FO-I or FOCE-I. It should be noted that the algorithm sometimes failed to converge possibly due to the sparse data or the complexity of the models in our analysis. Thus, estimation of distribution parameters (scalars or clustered Kpu) using SAEM method in NONMEM was also considered when FOCE-I failed. The FO-I or FOCE-I algorithm generally performed well in this work except in some cases where convergence failed or estimation was biased. SAEM-I/IMP could improve the estimation accuracy or convergence success at the expense of a longer running time and more parameter tuning (results not shown).

4.5.1. Compound 1: Diazepam

Rat

The PK data for diazepam in rat were extracted from literature studies (Table 4.2). Using a fixed blood clearance of 0.915L/h in rat (obtained from fitting a 3-compartment model), the concentration-time profiles were analysed with all the investigated simplified models. The estimated parameters are listed in **Table A3.9** and the fitted profiles of the best models are shown in **Figure A3.5-Figure A3.6**. All model parameters were estimated with high precision (*RSE*% < 50%) except for the Kpu of the central compartment in the lumped model where it was found to be 2730%. This lack of precision

is surprising but may be due to the insufficient flexibility of the model structure with the lung Kpu being highly influenced by the assumptions made around the other tissue kinetics and the non-applicability to diazepam i.e., not a compound with low Kpus. The SAEM algorithm led to more precise estimates (results not shown). Among the compartment models, several models could recapture an estimated Vss,b close to the Vss,b observed (0.91 L) in rat within- 20% error (models 3C and 3E), within 25% error (3D) and within 30% error (models 2C and 2D) (see **Table A3.9** for details). Based on the analysis of 124 compounds where it was found that 95% of Kpu predictions would agree with experimental values within a factor of 21 (see Appendix A3.4), models 3C and 3D have plausible estimated values for Kpus. Additionally, estimated rat Kpu and derived tissue concentrations were biologically plausible as the values obtained were close to tissue Kpu measurements in rat available from the literature (**Table A3.7**). Although differences exist between the estimated and observed Kpu values, tissue concentrations were still well predicted. Thus, the two models 3C and 3D were the best for fitting the diazepam data in rats (Table 4.4) and were subsequently selected for extrapolation to human (Figure 4.2). The simulated human PK profiles from models 3C and 3D showed a major improvement compared to the simulation using the traditional WBPBPK modelling approach (RMSE 4.9 and 10.2 vs 28). Similarly, the Vss,b in human estimated from the model 3C (132 L) and 3D (97 L) were within 1.1 and 1.5-fold error of the Vss,b observed (152 L) which is considerably better than the Vss,b estimated from the traditional WBPBPK modelling approach (41 L, 3.7-fold error). Thus, the approach using the 14 compartments with common scalars (especially models 3C and 3D) seemed more promising for translation and all the derived models from this approach were selected for application to the rest of the analysis.

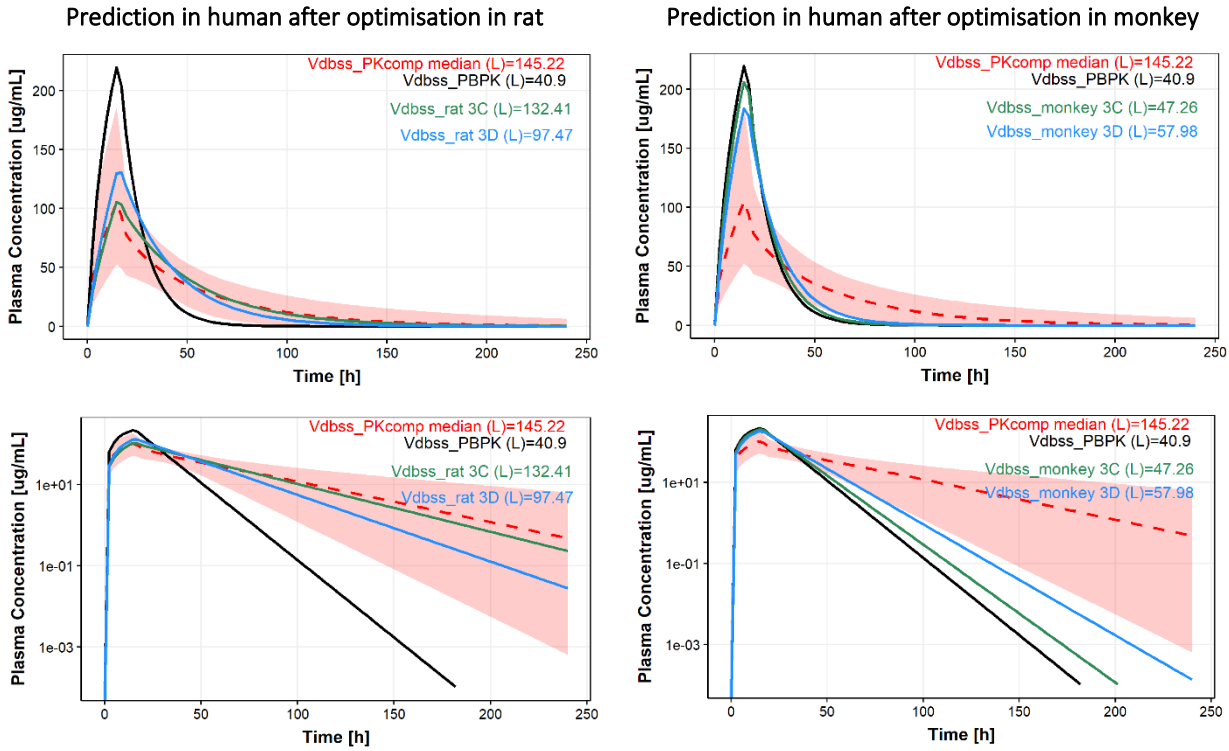
Table 4.5: Best models after fitting all models to diazepam data in rat and the associated tissue Kpu estimates

Model	Kpu													
	Adipose	Bone	Brain	Gut	Heart	Lung	Kidney	Liver	Muscle	Skin	Stomach	Spleen	Pancreas	RoB
3C	183	43.2	90.8	3.11	1.38	2.02	281	292	30.8	4.61	3.11	162	98.7	30.8
3D	14.6	41.6	87.2	58.8	26.1	38.3	341	353	29.6	87.5	58.8	196	94.8	29.6

Monkey

The PK data for diazepam in monkey were extracted from literature studies (Table 4.2). Using a fixed blood clearance of 9.97L/h in monkey (obtained from fitting a 2-compartment model), the concentration-time profiles were analysed with the approach of simplified models with scalars (models 3A-F). The estimated parameters are listed in **Table A3.10** and the fitted profiles of the best models are shown in **Figure A3.7**-**Figure A3.8**. No interindividual variability was estimated as there were only 2 subjects in the PK study. Models 3C and 3D presented parameter estimates with reasonable relative standard errors (*RSE*% < 50%). Models 3C and 3D recaptured an estimated Vss,b close to the Vss,b observed (11.1 L) within 20% error (see **Table A3.10** for details). Moreover, the obtained Kpu parameters of the two models were physiologically plausible, these models were the best for fitting the diazepam data in monkeys (Table 4.4) and were subsequently selected for extrapolation to human (Figure 4.2). The simulated human PK profiles from these selected models showed improvement compared to the simulation using the traditional WBPBPK modelling approach (RMSE 26 and 20 vs 28). The Vss,b in human estimated from models 3C (Vss,b=47 L) and 3D (Vss,b =57 L) were within 3.1-fold and 2.5-fold error of the Vss,b observed (Vss,b=152L), which were better than the Vss,b estimated from the traditional WBPBPK modelling approach (Vss,b=41L, 3.7-fold error).

Figure 4.2: Simulated human PK profiles of diazepam from the most suitable simplified PBPK models optimized in rat (left) and in monkey (monkey) *versus* traditional WBPBPK modelling approach



The dashed red line represents the median concentrations and the semi-transparent red field represents a simulation based 90% confidence interval for the median using the reference model (empirical 3-compartmental model). The solid lines represent the median concentrations using the traditional WBPBPK model approach (black), the model 3C (green) and the model 3D (light blue).

4.5.2. Compound 2: Midazolam

Rat

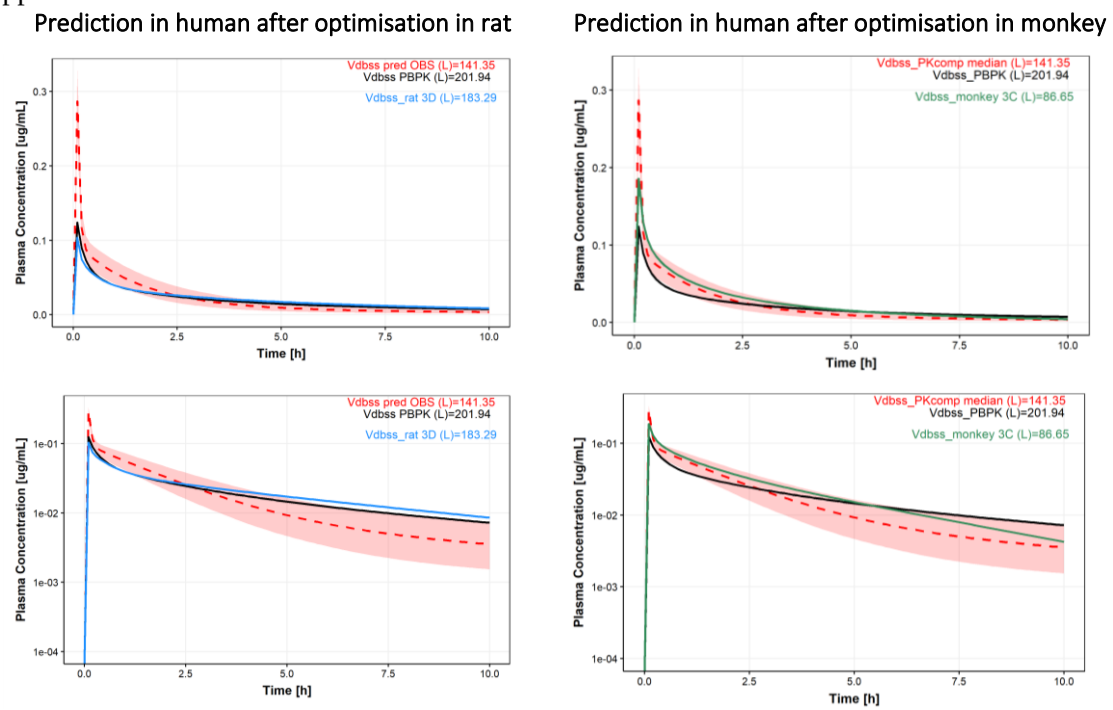
The PK data for midazolam in rat were extracted from literature studies (Table 4.2). Using a fixed blood clearance of 1.30 L/h (obtained from fitting a 2-compartment model), the concentration-time profiles were analysed with the models with scalars (models 3A-F). The estimated parameters are listed in **Table A3.11** and the fitted profiles of the best models are shown in **Figure A3.9**. In general, the RSE were low (<50%) with certain exceptions (see **Table A3.11** for details). Models 3C and 3D could capture an estimated Vss,b close to the Vss,b observed (0.64 L) within 20% error. Considering the different criteria (precision of estimates, physiological plausibility of Kpu estimates, and estimated Vss value), model 3D was the best for fitting the midazolam data in rats. Moreover, the estimated rat Kpu were biologically plausible as the values obtained were mostly within 2-fold of Kpu measurements in rat available from literature (**Table A3.8**).

Model 3D was subsequently selected for extrapolation to human (**Figure 4.3**). The simulated human PK profiles from model 3D was slightly better than the simulation using the traditional WBPBPK modelling approach (RMSE 0.021 vs 0.018). The Vss,b in human estimated from model 3D in rat (183 L, which is within a 1.3-fold error) was also closer to the Vss,b observed (141 L) compared to the Vss,b estimated from the traditional WBPBPK modelling approach (202 L, which is within a 1.5-fold error).

Monkey

The PK data for midazolam in monkey were extracted from literature studies (Table 4.2). Using a fixed blood clearance of 6.4 L/h (obtained from a 2-compartmental model), the concentration-time profiles were analysed with the models with scalars (models 3A-F). The estimated parameters are listed in **Table A3.12** and the fitted profiles of the best models are shown in **Figure A3.10**. In general, the RSE were low (<50%) with certain exceptions (see **Table A3.12** for details). Several models could capture an estimated Vss,b close to the Vss,b observed (8.8 L) within 20% error: (model 3D) and within 30% error (model 3C). The physiological plausibility of Kpu estimates of these 2 models was reasonable. Considering the different criteria (precision of estimates, physiological plausibility of Kpu estimates, and estimated Vss value), models 3C and 3D were the best for fitting the midazolam data in monkeys and were selected for extrapolation to human (**Figure 4.3**). It should be noted that the traditional WBPBPK modelling approach for midazolam gave good prediction in human (which was not the case for diazepam). The simulated human PK profiles from models 3C and 3D were comparable to the simulation using the traditional WBPBPK modelling approach (RMSE 0.018 and 0.019 vs 0.018). The Vss,b in human estimated from model 3C in monkey (131 L, which is within a 1.1-fold error) was similar to the Vss,b observed (141 L) and better than the Vss,b estimated from the traditional WBPBPK modelling approach (202 L, which is within a 1.5-fold error). The Vss,b in human estimated from model 3D in monkey was not better but still reasonable within 2.2-fold-error (302 L).

Figure 4.3: Simulated human PK profiles of midazolam from the most suitable simplified PBPK models optimized in rat (left) and in monkey (monkey) *versus* traditional WBPBPK modelling approach



The dashed red line represents the median concentrations and the semi-transparent red field represents a simulation based 90% confidence interval for the median using the reference model (empirical 3-compartmental model). The solid lines represent the median concentrations using the traditional WBPBPK model approach (black), the model 3C (green) and the model 3D (light blue).

4.5.3. Compound 3: Basmisanil

Rat

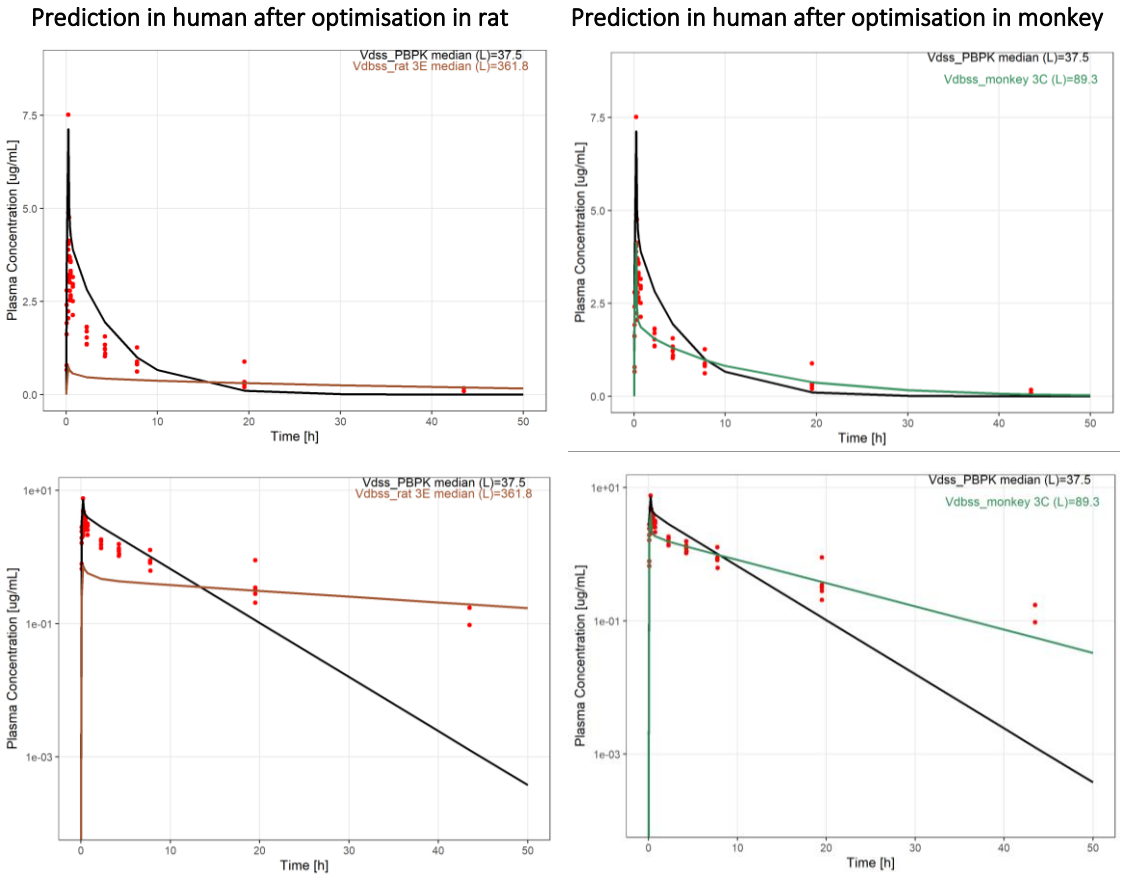
The PK data of basmisanil in rat were obtained from an internal PK study (Table 4.2). Using a fixed blood clearance of 0.824 L/h (obtained from a 2-compartment model), the concentration-time profiles were analysed with the approach of simplified models with scalars. The estimated parameters are listed in **Table A3.13**. No interindividual variability was estimated as there were only 2 subjects in the PK study. Model 3F did not converge. Although some models were able to produce parameter estimates with reasonable relative standard errors (RSE<50%), none of them provided an estimated Vss,b close to the Vss,b observed (0.91 L) within 20% error. The models' performance could be limited by the lack of measurements in the terminal phase of the PK data. In the absence of any additional data to improve the optimization in rat, the decision was to perform the prediction in human using the traditional PBPK model approach (**Figure 4.4**). Extrapolation should not be attempted even with the "best" worst model (model 3E) since this would cause incorrect predictions as illustrated in **Figure 4.4**. The simulated human

PK profiles from the traditional WBPBPK modelling approach was within a 3-fold error of the Vss,b in human (38 L vs 84 L). Therefore, the traditional WBPBPK modelling approach gave a good prediction in human for basmisanil.

Monkey

The PK data for basmisanil in monkey were obtained from a PK study of 3 subjects (Table 4.2) and using a fixed blood clearance of 3.82 L/h (obtained from a 2-compartment model), the concentration-time profiles were analysed with the approach of simplified models with scalars. When using FOCE-I, convergence failure was encountered with most of the models. The parameters estimated using SAEM are listed in **Table A3.14**. None of the models were able to produce parameter estimates with reasonable relative standard errors (RSE<50%), also, none of them provided an estimated Vss,b close to the Vss,b observed (24 L) within 20 or 30% error. This could be due to the heterogeneity of PK profiles in the 3 monkeys as the proportional error were very large (>68%). Similar to the case of basmisanil with rat data, the decision was to perform the prediction in human using the traditional PBPK model approach in the absence of any additional data to improve the optimization in monkey (**Figure 4.4**). Although the extrapolation using the "best" worst model (model 3C) could give good prediction in human as illustrated in **Figure 4.4**, there is no clear rationale and high uncertainty for this practice. Nonetheless, the traditional WBPBPK modelling approach gave a good prediction of PK profile in human for basmisanil within a 3-fold error of the Vss,b in human (38 L vs 84 L).

Figure 4.4: Simulated human PK profiles of basmisanil (IV infusion of 0.1mg for 15 min)



The red dots represent observed data. The solid lines represent the median concentrations using the traditional WBPBPK model approach (black), the model 3C (green). and the model 3E (brown).

4.6. Discussion

In this work, several approaches and PBPK models were presented for cross-species extrapolation (**Table 4.6**). The Kp prediction method used here was the Rodgers and Rowland method [20, 21], but other Kp methods could also be used [4, 59-61]. In general, they strongly rely on drug properties such LogP, pKa, fup and BP, which are expected to be well- characterized. A low confidence in the measurement of these values may subsequently prevent parameter estimation from *in vivo* data and increase uncertainty around the predictions in human [23].

Table 4.6: Assumptions, advantages, and disadvantages of the different simplified PBPK models for interspecies extrapolation

Models	Assumptions	Advantages	Disadvantages
Lumped 3 compartment model	Lumping based on similar tissue kinetics (blood-flows and volumes) and low tissue Kpus Similar lumped model structure between species Drug distributes homogenously among the lumped tissue at the same distribution rate Clearance implemented from plasma/blood	Easy to implement as closed form solution Possible to use computer-intensive methods for estimation Possible to derive plasma and lung profiles Meaningful use of plasma volume and prediction of drug distribution in groups of lumped tissues while only analysing plasma time-course data	May be difficult to extrapolate due to species differences in tissue kinetics Limited in use to compounds with relatively low Kpu values
14 compartment PBPK model with common Kpus	Grouping of tissue based on steady state similarities in drug partitioning Flow-limited distribution Identical model structure across species Species independence of Kpus	No assumption on tissue kinetics (except for lung) Better extrapolation potentially due to stronger similarity of tissue composition compared to tissue kinetics between species Possible to derive individual tissue profiles Applicable for a wider set of compounds	Longer running time due to model high dimensionality May be more difficult to estimate parameters with limited data
14 compartment PBPK model with common scalars	Grouping of tissues based on steady states similarities in drug partitioning Flow-limited distribution Identical model structure across species Species independence of scaling factors	No assumption on tissue kinetics (except for lung) Better extrapolation potentially due to stronger similarity of tissue composition compared to tissue kinetics between species Possible to derive individual tissue profiles Applicable for a wider set of compounds More flexibility by assuming similar bias from predicted Kpus (by R&R equations) across species	Long running time due to model high dimensionality May be more difficult to estimate parameters with limited data

In the kinetic lumping approach, one model structure (lumped model in man) was assumed to be common across species (**Table 4.6**). The lumped model was based on human as it is the target species for extrapolation in this work. This approach assigned similar tissue kinetics between species which potentially alter the mechanistic realism of the model in preclinical species. For most tissues, the ratio of blood flows to tissue volumes were similar between species except for adipose and skin (**Table A3.4**-**Table A3.6**). In human, skin and adipose volumes account for 3.9 % and 24 % of total tissue volume respectively, whereas they account for 19 % and 8.4 % respectively in rats, and for 11 %and 3 % in monkeys. Adipose tissue can be key for drug distribution in human due to its size and characteristics (lowest tissue water, high neutral lipids and low phospholipids) and thus having adipose as a separate compartment can help to better describe the PK in human. However, rats and monkeys are lean animals and adipose tissue do not represent such an important organ for drug distribution compared to other tissues. Optimization of the 'human' lumped model may limit its use and interpretation in preclinical species. Therefore, the cross-species extrapolation using this approach may not work well and be challenging especially for drugs that distribute largely in the adipose tissue, such as highly lipophilic compounds.

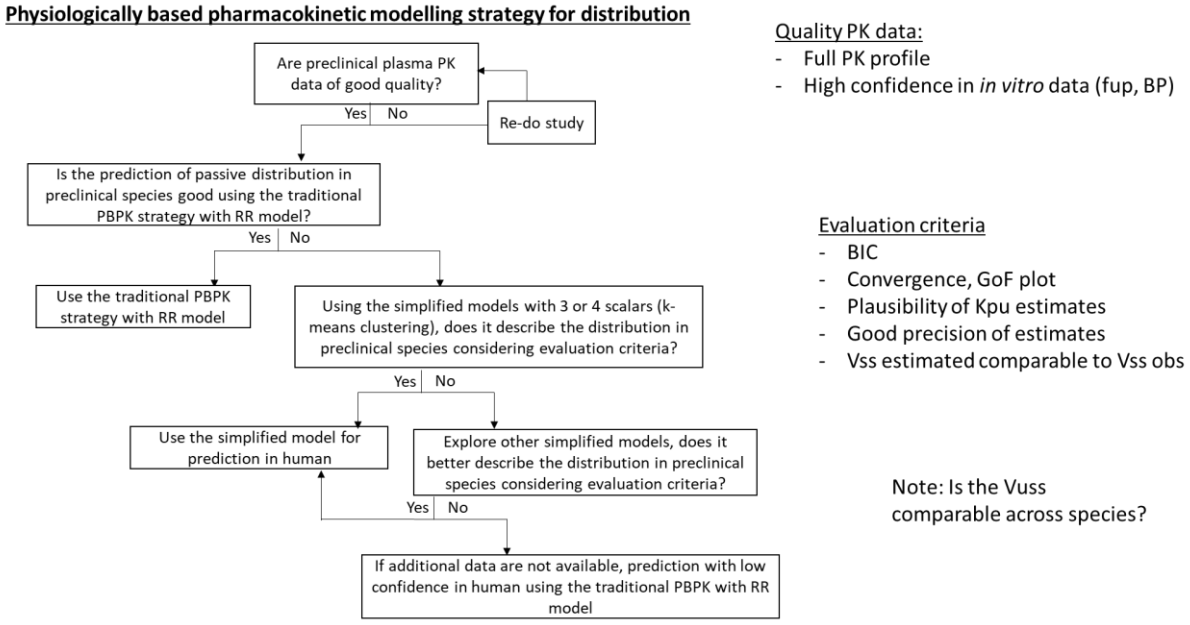
In the 14-compartment PBPK models approach with tissue commonalities, parameters could be separated into system- and drug-specific components which allowed the extrapolation of mechanistic knowledge from a reference species (here rat or monkey) to a target species (human) as in a traditional WBPBPK model. Species differences in organ size and blood flow rates were accounted for by extrapolation in all the models of this approach. In the models with common scalars, species differences in tissue composition are considered and different Kpu values are allowed for tissues whereas it is less flexible in the models with common Kpus (**Table 4.6**). Additionally, in the latter models, Kpus are assumed to be the same across species whereas in the former models it is assumed that the same bias from R&R Kpus exist across species. In the models with common scalars, the best models are able to provide physiologically relevant predictions of tissue distribution while analysing plasma profiles only. This was affirmed in the case of diazepam where actual tissue concentrations from rat studies were available in the literature (comparison can be found in Appendix A3.5).

Kp parameter values may be estimated from rat data and then used to predict human Kp values and drug disposition assuming Kpu are the same between species and

taking into account species differences [10]. Moreover, while some studies require plasma and various tissue concentrations in rats to estimate Kps and predict human PK [27, 28, 38, 62], the method proposed here only use plasma concentrations which are more readily available from animal studies.

The aim of this work was to propose a PBPK strategy for extrapolating drug distribution from preclinical species to human. Based on the findings obtained for these three compounds (lipophilic weak bases), a strategy is proposed in Figure 4.5.

Figure 4.5: Optimized PBPK modelling strategy for distribution



These three compounds were provided as examples on how human PK profiles could be predicted by optimizing simplified models in rats and monkeys. The predictive performance for extrapolation was compared retrospectively with the classical WBPBPK approach (with R&R Kpus) for diazepam and midazolam (weak bases). Sparse data and incomplete profiles resulted in parameter estimation issues for basmisanil and subsequently, it is not recommended to further extrapolate the "best" worst model as it could be inconsistent and impair predictions in human and the better choice is to use the traditional PBPK modelling approach with the R&R model (**Figure 4.5**). Generating good quality data should be a prerequisite for the extrapolation. When limited data are available at early sampling times, this could also affect the model optimization in preclinical species causing deviations in initial phase predictions in human. Nevertheless, the 14-compartment model assumes the lumped central compartment (blood and lungs) as the initial dilution space which increases accuracy and stability for fitting the initial

PK decline phases. This simplified PBPK approach reduced the number of parameters to be estimated from 16 (all tissues) to only three or four, resolving therefore parameter identifiability issues when trying to fit WBPBPK models to data (see Chapter 3). The insensitivity of one parameter for the overall PK profiles may explain the low precision in some of the models. Nevertheless, the high precision (low CV% values) for the parameter estimates in the best models in animals indicated good model performance, although fittings for some of the digitized data are only approximate. The presented approach can be combined with PBPK software such as Simcyp®, PK-Sim® or Gastroplus® by using the optimized parameters to derive the tissue Kpu values needed as inputs. For diazepam and midazolam, it was shown that as long as the volume of distribution (calculated using Eq. 4.2) and the clearance (here, it was fixed to the observed clearance value) were well characterized, the model described the overall profile in animals and would reasonably predict the concentration time profiles in human (better than the traditional WBPBPK modelling approach). Additionally, the models with scalars using k-means clustering generally performed well for fitting and extrapolation and can be tried first. In the case where these models do not describe the data well, other simplified PBPK models could be explored. If none of the models performed well in animals, as for basmisanil, the extrapolation could still be done using the traditional WBPBPK modelling approach although confidence in prediction accuracy will be low. It is however not recommended to use the poor fitting of simplified models to preclinical data, as the predictions in human are uncertain and could be poor. Variability and uncertainty in population physiology and drug parameters could potentially be integrated if they are known instead of only predicting an average human profile [63, 64].

This work has some limitations due to the assumptions and data sources. Relevant physiological data and evidence may be lacking to address the potential limitations of cross-species translation. Although monkey is generally considered to be the most suitable preclinical species, the case study using diazepam did not find greater similarity of monkey than rat for drug-distribution. This could be due to the lack of information regarding monkey tissue composition as the tissue composition data used for the monkey was a hybrid of rat, human and monkey data (**Table A3.3**). Therefore, better knowledge on species specific physiological parameters could improve the fitting and the prediction accuracy across species. Another explanation could be the interspecies difference in volume of distribution for diazepam. Rat Vuss,p (L/kg) was 4.6 times higher than human

Vuss,p (L/kg) whereas monkey Vuss,p (L/kg) was surprisingly 7.8 times higher. Besides insufficient data, this analysis also highlighted the need for good quality of PK data in animal studies to avoid parameter fitting issues and consequently the suitability of models for prediction in human. Low numbers of animals and variable methods for bioanalysis in the different preclinical studies might also have influence on the fitting of the models.

PBPK modelling is increasingly applied in drug development and regulatory review [65-69]. Recent regulatory guidance regarding PBPK modelling [70, 71] highlighted the need for a more systematic approach to establish confidence in PBPK models especially when estimating model parameters (middle-out approach). Overall, this work provided a rational and systematic strategy to predict human PK using preclinical data within a PBPK modelling strategy. The extrapolation is focused on selected best fitted models and plausible parameters which are not arbitrarily chosen thus bringing a more rational way to translate PBPK models for drug distribution between species. The strategy was applied to a dataset of 3 weak bases mostly due to the limited availability of PK data simultaneously in several species. The validation of this framework would require further assessment with a large dataset of compounds with diverse physicochemical properties. It should be noted that the predictive performance of the models in preclinical species were not validated with independent preclinical data as new data are generally not generated due to 3R principles. Following the development and verification of models in rats, it can be suggested to have an additional step for validating the predictive performance of these models using monkey data when they become available to increase confidence before predictions in human. Additionally, this strategy mainly focused on drug distribution which is only one key component of drug PK. Indeed, it did not consider challenges related to absorption and clearance [72, 73] by considering only intravenous administration and assuming a fitted clearance, which would have further increased the degree of complexity and potential uncertainty. For example, there were interspecies differences in plasma clearance which was very rapid and large in rat compared to man and monkey for diazepam whereas rat blood clearance was higher than hepatic blood flow rate for midazolam.

In this work, the PBPK modelling strategy was mainly focused on an application for interspecies extrapolation from animals to human. However, it could also be used for intra-species extrapolation from a base population to paediatrics or a special population

[74-77]. If some mechanisms are modified in the target population, this knowledge should be updated and incorporated in the model.

4.7. Conclusion

This work proposed a strategy to integrate preclinical data and optimize simplified PBPK models in order to successfully predict the distribution of small molecules in human. While fixing the clearance and using the approach of simplified PBPK models with common scalars, PK profiles could be well described in preclinical species and using the best models, plasma profiles were successfully predicted in human for two of the three tested compounds. Discrepancy for the third compound could be due to lack of data in preclinical species and may be addressed by generating appropriate PK data. In comparison with the traditional PBPK approach, the strategy proposed here provided an easy and systematic alternative for optimizing drug distribution in PBPK models and indicated better or similar accuracy of human drug distribution for diazepam and midazolam. This proof of concept was shown for diazepam and midazolam but the strategy can be generalised. The potential of this strategy for PBPK models of drug distribution could be applied to translation within species e.g., from adult to paediatric population.

4.8. References

[1] Reigner BG, Williams PE, Patel IH, Steimer JL, Peck C, van Brummelen P. An evaluation of the integration of pharmacokinetic and pharmacodynamic principles in clinical drug development. Experience within Hoffmann La Roche. Clin Pharmacokinet. 1997;33(2):142-52.

[2] Dedrick R, Bischoff KB, Zaharko DS. Interspecies correlation of plasma concentration history of methotrexate (NSC-740). Cancer Chemother Rep. 1970;54(2):95-101.

[3] Wajima T, Yano Y, Fukumura K, Oguma T. Prediction of human pharmacokinetic profile in animal scale up based on normalizing time course profiles. J Pharm Sci. 2004;93(7):1890-900.

[4] Poulin P, Jones RD, Jones HM, Gibson CR, Rowland M, Chien JY, et al. PHRMA CPCDC initiative on predictive models of human pharmacokinetics, part 5: prediction of plasma concentration-time profiles in human by using the physiologically-based pharmacokinetic modeling approach. J Pharm Sci. 2011;100(10):4127-57.

[5] Vuppugalla R, Marathe P, He H, Jones RD, Yates JW, Jones HM, et al. PhRMA CPCDC initiative on predictive models of human pharmacokinetics, part 4: prediction of plasma concentration-time profiles in human from in vivo preclinical data by using the Wajima approach. J Pharm Sci. 2011;100(10):4111-26.

[6] Bradshaw-Pierce EL, Eckhardt SG, Gustafson DL. A physiologically based pharmacokinetic model of docetaxel disposition: from mouse to man. Clin Cancer Res. 2007;13(9):2768-76.

[7] Hall C, Lueshen E, Mosat A, Linninger AA. Interspecies scaling in pharmacokinetics: a novel whole-body physiologically based modeling framework to discover drug biodistribution mechanisms in vivo. J Pharm Sci. 2012;101(3):1221-41.

[8] Thiel C, Schneckener S, Krauss M, Ghallab A, Hofmann U, Kanacher T, et al. A systematic evaluation of the use of physiologically based pharmacokinetic modeling for cross-species extrapolation. J Pharm Sci. 2015;104(1):191-206.

[9] Jones HM, Parrott N, Jorga K, Lave T. A novel strategy for physiologically based predictions of human pharmacokinetics. Clin Pharmacokinet. 2006;45(5):511-42.

[10] De Buck SS, Sinha VK, Fenu LA, Nijsen MJ, Mackie CE, Gilissen RA. Prediction of human pharmacokinetics using physiologically based modeling: a retrospective analysis of 26 clinically tested drugs. Drug Metab Dispos. 2007;35(10):1766-80.

[11] Jones HM, Gardner IB, Collard WT, Stanley PJ, Oxley P, Hosea NA, et al. Simulation of human intravenous and oral pharmacokinetics of 21 diverse compounds using physiologically based pharmacokinetic modelling. Clin Pharmacokinet. 2011;50(5):331-47.

[12] Zhang T, Heimbach T, Lin W, Zhang J, He H. Prospective Predictions of Human Pharmacokinetics for Eighteen Compounds. J Pharm Sci. 2015;104(9):2795-806.

[13] EFPIA MID3 Workgroup, Marshall SF, Burghaus R, Cosson V, Cheung SY, Chenel M, et al. Good Practices in Model-Informed Drug Discovery and Development: Practice, Application, and Documentation. CPT Pharmacometrics Syst Pharmacol. 2016;5(3):93-122.

[14] Zhao P, Zhang L, Grillo JA, Liu Q, Bullock JM, Moon YJ, et al. Applications of physiologically based pharmacokinetic (PBPK) modeling and simulation during regulatory review. Clin Pharmacol Ther. 2011;89(2):259-67.

[15] Jones H, Rowland-Yeo K. Basic concepts in physiologically based pharmacokinetic modeling in drug discovery and development. CPT Pharmacometrics Syst Pharmacol. 2013;2:e63.

[16] Tsamandouras N, Rostami-Hodjegan A, Aarons L. Combining the 'bottom up' and 'top down' approaches in pharmacokinetic modelling: fitting PBPK models to observed clinical data. Br J Clin Pharmacol. 2013;79(1):48-55.

[17] Li R, Barton HA, Yates PD, Ghosh A, Wolford AC, Riccardi KA, et al. A "middle-out" approach to human pharmacokinetic predictions for OATP substrates using physiologically-based pharmacokinetic modeling. J Pharmacokinet Pharmacodyn. 2014;41(3):197-209.

[18] Jamei M, Dickinson GL, Rostami-Hodjegan A. A framework for assessing inter-individual variability in pharmacokinetics using virtual human populations and integrating general knowledge of physical chemistry, biology, anatomy, physiology and genetics: A tale of 'bottom-up' vs 'top-down' recognition of covariates. Drug Metab Pharmacokinet. 2009;24(1):53-75.

[19] Margolskee A, Darwich AS, Pepin X, Aarons L, Galetin A, Rostami-Hodjegan A, et al. IMI - Oral biopharmaceutics tools project - Evaluation of bottom-up PBPK prediction success part 2: An introduction to the simulation exercise and overview of results. Eur J Pharm Sci. 2017;96:610-25.

[20] Rodgers T, Leahy D, Rowland M. Physiologically based pharmacokinetic modeling 1: predicting the tissue distribution of moderate-to-strong bases. J Pharm Sci. 2005;94(6):1259-76.

[21] Rodgers T, Rowland M. Physiologically based pharmacokinetic modelling 2: predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. J Pharm Sci. 2006;95(6):1238-57.

[22] Nestorov IA, Aarons LJ, Arundel PA, Rowland M. Lumping of whole-body physiologically based pharmacokinetic models. J Pharmacokinet Biopharm. 1998;26(1):21-46.

[23] Yau E, Olivares-Morales A, Gertz M, Parrott N, Darwich AS, Aarons L, et al. Global Sensitivity Analysis of the Rodgers and Rowland Model for Prediction of Tissue: Plasma Partitioning Coefficients: Assessment of the Key Physiological and Physicochemical Factors That Determine Small-Molecule Tissue Distribution. AAPS J. 2020;22(2):41.

[24] Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. 2018;46(D1):D1074-D82.

[25] Stillhart C, Parrott NJ, Lindenberg M, Chalus P, Bentley D, Szepes A. Characterising Drug Release from Immediate-Release Formulations of a Poorly Soluble Compound, Basmisanil, Through Absorption Modelling and Dissolution Testing. AAPS J. 2017;19(3):827-36.

[26] Diaz-Garcia JM, Oliver-Botana J, Fos-Galve D. Pharmacokinetics of diazepam in the rat: influence of a carbon tetrachloride-induced hepatic injury. J Pharm Sci. 1992;81(8):768-72.

[27] Igari Y, Sugiyama Y, Sawada Y, Iga T, Hanano M. Prediction of diazepam disposition in the rat and man by a physiologically based pharmacokinetic model. J Pharmacokinet Biopharm. 1983;11(6):577-93.

[28] Gueorguieva I, Aarons L, Rowland M. Diazepam pharamacokinetics from preclinical to phase I using a Bayesian population physiologically based pharmacokinetic model with informative prior distributions in WinBUGS. J Pharmacokinet Pharmacodyn. 2006;33(5):571-94.

[29] Klotz U. Effect of age on levels of diazepam in plasma and brain of rats. Naunyn Schmiedebergs Arch Pharmacol. 1979;307(2):167-9.

[30] Klotz U, Antonin KH, Bieck PR. Pharmacokinetics and plasma binding of diazepam in man, dog, rabbit, guinea pig and rat. J Pharmacol Exp Ther. 1976;199(1):67-73.

[31] Koyanagi T, Yamaura Y, Yano K, Kim S, Yamazaki H. Age-related pharmacokinetic changes of acetaminophen, antipyrine, diazepam, diphenhydramine, and ofloxacin in male cynomolgus monkeys and beagle dogs. Xenobiotica. 2014;44(10):893-901.

[32] Greenblatt DJ, Ehrenberg BL, Gunderman J, Locniskar A, Scavone JM, Harmatz JS, et al. Pharmacokinetic and electroencephalographic study of intravenous diazepam, midazolam, and placebo. Clin Pharmacol Ther. 1989;45(4):356-65.

[33] Greenblatt DJ, Allen MD, Harmatz JS, Shader RI. Diazepam disposition determinants. Clin Pharmacol Ther. 1980;27(3):301-12.

[34] Dhillon S, Richens A. Pharmacokinetics of diazepam in epileptic patients and normal volunteers following intravenous administration. Br J Clin Pharmacol. 1981;12(6):841-4.

[35] Klotz U, Avant GR, Hoyumpa A, Schenker S, Wilkinson GR. The effects of age and liver disease on the disposition and elimination of diazepam in adult man. J Clin Invest. 1975;55(2):347-59.

[36] Kaplan SA, Jack ML, Alexander K, Weinfeld RE. Pharmacokinetic profile of diazepam in man following single intravenous and oral and chronic oral administrations. J Pharm Sci. 1973;62(11):1789-96.

[37] Tsiros P, Bois FY, Dokoumetzidis A, Tsiliki G, Sarimveis H. Population pharmacokinetic reanalysis of a Diazepam PBPK model: a comparison of Stan and GNU MCSim. J Pharmacokinet Pharmacodyn. 2019.

[38] Langdon G, Gueorguieva I, Aarons L, Karlsson M. Linking preclinical and clinical whole-body physiologically based pharmacokinetic models with prior distributions in NONMEM. Eur J Clin Pharmacol. 2007;63(5):485-98.

[39] Kotegawa T, Laurijssens BE, Von Moltke LL, Cotreau MM, Perloff MD, Venkatakrishnan K, et al. In vitro, pharmacokinetic, and pharmacodynamic interactions of ketoconazole and midazolam in the rat. J Pharmacol Exp Ther. 2002;302(3):1228-37.

[40] Kuze J, Mutoh T, Takenaka T, Morisaki K, Nakura H, Hanioka N, et al. Separate evaluation of intestinal and hepatic metabolism of three benzodiazepines in rats with

cannulated portal and jugular veins: comparison with the profile in non-cannulated mice. Xenobiotica. 2009;39(11):871-80.

[41] Lau CE, Ma F, Wang Y, Smith C. Pharmacokinetics and bioavailability of midazolam after intravenous, subcutaneous, intraperitoneal and oral administration under a chronic food-limited regimen: relating DRL performance to pharmacokinetics. Psychopharmacology (Berl). 1996;126(3):241-8.

[42] Mandema JW, Tuk B, van Steveninck AL, Breimer DD, Cohen AF, Danhof M. Pharmacokinetic-pharmacodynamic modeling of the central nervous system effects of midazolam and its main metabolite alpha-hydroxymidazolam in healthy volunteers. Clin Pharmacol Ther. 1992;51(6):715-28.

[43] Nishimura T, Kato Y, Amano N, Ono M, Kubo Y, Kimura Y, et al. Species difference in intestinal absorption mechanism of etoposide and digoxin between cynomolgus monkey and rat. Pharm Res. 2008;25(11):2467-76.

[44] Tuk B, van Oostenbruggen MF, Herben VM, Mandema JW, Danhof M. Characterization of the pharmacodynamic interaction between parent drug and active metabolite in vivo: midazolam and alpha-OH-midazolam. J Pharmacol Exp Ther. 1999;289(2):1067-74.

[45] Ogasawara A, Kume T, Kazama E. Effect of oral ketoconazole on intestinal first-pass effect of midazolam and fexofenadine in cynomolgus monkeys. Drug Metab Dispos. 2007;35(3):410-8.

[46] Sakuda S, Akabane T, Teramura T. Marked species differences in the bioavailability of midazolam in cynomolgus monkeys and humans. Xenobiotica. 2006;36(4):331-40.

[47] Allonen H, Ziegler G, Klotz U. Midazolam kinetics. Clin Pharmacol Ther. 1981;30(5):653-61.

[48] Heizmann P, Eckert M, Ziegler WH. Pharmacokinetics and bioavailability of midazolam in man. Br J Clin Pharmacol. 1983;16 Suppl 1:43S-9S.

[49] Pentikainen PJ, Valisalmi L, Himberg JJ, Crevoisier C. Pharmacokinetics of midazolam following intravenous and oral administration in patients with chronic liver disease and in healthy subjects. J Clin Pharmacol. 1989;29(3):272-7.

[50] Smith MT, Eadie MJ, Brophy TO. The pharmacokinetics of midazolam in man. Eur J Clin Pharmacol. 1981;19(4):271-8.

[51] Thummel KE, O'Shea D, Paine MF, Shen DD, Kunze KL, Perkins JD, et al. Oral first-pass elimination of midazolam involves both gastrointestinal and hepatic CYP3A-mediated metabolism. Clin Pharmacol Ther. 1996;59(5):491-502.

[52] Tsunoda SM, Velez RL, von Moltke LL, Greenblatt DJ. Differentiation of intestinal and hepatic cytochrome P450 3A activity with use of midazolam as an in vivo probe: effect of ketoconazole. Clin Pharmacol Ther. 1999;66(5):461-71.

[53] Bjorkman S, Wada DR, Berling BM, Benoni G. Prediction of the disposition of midazolam in surgical patients by a physiologically based pharmacokinetic model. J Pharm Sci. 2001;90(9):1226-41.

[54] Bjorkman S, Rigemar G, Idvall J. Pharmacokinetics of midazolam given as an intranasal spray to adult surgical patients. Br J Anaesth. 1997;79(5):575-80.

[55] Bauer RJ. NONMEM Tutorial Part I: Description of Commands and Options, with Simple Examples of Population Analysis. CPT Pharmacometrics Syst Pharmacol. 2019.

[56] Liu X, Wang Y. Comparing the performance of FOCE and different expectation-maximization methods in handling complex population physiologically-based pharmacokinetic models. J Pharmacokinet Pharmacodyn. 2016;43(4):359-70.

[57] Bauer RJ. NONMEM Tutorial Part II: Estimation Methods and Advanced Examples. CPT Pharmacometrics Syst Pharmacol. 2019.

[58] Wang W, Hallow KM, James DA. A Tutorial on RxODE: Simulating Differential Equation Pharmacometric Models in R. CPT Pharmacometrics Syst Pharmacol. 2016;5(1):3-10.

[59] Assmus F, Houston JB, Galetin A. Incorporation of lysosomal sequestration in the mechanistic model for prediction of tissue distribution of basic drugs. Eur J Pharm Sci. 2017;109:419-30.

[60] Berezhkovskiy LM. Volume of distribution at steady state for a linear pharmacokinetic system with peripheral elimination. Journal of Pharmaceutical Sciences. 2004;93(6):1628-40.

[61] Peyret T, Poulin P, Krishnan K. A unified algorithm for predicting partition coefficients for PBPK modeling of drugs and environmental chemicals. Toxicol Appl Pharmacol. 2010;249(3):197-207.

[62] Xia B, Heimbach T, Lin TH, He H, Wang Y, Tan E. Novel physiologically based pharmacokinetic modeling of patupilone for human pharmacokinetic predictions. Cancer Chemother Pharmacol. 2012;69(6):1567-82.

[63] Price PS, Conolly RB, Chaisson CF, Gross EA, Young JS, Mathis ET, et al. Modeling interindividual variation in physiological factors used in PBPK models of humans. Crit Rev Toxicol. 2003;33(5):469-503.

[64] Krauss M, Tappe K, Schuppert A, Kuepfer L, Goerlitz L. Bayesian Population Physiologically-Based Pharmacokinetic (PBPK) Approach for a Physiologically Realistic Characterization of Interindividual Variability in Clinically Relevant Populations. PLoS One. 2015;10(10):e0139423.

[65] Jones HM, Chen Y, Gibson C, Heimbach T, Parrott N, Peters SA, et al. Physiologically based pharmacokinetic modeling in drug discovery and development: a pharmaceutical industry perspective. Clin Pharmacol Ther. 2015;97(3):247-62.

[66] Huang SM, Rowland M. The role of physiologically based pharmacokinetic modeling in regulatory review. Clin Pharmacol Ther. 2012;91(3):542-9.

[67] Zhuang X, Lu C. PBPK modeling and simulation in drug research and development. Acta Pharm Sin B. 2016;6(5):430-40.

[68] Luzon E, Blake K, Cole S, Nordmark A, Versantvoort C, Berglund EG. Physiologically based pharmacokinetic modeling in regulatory decision-making at the European Medicines Agency. Clin Pharmacol Ther. 2016.

[69] Shebley M, Sandhu P, Emami Riedmaier A, Jamei M, Narayanan R, Patel A, et al. Physiologically Based Pharmacokinetic Model Qualification and Reporting Procedures for Regulatory Submissions: A Consortium Perspective. Clin Pharmacol Ther. 2018;104(1):88-110.

[70] U.S. Food and Drug Administration, Physiologically Based Pharmacokinetic Analyses — Format and Content : Guidance for Industry 2018 [Available from: https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM531207.pdf; accessed May 2019]

[71] European Medicines Agency, Guideline on the qualification and reporting of physiologically based pharmacokinetic (PBPK) modelling and simulation 2018 [Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-reporting-physiologically-based-pharmacokinetic-pbpk-modelling-simulation_en.pdf; accessed May 2019]

[72] Chen Y, Jin JY, Mukadam S, Malhi V, Kenny JR. Application of IVIVE and PBPK modeling in prospective prediction of clinical pharmacokinetics: strategy and approach during the drug discovery phase with four case studies. Biopharm Drug Dispos. 2012;33(2):85-98.

[73] Miller NA, Reddy MB, Heikkinen AT, Lukacova V, Parrott N. Physiologically Based Pharmacokinetic Modelling for First-In-Human Predictions: An Updated Model Building Strategy Illustrated with Challenging Industry Case Studies. Clin Pharmacokinet. 2019.

[74] Jorga K, Chavanne C, Frey N, Lave T, Lukacova V, Parrott N, et al. Bottom-up Meets Top-down: Complementary Physiologically Based Pharmacokinetic and Population Pharmacokinetic Modeling for Regulatory Approval of a Dosing Algorithm of Valganciclovir in Very Young Children. Clin Pharmacol Ther. 2016;100(6):761-9.

[75] Lukacova V, Goelzer P, Reddy M, Greig G, Reigner B, Parrott N. A Physiologically Based Pharmacokinetic Model for Ganciclovir and Its Prodrug Valganciclovir in Adults and Children. AAPS J. 2016;18(6):1453-63.

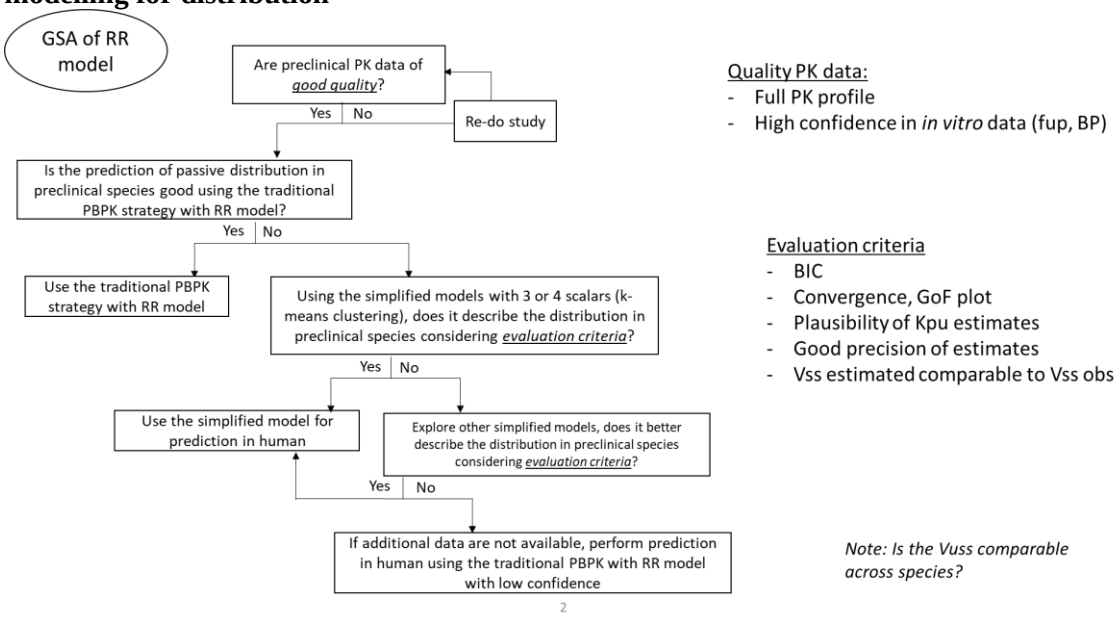
[76] Wendling T, Tsamandouras N, Dumitras S, Pigeolet E, Ogungbenro K, Aarons L. Reduction of a Whole-Body Physiologically Based Pharmacokinetic Model to Stabilise the Bayesian Analysis of Clinical Data. AAPS J. 2016;18(1):196-209.

[77] Wendling T, Dumitras S, Ogungbenro K, Aarons L. Application of a Bayesian approach to physiological modelling of mavoglurant population pharmacokinetics. J Pharmacokinet Pharmacodyn. 2015;42(6):639-57.

Chapter 5: Concluding remarks and future directions

This project focused on the development of a systematic framework to integrate preclinical data within PBPK models, especially in order to improve the prediction of human drug disposition during drug discovery and development programme. The PBPK modelling approach for prediction of human PK of small molecules has become more common with various studies validating this approach using large drug datasets and examples of successful predictions [1-6]. However, in drug discovery and development, there is often a need to optimise model parameters when bias exists between PBPK model predictions and observed data. Herein, a systematic approach has been proposed and evaluated for such purpose within a PBPK modelling framework (Figure 6). As stated in the regulatory guidelines for PBPK modelling, it is recommended to perform sensitivity analyses when developing and refining PBPK models, the framework started with a global sensitivity analysis to identify sensitive parameters involved in the drug distribution model [7, 8]. Due to the multidimensional nature of the PBPK models which limits estimation of the model parameters simultaneously, several simplified physiological models for distribution in PBPK models were investigated. These models were used for parameter estimation using data from preclinical species. Subsequently, the potential to extrapolate and improve human predictions incorporating pre-clinical data was evaluated. The framework proposed here has advantages and limitations which are going to be discussed in this chapter.

Figure 6: Proposed systematic framework for integrating preclinical data in PBPK modelling for distribution



5.1. Thesis findings

A Global Sensitivity Analysis (GSA) was performed in Chapter 2, on the Rodgers and Rowland (R&R) model which predicts tissue Kpus, the essential inputs for distribution in PBPK. Various GSA methods have been proposed and applied during PBPK modelling and simulation in recent years [9-13]. Some of the proposed GSA methods can be very sophisticated, allowing sampling of independent or correlated input parameters, and also possibly assesses qualitatively or quantitatively the variation of parameters on the model outputs [14]. In Chapter 2, the partial rank correlation coefficients method (PRCC) was chosen as it can deal with nonlinear relationships and monotonicity could be assumed between parameters and outputs [15]. The PRCC has been shown to be an efficient and robust methods without needing a very large sample size [15, 16]. It may be possible to use one of the more recently developed GSA methods which are able to better account for correlation between parameters, but these methods require data gathering to properly identify parameter correlations and are generally computationally expensive. In addition, the work in Chapter 2 was based on the R&R model which is the model most used for distribution. Several studies found that the R&R model was overall the most accurate model for predicting Kpu using datasets of drugs [17-19]. The R&R model have however some limitations as it does not take into account all the mechanisms that could exist (e.g., transporter activity, inaccurate prediction for highly lipophilic compounds, and relevance of lysosomal trapping for strong bases) [20,

21]. One evident finding from this GSA was that the predictions of the volume of distribution and tissue Kp values are very sensitive to drug parameters LogP and fup. Drug parameters are indeed important since differences in the distribution of drugs depends on their drug classes, binding, and lipophilicity (ability to interact with lipid constituents of tissues). Consistent experimental measurements of these values are essential as uncertainty in these drug parameters would largely impact the predicted Kpu values and thus the predicted drug PK profile. The most interesting finding was that tissue composition data in particular acidic phospholipids and extracellular proteins are fairly influential parameters although often overlooked compared to drug parameters. These results are consistent with a recent study where a standardized tissue composition data in human were used to predict tissue:plasma partition coefficients [22]. However, values from nonclinical species remain part of this standardized tissue composition database due to the limited availability of these values in human. Further research should be undertaken to investigate and characterize tissue composition data across species. Most of available tissue composition information are determined in rats although from different studies and strains of rat [23-25]. It should be noted that some efforts have been made to collect tissue composition data for several tissues from the literature in various species (mouse, rat, guinea pig, rabbit, beagle dog, pig, monkey, and human) [26, 27]. Due to the scarcity of data and limited knowledge of tissue composition data from the same source and species, a full species-specific tissue composition dataset is often composed of values averaged from a number of different studies and sourced from different species (and strains). This is especially true for the monkey tissue composition data used in Chapter 4, which was a hybrid of rat and human tissue composition data. Among the different tissue composition components, the GSA assessment (Chapter 2) suggests the highest chance to improve Kp predictions by focusing primarily on quantifying acidic phospholipids and extracellular proteins.

Prediction of profiles in human from preclinical and physicochemical data for a new chemical entity is an important part of drug discovery and development. Among the various approaches that are conventionally used, PBPK modelling is the most mechanistic approach by integrating species-specific and compound-specific parameters. Several studies demonstrated that animal PBPK models which could accurately predict animal steady state volume of distribution (Vss), resulted in accurate prediction of drug PK in human PBPK models [1, 2, 19, 28]. However, predictions from animal PBPK models

have also often showed mismatches between simulated and observed concentration profiles for some animals and compounds [2]. It is then necessary to optimize the PBPK models to improve the prediction accuracy in animals before performing simulations in humans. One of the issues that emerges is the complexity of whole-body PBPK model structure and the large number of parameters involved, which renders the optimization process challenging. In Chapter 3, several approaches were thus investigated to simplify the PBPK model structure and/or reduce the number of parameters in order to facilitate parameter estimation for PBPK modelling. A common approach to fit PK data is a classical compartmental analysis which usually results in empirical models of 1, 2 or 3 compartments. However, since these models are not physiological, their use for extrapolation from preclinical to clinical development is limited. Following this observation of model reduction to a few compartments, the first approach considered for simplifying PBPK models was to apply an approach similar to kinetic lumping. Lumping technique is a well-established method for simplifying PBPK models [29-32]. For example, it was reported that the WB-PBPK of 25 various compounds (moderate-to-strong bases, weak bases and acids) could be lumped into 1 to 5 compartments [32]. The model structure will generally change for each new compound as they are dependent on drug specific properties. A general lumped model was proposed to overcome the drug-specificity inherent to kinetic lumping. The proposed lumped model may however have limited application since the *a priori* lumping based on tissue volumes and blood flow rates required the assumptions of compounds with relatively low Kb values (Kb=1). On the other hand, these assumptions were not made in the approach where the structure of WB-PBPK models was conserved to 14 compartments. Two different cases were then considered: either tissues with similar compositions share common Kpu values or tissues share common empirical scalars reflecting the bias in the prediction of the Kpu by the R&R model. The first case is supported by evidence in the literature that correlates the Kp of different tissues, especially muscle, adipose, and skin, which were suggested to be representative of other tissue Kps [33, 34]. Here, the common Kp, or Kpu values could be considered as an average of several tissues, and thus not necessarily be comparable to particular experimental tissue Kps, or Kpus. In the second case, the assumption was to consider that some tissues have a common bias, reflected in a scalar that would adjust the Kp predictions from the R&R model. Empirical scaling factors are commonly used in practice during optimization of PBPK models, however the choice of universal or specific scalars for some tissues is not always consistent [35-37]. The second case provides a more

systematic and rational way to alleviate this concern and address the recommendations from regulatory guidelines [7, 8]. With diazepam as a model compound, the different approaches and models were used to fit PK data using a nonlinear mixed effect (NLME) approach. When fitting the data, all model parameters could be estimated including clearance, inter-individual variability on clearance, as well as Kpus or Kpu scalars. The results showed that these proposed models were flexible enough to fit data while keeping the physiological nature of the model. From a parameter optimization point of view, these mechanistic models are more informative than classical compartmental PK models. It is important to bear in mind that the selection of the model (mechanistic or empirical) would depend on its purpose and use to answer specific questions. In the second approach, the tissue grouping was mostly based on rat tissue composition. As mentioned in Chapter 2, reported tissue composition data include some level of variability and uncertainty due to their origin and this may affect the tissue grouping. Nevertheless, the concept of tissue grouping using clustering methods on tissue composition remains relevant. If new tissue composition becomes available, the same methodology can be applied, and tissue grouping may then be updated as well as associated models.

Chapter 4 can be seen as an application of the models/concepts developed in Chapter 3. The aim here was to evaluate the performance of the models previously investigated for translation of drug distribution from preclinical species (rat and monkey) to humans. The evaluation of the different PBPK models was performed for 3 test compounds. These were basic compounds with high lipophilicity, indicating a high membrane permeability, which is in accordance with the assumption that the passive diffusion is predominant for tissue distribution in these models. For translational purposes, the same model structure was assumed across species. The *a priori* lumped model was based on human physiology as human is the target species for extrapolation. This kinetic lumping approach assigned similar tissue kinetics between species which potentially alters the mechanistic realism of the model in preclinical species. Contrary to the kinetic lumping, the second approach based on steady state commonality in drug partitioning (common Kpus or common Kpu scalars) has the advantage to retain most of the organ structure of the WBPBPK model and is thus preferable. The main assumption of the models with common Kpus is that tissue composition is very similar across species and consequently Kpus are the same between species. It is consistent with a recent study which found no significant species differences of fraction unbound for tissue binding and

rat tissue fraction unbound can be a good predictor of other tissue binding and other species [38]. Whereas the main assumption of the models with common Kpu scalars is that empirical scaling factors can be derived from fitting animal PK data and these scaling factors are species independent. These latter models have more flexibility as tissues can have different Kpu values although having the same scalar. The results from the fitting to diazepam and midazolam preclinical data support the use of these more flexible models (especially with scalars obtained from the k-means clustering method) as they were generally more able to fit the available preclinical PK data and subsequently to predict drug concentrations in human. For basmisanil, the fitting of models to preclinical data was not successful possibly due to the quality of data, highlighting the need for good quality data for these purposes (i.e., well defined distribution and elimination phases). Although monkey could be expected to perform better as they are considered more similar to humans, it was not found that predictions from monkey would always be better than from rat. One explanation could be the limited *in vivo* data available in monkey in the study. Based on the findings obtained for these three compounds, a general strategy was proposed in the form of a decision-tree (Figure 6). One of the key points from this decision-tree relates specifically to the quality of preclinical PK data used: good quality data should be available or generated before considering any optimisation and extrapolation. Data quality could be improved by better sampling, optimal design, infusions instead of IV bolus, etc. Additionally, the use of several criteria allows to evaluate the performance of the simplified models for fitting the preclinical data and subsequently determine its suitability for extrapolation to human. Given the nature of these simplified models which are based on physiology, their use should be applicable to a large set of compounds. Further work should be undertaken with more compounds and different drug classes (acids, strong bases, neutral) in order to validate the proposed framework.

5.2. Future perspectives

A systematic framework was proposed for the optimisation and interspecies translation of PBPK models. It focused on the estimation of Kpu parameters, but estimation of other model parameters could also be investigated with the availability of different data. Further research might explore the combination of Bayesian estimation methods with this proposed PBPK modelling framework. And a similar approach could

be useful beyond the topic of interspecies translation of drug distribution such as translation of other PK processes or intraspecies extrapolation, for example.

5.2.1. Optimisation of drug and physiological parameters

This thesis focused on the estimation of PBPK parameters (especially Kpus) using preclinical IV data. The estimation of other PBPK model parameters could also be considered with the availability of different data. In the R&R model, Kpu depends on tissue composition and most of the parameters are based on the physicochemical properties of the compound except Ka, the association constant to acidic phospholipid content [24, 39]. Ka is calculated from the BP ratio, considering the known concentration of acidic phospholipid in blood cells and plasma protein binding. The affinity of a compound to acidic phospholipid is then assumed to be the same for all tissues of the body. Rodgers et al. predicted Kpus using tissue composition data and verified the Kpu predictions against experimental Kpu values which were steady state data following a constant rate infusion content [24, 39]. Given the availability of these data, the model can be fitted to all the tissue Kpu data to estimate a global value of Ka instead of using the derived value from the BP ratio (which can be used as an initial estimate). The parameter Ka could also be estimated using IV plasma data after taking into account renal clearance and intrinsic hepatic clearance. After substitution of Kpu in the PBPK model by the tissue composition model for each tissue, Ka would be the only parameter to estimate thereby vastly reducing the dimensionality of the WBPBPK model and associated estimation problem. The performance of these Kpu values calculated with the optimised Ka can then be compared to those that were estimated by the R&R model.

Another aspect to explore could be the optimisation of physiological parameters as many of these parameters are generally fixed. A further work might explore the fitting of a WBPBPK model to PK data to estimate the blood flow rates to each tissue (assuming perfusion rate limited distribution) rather than to assume literature values [40, 41]. The data from a study investigating PBPK modelling of beta-blockers (strong bases) in the rat following an IV bolus dose [42] could be used for this purpose. Additionally, seven racemic beta blockers were administered as cocktails of 3-4 compounds and individual enantiomers were measured in this study. The estimation of blood flow rates may also help to address the potential issue of a cardiovascular effect that could happen at the doses employed in the drug cocktail.

Additionally, the GSA from Chapter 2 showed that acidic phospholipid and extracellular protein are very sensitive parameters and should be well characterised. An alternative to experimental studies could be to optimize the values of these tissue composition data using the dataset of rat experimental Kps that have been collected from literature and used in Chapter 3 for deriving tissue groups. Tissue Kpu measured for strong basic compounds could inform the estimation of acidic phospholipids. Weak basic and neutral compounds would help the estimation of the tissue lipoprotein ratio while acidic compounds could help the estimation of the tissue to albumin ratio.

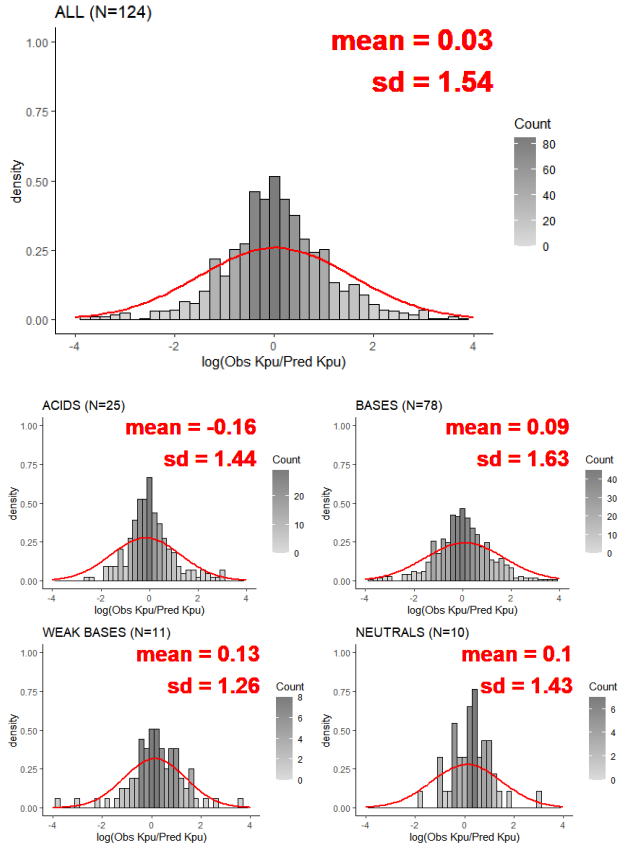
The findings from the optimization of drug and physiological parameters in a PBPK model would help to translate the prior knowledge and uncertainty associated with experimental methods and data into statistical prior distributions. The definition of appropriate prior distributions for the model parameters would aid to establish a Bayesian framework for PBPK modelling.

5.2.2. Bayesian hierarchical modelling

Overall, the present work is one of the first attempts to investigate a systematic framework for integrating drug preclinical data into PBPK modelling. A recent study proposed to optimize PBPK model parameters using available species-specific toxicokinetic data for an organic pollutant [43]. Such an approach which uses whole PK profiles (and not only PK parameters e.g., Vss or AUC) is especially of interest as it extracts most information out of the animal experiments and thus contributes to the 3R's principles. The fitting of PBPK models to preclinical data and subsequent extrapolation to humans allows the description of an average animal and the prediction of an average human profile. A natural progression of this work would be to integrate variability and uncertainty in population physiology and drug parameters. All system-related and drug-specific parameters are mechanistic in nature and prior information about their range can be extracted from anatomical and physiological literature, *in vitro* experiments, and previous published models. The use of Bayesian estimation methods could be optimal as it allows the integration of prior information [44-46]. Bayesian population PBPK has been applied successfully in physiological pharmacokinetic and toxicokinetic models [31, 43, 47-60]. The principle of the Bayesian approach is to define unknown parameters as random variables with probability distributions rather than unknown fixed variables [61]. The prior distribution of a parameter is a key part of Bayesian inference as it represents

the information about uncertain parameters that will be combined with the probability distribution of new data to yield the posterior distribution and the updated parameter estimates. A prior distribution for PBPK model parameters can be informed from the *in vitro* experiments, literature databases, and predictions derived from *in silico* models; here we focus on parameters that can be informed for predicting drug distribution. For example, the distribution of the accuracy of rat Kpu predictions from the R&R model (Figure 7) could be used to assign priors to the scalars in the models with common Kpu scalars. It was for example found that the ratio of log predicted to observed Kpu could follow a normal distribution for the different tissues.

Figure 7: Distribution of the log-transformed random error of observed Kpu/predicted Kpu overall and per drug class



Details of this analysis can be found in Appendix A3.3

Due to the limited data and the model sensitivity to noisy data or outliers, information from experimental data is often not sufficient to obtain estimates of parameters in a PBPK model. The Bayesian approach combined with population hierarchical modelling (if population data are available) can be considered as an alternative to the maximum likelihood approach [62]. The initial issue of a very high

dimensional parameter space in the Bayesian population PBPK approach could be addressed by the use of a simplified PBPK model similar to the maximum likelihood or 'classical' approach. The Bayesian population approach also provides posterior distributions for the individual and population parameters rather than single point estimates [63, 64]. Similar to the approach taken in Chapter 2, a GSA could help to identify sensitive parameters before Bayesian-PBPK model estimation as shown in a previous study [65]. Deriving credible prior distributions of relevant PBPK model parameters would be an important aspect. For PBPK models, it is sometimes difficult to translate the prior knowledge into statistical prior distribution. Indeed, most of the prior information regarding drug-related parameters is derived from *in vitro* experiments, literature databases, and *in silico* methods generating point estimates (e.g., mechanistic predictions of partition coefficients) or estimates with combined uncertainty and variability (e.g., predictions of intrinsic clearance from pooled human liver microsomes). Consequently, many experimental studies (in vitro and animal in vivo) focused on metabolism and excretion. PBPK models mechanistically integrate all available measures of metabolism and excretion in order to provide realistic estimates of human exposure in target organs. Various in vitro, in vivo, and in silico models are often employed throughout drug discovery and development to predict the metabolic disposition of new and existing drug molecules [66-70]. We could use the available data to derive priors of the *in vitro– in vivo* extrapolation (IVIVE) of metabolic clearance and also understand the potential reasons for mispredictions. Prior information for anatomical and physiological can be found in literature [71-76]. The estimated posterior probability distribution derived from Markov chain Monte Carlo (MCMC) analysis can then be used to perform PK simulations at population- or individual level [53, 55, 57].

5.2.3. Additional applications

In this thesis, a systematic framework for the translation of a PBPK model was primarily investigated for drug distribution. However, distribution is only one of the ADME processes that influence the PK of a drug. Elimination is a process that does not translate well from preclinical species due to species differences in enzyme and transporter expressions and activities. A framework similar to the one investigated herein could provide insights for the absorption process. The absorption is a complex process depending on the interplay between the drug/formulation properties and the

gastrointestinal tract (transit times, first pass effect). Both drug parameters and physiological parameters could influence the fraction absorbed (solubility, dissolution, permeability) and oral bioavailability [77]. Due to its multifactorial nature and oral bioavailability process, translation of drug absorption between species is often not well predicted [78, 79]. Fitting some of the absorption parameters in preclinical species may be informative and provide insights for human absorption. Future research could explore whether parameters such as solubility, dissolution, permeability, and clearance could be estimated systematically using preclinical PK data for several compounds. These parameters could then be assumed to be similar in human and permit the extrapolation of absorption to human. Additionally, the Bayesian approach could be considered in this systematic framework for absorption. This study may encounter limitations due to the availability of PK data following both an iv and oral dose in several species.

The present research focused on the use of simplified PBPK models for extrapolation from preclinical species to human. More broadly, these simplified models can have application beyond interspecies translation. When the simplified PBPK model is developed and selected, it could follow the compound even after clinical development. The insight gained from interspecies extrapolation will prove useful in expanding the application of these simplified models for intra-species translation such as extrapolation from adult to children or from healthy to disease populations.

5.3. Reference

[1] Jones HM, Gardner IB, Collard WT, Stanley PJ, Oxley P, Hosea NA, et al. Simulation of human intravenous and oral pharmacokinetics of 21 diverse compounds using physiologically based pharmacokinetic modelling. Clin Pharmacokinet. 2011;50(5):331-47.

[2] Jones HM, Parrott N, Jorga K, Lave T. A novel strategy for physiologically based predictions of human pharmacokinetics. Clin Pharmacokinet. 2006;45(5):511-42.

[3] Chen Y, Jin JY, Mukadam S, Malhi V, Kenny JR. Application of IVIVE and PBPK modeling in prospective prediction of clinical pharmacokinetics: strategy and approach during the drug discovery phase with four case studies. Biopharm Drug Dispos. 2012;33(2):85-98.

[4] De Buck SS, Sinha VK, Fenu LA, Nijsen MJ, Mackie CE, Gilissen RA. Prediction of human pharmacokinetics using physiologically based modeling: a retrospective analysis of 26 clinically tested drugs. Drug Metab Dispos. 2007;35(10):1766-80.

[5] Poulin P, Jones RD, Jones HM, Gibson CR, Rowland M, Chien JY, et al. PHRMA CPCDC initiative on predictive models of human pharmacokinetics, part 5: prediction of plasma concentration-time profiles in human by using the physiologically-based pharmacokinetic modeling approach. J Pharm Sci. 2011;100(10):4127-57.

[6] Miller NA, Reddy MB, Heikkinen AT, Lukacova V, Parrott N. Physiologically Based Pharmacokinetic Modelling for First-In-Human Predictions: An Updated Model Building Strategy Illustrated with Challenging Industry Case Studies. Clin Pharmacokinet. 2019.

[7] European Medicines Agency, Guideline on the qualification and reporting of physiologically based pharmacokinetic (PBPK) modelling and simulation 2018 [Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-reporting-physiologically-based-pharmacokinetic-pbpk-modelling-simulation_en.pdf; accessed May 2019]

[8] U.S. Food and Drug Administration, Physiologically Based Pharmacokinetic Analyses — Format and Content : Guidance for Industry 2018 [Available from: https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM531207.pdf; accessed May 2019]

[9] Melillo N, Aarons L, Magni P, Darwich AS. Variance based global sensitivity analysis of physiologically based pharmacokinetic absorption models for BCS I-IV drugs. J Pharmacokinet Pharmacodyn. 2019;46(1):27-42.

[10] Melillo N, Darwich AS, Magni P, Rostami-Hodjegan A. Accounting for inter-correlation between enzyme abundance: a simulation study to assess implications on global sensitivity analysis within physiologically-based pharmacokinetics. J Pharmacokinet Pharmacodyn. 2019.

[11] Liu D, Li L, Rostami-Hodjegan A, Bois FY, Jamei M. Considerations and Caveats when Applying Global Sensitivity Analysis Methods to Physiologically Based Pharmacokinetic Models. AAPS J. 2020;22(5):93.

[12] Daga PR, Bolger MB, Haworth IS, Clark RD, Martin EJ. Physiologically Based Pharmacokinetic Modeling in Lead Optimization. 1. Evaluation and Adaptation of GastroPlus To Predict Bioavailability of Medchem Series. Mol Pharm. 2018;15(3):821-30.

[13] Melillo N, Grandoni S, Cesari N, Brogin G, Puccini P, Magni P. Inter-compound and Intra-compound Global Sensitivity Analysis of a Physiological Model for Pulmonary Absorption of Inhaled Compounds. AAPS J. 2020;22(5):116.

[14] Saltelli A. Global sensitivity analysis : the primer. Chichester, England ; Hoboken, NJ: John Wiley; 2008. x, 292 p. p.

[15] Marino S, Hogue IB, Ray CJ, Kirschner DE. A methodology for performing global uncertainty and sensitivity analysis in systems biology. J Theor Biol. 2008;254(1):178-96.

[16] Saltelli A, Marivoet J. Non-parametric statistics in sensitivity analysis for model output: A comparison of selected techniques. Reliab Eng Syst Safe. 1990;28(2):229-53.

[17] Jones RD, Jones HM, Rowland M, Gibson CR, Yates JW, Chien JY, et al. PhRMA CPCDC initiative on predictive models of human pharmacokinetics, part 2: comparative assessment of prediction methods of human volume of distribution. J Pharm Sci. 2011;100(10):4074-89.

[18] Chan R, De Bruyn T, Wright M, Broccatelli F. Comparing Mechanistic and Preclinical Predictions of Volume of Distribution on a Large Set of Drugs. Pharmaceutical Research. 2018;35(4):87.

[19] Graham H, Walker M, Jones O, Yates J, Galetin A, Aarons L. Comparison of in-vivo and in-silico methods used for prediction of tissue: plasma partition coefficients in rat. J Pharm Pharmacol. 2012;64(3):383-96.

[20] Poulin P, Haddad S. Advancing prediction of tissue distribution and volume of distribution of highly lipophilic compounds from a simplified tissue-composition-based model as a mechanistic animal alternative method. J Pharm Sci. 2012;101(6):2250-61.

[21] Assmus F, Houston JB, Galetin A. Incorporation of lysosomal sequestration in the mechanistic model for prediction of tissue distribution of basic drugs. Eur J Pharm Sci. 2017;109:419-30.

[22] Utsey K, Gastonguay MS, Russell S, Freling R, Riggs MM, Elmokadem A. Quantification of the Impact of Partition Coefficient Prediction Methods on Physiologically Based Pharmacokinetic Model Output Using a Standardized Tissue Composition. Drug Metab Dispos. 2020;48(10):903-16.

[23] Musther H, Harwood MD, Yang J, Turner DB, Rostami-Hodjegan A, Jamei M. The Constraints, Construction, and Verification of a Strain-Specific Physiologically Based Pharmacokinetic Rat Model. J Pharm Sci. 2017;106(9):2826-38.

[24] Rodgers T, Leahy D, Rowland M. Physiologically based pharmacokinetic modeling 1: predicting the tissue distribution of moderate-to-strong bases. J Pharm Sci. 2005;94(6):1259-76.

[25] Rodgers T, Rowland M. Mechanistic approaches to volume of distribution predictions: understanding the processes. Pharm Res. 2007;24(5):918-33.

[26] Ruark CD, Hack CE, Robinson PJ, Mahle DA, Gearhart JM. Predicting passive and active tissue:plasma partition coefficients: interindividual and interspecies variability. J Pharm Sci. 2014;103(7):2189-98.

[27] Poulin P, Collet SH, Atrux-Tallau N, Linget JM, Hennequin L, Wilson CE. Application of the Tissue Composition-Based Model to Minipig for Predicting the Volume of Distribution at Steady State and Dermis-to-Plasma Partition Coefficients of

Drugs Used in the Physiologically Based Pharmacokinetics Model in Dermatology. J Pharm Sci. 2019;108(1):603-19.

[28] Mayumi K, Ohnishi S, Hasegawa H. Successful Prediction of Human Pharmacokinetics by Improving Calculation Processes of Physiologically Based Pharmacokinetic Approach. J Pharm Sci. 2019;108(8):2718-27.

[29] Nestorov IA, Aarons LJ, Arundel PA, Rowland M. Lumping of whole-body physiologically based pharmacokinetic models. J Pharmacokinet Biopharm. 1998;26(1):21-46.

[30] Pan S, Duffull SB. Automated proper lumping for simplification of linear physiologically based pharmacokinetic systems. J Pharmacokinet Pharmacodyn. 2019.

[31] Wendling T, Tsamandouras N, Dumitras S, Pigeolet E, Ogungbenro K, Aarons L. Reduction of a Whole-Body Physiologically Based Pharmacokinetic Model to Stabilise the Bayesian Analysis of Clinical Data. AAPS J. 2016;18(1):196-209.

[32] Pilari S, Huisinga W. Lumping of physiologically-based pharmacokinetic models and a mechanistic derivation of classical compartmental models. J Pharmacokinet Pharmacodyn. 2010;37(4):365-405.

[33] Richter WF, Starke V, Whitby B. The distribution pattern of radioactivity across different tissues in quantitative whole-body autoradiography (QWBA) studies. Eur J Pharm Sci. 2006;28(1-2):155-65.

[34] Bjorkman S. Prediction of the volume of distribution of a drug: which tissue-plasma partition coefficients are needed? J Pharm Pharmacol. 2002;54(9):1237-45.

[35] Shebley M, Sandhu P, Emami Riedmaier A, Jamei M, Narayanan R, Patel A, et al. Physiologically Based Pharmacokinetic Model Qualification and Reporting Procedures for Regulatory Submissions: A Consortium Perspective. Clin Pharmacol Ther. 2018;104(1):88-110.

[36] Nigade PB, Gundu J, Pai KS, Nemmani KVS, Talwar R. Prediction of volume of distribution in preclinical species and humans: application of simplified physiologically based algorithms. Xenobiotica. 2019;49(5):528-39.

[37] Mathew S, Tess D, Burchett W, Chang G, Woody N, Keefer C, et al. Evaluation of Prediction Accuracy for Volume of Distribution in Rat and Human Using In Vitro, In Vivo, PBPK and QSAR Methods. J Pharm Sci. 2021;110(4):1799-823.

[38] Ryu S, Tess D, Chang G, Keefer C, Burchett W, Steeno GS, et al. Evaluation of Fraction Unbound Across 7 Tissues of 5 Species. J Pharm Sci. 2020;109(2):1178-90.

[39] Rodgers T, Rowland M. Physiologically based pharmacokinetic modelling 2: predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. J Pharm Sci. 2006;95(6):1238-57.

[40] Kuwahira I, Moue Y, Ohta Y, Mori H, Gonzalez NC. Distribution of pulmonary blood flow in conscious resting rats. Respir Physiol. 1994;97(3):309-21.

[41] Igari Y, Sugiyama Y, Sawada Y, Iga T, Hanano M. Prediction of diazepam disposition in the rat and man by a physiologically based pharmacokinetic model. J Pharmacokinet Biopharm. 1983;11(6):577-93.

[42] Cheung SYA, Rodgers T, Aarons L, Gueorguieva I, Dickinson GL, Murby S, et al. Whole body physiologically based modelling of beta-blockers in the rat: events in tissues and plasma following an i.v. bolus dose. Br J Pharmacol. 2018;175(1):67-83.

[43] Chou WC, Lin Z. Bayesian evaluation of a physiologically based pharmacokinetic (PBPK) model for perfluorooctane sulfonate (PFOS) to characterize the interspecies uncertainty between mice, rats, monkeys, and humans: Development and performance verification. Environ Int. 2019;129:408-22.

[44] Gisleskog PO, Karlsson MO, Beal SL. Use of prior information to stabilize a population data analysis. Journal of Pharmacokinetics and Pharmacodynamics. 2002;29(5-6):473-505.

[45] Walley R, Sherington J, Rastrick J, Detrait E, Hanon E, Watt G. Using Bayesian analysis in repeated preclinical in vivo studies for a more effective use of animals. Pharm Stat. 2016;15(3):277-85.

[46] Davis JL, Tornero-Velez R, Setzer RW. Computational Approaches for Developing Informative Prior Distributions for Bayesian Calibration of PBPK Models. Parameters for Pesticide QSAR and PBPK/PD Models for Human Risk Assessment. ACS Symposium Series. 1099: American Chemical Society; 2012. p. 291-306.

[47] Bois FY, Gelman A, Jiang J, Maszle DR, Zeise L, Alexeef G. Population toxicokinetics of tetrachloroethylene. Arch Toxicol. 1996;70(6):347-55.

[48] Bois FY, Jackson ET, Pekari K, Smith MT. Population toxicokinetics of benzene. Environ Health Perspect. 1996;104 Suppl 6:1405-11.

[49] Bois FY. Statistical analysis of Fisher et al. PBPK model of trichloroethylene kinetics. Environ Health Perspect. 2000;108 Suppl 2:275-82.

[50] Jonsson F, Bois FY, Johanson G. Assessing the reliability of PBPK models using data from methyl chloride-exposed, non-conjugating human subjects. Arch Toxicol. 2001;75(4):189-99.

[51] Gueorguieva I, Aarons L, Rowland M. Diazepam pharamacokinetics from preclinical to phase I using a Bayesian population physiologically based pharmacokinetic model with informative prior distributions in WinBUGS. J Pharmacokinet Pharmacodyn. 2006;33(5):571-94.

[52] Langdon G, Gueorguieva I, Aarons L, Karlsson M. Linking preclinical and clinical whole-body physiologically based pharmacokinetic models with prior distributions in NONMEM. Eur J Clin Pharmacol. 2007;63(5):485-98.

[53] Yang Y, Xu X, Georgopoulos PG. A Bayesian population PBPK model for multiroute chloroform exposure. J Expo Sci Environ Epidemiol. 2010;20(4):326-41.

[54] Krauss M, Burghaus R, Lippert J, Niemi M, Neuvonen P, Schuppert A, et al. Using Bayesian-PBPK modeling for assessment of inter-individual variability and subgroup stratification. In Silico Pharmacol. 2013;1:6.

[55] Krauss M, Tappe K, Schuppert A, Kuepfer L, Goerlitz L. Bayesian Population Physiologically-Based Pharmacokinetic (PBPK) Approach for a Physiologically Realistic Characterization of Interindividual Variability in Clinically Relevant Populations. PLoS One. 2015;10(10):e0139423.

[56] Tsamandouras N, Dickinson G, Guo Y, Hall S, Rostami-Hodjegan A, Galetin A, et al. Development and Application of a Mechanistic Pharmacokinetic Model for Simvastatin and its Active Metabolite Simvastatin Acid Using an Integrated Population PBPK Approach. Pharm Res. 2015;32(6):1864-83.

[57] Wendling T, Dumitras S, Ogungbenro K, Aarons L. Application of a Bayesian approach to physiological modelling of mavoglurant population pharmacokinetics. J Pharmacokinet Pharmacodyn. 2015;42(6):639-57.

[58] Tsiros P, Bois FY, Dokoumetzidis A, Tsiliki G, Sarimveis H. Population pharmacokinetic reanalysis of a Diazepam PBPK model: a comparison of Stan and GNU MCSim. J Pharmacokinet Pharmacodyn. 2019.

[59] Lang J, Vincent L, Chenel M, Ogungbenro K, Galetin A. Simultaneous Ivabradine Parent-Metabolite PBPK/PD Modelling Using a Bayesian Estimation Method. AAPS J. 2020;22(6):129.

[60] Yokley K, Tran HT, Pekari K, Rappaport S, Riihimaki V, Rothman N, et al. Physiologically-based pharmacokinetic modeling of benzene in humans: a Bayesian approach. Risk Anal. 2006;26(4):925-43.

[61] Gelman A. Bayesian data analysis. Third edition. ed. Boca Raton: CRC Press; 2014. xiv, 661 pages p.

[62] Gelman A, Bois F, Jiang J. Physiological Pharmacokinetic Analysis Using Population Modeling and Informative Prior Distributions. Journal of the American Statistical Association. 1996;91(436):1400-12.

[63] Wakefield JC, Smith AFM, Racine-Poon A, Gelfand AE. Bayesian Analysis of Linear and Non-Linear Population Models by Using the Gibbs Sampler. Journal of the Royal Statistical Society Series C (Applied Statistics). 1994;43(1):201-21.

[64] Wakefield J. The Bayesian Analysis of Population Pharmacokinetic Models. Journal of the American Statistical Association. 1996;91(433):62-75.

[65] Hsieh NH, Reisfeld B, Bois FY, Chiu WA. Applying a Global Sensitivity Analysis Workflow to Improve the Computational Efficiencies in Physiologically-Based Pharmacokinetic Modeling. Front Pharmacol. 2018;9:588.

[66] Proctor NJ, Tucker GT, Rostami-Hodjegan A. Predicting drug clearance from recombinantly expressed CYPs: intersystem extrapolation factors. Xenobiotica. 2004;34(2):151-78.

[67] Ito K, Houston JB. Prediction of human drug clearance from in vitro and preclinical data using physiologically based and empirical approaches. Pharm Res. 2005;22(1):103-12.

[68] Obach RS, Baxter JG, Liston TE, Silber BM, Jones BC, MacIntyre F, et al. The prediction of human pharmacokinetic parameters from preclinical and in vitro metabolism data. J Pharmacol Exp Ther. 1997;283(1):46-58.

[69] Barter ZE, Bayliss MK, Beaune PH, Boobis AR, Carlile DJ, Edwards RJ, et al. Scaling factors for the extrapolation of in vivo metabolic drug clearance from in vitro data: reaching a consensus on values of human microsomal protein and hepatocellularity per gram of liver. Curr Drug Metab. 2007;8(1):33-45.

[70] De Buck SS, Sinha VK, Fenu LA, Gilissen RA, Mackie CE, Nijsen MJ. The prediction of drug metabolism, tissue distribution, and bioavailability of 50 structurally diverse compounds in rat using mechanism-based absorption, distribution, and metabolism prediction tools. Drug Metab Dispos. 2007;35(4):649-59.

[71] Leggett RW, Williams LR. Suggested reference values for regional blood volumes in humans. Health Phys. 1991;60(2):139-54.

[72] Valentin J. Basic anatomical and physiological data for use in radiological protection: reference values. ICRP Publication 89. Ann ICRP. 2002;32(3-4):1-277.

[73] Davies B, Morris T. Physiological parameters in laboratory animals and humans. Pharm Res. 1993;10(7):1093-5.

[74] Kuwahira I, Gonzalez NC, Heisler N, Piiper J. Changes in regional blood flow distribution and oxygen supply during hypoxia in conscious rats. J Appl Physiol (1985). 1993;74(1):211-4.

[75] Kuwahira I, Gonzalez NC, Heisler N, Piiper J. Regional blood flow in conscious resting rats determined by microsphere distribution. J Appl Physiol (1985). 1993;74(1):203-10.

[76] Brown RP, Delp MD, Lindstedt SL, Rhomberg LR, Beliles RP. Physiological parameter values for physiologically based pharmacokinetic models. Toxicol Ind Health. 1997;13(4):407-84.

[77] Martinez MN, Amidon GL. A mechanistic approach to understanding the factors affecting drug absorption: a review of fundamentals. J Clin Pharmacol. 2002;42(6):620-43.

[78] Cao X, Gibbs ST, Fang L, Miller HA, Landowski CP, Shin HC, et al. Why is it challenging to predict intestinal drug absorption and oral bioavailability in human using rat model. Pharm Res. 2006;23(8):1675-86.

[79] Jones CR, Hatley OJ, Ungell AL, Hilgendorf C, Peters SA, Rostami-Hodjegan A. Gut Wall Metabolism. Application of Pre-Clinical Models for the Prediction of Human Drug Absorption and First-Pass Elimination. AAPS J. 2016;18(3):589-604.

Appendix A1: Supplementary material for Chapter 2

A1.1. Details of the partial rank correlation coefficient (PRCC) analysis

The global sensitivity analysis (GSA) was performed using R v. 3.4.2 [1] and combined the partial rank correlation coefficient (PRCC) with the Latin hypercube sampling (LHS) scheme [2, 3]. In the random sampling scheme, the whole range of each parameter is divided into N equal probability segments, each of which is sampled once. Monte–Carlo simulations (N = 10000) were performed to randomly generate independent simulated values of drug variables from their respective uniform distribution (Table 2.1). A matrix that consists of N rows for the number of simulations (sample size) and of k columns corresponding to the number of varied parameters (see Figure below) was generated. And using each combination of generated drug parameter values (each row of the matrix) N model solutions were calculated. The model outputs of interest (here, tissues Kpu outputs) were collected for each model simulation.

A correlation provides a measure of the strength of a linear association between an input and an output. Using the residuals obtained from the regression procedure, partial correlation characterises the linear relationship between the input parameters and the outputs after discounting the linear effects of the inputs on the outcome measures [4]. A Pearson's correlation coefficient (r_{xy}) between two variables x and y is given by:

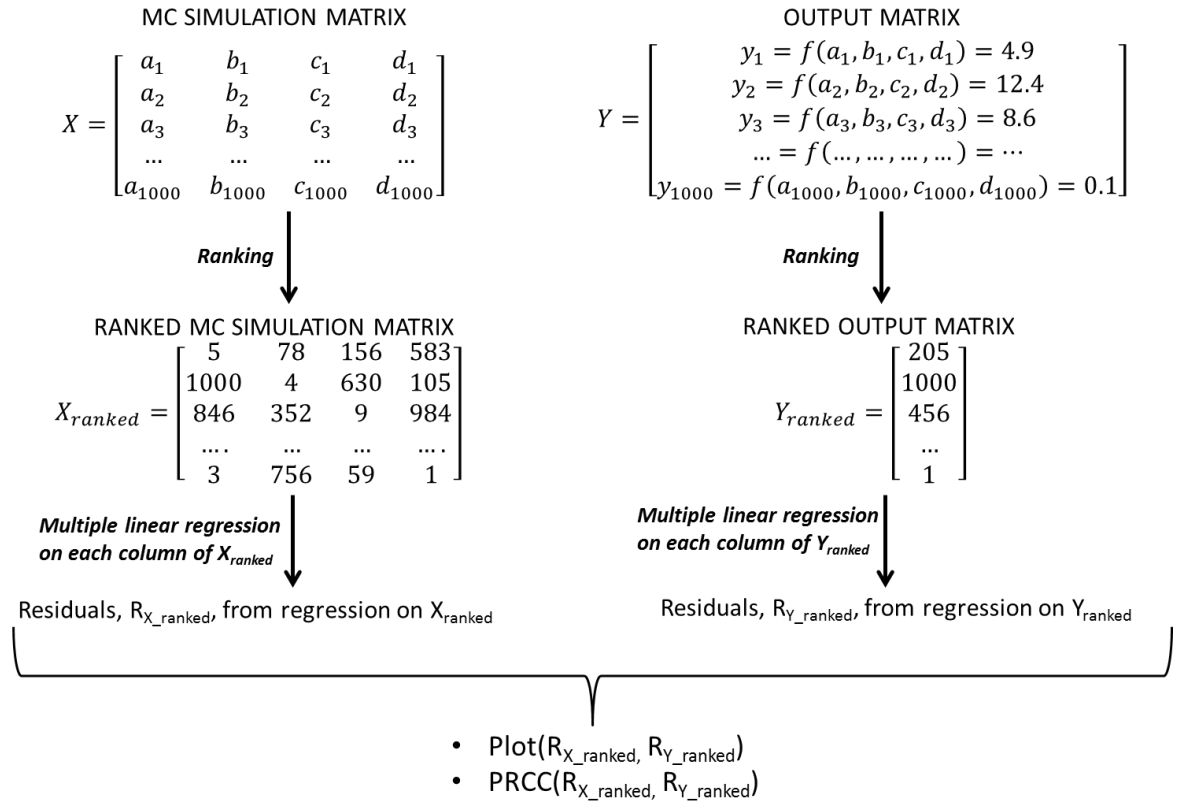
$$r_{xy} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \times \sum_i (y_i - \bar{y})^2}} \qquad \textbf{Eq. A1.1}$$

where x_i and y_i are the set of paired sampled data, and $\bar{x}$ and $\bar{y}$ are the respective sample means. PRCC performs a partial correlation on rank-transformed data. The transformation usually results in uniform residuals (x_j-$\hat{x}_j$) and (y-$\hat{y}_j$) for the transformed variables, where x_j is the rank transformed sampled j^{th} parameter, and y is the rank transformed output variable. $\hat{x}_j$ and $\hat{y}_j$ are built for k samples following two linear regression models: $\hat{x_j} = c_0 + \sum_{\substack{p=1 \\ p \neq j}}^{k} c_p x_p$ and $\hat{y} = b_0 + \sum_{\substack{p=1 \\ p \neq j}}^{k} b_p x_p$.

A Pearson correlation coefficient for the residuals from those two regression models gives the PRCC value for that specific parameter [4] (see **Figure A1.1** below). The measure of the PRCC indicates the importance of the uncertainty in estimating the parameter value to the imprecision in the output variable prediction [5] which allows a classification of parameters. The closer the PRCC value is to +1 or -1, the stronger the input parameter influences the outcome measure. The sign of the PRCC indicates the

qualitative relationship between the input parameter and the output variable. A positive PRCC value indicates that an increase in the input parameter value leads to an increase in the predicted output variable whereas a negative value indicates that it leads to a decrease in the predicted output variable. The significance of a non-zero PRCC value was tested using a two-sided Student's t-test [4]. As the number of tests performed is large, a multiple test correction was applied. The Bonferroni correction provides the most conservative approach to control for false positives multiplying the p-value of each PRCC by the number of tests performed [6]. If the corrected p-value is still below 0.001 for significance, then the PRCC is significantly different from 0. Otherwise, the PRCC will be considered not significant.

Figure A1.1. Scheme of sensitivity analysis performed with Monte Carlo simulation and PRCC methods



A1.2. Dependency between LogP and fu_p

Model input variables are typically assumed to be independent for practical reasons as non-independent inputs samples are more complex to generate and can need a very high sample size to compute sensitivity measures. However, the assumption of independence among input variables may not be appropriate given the nature of the relationship between lipophilicity and plasma protein binding. Consequently, several degrees of dependency between LogP and fu_p were considered when sampling the LogP and fu_p as drug inputs:

- Independence of LogP and fu_p: LogP and fu_p were each sampled independently from its defined uniform distribution with the LHS method.
- Linear relationship between LogP and fu_p while investigating different correlation coefficient ρ=-0.3, -0.5 and -0.9. The LHS method assumes that the sampling is performed independently for each parameter and introducing correlations between parameters other than in the special case of Gaussian distributions is not trivial. Iman and Conover developed a procedure to impose correlations on sampled values based on the ranks of the variables instead of using the values of the variables [7]. LogP and fu_p were generated using LHS from a bivariate normal distribution with statistical dependence between these two variables (correlation matrix of ρ) and each having a normal marginal distribution following Iman and Conover's method implemented in the R package "*EnvStats"* [8]. The variables were then transformed by their cumulative distribution function to get uniform distributions in the interval [0, 1]. The marginal distributions with the limits defined in **Table A1.2** were finally applied to the uniformly transformed variables to get variables uniformly distributed and correlated with the desired ρ.
- Nonlinear relationship between LogP/LogD and fu_p (ρ =-0.8 for neutral basic drugs, ρ =0.5 for acids) [9]. For example, LogP and pKa were each sampled independently from its defined uniform distribution with the LHS method as the relationships described by Yamazaki and Kanaoka were between LogD and plasma protein binding for basic, neutral, and acidic compounds. LogD was then calculated from LogP and pKa as follows (Eq. A1.2):

$$LogD_{pH\ 7.4} = LogP - log10(1 + 10^{7.4-pKa}) \qquad \textbf{Eq. A1.2}$$

For neutral and basic drugs, fu_p is calculated using the Eq. A1.3 [9]:

$$fu_p = (1 - \frac{0.5578 \cdot exp(logD_{pH\ 7.4}) + 0.0188}{0.5578 \cdot exp(logD_{pH\ 7.4}) + 1.0188}$$ **Eq. A1.3**

And for acidic drugs, fu_p is calculated by the Eq. A1.4 [9]:

$$fu_p = (1 - \frac{0.3127 \cdot exp(logD_{pH\ 7.4}) + 0.5121}{0.3127 \cdot exp(logD_{pH\ 7.4}) + 1.5121}$$ **Eq. A1.4**

When dealing with correlated factors, Saltelli et al. advised that dependencies be treated as explicit relationships with a noise term [10], and so fu_p is then a function of LogP-pKa and a noise factor. This noise factor following a standard normal distribution is added to the calculated fu_p which has been logit-transformed, the sum is then passed through an inverse-logit function in order to ensure that fu_p has logical constraints between 0 and 1 [11].

The sample size of the GSA for the different degrees of dependency evaluated is chosen to be the same (N=10000).

A1.3. Details of Incorporation of variability/uncertainty in physiological parameters

The sensitivity to physiological parameters was explored by incorporating 30% variability/uncertainty alternately on different biological parameters:

- Variability/uncertainty on all fractional tissue lipid volumes, fractional tissue water volumes, as well as acid phospholipid concentrations and tissue:plasma proteins ratios
- Variability/uncertainty only on fractional tissue lipid volumes and fractional tissue water volumes
- Variability/uncertainty only on acid phospholipid concentrations and tissue:plasma proteins ratios

Population distributions of fractional tissue volumes that have means matching the typical average values used for tissue composition-based models [12] and a CV of 30 % (30 % CV was selected here as a middle case) were generated.

Based on **Table A1.2**, the fractional parameter (f_{NL}, f_{NP}, f_{EW} and f_{IW}) were assumed to follow logistic-normal distributions [13]. However, the sum of fractional tissue volumes given in **Table A1.2** did not equal to 1, and a normalisation was applied to each fractional volume to keep the overall proportion. The j normalised fractional tissue volumes per tissue were assumed to follow a (j+1)-dimensional logistic-normal distribution which is derived after the transformation of a standard (j+1)-multivariate normal distribution with mean vector M and variance-covariance matrix Σ. Examples of how to generate samples from a logistic-normal distribution were previously reported [11, 13]. It was assumed that $F = [f_1, f_2, ..., f_j]^T \sim N_j(M, \Sigma)$, where F is a j-dimensional vector that follows a standard multivariate normal distribution with mean vector M defined as a null-vector of length j, and variance-covariance matrix Σ defined as a j-scalar matrix of 0.0862 ($=0.30^2$ in order to have 30 % CV). The vector F is updated by adding an additional zero element, so that $f_{j+1} = 0$. Then applying the logistic transformation (Eq. A1.5) on the updated vector F, we derived a (j+1)-dimensional vector $\Theta = [\theta_1, \theta_2, \ldots, \theta_{j+1}]^T$ following the logistic-normal distribution.

$$\theta_i = \frac{e^{f_i}}{\sum_{i=1}^{j+1} e^{f_i}} \qquad \textbf{Eq. A1.5}$$

For plasma, the two normalised fractional parameters (f_{NL} and f_{NP}) were assumed to follow a two-dimensional logistic normal distribution. For adipose, bone, brain, gut, heart, kidney, liver, lung muscle, skin, and spleen, four normalised fractional parameters

(f_{NL}, f_{NP}, f_{EW} and f_{IW}) were assumed to follow a four-dimensional logistic normal distribution. And for blood cells, the three normalised fractional parameters (f_{NL}, f_{NP} and f_{IW}) were assumed to follow a three-dimensional logistic normal distribution. With this approach, the normalised fractional parameters will always be constrained to be between 0 and 1 and their sum equal to 1. Here, M and Σ parameters were fixed to generate population distributions of fractional tissue volumes that have means matching the average values and a CV of 30% in the logistic domain.

Due to the high dimensionality coming from the 4^{th} or 3^{rd} order vector, evaluating the integral for the calculation of mean and variance of the logistic normal is challenging. A solution is to primarily simulate in R some initial data that agree with the constraints that all the fractional parameters should always be between 0 and 1 and their sum equal to 1. Assume ϕ being the vector of fractions and the sum of elements of ϕ equals to 1. A multivariate logistic vector of j-1 dimensions was found so that its multivariate inverse logit formulation (MVIL) has a mean (μ) equal to 1 summarised as MVIL(μ)=1. This was obtained by minimizing the sum of squares of the absolute value of |MVIL(μ)-ϕ | (using the function *nlminb* in R). A total of 1000 samples were generated with (j+1)-dimensional logistic-normal distribution for the normalised fractional volumes per tissue and these values were then multiplied by the sum of fractional volumes at individual tissue level so that the sampled fractional tissue volumes were in the original scale.

Also, acid phospholipids concentration, albumin and lipoprotein ratio, albumin ratio and lipoprotein ratio were sampled from a normal distribution N(μ, σ) where μ is the average value given in Table II and σ is the variance of the associated normal in order to generate distributions of these physiological parameters with mean matching the average values and a CV of 30%.

Finally, Kpu values were calculated based on Rodgers-Rowland equations [14, 15] for a hypothetical compound for each class (neutral, acid, weak base, strong base) and according to several scenarios of LogP and fu_p (Table 2.3) using three different assumptions: either both sampled fractional tissue volumes and sampled values of other physiological parameters, or only sampled fractional tissue volumes and average values of physiological parameters, or else average values of fractional tissue volumes and sampled values of other physiological parameters. BP is commonly assumed to be one for all compounds except ionised acids where a value of one minus haematocrit is generally used as an approximation when the BP data is not measured *in vitro/ ex vivo* (although BP is not involved in the Kpu equation for acids).

Table A1.1: Tissue blood flow rates (Q) and volumes (V) for a reference man of 70 kg

	Blood flow rate (Fraction of Qc)	Volume (Fraction of BW)	Density
Adipose	0.05	0.214	0.96
Bone	0.05	0.143	1.92
Brain	0.12	0.020	1.04
Gut	0.09	0.017	1.04
Heart	0.04	0.005	1.04
Kidney	0.19	0.004	1.04
Liver (arterial)	0.065	0.026	1.08
Lung	1	0.008	1.04
Muscle	0.17	0.4	1.04
Pancreas	0.01	0.001	1.04
Skin	0.05	0.037	1.04
Spleen	0.03	0.003	1.04
Rest of body	0.325	0.043	1.04
Plasma		0.046	1.04
Blood	1	0.079	1.04
-Arterial blood	1	0.020	1.04
-Venous blood	1	0.059	1.04

Tissue blood flow data are from [16, 17]
Qc is used to describe total cardiac output. Qc (L/min) was calculated as function of BW [18]:
Qc (L/h) =0.06×187 × BW0.81
Fraction of body weight data is from [19].
The value of 1.040 was used for all tissue density except for adipose, bone [19] and liver [20].

Table A1.2: Physiological input parameters used for tissue composition-based models in humans [12]

	Water (f_w)	Neutral lipids (f_{NL})	Neutral phospho-lipids (f_{ph})	Extra-cellular water (f_{EW})	Intra-cellular water (f_{IW})	Acidic phospho lipids (mg/g) (Concentration)	Albumin ratio	Lipoprotein ratio
Plasma	0.95	0.0032	0.0021					
Adipose	0.15	0.79	0.002	0.135	0.017	0.4	0.049	0.068
Bone	0.45	0.074	0.0011	0.1	0.346	0.67	0.1	0.05
Brain	0.78	0.051	0.0565	0.162	0.62	0.4	0.048	0.041
Gut	0.76	0.0487	0.0163	0.282	0.475	2.41	0.158	0.141
Heart	0.78	0.0115	0.0166	0.32	0.456	2.25	0.157	0.16
Kidney	0.76	0.0207	0.0162	0.273	0.483	5.03	0.13	0.137
Liver	0.73	0.0348	0.0252	0.161	0.573	4.56	0.086	0.161
Lung	0.78	0.003	0.009	0.336	0.446	3.91	0.212	0.168
Muscle	0.71	0.022	0.0072	0.079	0.63	2.42	0.064	0.059
Skin	0.67	0.0284	0.0111	0.382	0.291	1.32	0.277	0.096
Spleen	0.79	0.0201	0.0198	0.207	0.579	3.18	0.097	0.207
Thymus	0.78	0.0168	0.0092	0.15	0.626	2.3	0.075	0.075
Blood cells	0.63	0.0012	0.0033		0.603	0.57		

Table A1.3: Physiological input parameters used for tissue composition-based models on rats [21]

	Water (f_w)	Neutral lipids (f_{NL})	Neutral phospho-lipids (f_{NP})	Extra-cellular water (f_{EW})	Intra-cellular water (f_{IW})	Acidic phospho lipids (mg/g) (Concentration)	Albumin ratio	Lipoprotein ratio
Plasma	0.96	0.00147	0.00083	1				
Adipose	0.12	0.853	0.002	0.175	0.017	0.4	0.049	0.068
Bone	0.446	0.0273	0.0027	0.42	0.346	0.67	0.1	0.05
Brain	0.788	0.0392	0.0533	0.162	0.62	0.4	0.048	0.041
Gut	0.749	0.0292	0.0138	0.39	0.475	2.41	0.158	0.141
Heart	0.779	0.014	0.0118	0.156	0.456	2.25	0.157	0.16
Kidney	0.771	0.0123	0.0284	0.346	0.483	5.03	0.13	0.137
Liver	0.705	0.0138	0.0303	0.159	0.573	4.56	0.086	0.161
Lung	0.79	0.0219	0.014	0.484	0.446	3.91	0.212	0.168
Muscle	0.756	0.01	0.009	0.115	0.63	2.42	0.064	0.059
Skin	0.651	0.0239	0.018	0.462	0.291	1.32	0.277	0.096
Spleen	0.771	0.0077	0.0136	0.264	0.579	3.18	0.097	0.207
Thymus					0.626	2.3	0.075	0.075
Blood cells					0.603	0.57		

A1.4. Supplementary figures

Figure A1.2: Scatterplots of adipose Kpu vs LogP or fu_p for neutrals with different degrees of correlation between LogP and fu_p (0, -0.3, -0.5, -0.9, nonlinear)

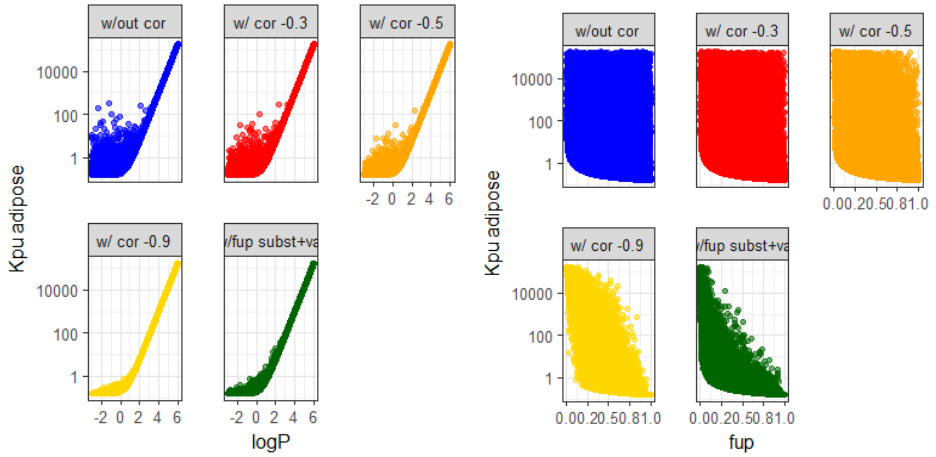


Figure A1.3: Scatterplots of adipose Kpu vs LogP, fu_p or pKa for acids with different degrees of correlation between LogP and fu_p (0, -0.3, -0.5, -0.9, nonlinear)

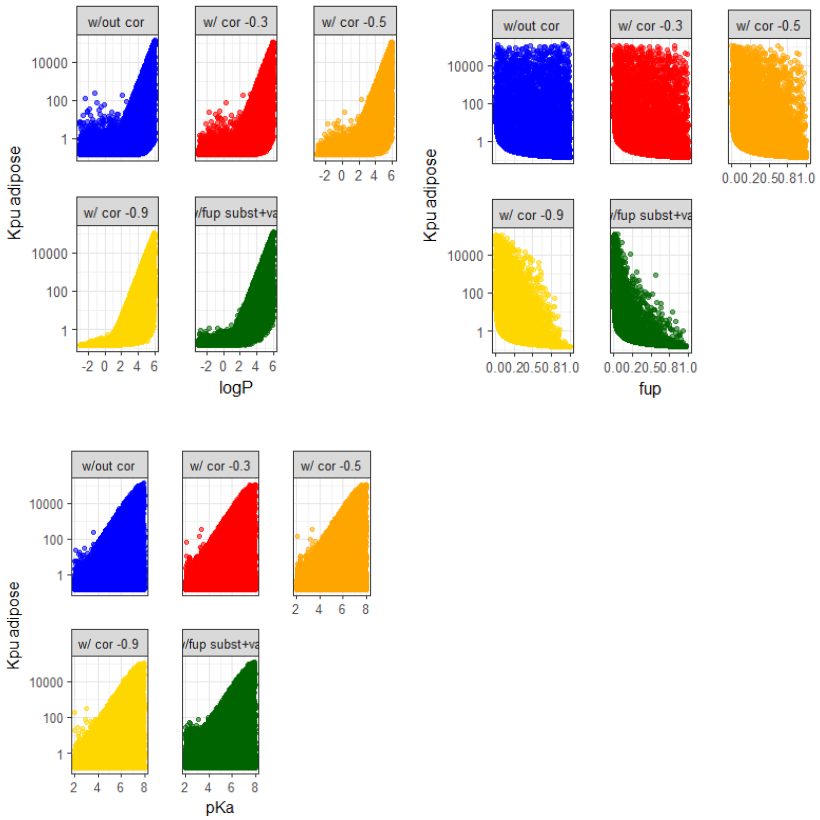


Figure A1.4: Scatterplots of adipose Kpu vs LogP, fu_p or pKa for weak bases with different degrees of correlation between LogP and fu_p (0, -0.3, -0.5, -0.9, nonlinear)

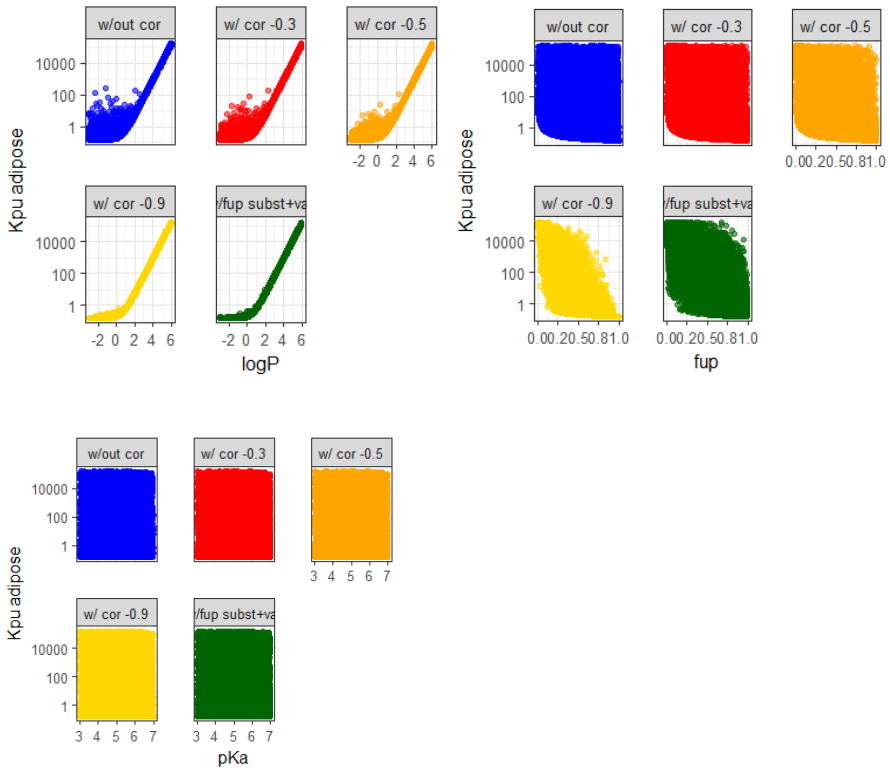


Figure A1.5: Scatterplots of adipose Kpu vs LogP, fu_p, pKa or BP for strong bases with different degrees of correlation between LogP and fu_p (0, -0.3, -0.5, -0.9, nonlinear)

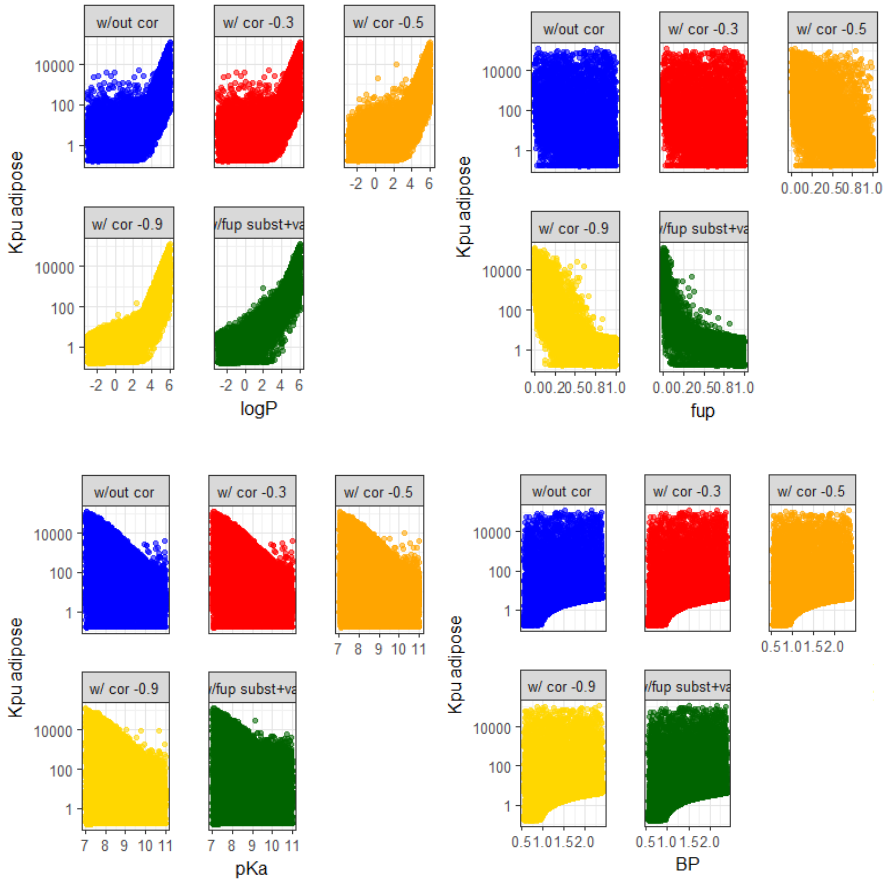
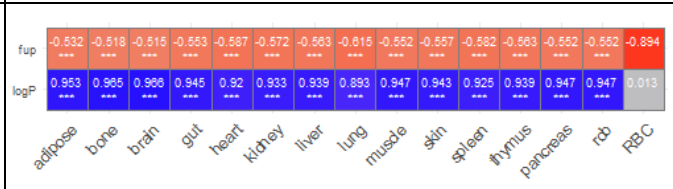
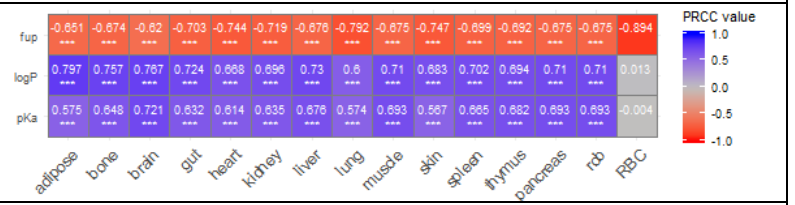
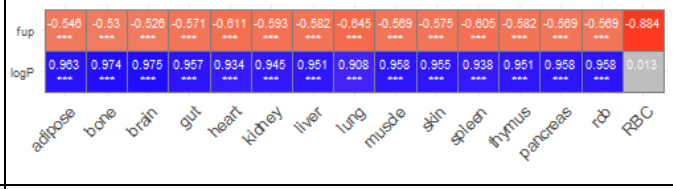
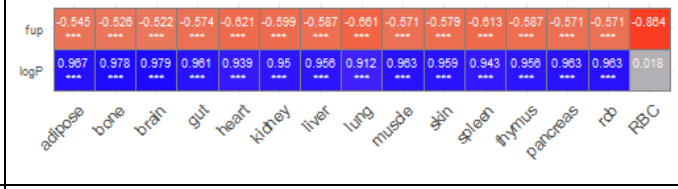
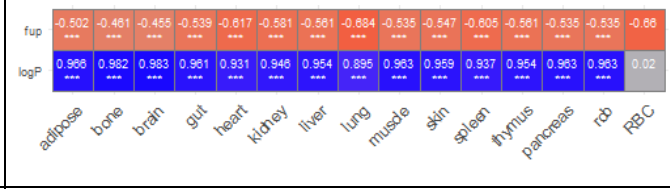
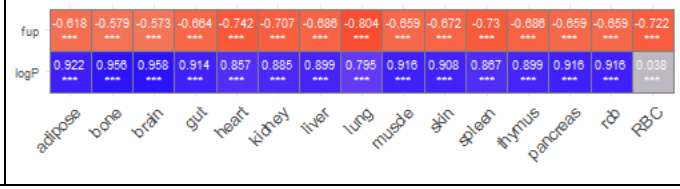
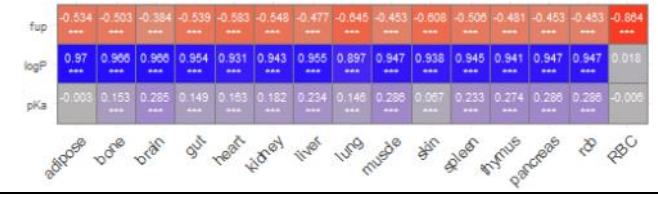
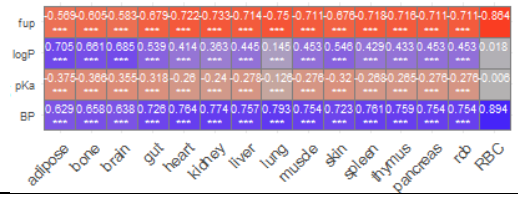
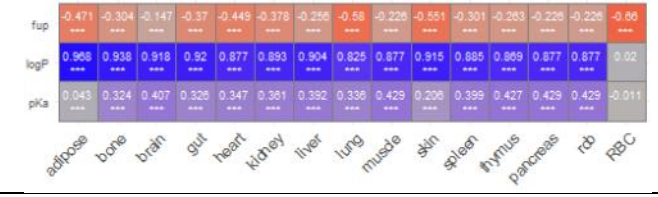
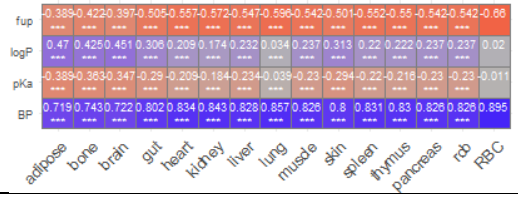
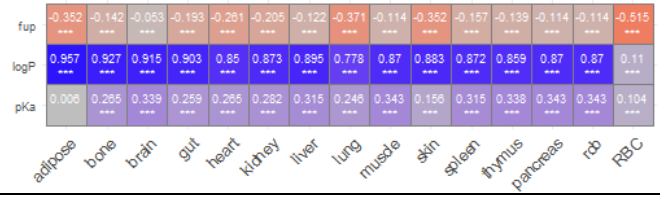
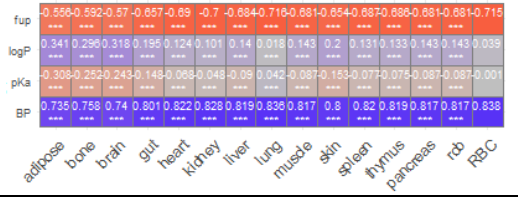


Figure A1.6: PRCC of tissue Kpus for neutrals and acids with different relationship between LogP and fraction unbound in plasma (fu_p)

LogP and fu_p relationship	Neutral	Acid
Independent		
Correlation of ρ=-0.3		
Correlation of ρ=-0.5		
Correlation of ρ=-0.9		
Nonlinear with fu_p (Yamazaki [9])		

*, **, *** *signifies a PRCC significantly different from zero: p<0.05, <0.01, 0.001 respectively*

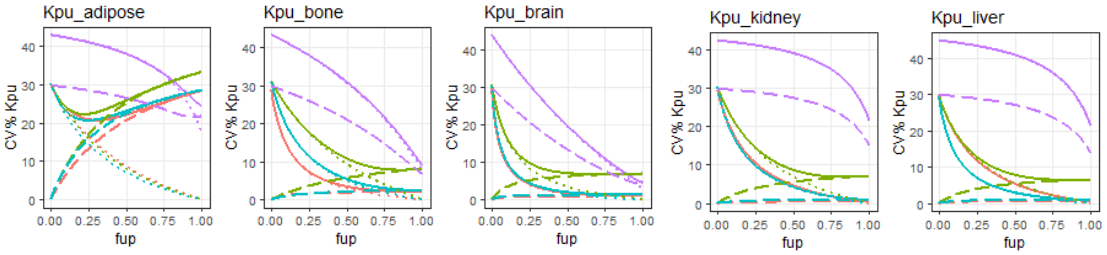
Figure A1.7: PRCC of tissue Kpus for each weak and strong bases with different relationship between LogP and fraction unbound in plasma (fu_p)

LogP and fu_p relationship	Weak base	Strong base
Independent		
Correlation of ρ=-0.3		
Correlation of ρ=-0.5		
Correlation of ρ=-0.9		
Nonlinear with fu_p (Yamazaki [9])		

*, **, *** signifies a PRCC significantly different from zero: p<0.05, <0.01, 0.001 respectively

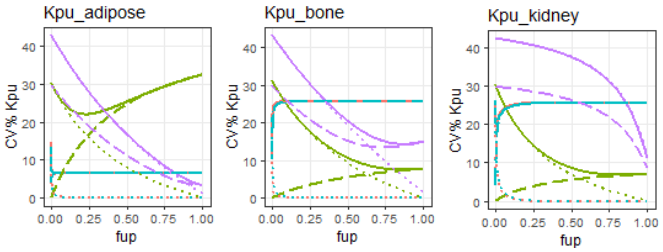
Figure A1.8: Effect of inputting CV30% uncertainty simultaneously in all tissue composition data (solid line) or individually in fractional tissue volumes (dashed line) and/or all tissue acidic phospholipids/extracellular protein ratios (dotted line) when varying fu_p or LogP on typical tissue Kpus

A. LogP= -0.3 and fu_p=[0.001:1]



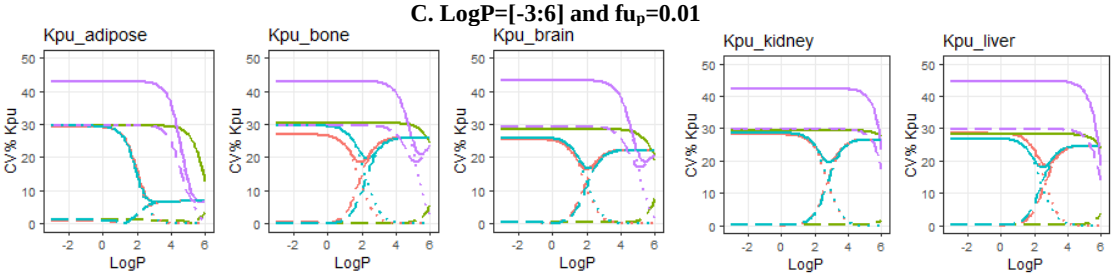
Other tissues Kpu: skin behaves similarly to bone; spleen behaves similarly to liver; gut, heart, lung, muscle, thymus, pancreas, rest of body behave similarly to kidney

B. LogP=3 and fu_p=[0.001:1]



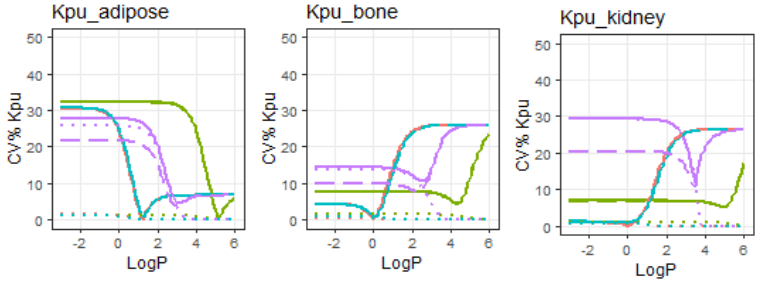
Other tissues Kpu: brain, skin behave similarly to bone; gut, heart, liver, lung, muscle, spleen, thymus, pancreas, rest of body behave similarly to kidney;

C. LogP=[-3:6] and fu_p=0.01



Other tissues Kpu: skin behaves similarly to bone; muscle, skin, spleen, thymus, pancreas, rest of body behave similarly to liver; gut, heart, lung behave similarly to kidney;

D. LogP=[-3:6] and fu_p=0.9



Other tissues Kpu: brain, skin, muscle, gut, heart, thymus, pancreas, rest of body behave similarly to bone; liver, lung, spleen behave similarly to kidney;

<u>Hypothetical compounds</u>: neutral (red line), acidic (green line), weakly basic (blue v) and strongly basic (purple line)
<u>4 case scenarios</u>: hydrophilic (A), lipophilic (B), highly bound (C), lowly bound (D) compound

A1.5. References

[1] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria2018.

[2] McKay MD, Beckman RJ, Conover WJ. Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. Technometrics. 1979;21(2):239-45.

[3] Blower SM, Dowlatabadi H. Sensitivity and Uncertainty Analysis of Complex-Models of Disease Transmission - an Hiv Model, as an Example. Int Stat Rev. 1994;62(2):229-43.

[4] Marino S, Hogue IB, Ray CJ, Kirschner DE. A methodology for performing global uncertainty and sensitivity analysis in systems biology. J Theor Biol. 2008;254(1):178-96.

[5] Hamby DM. A review of techniques for parameter sensitivity analysis of environmental models. Environ Monit Assess. 1994;32(2):135-54.

[6] Abdi H. The Bonferonni and Šidák Corrections for Multiple Comparisons. Encyclopedia of measurement and statistics. 2007;3.

[7] Iman RL, Conover WJ. A Distribution-Free Approach to Inducing Rank Correlation among Input Variables. Commun Stat B-Simul. 1982;11(3):311-34.

[8] Millard SP. EnvStats: An R Package for Environmental Statistics: Springer, New York; 2013.

[9] Yamazaki K, Kanaoka M. Computational prediction of the plasma protein-binding percent of diverse pharmaceutical compounds. J Pharm Sci. 2004;93(6):1480-94.

[10] Saltelli A. Global sensitivity analysis : the primer. Chichester, England ; Hoboken, NJ: John Wiley; 2008. x, 292 p. p.

[11] Tsamandouras N, Wendling T, Rostami-Hodjegan A, Galetin A, Aarons L. Incorporation of stochastic variability in mechanistic population pharmacokinetic models: handling the physiological constraints using normal transformations. J Pharmacokinet Pharmacodyn. 2015;42(4):349-73.

[12] Poulin P, Jones RD, Jones HM, Gibson CR, Rowland M, Chien JY, et al. PHRMA CPCDC initiative on predictive models of human pharmacokinetics, part 5: prediction of plasma concentration-time profiles in human by using the physiologically-based pharmacokinetic modeling approach. J Pharm Sci. 2011;100(10):4127-57.

[13] Aitchison J, Shen SM. Logistic-Normal Distributions - Some Properties and Uses. Biometrika. 1980;67(2):261-72.

[14] Rodgers T, Leahy D, Rowland M. Physiologically based pharmacokinetic modeling 1: predicting the tissue distribution of moderate-to-strong bases. J Pharm Sci. 2005;94(6):1259-76.

[15] Rodgers T, Rowland M. Physiologically based pharmacokinetic modelling 2: predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. J Pharm Sci. 2006;95(6):1238-57.

[16] Williams LR, Leggett RW. Reference values for resting blood flow to organs of man. Clin Phys Physiol Meas. 1989;10(3):187-217.

[17] Valentin J. Basic anatomical and physiological data for use in radiological protection: reference values. ICRP Publication 89. Ann ICRP. 2002;32(3-4):1-277.

[18] Nestorov I. Modelling and simulation of variability and uncertainty in toxicokinetics and pharmacokinetics. Toxicol Lett. 2001;120(1-3):411-20.

[19] Brown RP, Delp MD, Lindstedt SL, Rhomberg LR, Beliles RP. Physiological parameter values for physiologically based pharmacokinetic models. Toxicol Ind Health. 1997;13(4):407-84.

[20] Heinemann A, Wischhusen F, Puschel K, Rogiers X. Standard liver volume in the Caucasian population. Liver Transpl Surg. 1999;5(5):366-8.
[21] Poulin P, Theil FP. Prediction of pharmacokinetics prior to in vivo studies. 1. Mechanism-based prediction of volume of distribution. J Pharm Sci. 2002;91(1):129-56.

Appendix A2: Supplementary material for Chapter 3

A2.1. Supplementary Tables

Table A2.1: Physiological parameter values for a reference man (70kg)[1-4] and rat (250g)(Brown et al., 1997; Kuwahira et al., 1994)

Tissues	Human		Rat	
	Blood flows (L/min)	Volume (L)	Blood flows (ml/min)	Volume (ml)
Adipose	0.292	15.619	5.816	19.792
Bone	0.292	5.210	10.136	5.401
Brain	0.701	1.346	1.662	1.370
Gut	0.526	1.010	9.139	5.385
Heart	0.234	0.316	4.071	0.793
Kidney	1.109	0.296	11.715	1.755
Liver	1.489	1.666	12.546	8.472
Lung	5.839	0.512	83.085	1.202
Muscle	0.993	26.923	23.098	97.188
Pancreas	0.058	0.094	1.496	0.769
Skin	0.292	2.497	4.819	45.745
Spleen	0.175	0.175	0.831	0.481
Stomach	0.058	0.141	1.080	1.106
RoB	0.730	2.914	7.228	27.938
Arterial	5.839	1.329	83.085	6.538
Venous	5.839	3.988	83.085	13.077

RoB: rest of body

Table A2.2: Tissue composition data for human (Poulin et al., 2011)

Tissue	Vw	vNL	vNP	fEW	fIW	cAP	ALR	AR	LPR
plasma	0.95	0.0032	0.0021	NaN	NaN	NaN	NaN	NaN	NaN
adipose	0.15	0.79	0.002	0.135	0.017	0.4	0.15	0.049	0.068
bone	0.45	0.074	0.0011	0.1	0.346	0.67	0.5	0.1	0.05
brain	0.78	0.051	0.0565	0.162	0.62	0.4	0.5	0.048	0.041
gut	0.76	0.0487	0.0163	0.282	0.475	2.41	0.5	0.158	0.141
heart	0.78	0.0115	0.0166	0.32	0.456	2.25	0.5	0.157	0.16
kidneys	0.76	0.0207	0.0162	0.273	0.483	5.03	0.5	0.13	0.137
liver	0.73	0.0348	0.0252	0.161	0.573	4.56	0.5	0.086	0.161
lungs	0.78	0.003	0.009	0.336	0.446	3.91	0.5	0.212	0.168
muscle	0.71	0.022	0.0072	0.079	0.63	2.42	0.5	0.064	0.059
pancreas	0.641	0.0403	0.009	0.12	0.664	1.67	NaN	0.06	0.06
skin	0.67	0.0284	0.0111	0.382	0.291	1.32	0.5	0.277	0.096
spleen	0.79	0.0201	0.0198	0.207	0.579	3.18	0.5	0.097	0.207
thymus	0.78	0.0168	0.0092	0.15	0.626	2.3	0.5	0.075	0.075
RBC	0.63	0.0012	0.0033	NaN	0.603	0.57	NaN	NaN	NaN

Vw: fractional volume of water; vNL: fractional volume of neutral lipids; vNP: fractional volume of neutral phospholipids; fEW: fractional volume of extracellular water; fIW: fractional volume of intracellular water; cAP: acid phospholipid concentrations (mg/g); ALR: albumin and lipoprotein ratio; AR: albumin ratio; LPR: lipoprotein ratio; RBC: red blood cells

Table A2.3: Tissue composition data for rat [5]

Tissue	Vw	vNL	vNP	fEW	fIW	cAP	AR	LPR
plasma	0.96	0.0023	0.0013	0.945	NaN	0.057	NaN	NaN
adipose	0.144	0.853	0.0016	0.135	0.017	0.4	0.049	0.068
bone	0.417	0.0174	0.0016	0.1	0.346	0.67	0.1	0.05
brain	0.753	0.0391	0.0015	0.162	0.62	0.4	0.048	0.041
gut	0.738	0.0375	0.0124	0.282	0.475	2.41	0.158	0.141
heart	0.568	0.0135	0.0106	0.32	0.456	2.25	0.157	0.16
kidneys	0.672	0.0121	0.024	0.273	0.483	5.03	0.13	0.137
liver	0.642	0.0135	0.0238	0.161	0.573	4.56	0.086	0.161
lungs	0.574	0.0215	0.0123	0.336	0.446	3.91	0.212	0.168
muscle	0.726	0.01	0.0072	0.118	0.63	1.53	0.064	0.059
pancreas	0.641	0.0403	0.009	0.12	0.664	1.67	0.06	0.06
skin	0.658	0.0603	0.0044	0.382	0.291	1.32	0.277	0.096
spleen	0.562	0.0071	0.0107	0.207	0.579	3.18	0.097	0.207
thymus	0.752	0.0168	0.0092	0.15	0.626	2.3	0.075	0.075
RBC	0.6	0.0017	0.0029	NaN	0.603	0.5	NaN	NaN

Vw: fractional volume of water; vNL: fractional volume of neutral lipids; vNP: fractional volume of neutral phospholipids; fEW: fractional volume of extracellular water; fIW: fractional volume of intracellular water; cAP: acid phospholipid concentrations (mg/g); AR: albumin ratio; LPR: lipoprotein ratio; RBC: red blood cells

Table A2.4: Clinical PK studies of diazepam available

Study	Number of PK profiles	IV dose (length of infusion)	Reported total clearance (s.d.)	Reported volume of distribution (s.d)	Number of subjects and characteristics	fup and BP used in study	Reference
1	1 average	0.15mg/kg (1min)	2.1 (0.3) L/h (plasma)	87 (2) L	6 males and 5 females	NA	[6]
2	4 average	0.15mg/kg (0.375min)	1.3 to 1.7 L/h (plasma)	80 to 161 L	1 young male, 1 elderly male, 1 young female, 1 elderly female	fup: 0.009 to 0.027	[7]
3	1 average	10mg (2min)	1.2 (0.4) L/h (plasma)	63 (19) L	6 males	NA	[8]
4	1 individual	0.1mg/kg (2min)	1.6 (0.5) L/h (plasma)	66 (13) L	1 normal individual	fup: 0.026 (0.01); BP: 0.58 (0.15)	[9]
5	1 average	0.1mg/kg (2min)	1.6 (0.2) L/h (plasma)	79 (20) L	10 young males	fup: 0.032 (0.008); BP:0.58 (0.11)	[10]
6	4 individual	10mg (bolus)	3.3 to 4.0 L/h (blood)	105 to 174 L	4 healthy males	NA	[11]
7	23 individual	0.15mg/kg (0.375min)	1.8 L/h (plasma)	NA	12 healthy males/females and 11 healthy females	fup: 0.015; BP:0.65	[7, 12, 13]

Table A2.5: Rat PK studies of diazepam available

Study	Number of PK profiles	IV dose (length of infusion)	Reported total clearance (s.d.)	Reported volume of distribution (s.d)	Number of subjects and characteristics	fup and BP used in study	Reference
1	1 average	4mg/kg (bolus);	1.1 (0.2) L/h (plasma)	1.3 (0.2) L	10 male Wistar	fup: 0.036 (0.009); BP:1.06 (0.04)	[14]
2	1 average	1.2mg/kg (bolus)	1.2 (0.03) L/h (plasma)	1.4 (0.1) L	4 male Wistar	fup:0.14 (0.003); BP:1.037 (0.007)	[15, 16]
3	1 average	1mg (5min infusion)	400 ml/min (intrinsic clearance)	NA	24 male Sprague-Dawley	fup:0.15; BP:1	[17, 18]
4	2 average	5mg/kg (bolus)	1.1 L/h for middle-aged rat, 3.3 L/h for old rats (plasma)	1.2 for middle-aged rat, 13 L for old rats	5 middle-aged male Wistar and 5 old male Wistar	NA	[19]
5	1 average	5mg/kg (bolus)	3.9 (0.5) L/h (blood)	1.1 (0.2) L	3 male Wistar	fup:0.137 (0.011); BP:0.38 (0.07)	[9]

A2.2. Closed form solutions of the lumped 3 compartment model and the 14 compartmental PBPK model

A2.2.1. Equations for the lumped PBPK model (3 compartments)

This model assumes that drug very quickly reach a quasi-steady state in the central compartment (lung, arterial, venous). The differential equation (Eq. A2.1) for the drug concentration in this compartment is:

$$\frac{dC_{central}}{dt} = \left(\frac{Q_{p1} \cdot \mathrm{C_{p1}}}{Kb_{p1}} + \frac{Q_{p2} \cdot \mathrm{C_{p2}}}{Kb_{p2}} - Q_{central} \cdot \frac{C_{central}}{Kb_{central}} - CL_{blood} \cdot \frac{C_{central}}{Kb_{central}} \right) \cdot \frac{1}{V_{central}} \quad \textbf{Eq. A2.1}$$

Where $C_{central}$= $C_{arterial}$+C_{venous}+C_{lung}, C_{p1}, C_{p2} are the venous blood total drug concentration in the central, the peripheral 1 and peripheral 2 compartments, respectively. The blood flow $Q_{central}$ and the volume $V_{central}$ of the central compartment are defined as $Q_{central}$=cardiac output and $V_{central}$= $V_{arterial}$+V_{venous}+ V_{lung} ; $Kb_{central}$ is defined as the blood tissue-to-plasma partitioning coefficient of the central compartment and corresponding to ($V_{arterial}$+V_{venous}+V_{lung} Kb_{lung})/(V_{lung}+$V_{arterial}$+V_{venous}); Q_{p1} and Kb_{p1} are the blood flow and blood tissue-to-plasma partitioning coefficient of all tissues in the lumped peripheral 1 compartment; Kb_{p1} is weighted by the following volume (ΣV_i-V_{liver}-V_{kidney}) + V_{liver} (1-CL_H/Q_{liver}) + V_{kidney} (1-CL_R/Q_{renal}))/ΣV_i where i includes tissue lumped in the peripheral 1 compartment; Q_{p2} and Kb_{p2} are the blood flow and blood tissue-to-plasma partitioning coefficient of all the tissues in the lumped peripheral 2 compartment; CL_{blood} is the total blood clearance and corresponds here to the sum of the hepatic and renal clearance (CL_H and CL_R, respectively).

Similar to a classical 3 compartment model with an infusion dose at time t_{dose} and length of infusion Tinf, the total drug concentration C(t) in the central compartment (lung, arterial, venous) of the lumped model can be described as closed form solutions (Eq. A2.2-Eq. A2.3):

If t-t_{dose}≤ Tinf,

$$C(t) = \frac{Dose}{Tinf} \cdot \left[\frac{A}{\alpha} \cdot \left(1 - e^{-\alpha \cdot (t - t_{dose})}\right) + \frac{B}{\beta} \cdot \left(1 - e^{-\beta \cdot (t - t_{dose})}\right) + \frac{C}{\gamma} \cdot \left(1 - e^{-\gamma \cdot (t - t_{dose})}\right) \right] \quad \textbf{Eq. A2.2}$$

If t-t_{dose}>Tinf,

$$C(t) = \frac{Dose}{Tinf} \cdot \left[\frac{A}{\alpha} \cdot \left(1 - e^{-\alpha \cdot (t - t_{dose} - Tinf)}\right) + \frac{B}{\beta} \cdot \left(1 - e^{-\beta \cdot (t - t_{dose} - Tinf)}\right) + \frac{C}{\gamma} \cdot \left(1 - e^{-\gamma \cdot (t - t_{dose} - Tinf)}\right) \right] \quad \textbf{Eq. A2.3}$$

With the variables α, β, γ, A, B and C defined as following:

$A = \frac{1}{V} \cdot \frac{k_{21}-\alpha}{\alpha-\beta} \cdot \frac{k_{31}-\alpha}{\alpha-\gamma}$, $B = \frac{1}{V} \cdot \frac{k_{21}-\beta}{\beta-\alpha} \cdot \frac{k_{31}-\beta}{\beta-\gamma}$, $C = \frac{1}{V} \cdot \frac{k_{21}-\gamma}{\gamma-\beta} \cdot \frac{k_{31}-\gamma}{\gamma-\alpha}$

$a_0 = k \cdot k_{21} \cdot k_{31}$, $a_1 = k \cdot k_{31} + k_{21} \cdot k_{31} + k_{21} \cdot k_{13} + k \cdot k_{21} + k_{31} \cdot k_{21}$, $a_2 = k + k_{12} + k_{13} + k_{21} + k_{31}$

$p = a_1 - \frac{a_2^2}{3}$, $q = 2 \cdot \frac{a_2^3}{27} - \frac{a_1 \cdot a_2}{3} + a_0$,

$r_1 = \sqrt{-\left(\frac{p^3}{27}\right)}$, $r_2 = 2 \cdot r_1^{1/3}$, $\Phi = \frac{\arccos\left(-\frac{q}{2 \cdot r_1}\right)}{3}$

$\alpha = -\left(\cos(\Phi) \cdot r_2 - \frac{a_2}{3}\right)$, $\beta = -\left(\cos\left(\Phi + \frac{2 \cdot \pi}{3}\right) \cdot r_2 - \frac{a_2}{3}\right)$, $\gamma = -\left(\cos\left(\Phi + \frac{4 \cdot \pi}{3}\right) \cdot r_2 - \frac{a_2}{3}\right)$

The general 3-compartment linear mammillary model is available in NONMEM and parameterized in micro-constants k, k_{12}, k_{21}, k_{13}, and k_{31} (TRANS 1 subroutine) or with CL, V_1, Q_2, V_2, Q_3 and V_3 (TRANS 4 subroutine). These parameters are defined as following (Eq. A2.4-Eq. A2.8):

$$k_{12} = \frac{Q_2}{(V_1 - V_{arterial} - V_{venous}) \cdot Kb_1 + V_{arterial} + V_{venous}} \quad \textbf{Eq. A2.4}$$

$$k_{21} = \frac{Q_2}{Kb_2 \cdot \left[V_2 - V_{liver} \cdot \frac{CL_H}{Q_{liver}} - V_{kid} \cdot \frac{CL_R}{Q_{kidney}}\right]} \quad \textbf{Eq. A2.5}$$

$$k_{31} = \frac{Q_3}{V_3 \cdot Kb_3} \quad \textbf{Eq. A2.6}$$

$$k_{13} = \frac{Q_3}{(V_1 - V_{arterial} - V_{venous}) \cdot Kb_1 + V_{arterial} + V_{venous}} \quad \textbf{Eq. A2.7}$$

$$k = \frac{CL_{blood}}{(V_1 - V_{arterial} - V_{venous}) \cdot Kb_1 + V_{arterial} + V_{venous}} \quad \textbf{Eq. A2.8}$$

Where Kb_1, Kb_2 and Kb_3 are the blood tissue-to-plasma partitioning coefficient of the central, lumped peripheral 1 compartment and lumped peripheral 2 compartments; CL_{blood} is the total blood clearance.

For the lumped model in man, the following blood flows (Q) and volumes (V) are used:

$Q_2 = Q_{muscle} + Q_{bone} + Q_{skin} + Q_{brain} + Q_{RoB} + Q_{portal\,vein} + Q_{heart} + Q_{kidney}$, $Q_3 = Q_{adipose}$, $Q_1 = Q_2 + Q_3$

$V_1 = V_{lung} + V_{art} + V_{ven}$, $V_2 = V_{muscle} + V_{bone} + V_{skin} + V_{brain} + V_{RoB} + V_{liver} + V_{stomach} + V_{gut} + V_{pancreas} + V_{spleen} + V_{heart} + V_{kidney}$, $V_3 = V_{adipose}$

For the lumped model in rat, the following blood flows (Q) and volumes (V) are used:

$Q_2 = Q_{muscle} + Q_{bone} + Q_{adipose} + Q_{brain} + Q_{RoB} + Q_{portal\,vein} + Q_{heart} + Q_{kidney}$, $Q_3 = Q_{skin}$, $Q_1 = Q_2 + Q_3$

$V_1 = V_{lung} + V_{art} + V_{ven}$, $V_2 = V_{muscle} + V_{bone} + V_{adipose} + V_{brain} + V_{RoB} + V_{liver} + V_{stomach} + V_{gut} + V_{pancreas} + V_{spleen} + V_{heart} + V_{kidney}$, $V_3 = V_{skin}$

A2.2.2. Equations for 14 compartmental PBPK model

For the central (arterial, venous blood and lung compartment) and renal elimination form part of the central compartment, the rate equation is defined as (Eq. A2.9):

$$V_{central}\frac{dC_{central}}{dt} = \left(\sum Q_i \cdot \frac{\mathrm{C_i}}{Kb_i}\right) - (Q_{central} + CL_R) \cdot C_{central} \qquad \textbf{Eq. A2.9}$$

where $C_{central}$ is the concentration of the central compartment. The blood flow $Q_{central}$ and the volume $V_{central}$ of the central compartment are defined as $Q_{central}$=cardiac output and $V_{central}$= $V_{arterial}$+V_{venous}+ V_{lung}; $Kb_{central}$ is defined as the blood tissue-to-plasma partitioning coefficient of the central compartment and corresponding to ($V_{arterial}$+V_{venous}+V_{lung} Kb_{lung})/(V_{lung}+$V_{arterial}$+V_{venous}). CL_R is the renal clearance. The sum $\sum Q_i \cdot \frac{C_i}{Kb_i}$ includes all the ith tissues except the stomach, gut, pancreas and spleen tissues.

For each ith non-eliminating tissues, the rate equation is defined as (Eq. A2.10):

$$V_i \cdot \frac{dC_i}{dt} = Q_i \cdot \left(\mathrm{C_{central}} - \frac{C_i}{Kb_i}\right) \qquad \textbf{Eq. A2.10}$$

where C_i, V_i and Q_i are the total drug concentration, the volume and blood flow of the ith-tissue. For each ith-tissue, the tissue-to-blood partition coefficient, Kb, represents the ratio of steady state tissue drug concentration to the steady state blood concentration.

For the liver, the rate equation is defined as (Eq. A2.11):

$$V_{liver} \cdot \frac{dC_{liver}}{dt} = Q_{ha} \cdot \mathrm{C_{central}} + \sum Q_{splan,i} \cdot \frac{C_{splan,i}}{Kb_{splan,i}} - Q_{liver} \cdot \frac{C_{liver}}{Kb_{liver}} - CLint_{liver} \cdot fu_b \cdot \frac{C_{liver}}{Kb_{liver}} \qquad \textbf{Eq. A2.11}$$

where the $Q_{splan,i}$, $C_{splan,i}$ and $Kb_{splan,i}$ are the concentration, the blood flow, the volume and the blood partition coefficient of the ith splanchnic organs (stomach, gut, pancreas and spleen); Q_{ha} is the blood flow from the hepatic artery; C_{liver}, Q_{liver}, V_{liver}, Kb_{liver} are the concentration, the blood flow, the volume and the blood partition coefficient of the liver; CLint is the hepatic intrinsic clearance and fu_b is the fraction unbound in blood.

A2.3. Clustering analysis

Clustering algorithms use the distance between observations in order to separate them into different groups. Euclidean distance is the most commonly used method to measure the distance. It is calculated using the square of the difference between x and y coordinates of the two points a and b in a 2-dimensional space:

$$\text{Euclidean distance}(a, b) = \sqrt{(a_x + b_x)^2 + \left(a_y + b_y\right)^2} \qquad \textbf{Eq. A2.12}$$

Note, if two variables do not have the same units, one may have more weight in the calculation of the Euclidean distance than the other. Therefore, it is preferable to scale the data such that variables are independent of their unit.

A2.3.1. K-means clustering

The K-means clustering method aims to assign observations to k groups. After pre-specifying a number of k clusters with randomly selected centroids (cluster means) representing each cluster, the algorithm minimizes the squared error between each observation and the centroid by calculating the distance of all data points to the centroids and then assign the data points to the closest cluster. During each iteration, it updates the centroids of each cluster by taking the mean of all data in the new clusters. The iterative process continues until all points converge (convergence criterion met) and cluster centres stop moving. This algorithm is highly computing-efficient and produce tight clusters, but it is sensitive to initial seeds and outliers. The effect of outliers can greatly skew the central measure, especially when the mean is used. Additionally, k-means clustering method is found to work well when the shape of the clusters is spherically shaped (assuming uncorrelated variables) but not so well when clusters are globular or of varying sizes and density, which is more realistic and common and may be our case.

The k-means algorithm aims to minimize the within cluster sum of squared errors (SSE) which is equal to the sum of pairwise squared Euclidean distances:

$$SSE = \sum_{k=1}^{K} \sum_{x \in C_k} (x - c_k)^2 \qquad \textbf{Eq. A2.13}$$

Where k is the index specifying a given cluster, c_k is the central mean, and x are the observations belonging to that central mean, c_k are the observations that are part of the cluster cantered by the particular centroid c_k.

As k-means algorithm uses a random set of initial centroids, the clustering results may vary every time. By running the algorithm several times with different initial centres and taking the most optimal one, it can limit this issue and stabilize the results. More details on k-means clustering can be found in [20].

A2.3.2. Hierarchical clustering

Contrary to the k-means clustering algorithm, the hierarchical clustering is not as sensitive to outliers and does not require a pre-specified number of clusters nor randomly select initial centroids. Hierarchical clustering builds a hierarchy of clusters, based on the measure of closeness/similarity between the observations and hence is more informative than k-means clustering. Before applying a clustering algorithm, dissimilarity (or similarity) measures between elements needs to be derived as a basis for comparing the different observations. In the agglomerative clustering (most common type of hierarchical clustering), small clusters are merged into larger ones: each data point in the beginning stands for an individual cluster on their own, and then the two closest clusters are merged into a new cluster until only one cluster is left. The decision of merging two clusters is based on the distance between observations and the distance between clusters (linkage). Calculating the distance matrix between observations can be done by calculating the Euclidean distance. Then for calculating the linkage, Ward's method [21] is the most popular as it usually creates compact, homogeneous clusters, by joining clusters that minimally increase a given measure of heterogeneity. This measure is the sum of the square of the distance between two clusters c1 and c2:

$$D_{c1,c2} = \frac{n_{c1} \cdot n_{c2}}{n_{c1} + n_{c2}} \parallel \bar{x}_{c1} - \bar{x}_{c2} \parallel^2 \qquad \textbf{Eq. A2.14}$$

Where $\bar{x}_{ck}$ is the centre of cluster ck (with k, the index specifying a given cluster), and n is the number of observations in it. At first, the sum of squares starts out at zero (because every point is in its own cluster) and then grows as clusters are merged.

Hierarchical clustering provides an output as a tree-like diagram called a dendrogram. By observing, the number of clusters, that best depict different groups, can be chosen (number of vertical lines in the dendrogram cut by a horizontal line that can transverse the maximum vertical distance vertical without intersecting a cluster). Hierarchical clustering is suitable for datasets with arbitrary shape and attributes of arbitrary type. The hierarchical relationship among clusters is easily detected, and the scalability is in general relatively high. More details on hierarchical clustering can be found in [22].

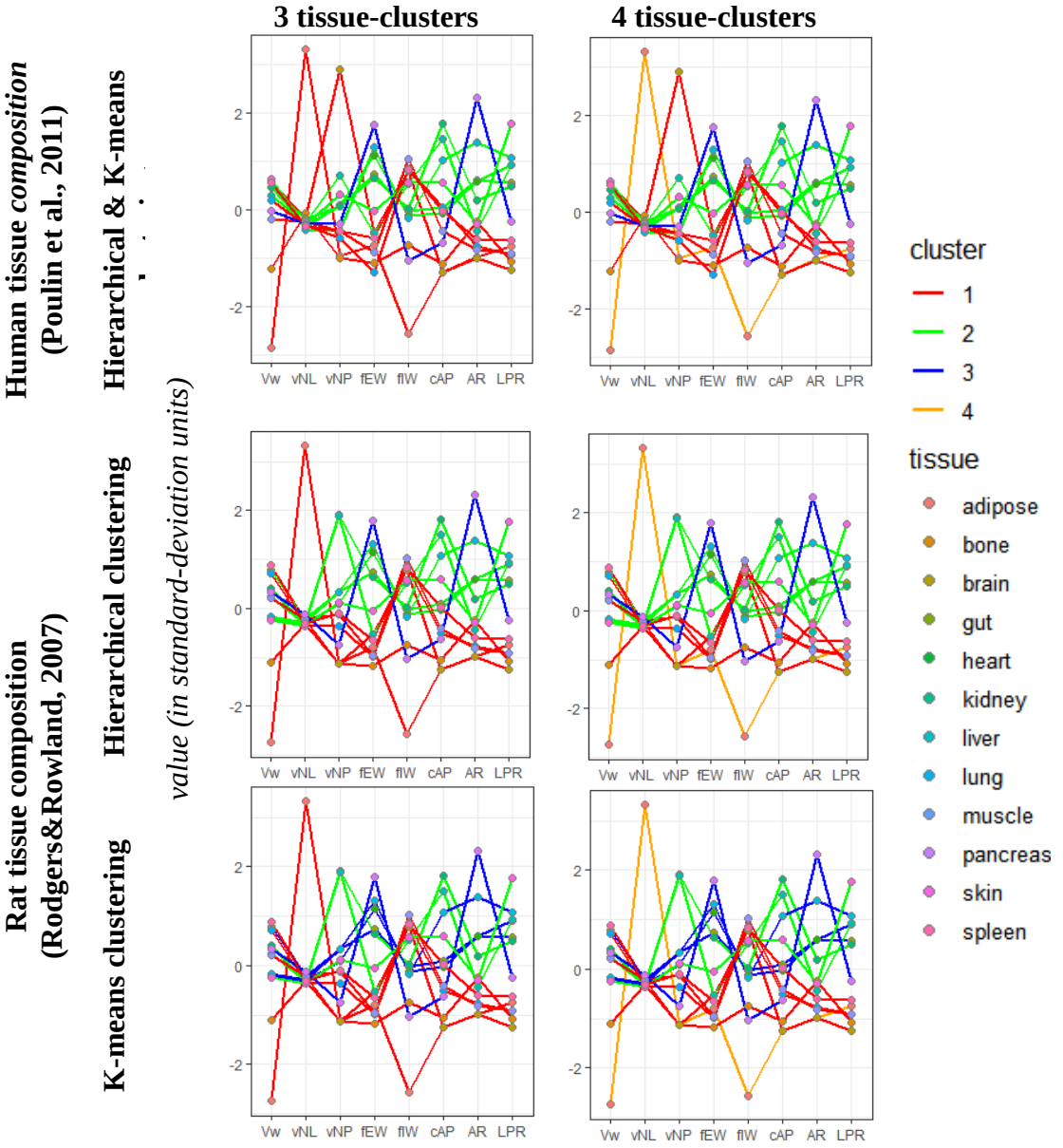
A2.4. Determination of plasma protein binding and blood to plasma partitioning ratio

Plasma protein binding was determined by equilibrium dialysis using the 96-well equilibrium dialysis device (HTD-96b, HTDialysis)[23, 24]. Blank human and rat plasma (K2-EDTA, mixed gender, pooled) spiked with diazepam was dialysed against phosphate buffer (133 mM pH=7.4) for 5 hours in triplicates. Prior to centrifugation, final dialysate and plasma samples were matched with blank matrices (90/10 buffer/plasma v/v) and proteins were precipitated by addition of acetonitrile (containing the internal standard, IS) in 3 volumes. Supernatant of the centrifuged samples were analysed by LC-MS/MS. The fup value was calculated as the ratio of the peak area ratio of diazepam:IS in the final buffer sample to the peak area ratio of diazepam:IS in the final plasma sample.

The blood to plasma partitioning ratio was determined by spiking fresh male human or rat blood with diazepam. After a 30-minute incubation, an aliquot of blood was collected and plasma was produced from the rest of the blood. Final plasma and blood samples were matched with blank matrices (50/50 blood/plasma), mixed with acetonitrile (containing the internal standard), centrifuged and the supernatants were analysed by LC-MS/MS following appropriate dilution. BP was calculated as the ratio of the peak area ratio in spiked blood to the peak area ratio in spiked plasma at 0.5h.

A2.5. Cluster analysis of tissues based on tissue composition

Figure A2.1: Characteristics of clustered tissues in terms of tissue composition with the different clustering methods in human and rat (tissue composition data are standardised)



Each vertical bar represents a tissue component and each line represent a tissue that is grouped in one of the 3 or 4 clusters. Distinct clusters are shown in different coloured solid lines.
Vw: fractional volume of water; vNL: fractional volume of neutral lipids; vNP: fractional volume of neutral phospholipids; fEW: fractional volume of extracellular water; fIW: fractional volume of intracellular water; cAP: acid phospholipid concentrations (mg/g); AR: albumin ratio; LPR: lipoprotein ratio; RBC: red blood cells

A2.6. Clustering analysis based on imputed dataset of rat steady state Kpus

Table A2.6: Percentage of missing data in total and per tissue in the Kpu dataset

Total	Pancreas	Bone	Thymus	Spleen	Gut	Skin	Adipose	Liver	Lung	Brain	Heart	Kidney	Muscle
48	85	81	80	74	60	56	47	27	24	23	23	23	17

A2.7. Fits of diazepam data in human

A2.7.1. For the empirical 2 compartment model

Figure A2.2: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the empirical 2 compartment model

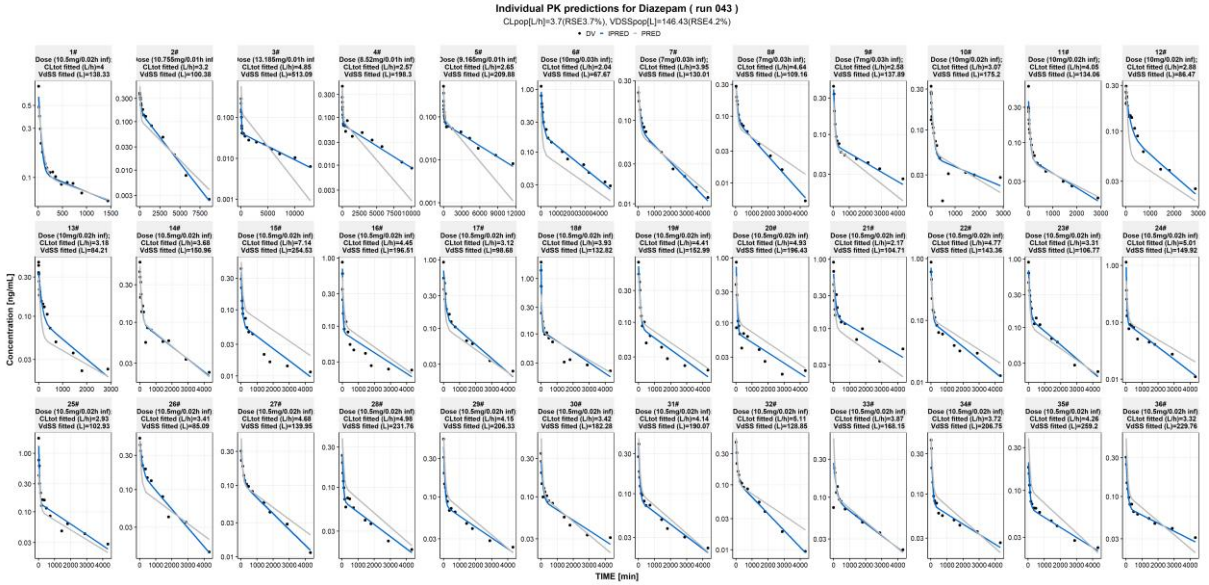
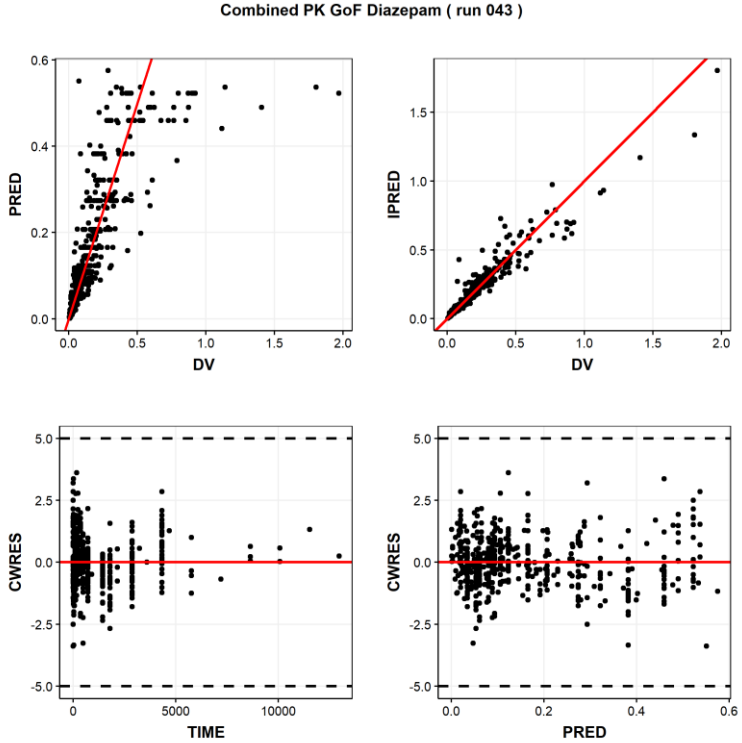


Figure A2.3: Basic goodness-of-fits plots for the empirical 2 compartment model



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.7.2. For the lumped 3 compartment model (man)

Figure A2.4: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the lumped 3-compartment model (man)

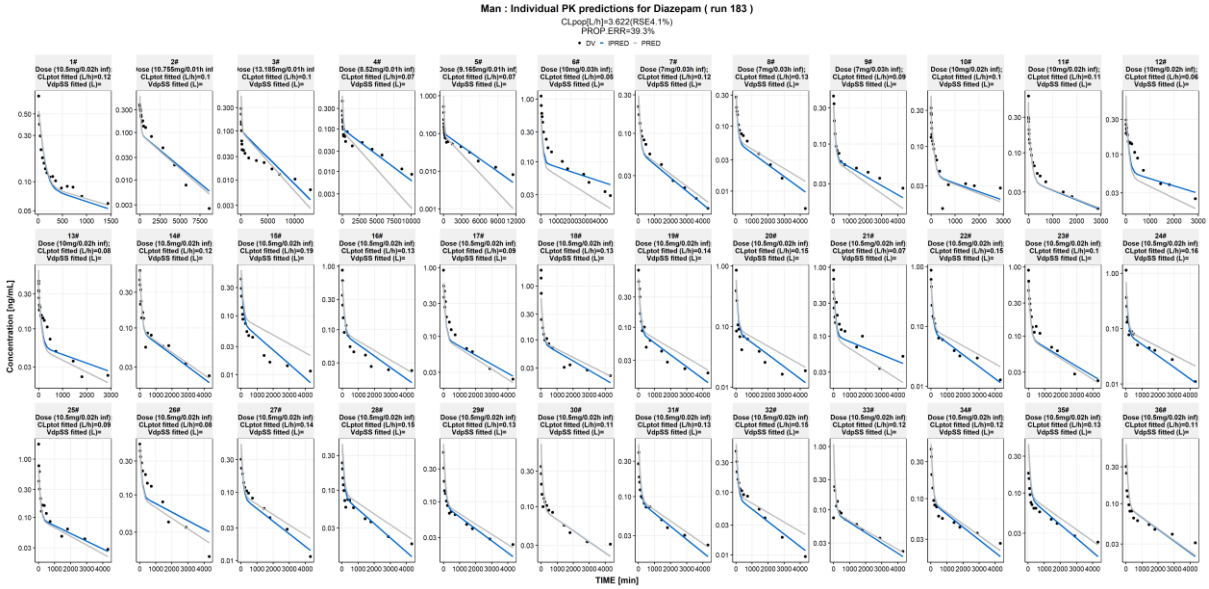
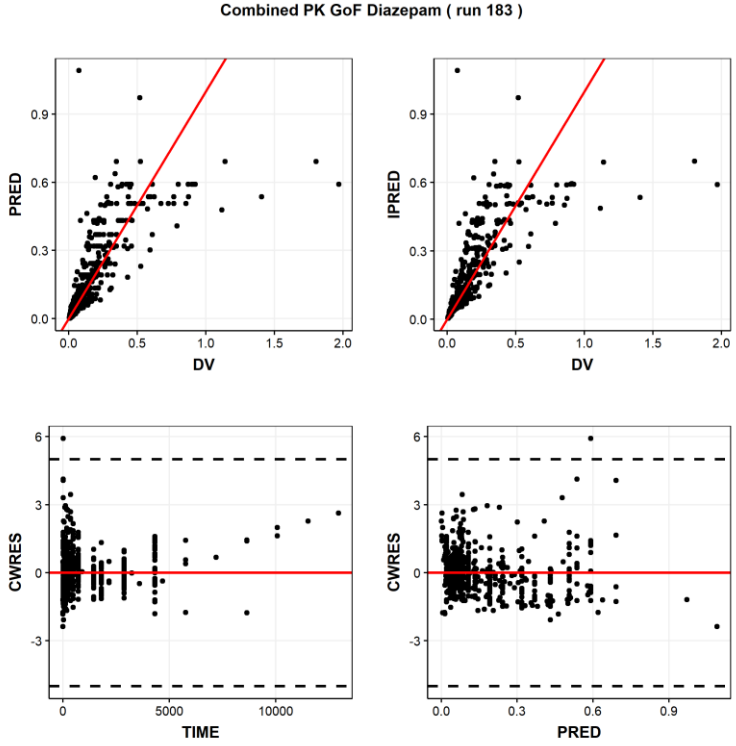


Figure A2.5: Basic goodness-of-fits plots for the lumped 3 compartment model (man)



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.7.3. For the PBPK model with 3 common Kpus model

Figure A2.6: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 3 common Kpus model

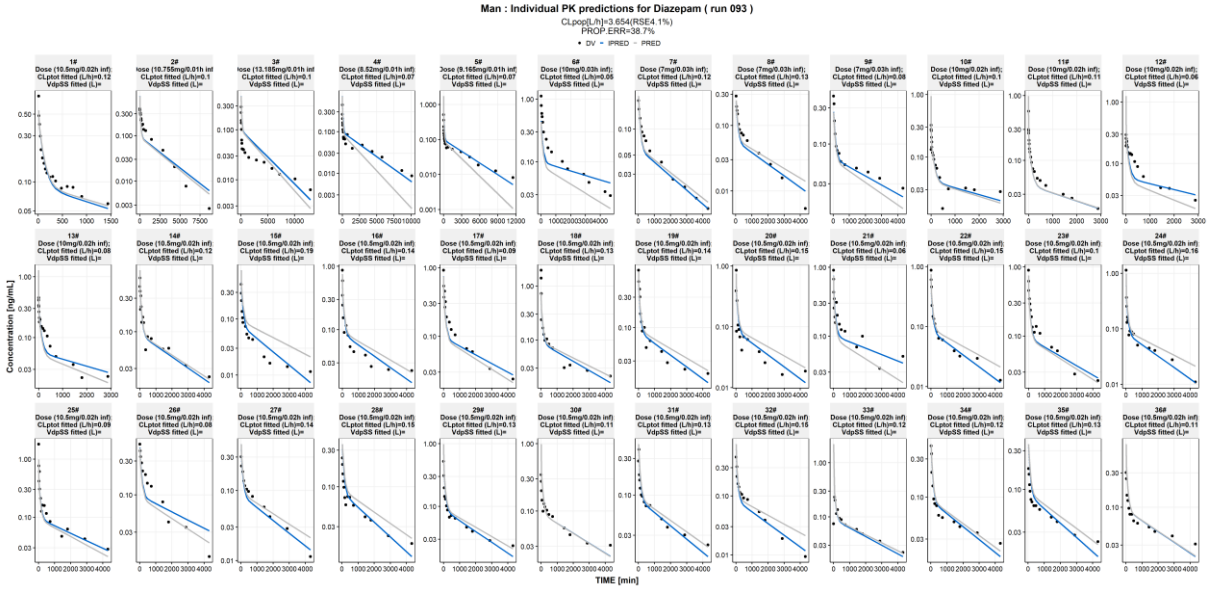
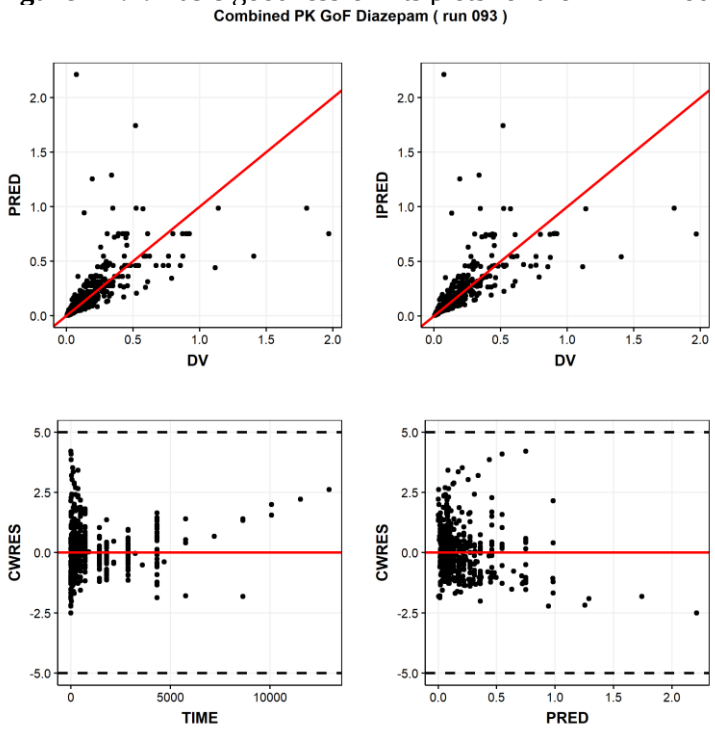


Figure A2.7: Basic goodness-of-fits plots for the PBPK model with 3 common Kpus model



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.7.4. For the PBPK model with 4 common Kpus model

Figure A2.8: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 4 common Kpus model

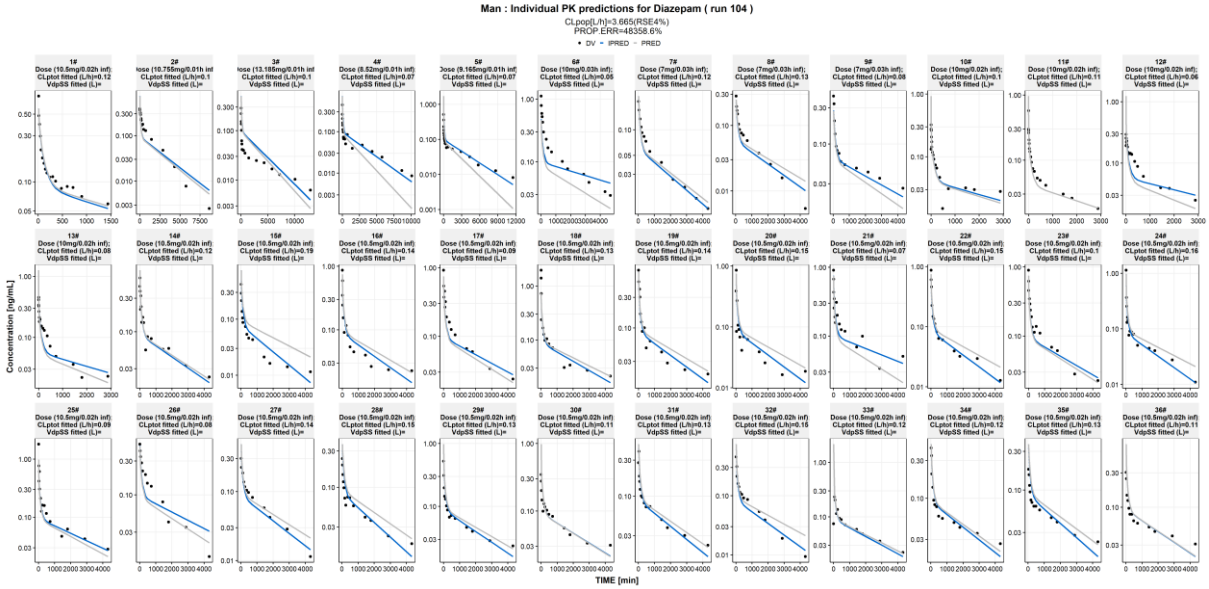
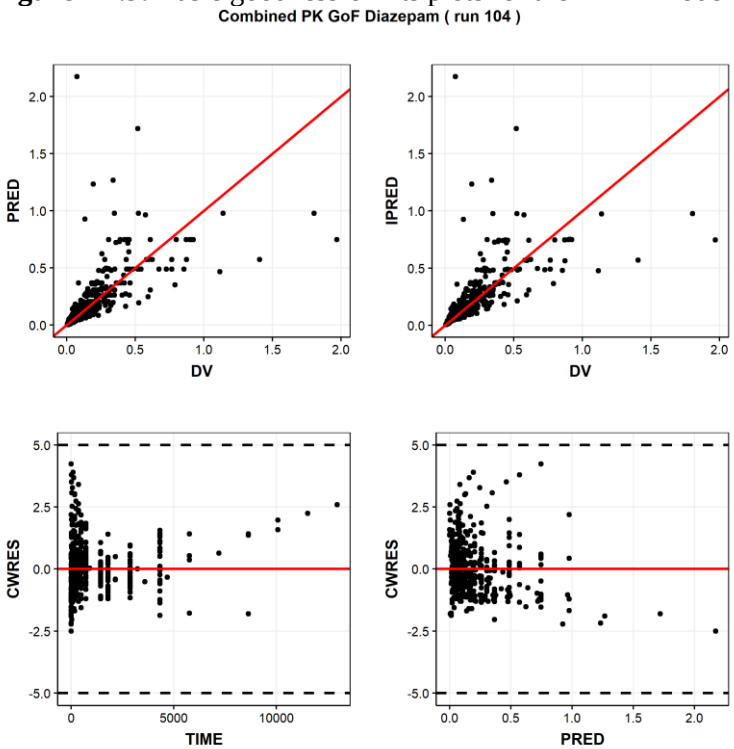


Figure A2.9: Basic goodness-of-fits plots for the PBPK model with 4 common Kpus model



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.7.5. For the PBPK model with 3 scalars model

Figure A2.10: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 3 scalars model

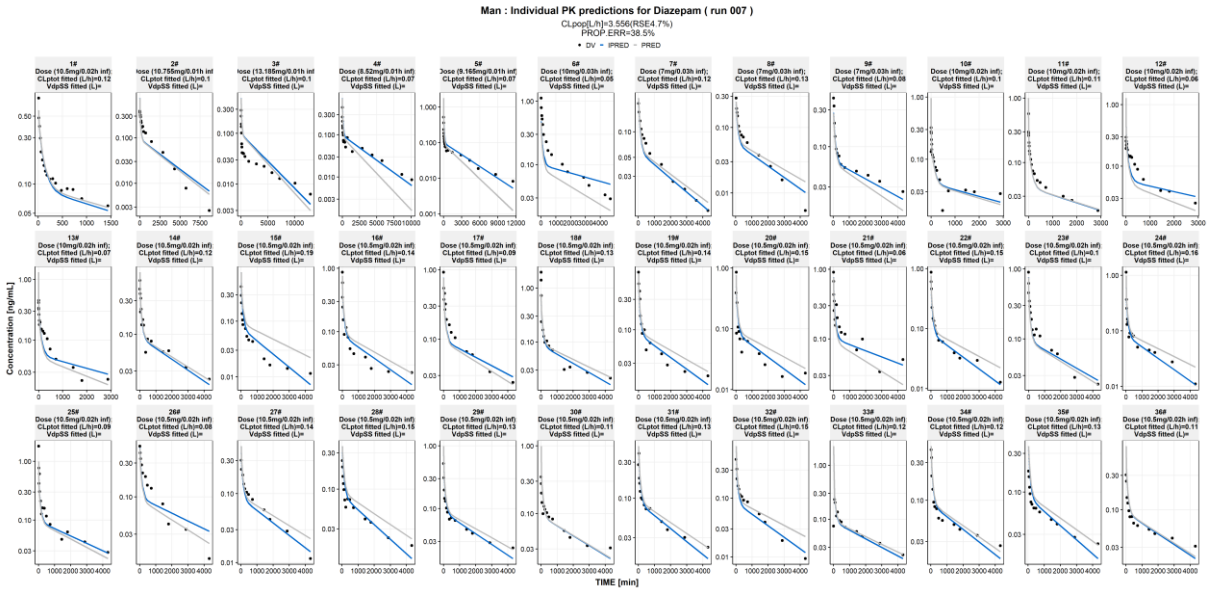
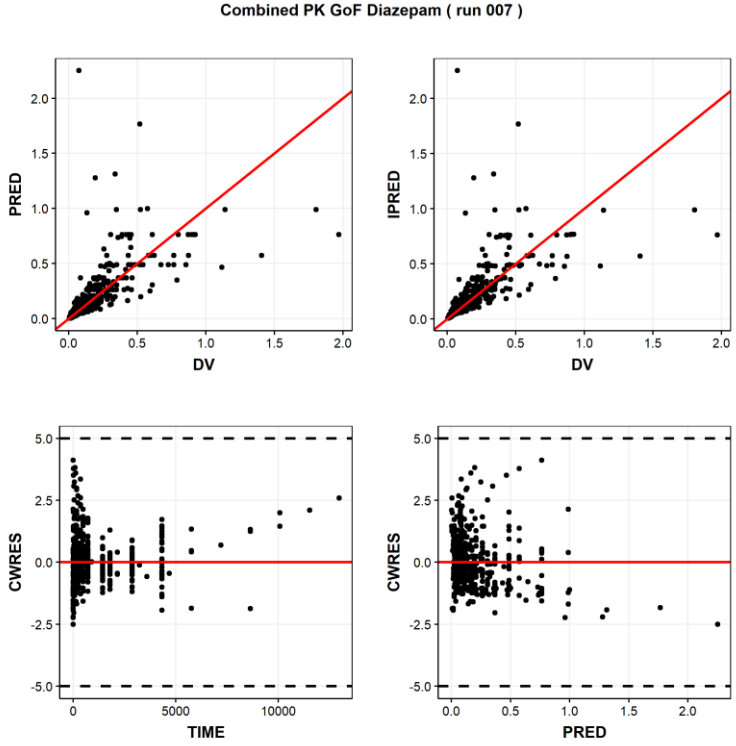


Figure A2.11: Basic goodness-of-fits plots for the PBPK model with 3 scalars model



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.7.6. For the PBPK model with 4 scalars model

Figure A2.12: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 4 scalars model

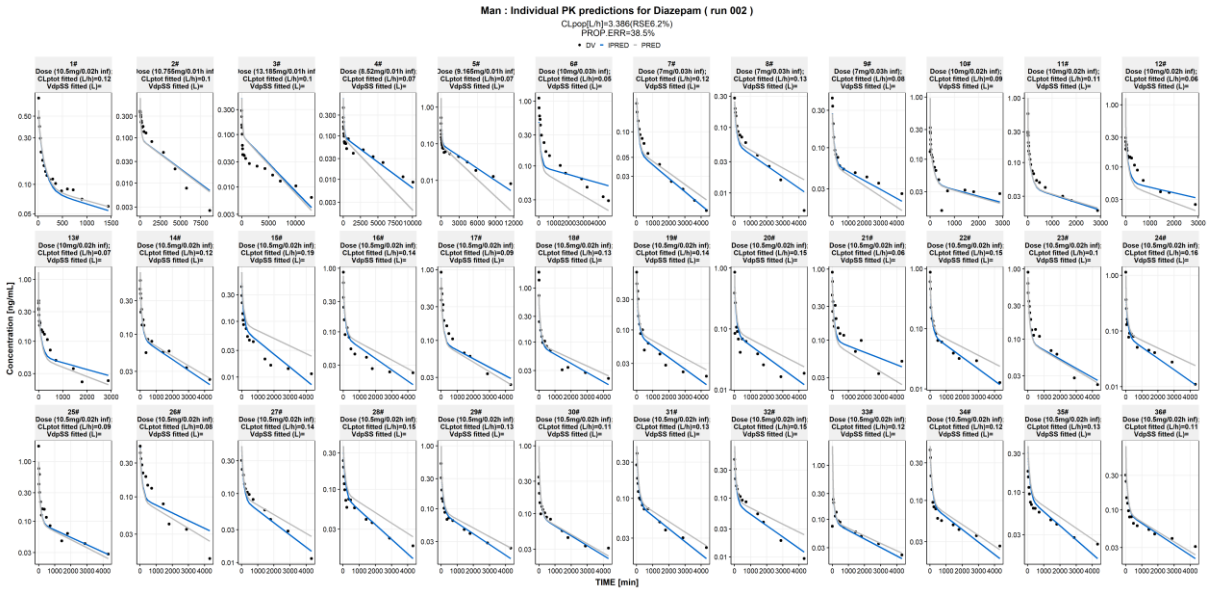
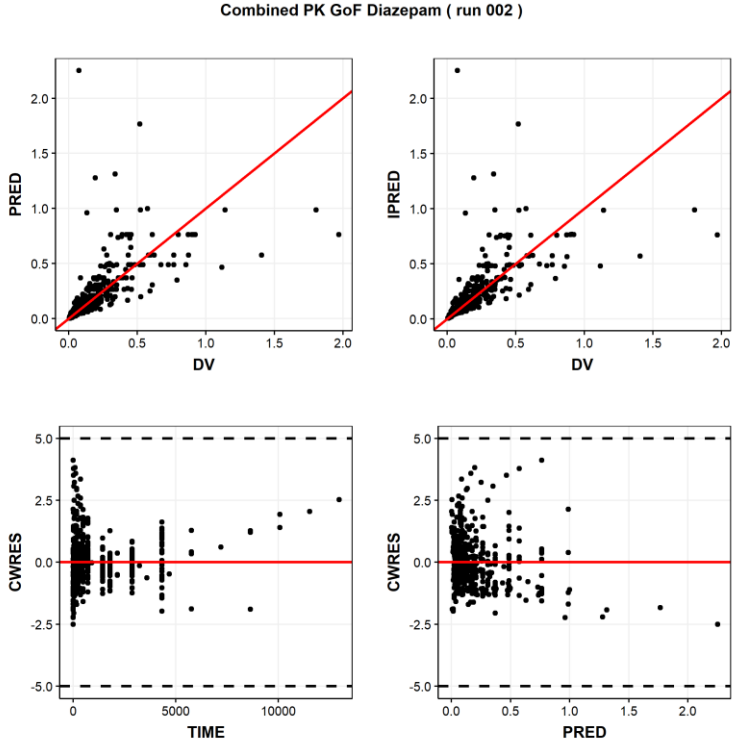


Figure A2.13: Basic goodness-of-fits plots for the PBPK model with 4 scalars model



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.8. Fits of diazepam data in rat

A2.8.1. For the empirical 3 compartment model

Figure A2.14: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the empirical 3 compartment model

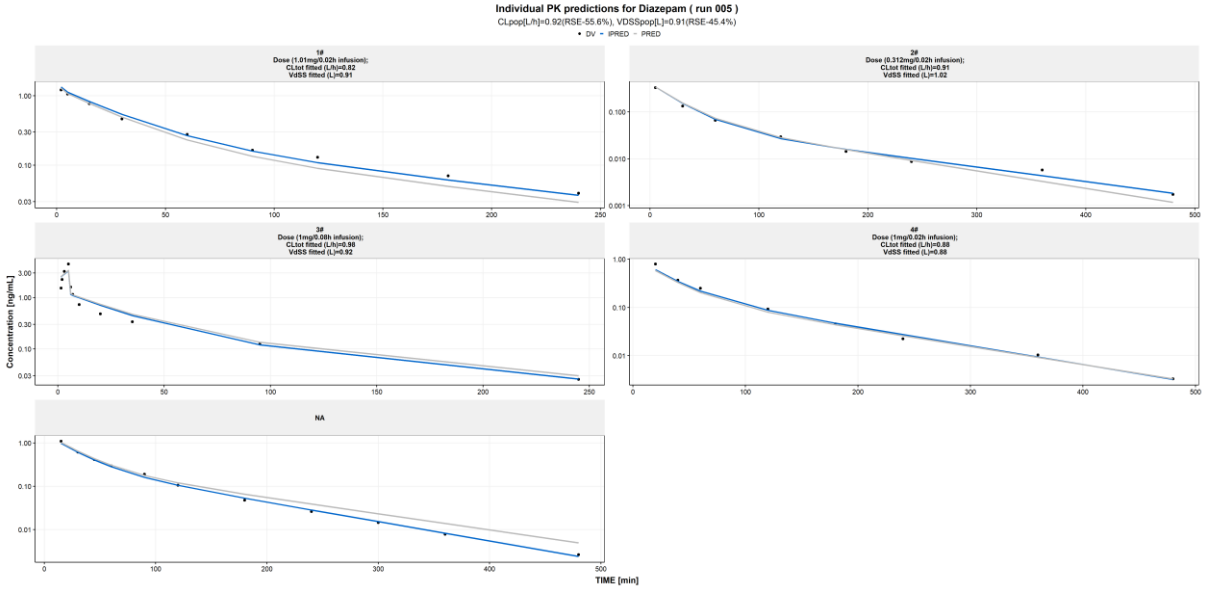
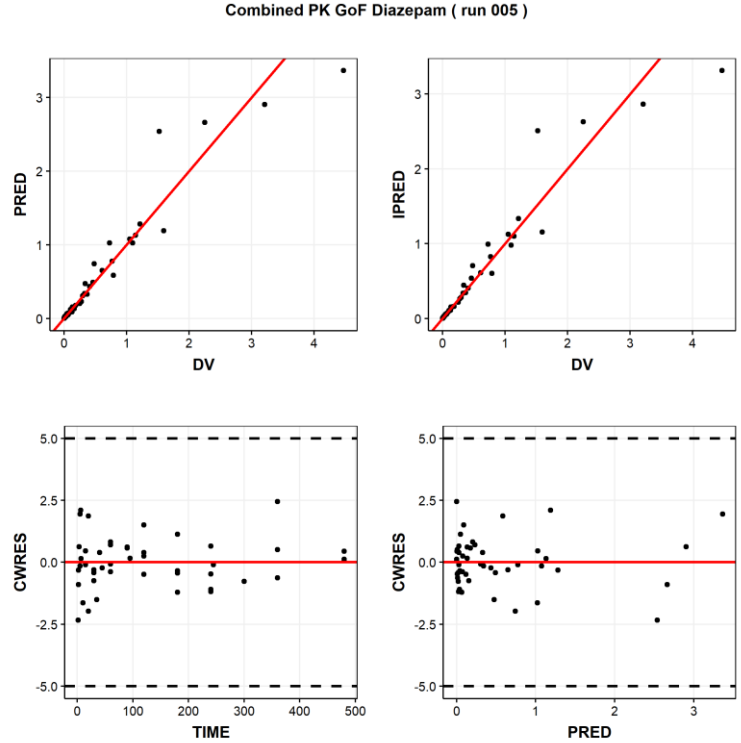


Figure A2.15: Basic goodness-of-fits plots for the empirical 3 compartment model



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.8.2. For the lumped 3 compartment model (rat)

Figure A2.16: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the empirical 3 compartment model

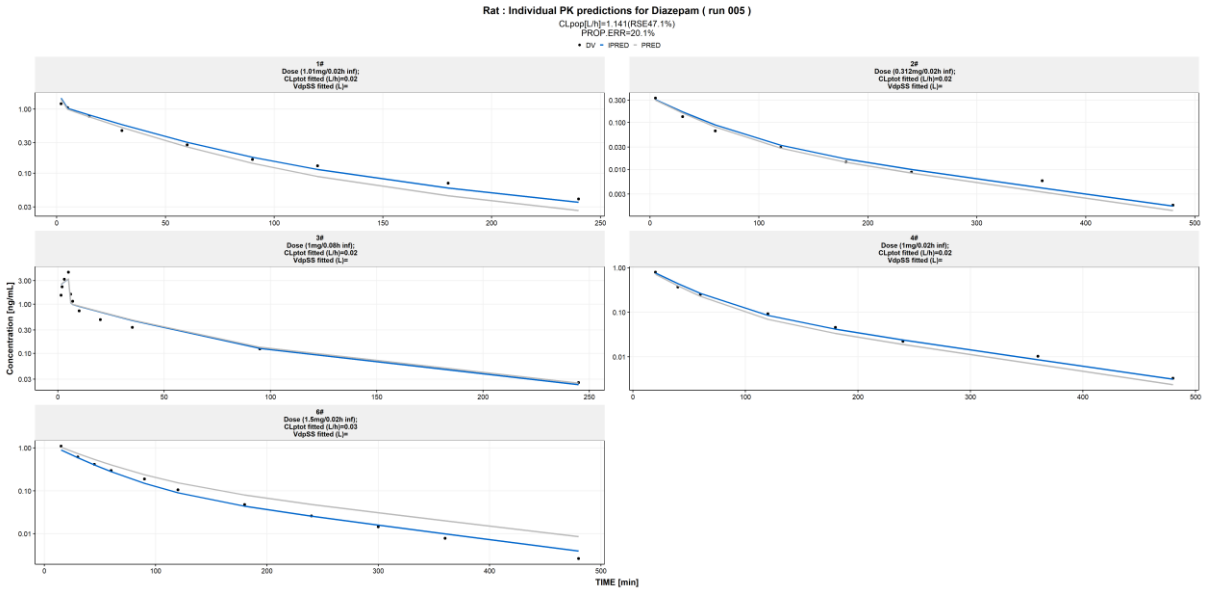
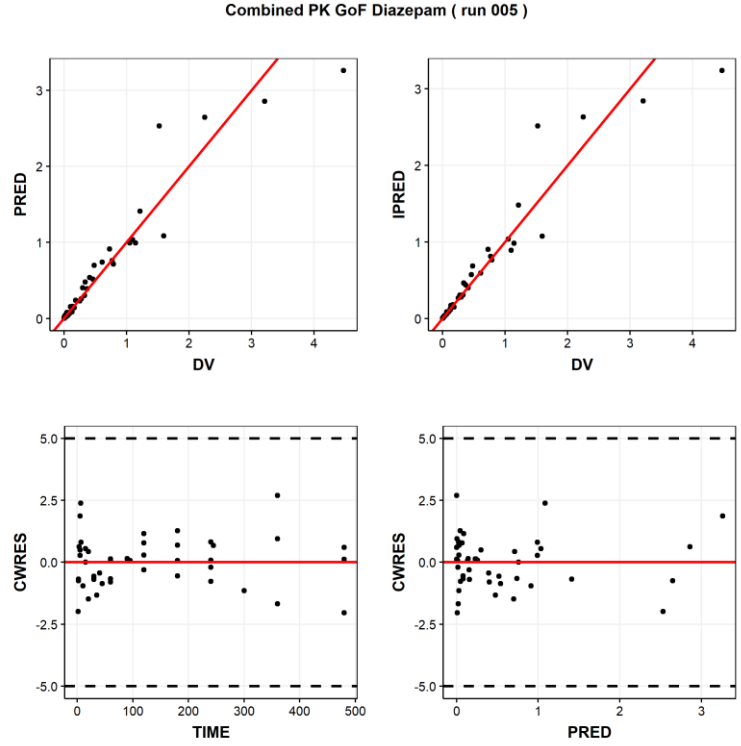


Figure A2.17: Basic goodness-of-fits plots for the empirical 3 compartment model



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.8.3. For the PBPK model with 3 common Kpus model (Hierarchical clustering)

Figure A2.18: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 3 common Kpus model (Hierarchical clustering)

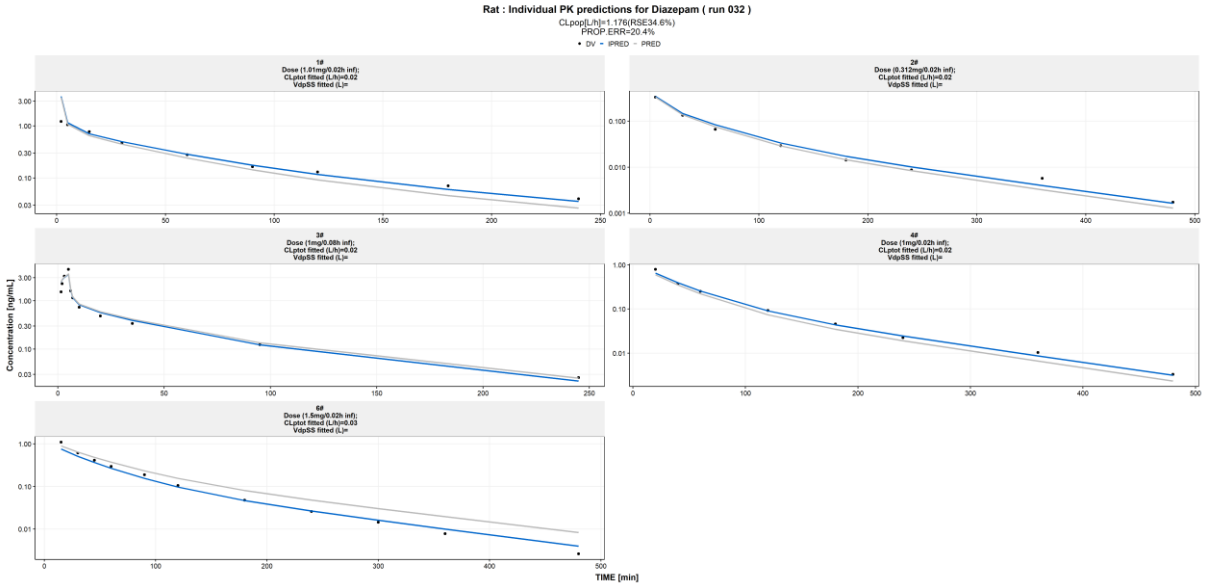
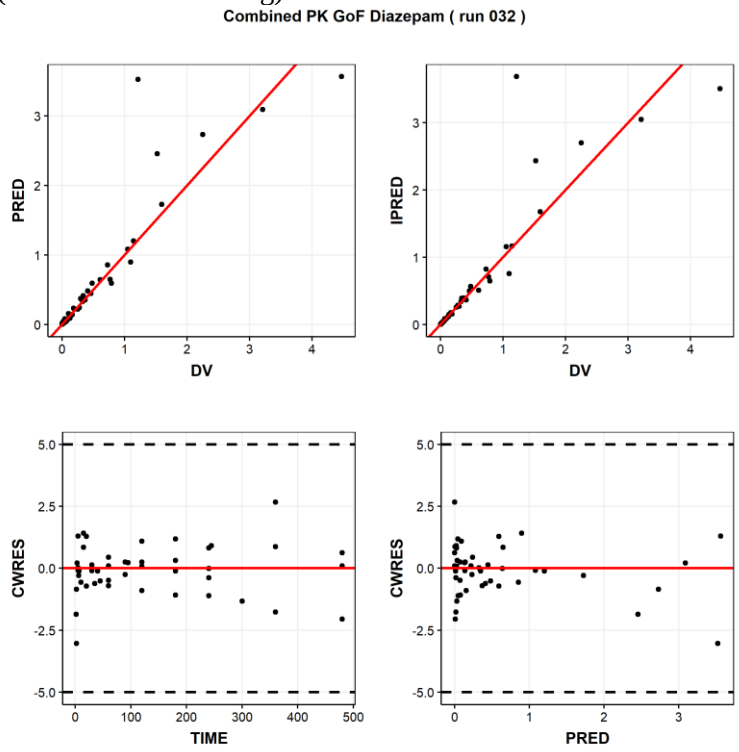


Figure A2.19: Basic goodness-of-fits plots for the PBPK model with 3 common Kpus model (Hierarchical clustering)



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.8.4. For the PBPK model with 4 common Kpus model (Hierarchical clustering)

Figure A2.20: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 4 common Kpus model (Hierarchical clustering)

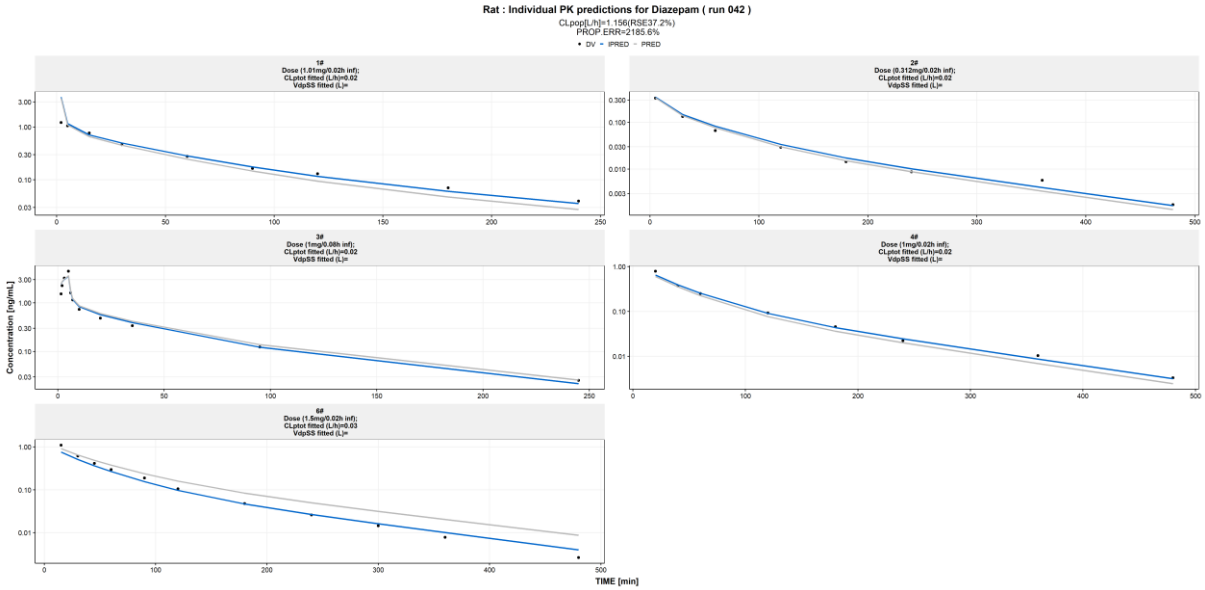
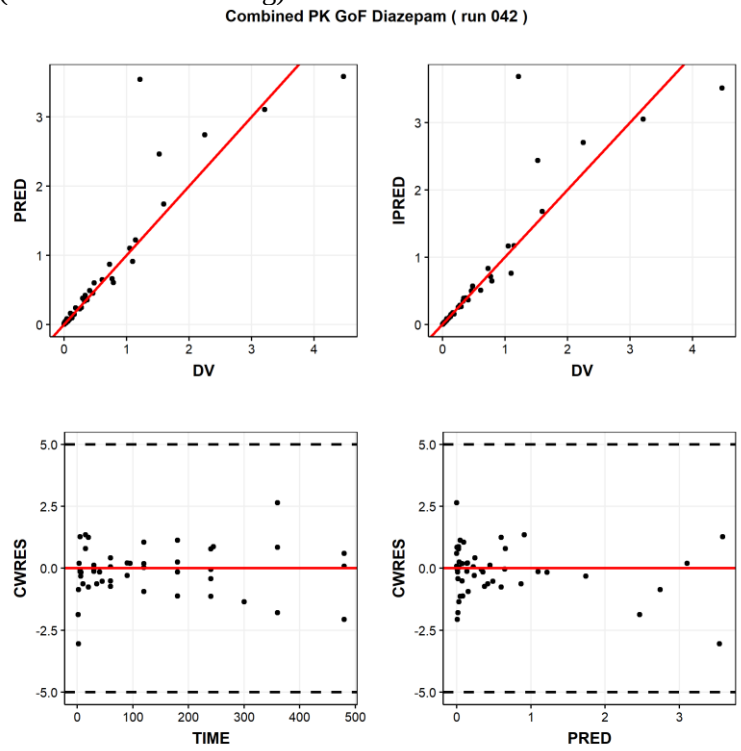


Figure A2.21: Basic goodness-of-fits plots for the PBPK model with 4 common Kpus model (Hierarchical clustering)



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.8.5. For the PBPK model with 3 common Kpus model (K-means clustering)

Figure A2.22: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 3 common Kpus model (K-means clustering)

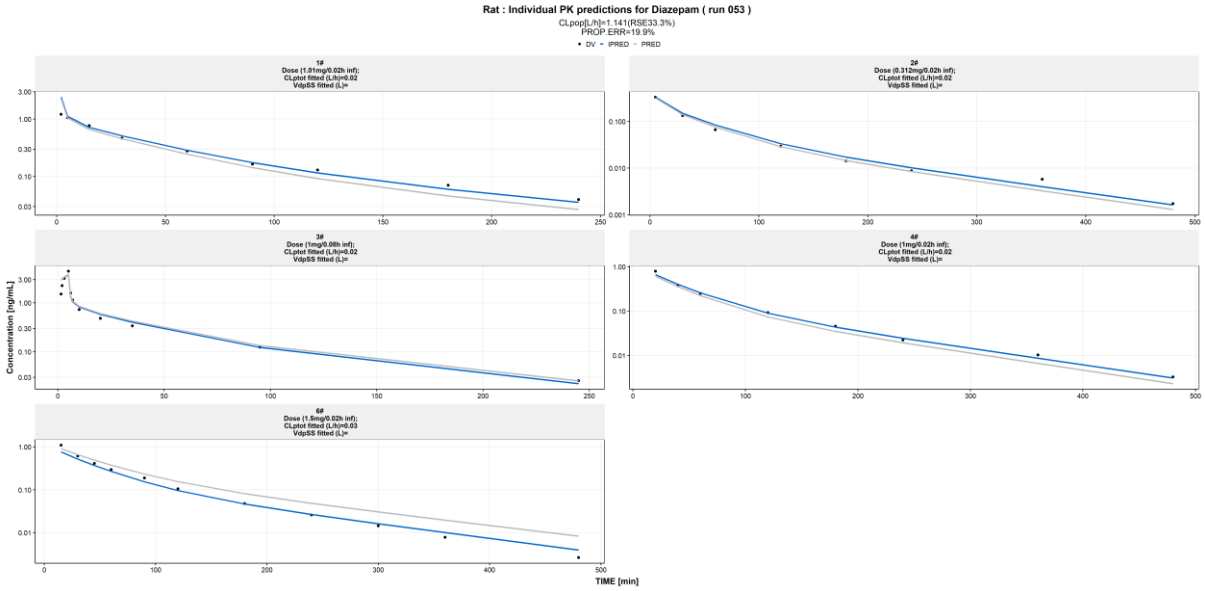
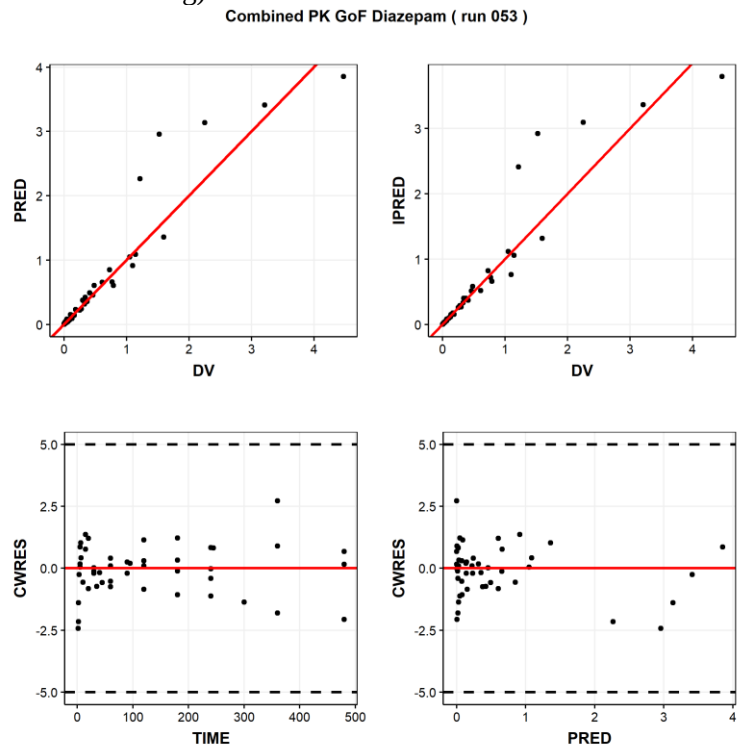


Figure A2.23: Basic goodness-of-fits plots for the PBPK model with 3 common Kpus model (K-means clustering)



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.8.6. For the PBPK model with 4 common Kpus model (K-means clustering)

Figure A2.24: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 4 common Kpus model (K-means clustering)

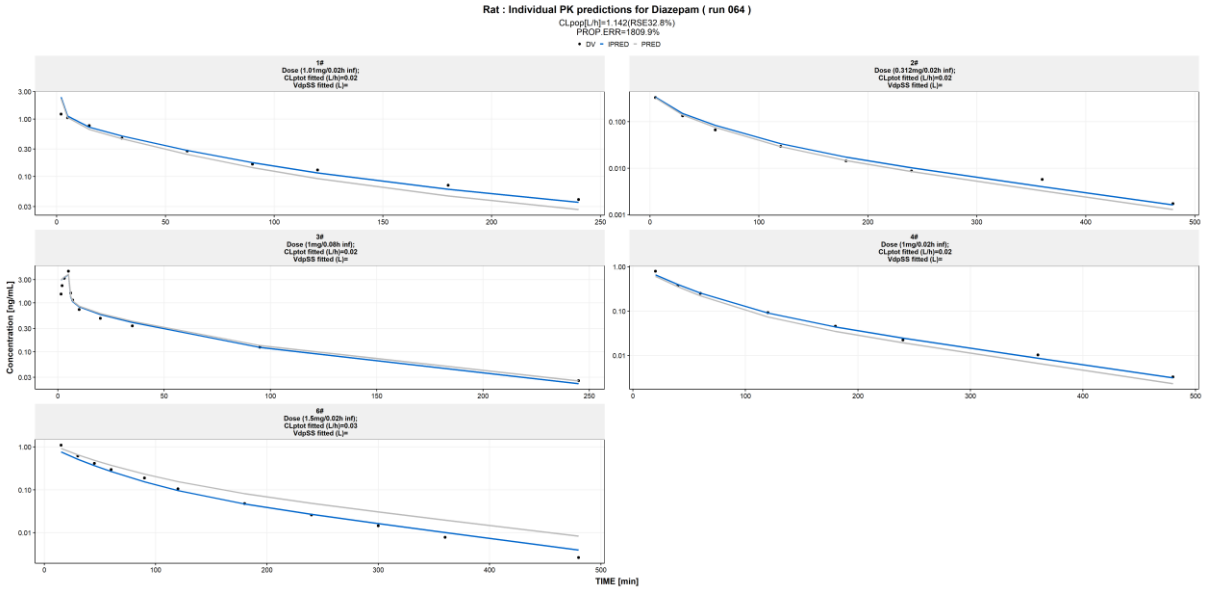
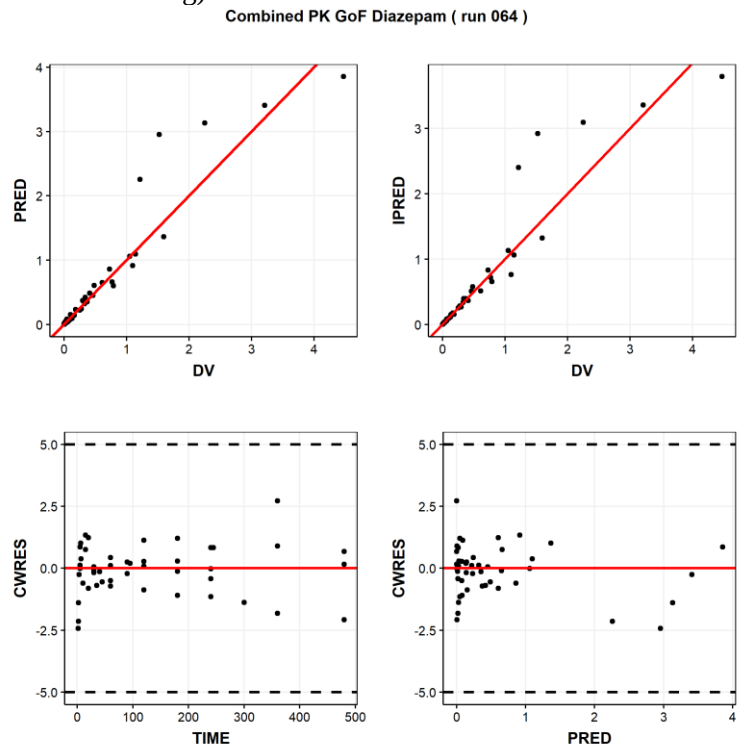


Figure A2.25: Basic goodness-of-fits plots for the PBPK model with 4 common Kpus model (K-means clustering)



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.8.7. For the PBPK model with 3 common Kpus model (clustering on steady state Kpus)

Figure A2.26: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 3 common Kpus model (clustering on steady state Kpus)

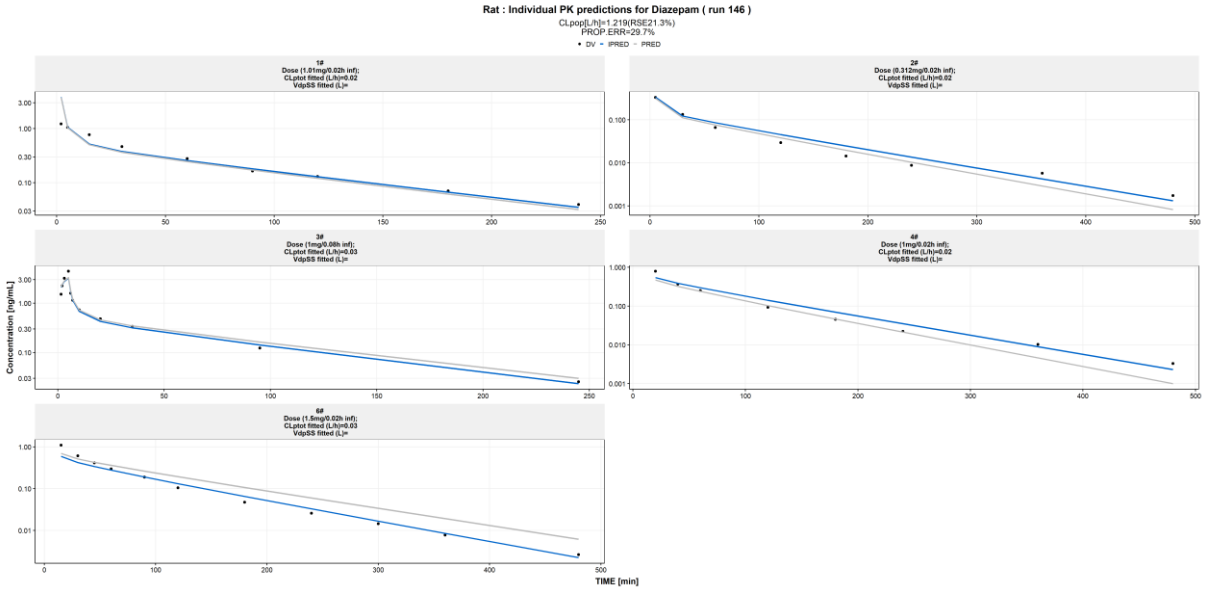
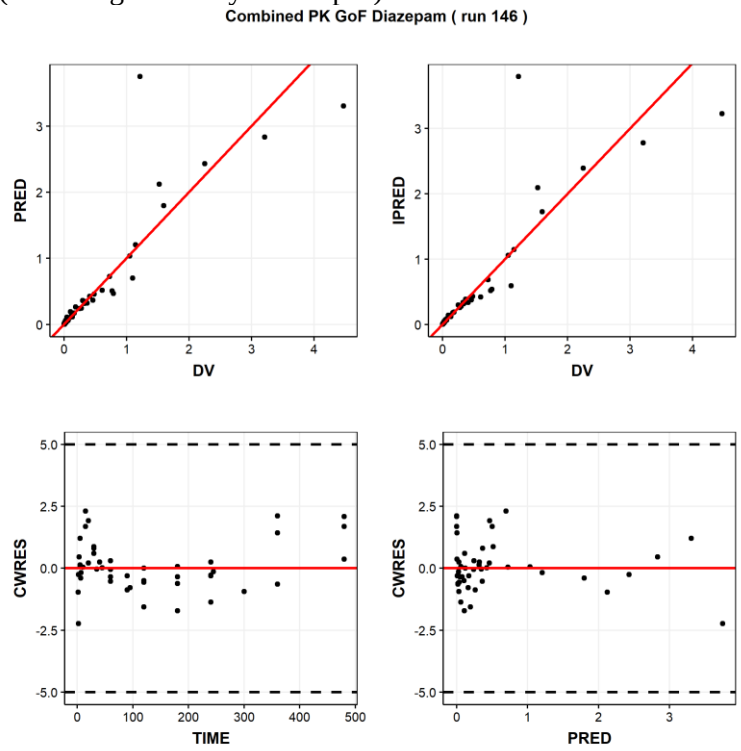


Figure A2.27: Basic goodness-of-fits plots for the PBPK model with 3 common Kpus model (clustering on steady state Kpus)



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.8.8. For the PBPK model with 4 common Kpus model (clustering on steady state Kpus)

Figure A2.28: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 4 common Kpus model (clustering on steady state Kpus)

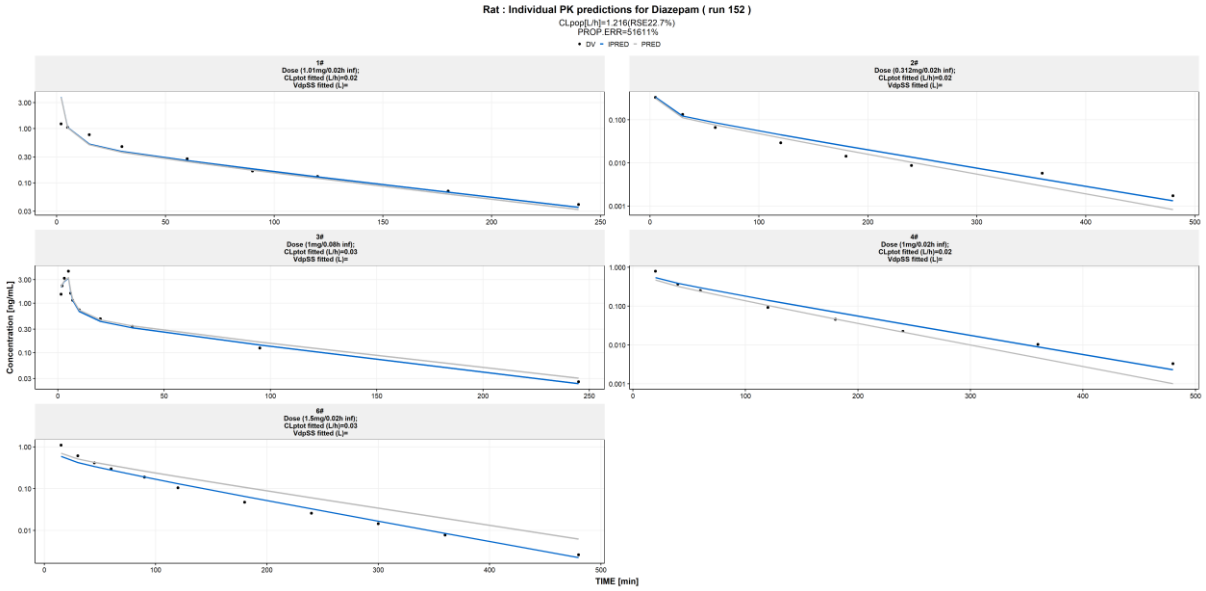
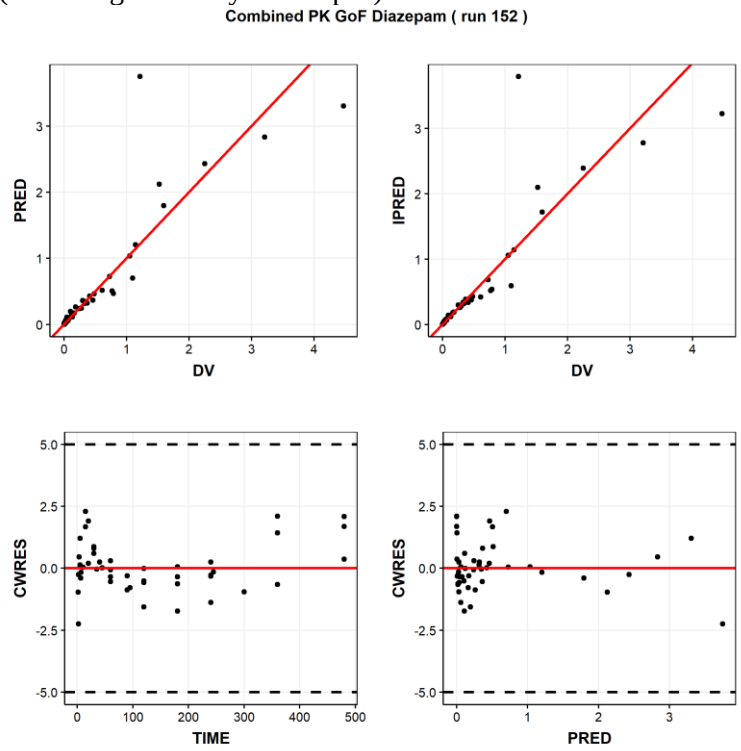


Figure A2.29: Basic goodness-of-fits plots for the PBPK model with 4 common Kpus model (clustering on steady state Kpus)



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.8.9. For the PBPK model with 3 common scalars model (Hierarchical clustering)

Figure A2.30: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 3 scalars model (Hierarchical clustering)

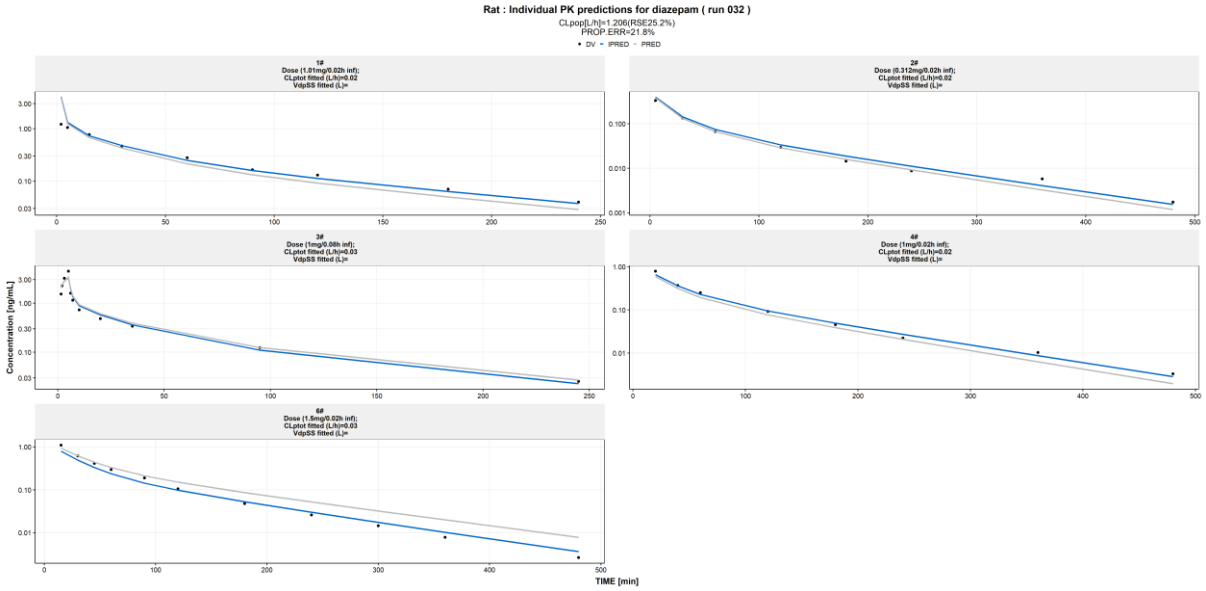
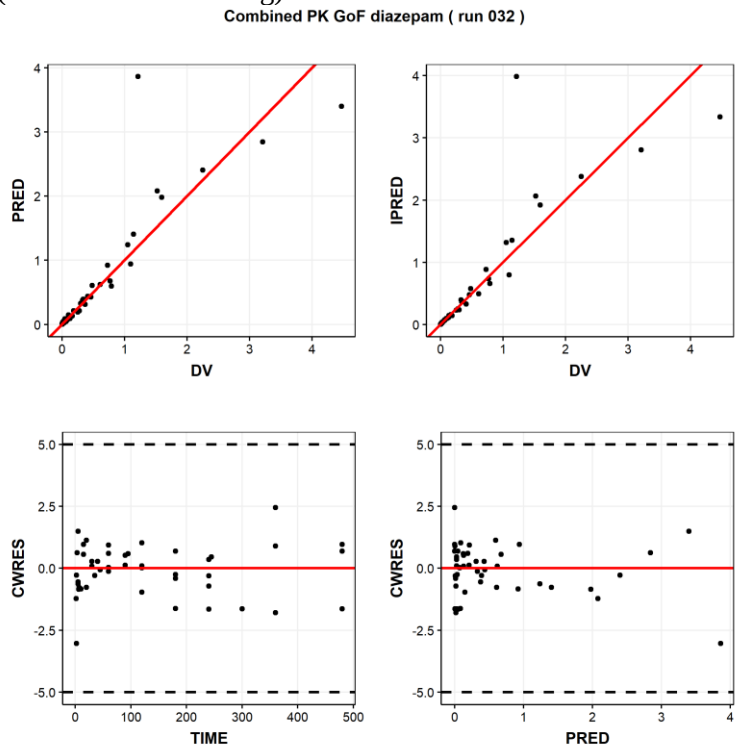


Figure A2.31: Basic goodness-of-fits plots for the PBPK model with 3 scalars model (Hierarchical clustering)



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.8.10. For the PBPK model with 4 common scalars model (Hierarchical clustering)

Figure A2.32: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 4 scalars model (Hierarchical clustering)

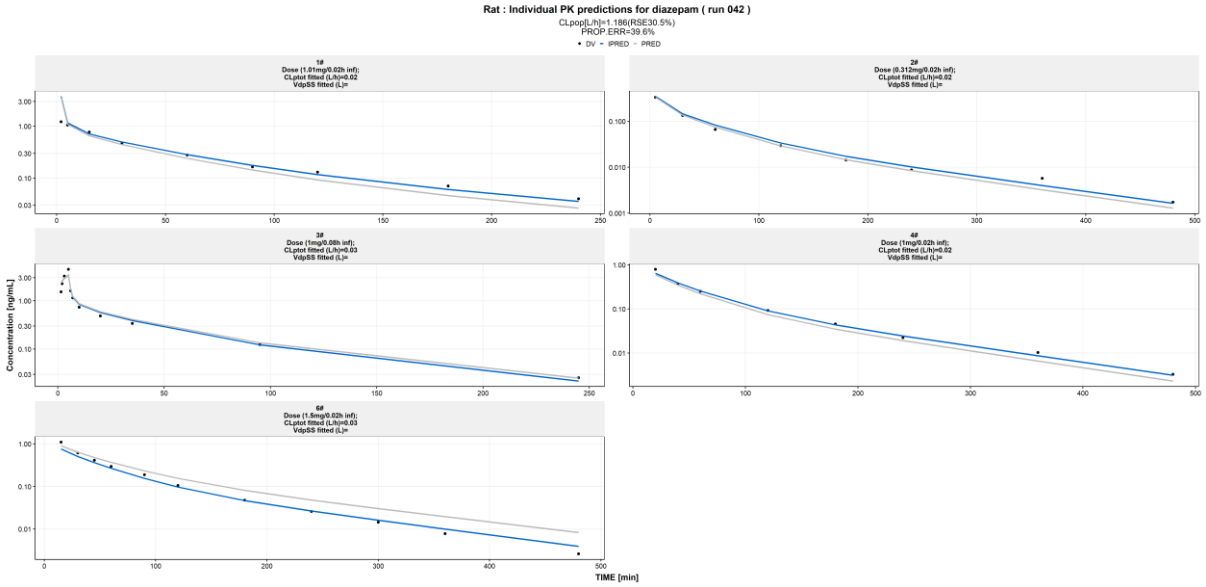
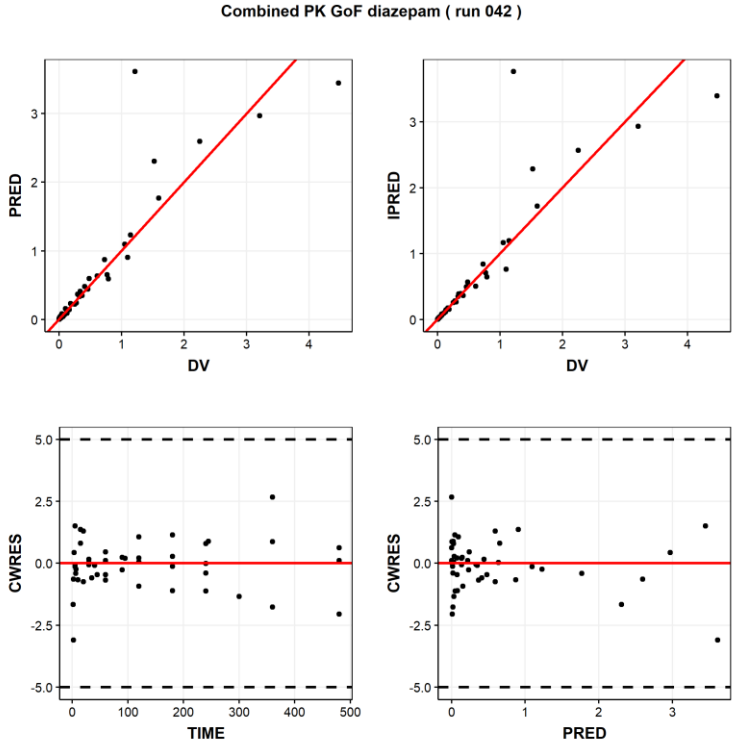


Figure A2.33: Basic goodness-of-fits plots for the PBPK model with 4 scalars model (Hierarchical clustering)



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.8.11. For the PBPK model with 3 common scalars model (K-means clustering)

Figure A2.34: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 3 scalars model (K-means clustering)

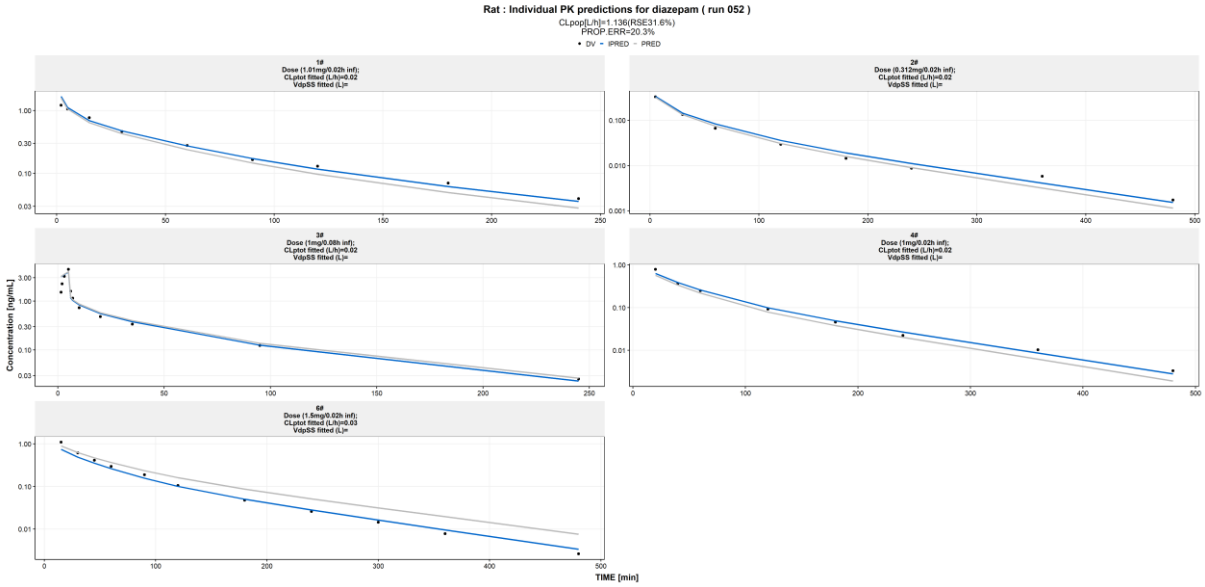
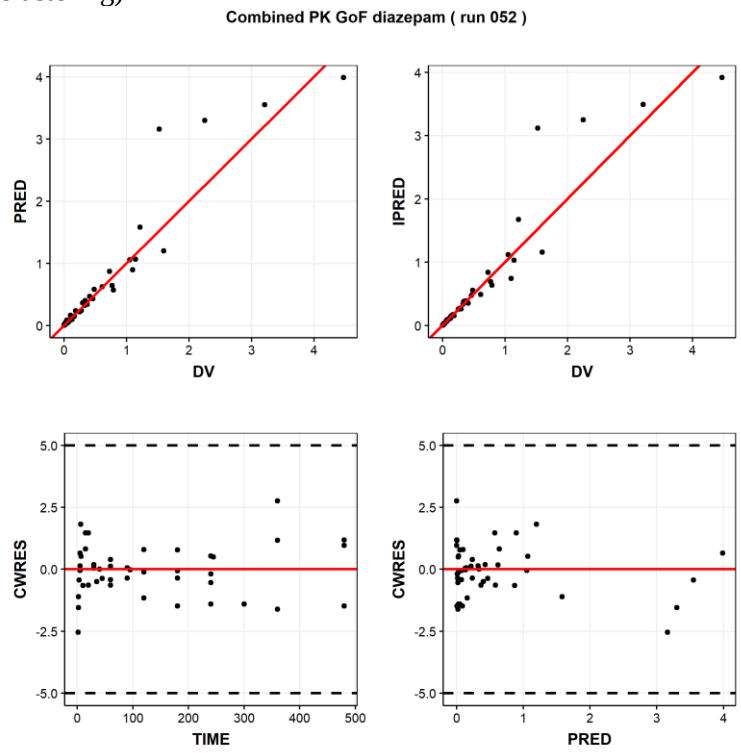


Figure A2.35: Basic goodness-of-fits plots for the PBPK model with 3 scalars model (K-means clustering)



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.8.12. For the PBPK model with 4 common scalars model (K-means clustering)

Figure A2.36: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 4 scalars model (K-means clustering)

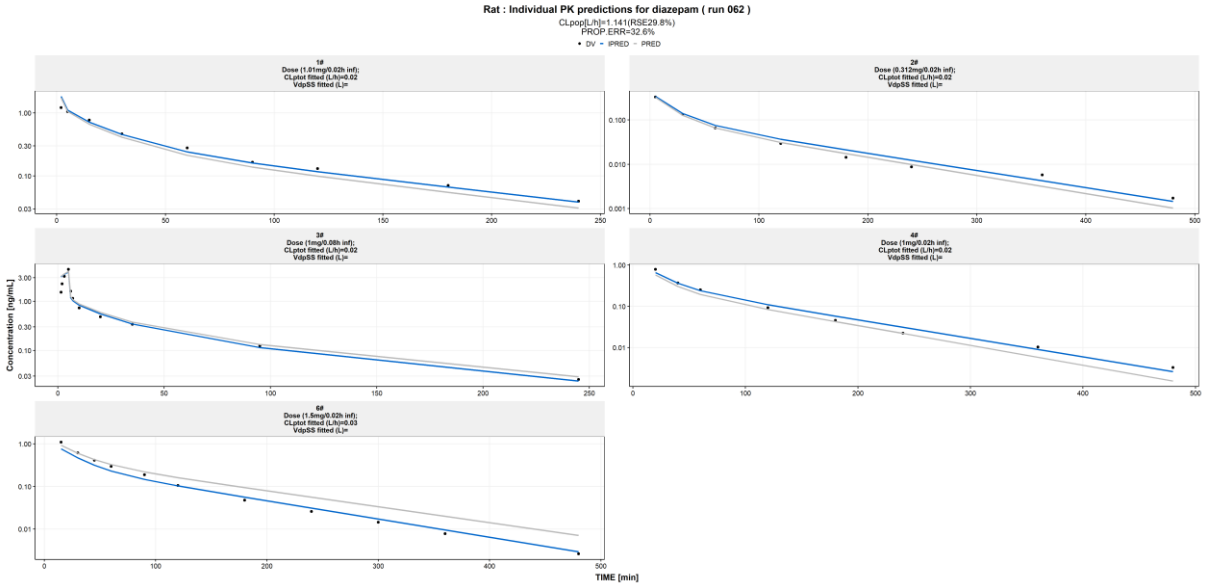
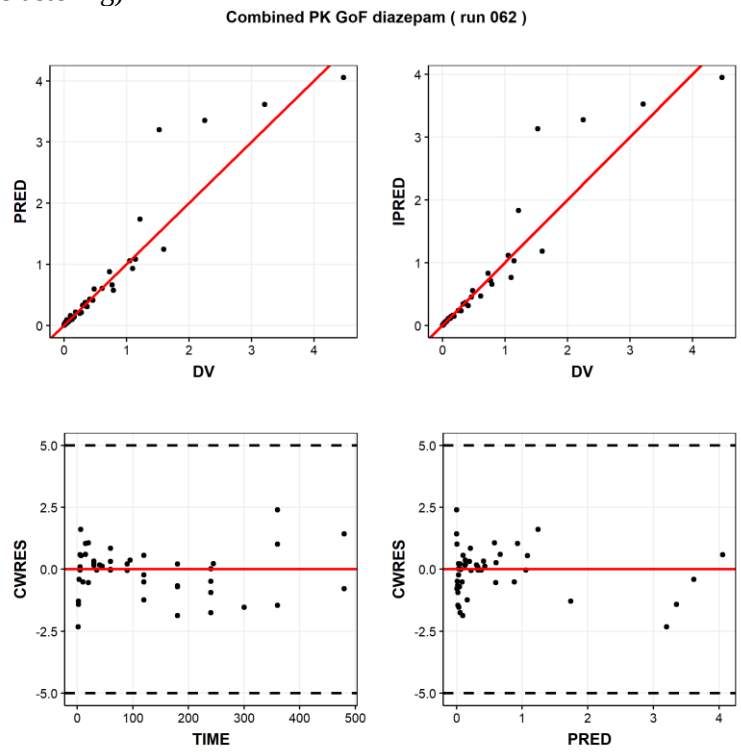


Figure A2.37: Basic goodness-of-fits plots for the PBPK model with 4 scalars model (K-means clustering)



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.8.13. For the PBPK model with 3 common scalars model (clustering on steady state Kpus)

Figure A2.38: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 3 scalars model (clustering on steady state Kpus)

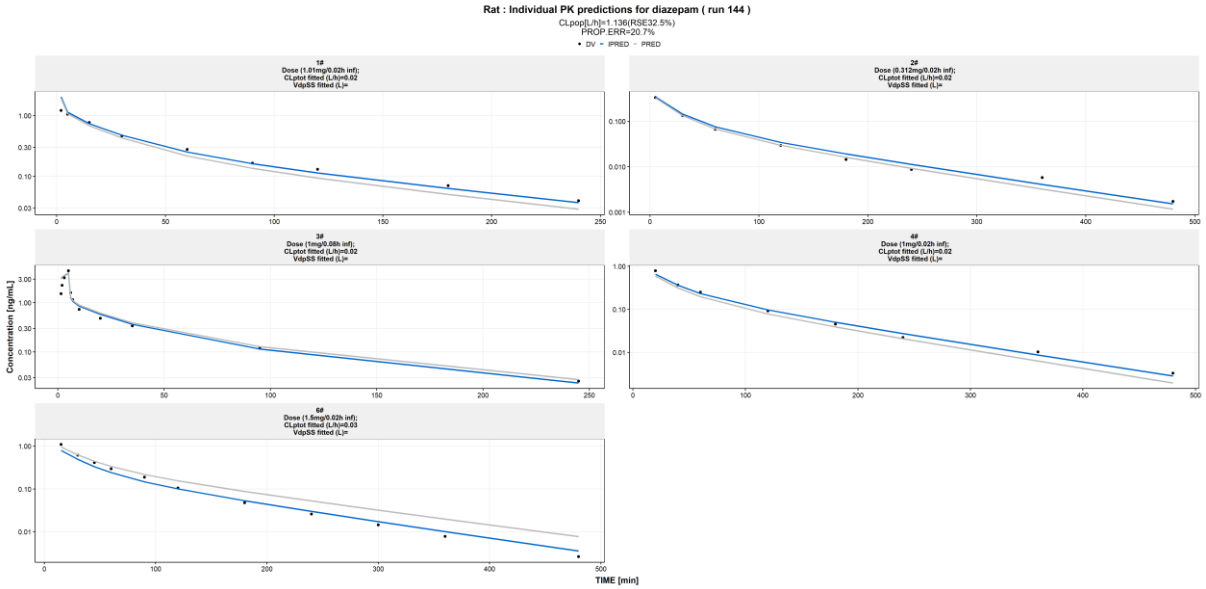
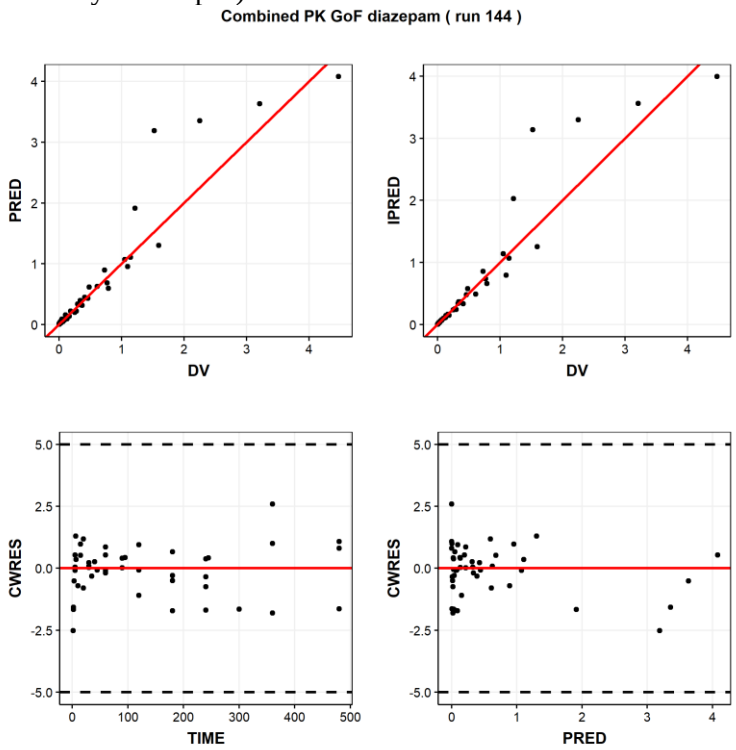


Figure A2.39: Basic goodness-of-fits plots for the PBPK model with 3 scalars model (clustering on steady state Kpus)



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.8.14. For the PBPK model with 4 common scalars model (clustering on steady state Kpus)

Figure A2.40: Plots of the observations (*black circles*), population predictions (*solid grey lines*) and individual predictions (*solid blue lines*) *versus* time for the PBPK model with 4 scalars model (clustering on steady state Kpus)

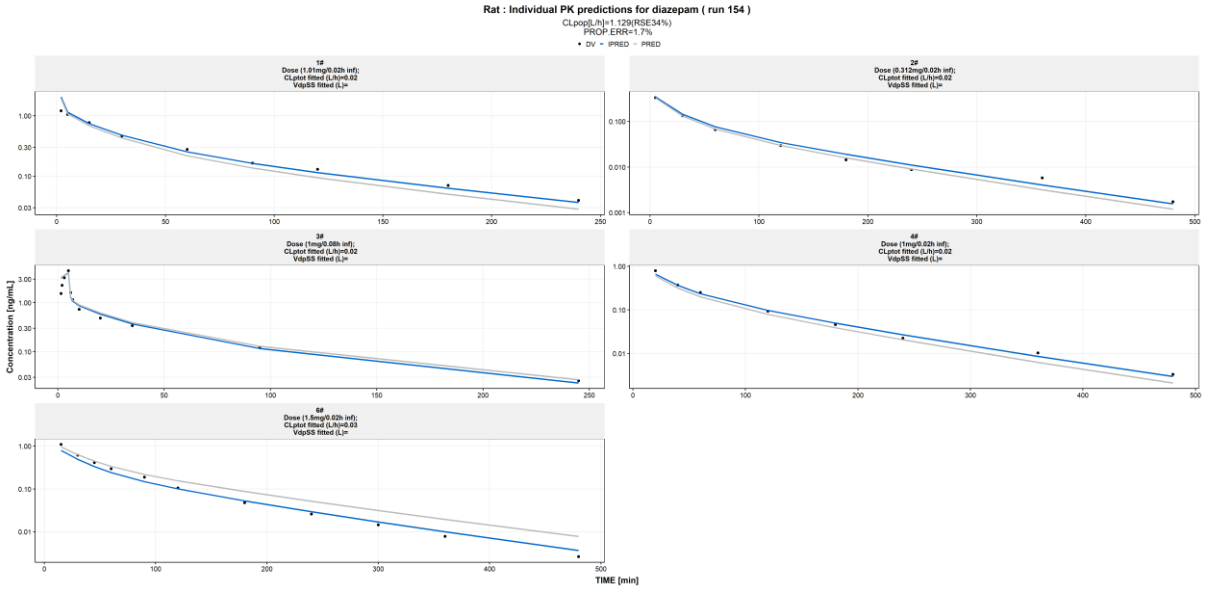
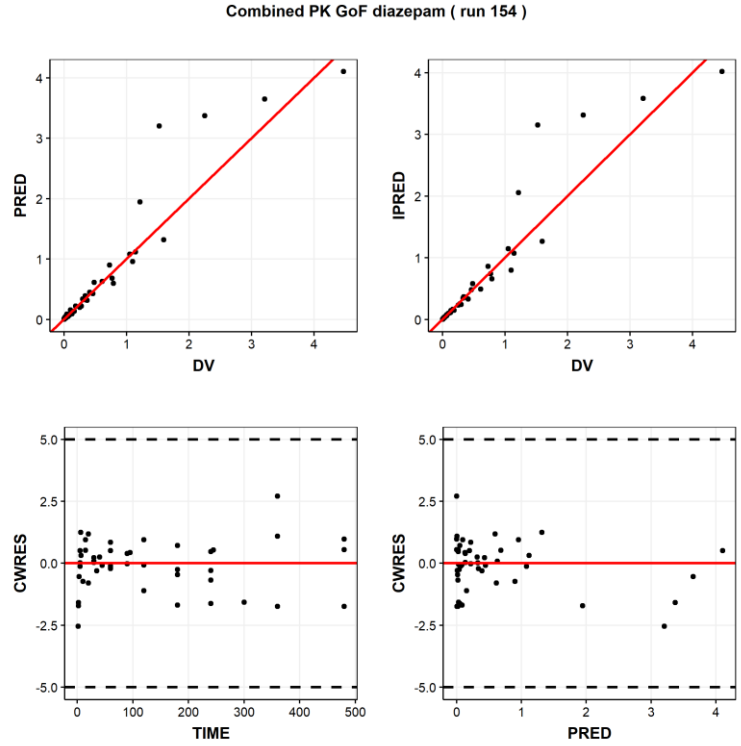


Figure A2.41: Basic goodness-of-fits plots for the PBPK model with 3 scalars model (clustering on steady state Kpus)



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A2.9. References

[1] Valentin J. Basic anatomical and physiological data for use in radiological protection: reference values. ICRP Publication 89. Ann ICRP. 2002;32(3-4):1-277.

[2] Williams LR, Leggett RW. Reference values for resting blood flow to organs of man. Clin Phys Physiol Meas. 1989;10(3):187-217.

[3] Nestorov I. Modelling and simulation of variability and uncertainty in toxicokinetics and pharmacokinetics. Toxicol Lett. 2001;120(1-3):411-20.

[4] Brown RP, Delp MD, Lindstedt SL, Rhomberg LR, Beliles RP. Physiological parameter values for physiologically based pharmacokinetic models. Toxicol Ind Health. 1997;13(4):407-84.

[5] Rodgers T, Rowland M. Mechanistic approaches to volume of distribution predictions: understanding the processes. Pharm Res. 2007;24(5):918-33.

[6] Greenblatt DJ, Ehrenberg BL, Gunderman J, Locniskar A, Scavone JM, Harmatz JS, et al. Pharmacokinetic and electroencephalographic study of intravenous diazepam, midazolam, and placebo. Clin Pharmacol Ther. 1989;45(4):356-65.

[7] Greenblatt DJ, Allen MD, Harmatz JS, Shader RI. Diazepam disposition determinants. Clin Pharmacol Ther. 1980;27(3):301-12.

[8] Dhillon S, Richens A. Pharmacokinetics of diazepam in epileptic patients and normal volunteers following intravenous administration. Br J Clin Pharmacol. 1981;12(6):841-4.

[9] Klotz U, Antonin KH, Bieck PR. Pharmacokinetics and plasma binding of diazepam in man, dog, rabbit, guinea pig and rat. J Pharmacol Exp Ther. 1976;199(1):67-73.

[10] Klotz U, Avant GR, Hoyumpa A, Schenker S, Wilkinson GR. The effects of age and liver disease on the disposition and elimination of diazepam in adult man. J Clin Invest. 1975;55(2):347-59.

[11] Kaplan SA, Jack ML, Alexander K, Weinfeld RE. Pharmacokinetic profile of diazepam in man following single intravenous and oral and chronic oral administrations. J Pharm Sci. 1973;62(11):1789-96.

[12] Tsiros P, Bois FY, Dokoumetzidis A, Tsiliki G, Sarimveis H. Population pharmacokinetic reanalysis of a Diazepam PBPK model: a comparison of Stan and GNU MCSim. J Pharmacokinet Pharmacodyn. 2019.

[13] Langdon G, Gueorguieva I, Aarons L, Karlsson M. Linking preclinical and clinical whole-body physiologically based pharmacokinetic models with prior distributions in NONMEM. Eur J Clin Pharmacol. 2007;63(5):485-98.

[14] Diaz-Garcia JM, Oliver-Botana J, Fos-Galve D. Pharmacokinetics of diazepam in the rat: influence of a carbon tetrachloride-induced hepatic injury. J Pharm Sci. 1992;81(8):768-72.

[15] Igari Y, Sugiyama Y, Sawada Y, Iga T, Hanano M. Prediction of diazepam disposition in the rat and man by a physiologically based pharmacokinetic model. J Pharmacokinet Biopharm. 1983;11(6):577-93.

[16] Igari Y, Sugiyama Y, Sawada Y, Iga T, Hanano M. Tissue distribution of 14C-diazepam and its metabolites in rats. Drug Metab Dispos. 1982;10(6):676-9.

[17] Gueorguieva I, Nestorov IA, Murby S, Gisbert S, Collins B, Dickens K, et al. Development of a whole body physiologically based model to characterise the pharmacokinetics of benzodiazepines. 1: Estimation of rat tissue-plasma partition ratios. J Pharmacokinet Pharmacodyn. 2004;31(4):269-98.

[18] Gueorguieva I, Aarons L, Rowland M. Diazepam pharamacokinetics from preclinical to phase I using a Bayesian population physiologically based pharmacokinetic model with informative prior distributions in WinBUGS. J Pharmacokinet Pharmacodyn. 2006;33(5):571-94.

[19] Klotz U. Effect of age on levels of diazepam in plasma and brain of rats. Naunyn Schmiedebergs Arch Pharmacol. 1979;307(2):167-9.

[20] Hartigan JA, Wong MA. Algorithm AS 136: A K-Means Clustering Algorithm. Journal of the Royal Statistical Society Series C (Applied Statistics). 1979;28(1):100-8.

[21] Ward JH. Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association. 1963;58(301):236-44.

[22] Johnson SC. Hierarchical clustering schemes. Psychometrika. 1967;32(3):241-54.

[23] Zamek-Gliszczynski MJ, Ruterbories KJ, Ajamie RT, Wickremsinhe ER, Pothuri L, Rao MV, et al. Validation of 96-well equilibrium dialysis with non-radiolabeled drug for definitive measurement of protein binding and application to clinical development of highly-bound drugs. J Pharm Sci. 2011;100(6):2498-507.

[24] Banker MJ, Clark TH, Williams JA. Development and validation of a 96-well equilibrium dialysis apparatus for measuring plasma protein binding. J Pharm Sci. 2003;92(5):967-74.

Appendix A3: Supplementary material for Chapter 4

A3.1. Experimental data of diazepam and midazolam

A3.1.1. Protein Binding Studies

Plasma protein binding was determined by equilibrium dialysis using the 96-well equilibrium dialysis device (HTD-96b, HTDialysis) [1, 2]. Blank human and rat plasma (K2-EDTA, mixed gender, pooled) spiked with the studied compound was dialysed against phosphate buffer (133 mM pH=7.4) for 5 hours in triplicate. Prior to centrifugation, final dialysate and plasma samples were matched with blank matrices (90/10 buffer/plasma v/v) and proteins were precipitated by addition of acetonitrile (containing the internal standard, IS) in 3 volumes. Supernatant of the centrifuged samples were analysed by LC-MS/MS. The fup value was calculated as the ratio of the peak area ratio of compound:IS in the final buffer sample to the peak area ratio of compound:IS in the final plasma sample.

A3.1.2. Blood/Plasma Partitioning

The blood to plasma partitioning ratio was determined by spiking fresh male human or rat blood with the studied compound. After a 30-minute incubation, an aliquot of blood was collected and plasma was produced from the rest of the blood. Final plasma and blood samples were matched with blank matrices (50/50 blood/plasma), mixed with acetonitrile (containing the internal standard), centrifuged and the supernatants were analysed by LC-MS/MS following appropriate dilution. BP was calculated as the ratio of the peak area ratio in spiked blood to the peak area ratio in spiked plasma at 0.5h.

A3.2. Experimental data of basmisanil

All animal studies were performed according to the Swiss animal welfare regulations and in accordance with applicable SOPs and guidelines for the care and use of laboratory animals.

A3.2.1. Rat PK study

Basmisanil was administered intravenously as a solution in 33% NMP/26 % Hydroxypropyl g-Cyclodextrine (2 mL/kg body weight) at a mean dose level of 5 mg/kg to two male Wistar rats. The rats (Wistar strain, body weight 200-277 g) were obtained from Biological Research Laboratories, Füllinsdorf, Switzerland. Food was given ad libitum. The animals had free access to tap water during the whole study period. Blood

samples (approx. 0.5 mL per time point) were collected at 0.083, 0.25, 0.5, 0.75, 1.5, 3, 6 and 24 hours after intravenous. Collection tubes contained EDTA as anticoagulant and stabilizer. After centrifugation, plasma was removed and stored deep-frozen at approximately -20 °C until analysis. Concentrations in plasma were determined by HPLC with tandem mass spectrometric detection. The lower limit of quantification was 1 ng/mL, using 50µL of plasma specimen.

A3.2.2. Monkey PK study

Basmisanil was administered intravenously as a solution in 31.5% polyethylene glycol 400 /29 % propylene glycol (0.25 mL/kg body weight) at a mean dose level of 1 mg/kg to 3 Cynomolgus monkey. Cynomolgus monkeys (11.8 kg, 11.5 kg and 10.5 kg in body weight) were from a colony kept at F. Hoffmann-La Roche in Basel/ Switzerland. The animals had free access to food and tap water during the whole study period. Blood samples (approx. 0.5 mL per time point) were collected at 0.083, 0.25, 0.5, 0.75, 1.5, 3, 6 and 24 hours after intravenous administration. Collection tubes contained EDTA as anticoagulant and stabilizer. After centrifugation, plasma was removed and stored deep-frozen at approximately -20 °C until analysis. Concentrations in plasma were determined by HPLC with tandem mass spectrometric detection. The lower limit of quantification was 1 ng/mL, using 50µL of plasma specimen.

A3.2.3. Clinical PK study

PK data were collected from a Phase 1, single-centre, single-cohort, open-label, multiple dose study in 6 healthy male subjects was performed (ClinicalTrials.gov identifier NCT01684891). The clinical study was reviewed and approved by an independent ethics committee (Stichting Beoordeling Ethiek Bio-Medisch Onderzoek, Assen, NL). Only data related to the first dose administration were used in this work. On the mornings of Days 1 (and 28), an iv dose containing 0.1 mg of [^{13}C]-labelled basmisanil was given as an infusion over 15 minutes. Serial venous blood samples for pharmacokinetic assessments were collected. In addition, urine and faeces were collected according to predefined collection time points and intervals. Stopping criteria for urine and faeces collection: <1% of the administered dose excreted within 24 hours in faeces and urine on two consecutive days. Plasma concentrations of unlabelled basmisanil and its primary

metabolite, and of the stable isotopes [^{13}C]-basmisanil and [^{13}C]-metabolite were measured by a specific liquid chromatography/tandem mass spectrometry (LC/MS-MS) method. Total ([^{14}C]) radioactivity concentrations were determined in whole blood, plasma, urine and faeces by liquid scintillation counting. Six healthy male subjects were included. All subjects completed the study as planned and were included in all analyses. The parent compound was not detected in urine.

A3.2.4. Protein Binding Studies

Equilibrium dialysis studies were carried out as described by Banker et al [3]. Briefly, the dialysis sides of the apparatus were loaded with 0.15 mL phosphate buffer (mixture of 0.133 M potassium-di-hydrogen-phosphate and 0.133 M di-sodium-hydrogen-phosphate, pH 7.5). The same volume of diluted plasma spiked with different concentrations of test compound was pipetted into the sample side of each well. After the apparatus was loaded, the dialysis unit was sealed with an adhesive cover. The unit was incubated at 37°C for 5 hours, sufficient time to ensure equilibrium was reached. At the end of dialysis, the plasma and buffer samples were retrieved and the drug concentrations quantified by LSC.

The free fraction fu in plasma was calculated as follows: fu = (CB/CPe) where CB and CPe are the concentrations in buffer and plasma at the end of the dialysis, respectively.

Concentrations of [^{14}C]-basmisanil were determined in buffer and plasma by measuring radioactivity on a liquid scintillation counter with on-line quench correction by means of an external standard. For all samples, aliquots of 50 µL were added to 4.5 mL of scintillation liquid in counter vials.

A3.2.5. Blood/Plasma Partitioning

Aliquots (3.0 mL to 9.0 mL) of freshly drawn blood were centrifuged at low speed (600x g) for 10 minutes to generate a small erythrocyte-free plasma layer, and equilibrated at 37°C. The samples were spiked with aliquots of [^{14}C]-basmisanil solutions in DMSO, added to the erythrocyte-free plasma layer. This procedure was used to avoid the haemolysis otherwise observed when blood is spiked directly with buffer. Samples were immediately mixed at the desired constant temperature on a Coulter Mixer™ (Coulter Electronics, UK). The concentrations of [^{14}C]-basmisanil in blood ranged from ~10 to 10,000 ng/mL. After 5, 15 and 30 min, aliquots were removed, centrifuged at the same

temperature (3,000x g for 3 min), and concentrations of [^{14}C]-basmisanil in plasma and whole blood determined by LSC.

Reversibility of the distribution was established as follows: at the end of the equilibration phase, the plasma was removed and weighed, and the red blood cells resuspended in equal volumes of fresh plasma at the same temperature [2].

The haematocrit (H) was determined in the freshly drawn blood using a haematocrit centrifuge and haematocrit reader (Haemofuge™, Heraeus, Switzerland).

The blood/plasma concentration ratio (λ) and the fraction of drug in the erythrocytes (fE) can be calculated from:

$\lambda = CW/CP$ and $fE = QE/QW = (\lambda+H-1)/\lambda$

Here, CW and CP are the drug concentrations in whole blood and plasma, QE and QW are the amount of drug in the erythrocyte compartment and in blood.

Drug concentrations in whole blood were determined (in triplicate) after digestion of samples (50 µL) with Soluene 350® (Perkin Elmer, Order No. 6003038)/Isopropanol 1:1, bleaching with H_2O_2, and addition of Ultima-Gold® scintillation fluid (16 mL). For plasma, 50 µL aliquots were mixed with 4.5 mL scintillation fluid (triplicate determinations). The radioactivity was quantified in a liquid scintillation counter with online quench correction by means of an external standard.

A3.3. Biological plausibility of tissue to plasma unbound partitioning (Kpu) values

A number of equations have been developed to predict tissue to plasma partition coefficients (Kp) as they are widely used to describe drug distribution in PBPK models [3-10]. A study by Graham et al. compared the predictive performance between multiple tissue:plasma partition coefficient prediction methods with experimental rat partition coefficients and showed the Rodgers and Rowland (R&R) model generally made the most accurate predictions of Kps and Vss [11]. It was reported that 77.3% of the Kp predicted by R&R model were within 3-fold of experimentally determined Kp values. The two mechanistic equations proposed by Rodgers et al. [8, 9] for predictions of unbound partition coefficients (Kpu) require various input parameters such as physicochemical properties and *in vitro* data, as well as tissue composition data both of which can be accompanied by uncertainty and/or variability and thus affecting the predicted Kpu values [12].

Consequently, since it can be assumed that the true tissue to plasma unbound value (Kpu) is not equal to the R&R predicted tissue to plasma unbound partition coefficient ($\widehat{Kpu_T}$), a random error factor or bias (RE) can be considered as follows:

$$\frac{\boldsymbol{Kpu_T}}{\boldsymbol{\widehat{Kpu_T}}} = \boldsymbol{RE} \qquad \textbf{Eq. A3.1}$$

A log-normal distribution was assumed for RE where RE is expected to be 1 if Kpu predictions from the R&R equations are not systematically biased. Similarly, $\log\left(\frac{Kpu_T}{\widehat{Kpu_T}}\right)$ was assumed to follow a normal distribution with mean μ and standard deviation σ.

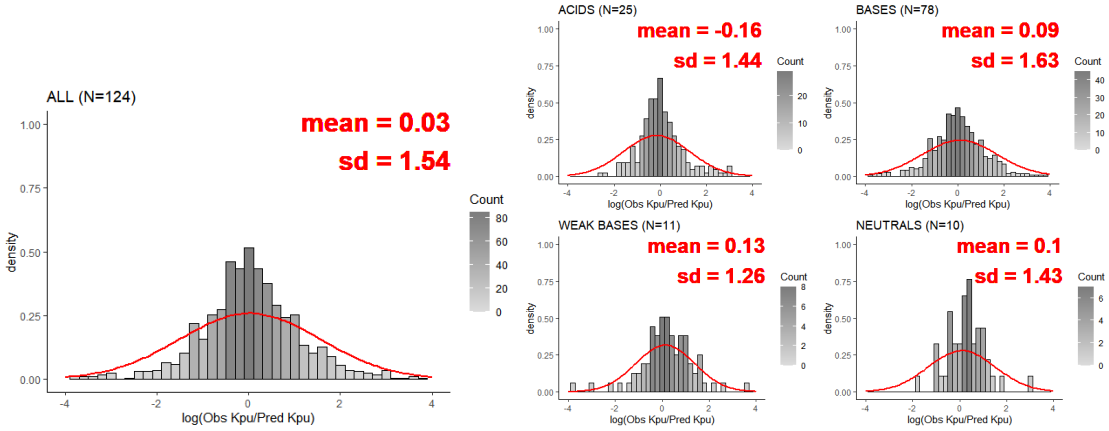
Fitting of data was performed in order to estimate the parameters of the normal distribution $N(\mu,\sigma^2)$. The dataset was rat Kps collected from the literature. In this dataset of 124 compounds (78 strong bases, 10 weak bases, 25 acids and 10 neutrals), 48% of the Kp data was missing.

Experimental steady state tissue to plasma partition coefficients (Kp) can be converted to tissue to unbound plasma partition coefficients (Kpu) by considering the free fraction in plasma (fu_p) of each compound:

$$\boldsymbol{Kp = Kpu \cdot fu_p} \qquad \textbf{Eq. A3.2}$$

After log-transforming the ratio of experimental steady state Kpu data to the R&R predicted Kpu, a normal distribution was fitted to the dataset of $\log\left(\frac{Kpu_T}{\widehat{Kpu_T}}\right)$. Results of the fittings are shown in Figure A3.1**.**

Figure A3.1: Distribution of the log-transformed random error of observed Kpu/predicted Kpu overall and per drug class



Using the fitted normal distribution, it is possible to derive 95% confidence interval of $\log\left(\frac{Kpu_T}{\widehat{Kpu_T}}\right)$ and subsequently $\frac{Kpu_T}{\widehat{Kpu_T}}$. Overall, Kpu predictions were within a 13-fold-error of the true value with 90% probability (21-fold-error with 95% probability). For acidic compounds, 90% of Kpu predictions agreed with experimental values within a factor of 9 (95% within a factor 14). For basic compounds, 90% of Kpu predictions agreed with experimental values within a factor of 16 (95% within a factor 27). For weakly basic compounds, 90% of Kpu predictions agreed with experimental values within a factor of 9 (95% within a factor 13). For basic compounds, 90% of Kpu predictions agreed with experimental values within a factor of 12 (95% within a factor 18).

Therefore, this analysis gave an indication about physiological plausibility of the Kpu values estimated.

A3.4. Physiological data

Table A3.1: Tissue composition data for rat [13]

Tissue	Vw	vNL	vNP	fEW	fIW	cAP	AR	LPR
plasma	0.96	0.0023	0.0013	0.945	NaN	0.057	NaN	NaN
adipose	0.144	0.853	0.0016	0.135	0.017	0.4	0.049	0.068
bone	0.417	0.0174	0.0016	0.1	0.346	0.67	0.1	0.05
brain	0.753	0.0391	0.0015	0.162	0.62	0.4	0.048	0.041
gut	0.738	0.0375	0.0124	0.282	0.475	2.41	0.158	0.141
heart	0.568	0.0135	0.0106	0.32	0.456	2.25	0.157	0.16
kidneys	0.672	0.0121	0.024	0.273	0.483	5.03	0.13	0.137
liver	0.642	0.0135	0.0238	0.161	0.573	4.56	0.086	0.161
lungs	0.574	0.0215	0.0123	0.336	0.446	3.91	0.212	0.168
muscle	0.726	0.01	0.0072	0.118	0.63	1.53	0.064	0.059
pancreas	0.641	0.0403	0.009	0.12	0.664	1.67	0.06	0.06
skin	0.658	0.0603	0.0044	0.382	0.291	1.32	0.277	0.096
spleen	0.562	0.0071	0.0107	0.207	0.579	3.18	0.097	0.207
thymus	0.752	0.0168	0.0092	0.15	0.626	2.3	0.075	0.075
RBC	0.6	0.0017	0.0029	NaN	0.603	0.5	NaN	NaN

Vw: fractional volume of water; vNL: fractional volume of neutral lipids; vNP: fractional volume of neutral phospholipids; fEW: fractional volume of extracellular water; fIW: fractional volume of intracellular water; cAP: acid phospholipid concentrations (mg/g); AR: albumin ratio; LPR: lipoprotein ratio; RBC: red blood cells

Table A3.2: Tissue composition data for human (Poulin et al., 2011)

Tissue	Vw	vNL	vNP	fEW	fIW	cAP	ALR	AR	LPR
plasma	0.95	0.0032	0.0021	NaN	NaN	NaN	NaN	NaN	NaN
adipose	0.15	0.79	0.002	0.135	0.017	0.4	0.15	0.049	0.068
bone	0.45	0.074	0.0011	0.1	0.346	0.67	0.5	0.1	0.05
brain	0.78	0.051	0.0565	0.162	0.62	0.4	0.5	0.048	0.041
gut	0.76	0.0487	0.0163	0.282	0.475	2.41	0.5	0.158	0.141
heart	0.78	0.0115	0.0166	0.32	0.456	2.25	0.5	0.157	0.16
kidneys	0.76	0.0207	0.0162	0.273	0.483	5.03	0.5	0.13	0.137
liver	0.73	0.0348	0.0252	0.161	0.573	4.56	0.5	0.086	0.161
lungs	0.78	0.003	0.009	0.336	0.446	3.91	0.5	0.212	0.168
muscle	0.71	0.022	0.0072	0.079	0.63	2.42	0.5	0.064	0.059
pancreas	0.641	0.0403	0.009	0.12	0.664	1.67	NaN	0.06	0.06
skin	0.67	0.0284	0.0111	0.382	0.291	1.32	0.5	0.277	0.096
spleen	0.79	0.0201	0.0198	0.207	0.579	3.18	0.5	0.097	0.207
thymus	0.78	0.0168	0.0092	0.15	0.626	2.3	0.5	0.075	0.075
RBC	0.63	0.0012	0.0033	NaN	0.603	0.57	NaN	NaN	NaN

Vw: fractional volume of water; vNL: fractional volume of neutral lipids; vNP: fractional volume of neutral phospholipids; fEW: fractional volume of extracellular water; fIW: fractional volume of intracellular water; cAP: acid phospholipid concentrations (mg/g); ALR: albumin and lipoprotein ratio; AR: albumin ratio; LPR: lipoprotein ratio; RBC: red blood cells

Table A3.3: Tissue composition data for monkey [8, 14, 15]

Tissue	Vw	vNL	vNP	fEW	fIW	cAP**	AR**	LPR**
plasma	0.95*	0.002	0.0021*	0.945**	NaN	0.057	NaN	NaN
adipose	0.15*	0.79*	0.002*	0.141**	0.039**	0.4	0.049	0.068
bone	0.45*	0.074*	0.0011*	0.3**	0.31**	0.67	0.1	0.05
brain	0.804	0.063	0.0471	0.166558	0.637442	0.4	0.048	0.041
gut	0.76*	0.0487*	0.0163*	0.282**	0.475**	2.41	0.158	0.141
heart	0.794	0.015	0.005	0.327423	0.466577	2.25	0.157	0.16
kidneys	0.76*	0.009	0.0458	0.308**	0.516**	5.03	0.13	0.137
liver	0.772	0.006	0.0148	0.169335	0.602665	4.56	0.086	0.161
lungs	0.826	0.003*	0.009*	0.354905	0.471095	3.91	0.212	0.168
muscle	0.791	0.006	0.0121	0.088137	0.702863	1.53	0.064	0.059
pancreas	0.641*	0.0403*	0.009*	0.12**	0.664**	1.67	0.06	0.06
skin	0.758	0.0284*	0.0111*	0.430247	0.327753	1.32	0.277	0.096
spleen	0.78*	0.0168*	0.0092*	0.207**	0.579**	3.18	0.097	0.207
thymus	0.752*	0.0168*	0.0092*	0.15**	0.626**	2.3	0.075	0.075
RBC	0.698	0.002	0.0037	NaN	0.698	0.5	NaN	NaN

Vw: fractional volume of water; vNL: fractional volume of neutral lipids; vNP: fractional volume of neutral phospholipids; fEW: fractional volume of extracellular water; fIW: fractional volume of intracellular water; cAP: acid phospholipid concentrations (mg/g); ALR: albumin and lipoprotein ratio; AR: albumin ratio; LPR: lipoprotein ratio; RBC: red blood cells
**Human source; **Rat source*

Table A3.4: Blood flow and volumes data for a reference rat (250g)[16, 17]

Tissues	Blood flows (ml/min)	Volume (ml)
Lung	83.09	1.20
Arterial	83.09	6.54
Kidney	11.72	1.76
Venous	83.09	13.08
Heart	4.07	0.79
Pancreas	1.50	0.77
Bone	10.146	5.40
Spleen	0.83	0.48
Gut	9.14	5.39
Liver	12.55	8.47
Brain	1.66	1.37
Stomach	1.08	1.11
Adipose	5.82	19.79
RoB	7.23	27.94
Muscle	23.10	97.19
Skin	4.82	45.75

Table A3.5: Blood flow and volumes data for a reference man (70kg)[17-20]

Tissues	Blood flows (L/min)	Volume (L)
Lung	5.84	0.51
Arterial	5.84	1.33
Kidney	1.11	0.30
Venous	5.84	3.99
Spleen	0.18	0.18
Heart	0.23	0.32
Liver	1.11	1.67
Pancreas	0.06	0.09
Brain	0.70	1.35
Gut	0.53	1.01
Stomach	0.06	0.14
RoB	0.73	2.91
Skin	0.29	2.50
Bone	0.29	5.21
Muscle	0.99	26.92
Adipose	0.29	15.62

Table A3.6: Blood flow and volumes data for a monkey (5kg) [17, 21-23]

Tissues	Blood flows (mL/min)	Volume (mL)
Lung	785.8	41.9
Arterial	785.8	86.1
Kidney	99.8	18.1
Venous	785.8	172.2
Heart	43.2	18.4
Pancreas	7.1	9.6
Bone	98.2	386.9
Spleen	14.9	3.3
Gut	90.4	184.6
Liver	123.4	118.5
Brain	51.9	81.6
Stomach	11.0	23.8
Adipose	19.6	131.0
RoB	121.8	130.7
Muscle	154.8	2535.5
Skin	33.8	487.0

A3.5. Comparison of tissue concentrations between experimental and predicted values after optimization of simplified models using blood or plasma data

Following fitting of the simplified PBPK models, new Kpu estimates were obtained either from the approach with common Kpus or common scalars. It is then reasonable to hypothesize that using the new Kpu values, the models can describe tissue profiles besides plasma profiles. This is indeed an advantage over a compartmental model approach which can fit plasma data but is not highly informative about the drug tissue distribution. Furthermore, this simplified PBPK model approach is able to provide prediction of drug distribution in each tissue instead of in groups of lumped tissues as in the kinetically lumped models. The question of whether these predictions of drug distributions from these simplified PBPK models are then physiologically relevant needs to be considered.

A3.5.1. Example of diazepam in rat

Measured tissue concentrations and Kpus were available in rat for diazepam [24, 25]. Thus, a comparison of tissue concentrations and Kpus could be made with the best suited models optimized in rat for diazepam. Results are shown in Table A3.7 and Figure A3.2.

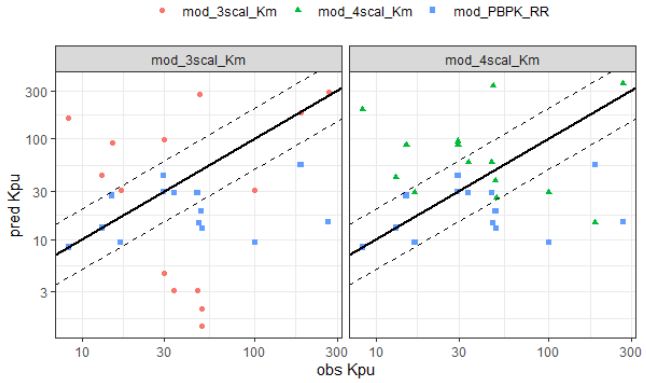
Tissue concentration seemed to be well predicted using the optimized Kpu in tissues like adipose and muscle (Figure A3.3). The model with common scalars using clustering on K-means data could gave good predictions of concentration in lung, heart, liver.

Table A3.7: Comparison of measured and optimised rat Kpus of diazepam

Kpu	Experimental Kpus [1]	Experimental Kpus [2]	Model 3C	Model 3D
Lung		23.29	2.02	38.3
Gut	37.6	14.29	3.11	58.8
Stomach	58.73		3.11	588
Pancreas			98.7	94.8
Liver	29	32.14	292	353
Bone	0	0	43.2	41.6
Brain	18.4	7.29	90.8	87.2
Heart	50.8	15.64	1.38	26.1
Kidney	41.67	16.43	281	341
Skin	35.4	24.07	4.63	87.5
Muscle	18.33	9.79	30.8	29.6
Adipose	235.33	92.14	183	14.6
RoB		28.57	30.8	29.6
Testis	57.2	21.43		
Spleen		12.86	162	196

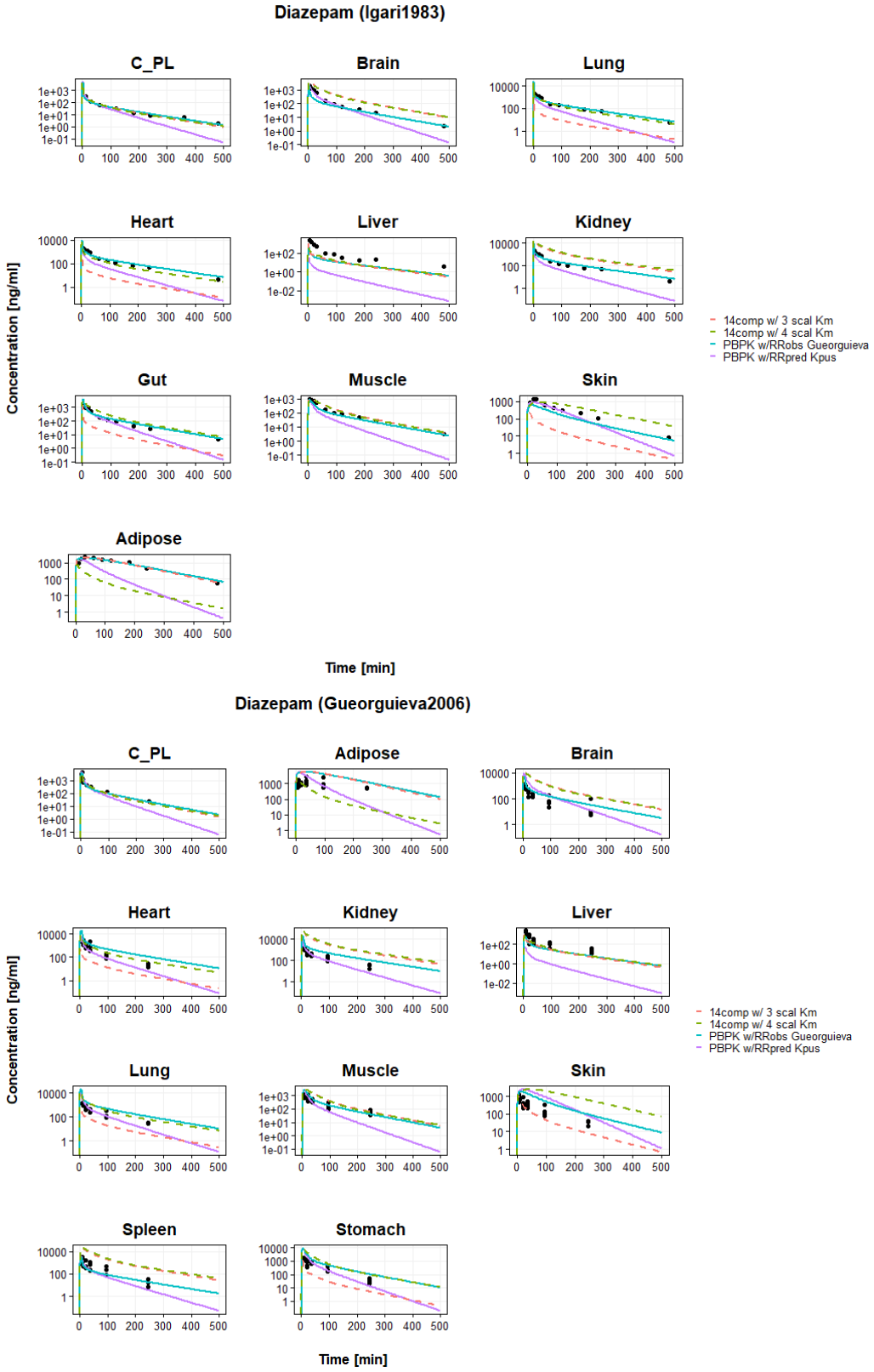
[1] *[26](area method);* [2] *[27]*

Figure A3.2: Comparison of measured [26] and optimised rat Kpus of diazepam vs R&R predicted Kpu



Dashed-lines represent 2-fold error. Mod_3scal_Km=model 3C and Mod_4scal_Km=model 3D

Figure A3.3: Comparison of measured and predicted rat tissue concentration of diazepam



A3.5.2. Example of midazolam in rat

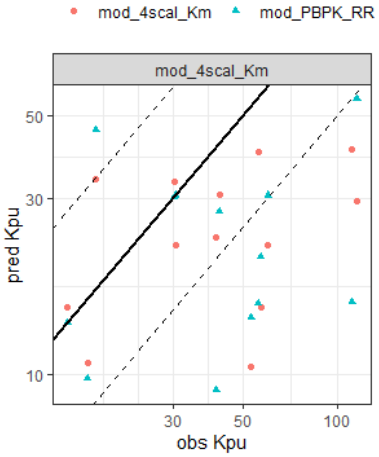
Measured Kpus were available in rat for midazolam [28]. A comparison of tissue Kpus could be made with the best suited models optimized in rat for midazolam. Results are shown in **Table A3.8** and **Figure A3.4**.

Table A3.8: Comparison of measured and optimised rat Kpus of midazolam

	Experimental Kpus [1]	Model 3D
Lung	**57**	**15.3**
Gut	**29-60**	**22.4**
Stomach	**97**	**22.4**
Liver	**111**	**40.5**
Brain	**42**	**30.5**
Heart	**53**	**10.5**
Kidney	**56**	**40**
Skin	**17**	**33.6**
Muscle	**16**	**10.8**
Adipose	**115**	**29.5**
Spleen	**41**	**23.5**

[1] *[28]*

Figure A3.4: Comparison of measured [28] and optimised rat Kpus of midazolam vs R&R predicted Kpu



Dashed-lines represent 2-fold error

A3.6. Performance of the best models

A3.6.1. Diazepam

Fitting of the investigated models in preclinical species

Table A3.9: Parameter estimates of the investigated model for diazepam in rat

Models	BIC	Kpu1/SF1 (RSE%)	Kpu2 /SF2 (RSE%)	Kpu3/SF3 (RSE%)	Kpu4/SF4 (RSE%)	IIV CL as %CV (RSE%)	Prop err	Vss,b estimated [L]	Vss within 20% (YYY), 25% (YY), 30% (Y)
1	-209.012	40.3 (2730)	26.4 (2.5)	199.5 (12.5)	N/A	26.8% (25.4)	19.5% (14.5)	1.11	YY
2A	-208.136	28.4 (12)	307 (45.9)	88.2 (13.9)	N/A	29.2% (21.8)	20.3% (8.6)	1.34	N/A
2B	-169.673	29.4 (12.7)	309 (44.1)	88.1 (13.9)	23.1 (7.1)	28.8% (21.8)	20.2% (8.6)	1.34	N/A
2C	-211.201	27.3 (11.6)	427 (19.5)	87.3 (13.5)	N/A	26.4% (20.8)	19.8% (7.1)	1.14	Y
2D	-172.823	29.1 (12.6)	438 (11.4)	87.3 (13.6)	17.8 (12.9)	26.9% (20.6)	19.8% (7.1)	1.15	Y
2E	-169.818	47.0 (6.8)	334 (7.3)	424 (59.1)	N/A	33.6% (33.6)	29.3% (9.9)	1.22	N/A
2F	-135.196	46.9 (6.8)	334 (8.4)	3.11 (24.2)	505 (69.5)	31.6% (19.7)	29.2% (9.9)	1.3	N/A
3A	-198.144	1.92 (28.9)	27.3 (20.7)	1.77 (18.7)	N/A	30.6 (18.9)	21.5 (7.8)	1.71	N/A
3B	-169.725	3.20 (11.7)	19.5 (31.9)	2.03 (13.9)	0.394 (7.7)	29.8 (21.1)	20.2 (8.8)	1.56	N/A
3C	-206.274	3.33 (9.5)	19.5 (27.6)	0.11 (23.8)	N/A	26.2 (21)	20.1 (9.2)	1.06	YYY
3D	-173.892	3.2 (11.2)	23.5 (23.7)	2.02 (13.8)	0.265 (20.9)	26 (21.7)	19.6 (8.8)	1.11	YY
3E	-204.291	1.77 (11.5)	42.9 (18.7)	0.894 (5.8)		26.3 (20.3)	20.5 (6.8)	0.99	YYY
3F	-170.095	1.76 (11.4)	42.4 (18.8)	757 (17.2)	0.009 (10600)	25.9 (21.2)	20.4 (7.4)	1.86	N/A

*Estimates were reported with relative standard error (RSE) in brackets. Relative standard errors (RSEs) were calculated as: (standard error/estimate)*100. To obtain the RSEs in the domain of the reported original parameter instead of the log-transformed domain, normal/log-normal reverse algebra was applied (see A3.7). Intra-individual variability (IIV) was expressed as Coefficient of variation (% CV) which was calculated as:* $\sqrt{(e^{\omega^2} - 1)} \cdot 100$

Abbreviations for model are defined in Table 4.3.

Table A3.10: Parameter estimates of the investigated model for diazepam in monkey

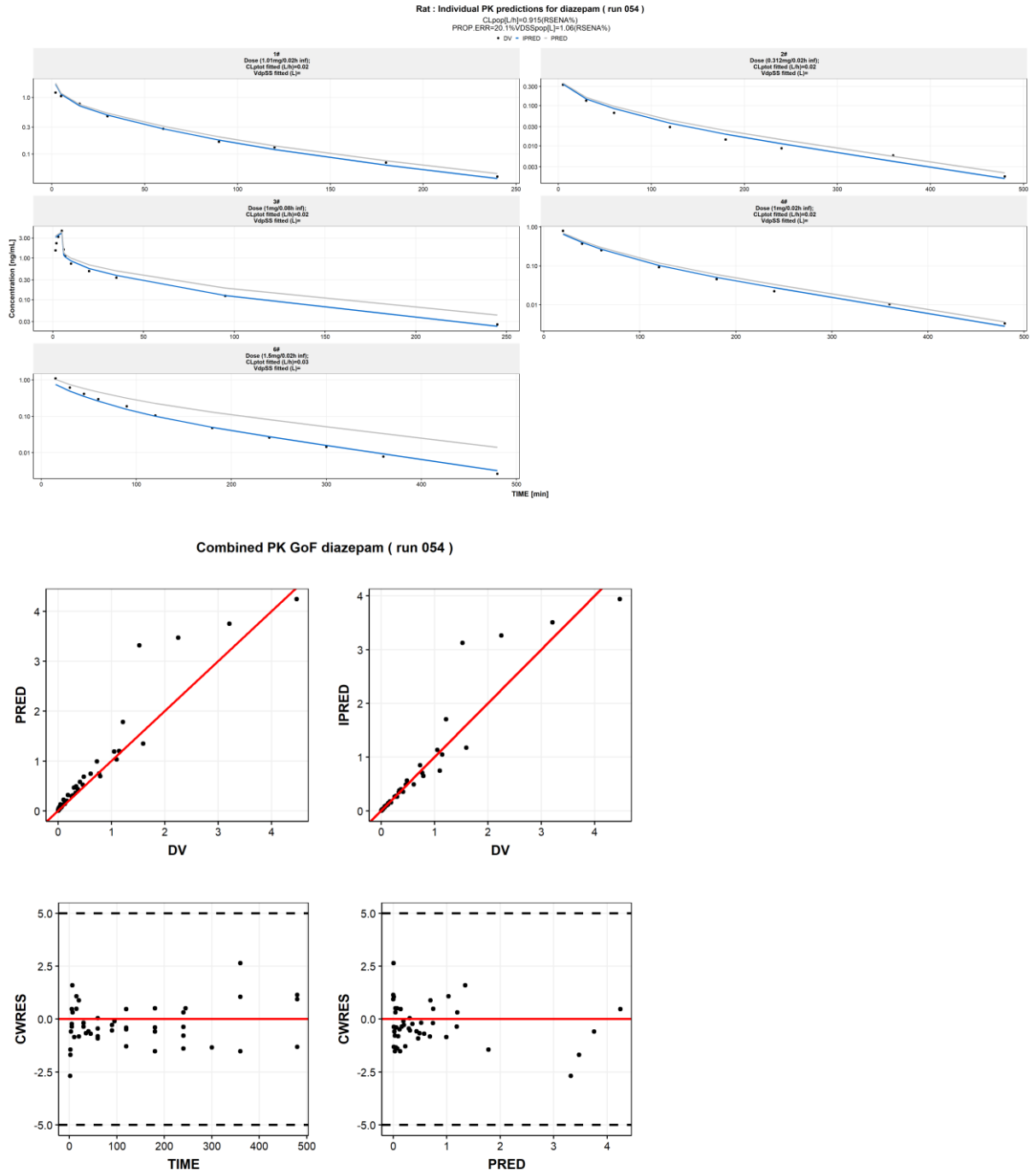
Models	BIC	SF1 (RSE%)	SF2 (RSE%)	SF3 (RSE%)	SF4 (RSE%)	Prop err	Vss,b est [L/kg]	Vss within 20% (YYY), 25% (YY), 30% (Y)
3A	**-140.543**	**0.153 (52.9)**	**44..5 (100)**	**3.34 (3.9)**	**N/A**	**51.4% (14.9)**	**50.9**	**N/A**
3B	**-137.609**	**0.289 (20.7)**	**21.0 (94.2)**	**3.30 (4.4)**	**1.20 (10.4)**	**55.8% (4.6)**	**28.8**	**N/A**
3C	**-138.918**	**0.500 (23.9)**	**12.9 (10.9)**	**3.47 (7.4)**	**N/A**	**64.3% (6.5)**	**12.17**	**YYY**
3D	**-136.666**	**0.425 (26.6)**	**14.8 (8.9)**	**3.38 (4.4)**	**1.20 (25.4)**	**62.8% (6.6)**	**12.16**	**YYY**
3E	**-132.007**	**1.23 (5.7)**	**0.453 (5.7)**	**121 (10.0)**	**N/A**	**63.2% (7.3)**	**14.9**	**N/A**
3F	**-129.234**	**1.23 (5.7)**	**0.921 (34.7)**	**8.33 (148)**	**120 (10.5)**	**63.2% (7.3)**	**14.8**	**N/A**

*Estimates were reported with relative standard error (RSE) in brackets. Relative standard errors (RSEs) were calculated as: (standard error/estimate)*100. To obtain the RSEs in the domain of the reported original parameter instead of the log-transformed domain, normal/log-normal reverse algebra was applied (see A3.7). Intra-individual variability (IIV) was expressed as Coefficient of variation (% CV) which was calculated as:* $\sqrt{(e^{\omega^2} - 1)} \cdot 100$

Abbreviations for model are defined in Table 4.3.

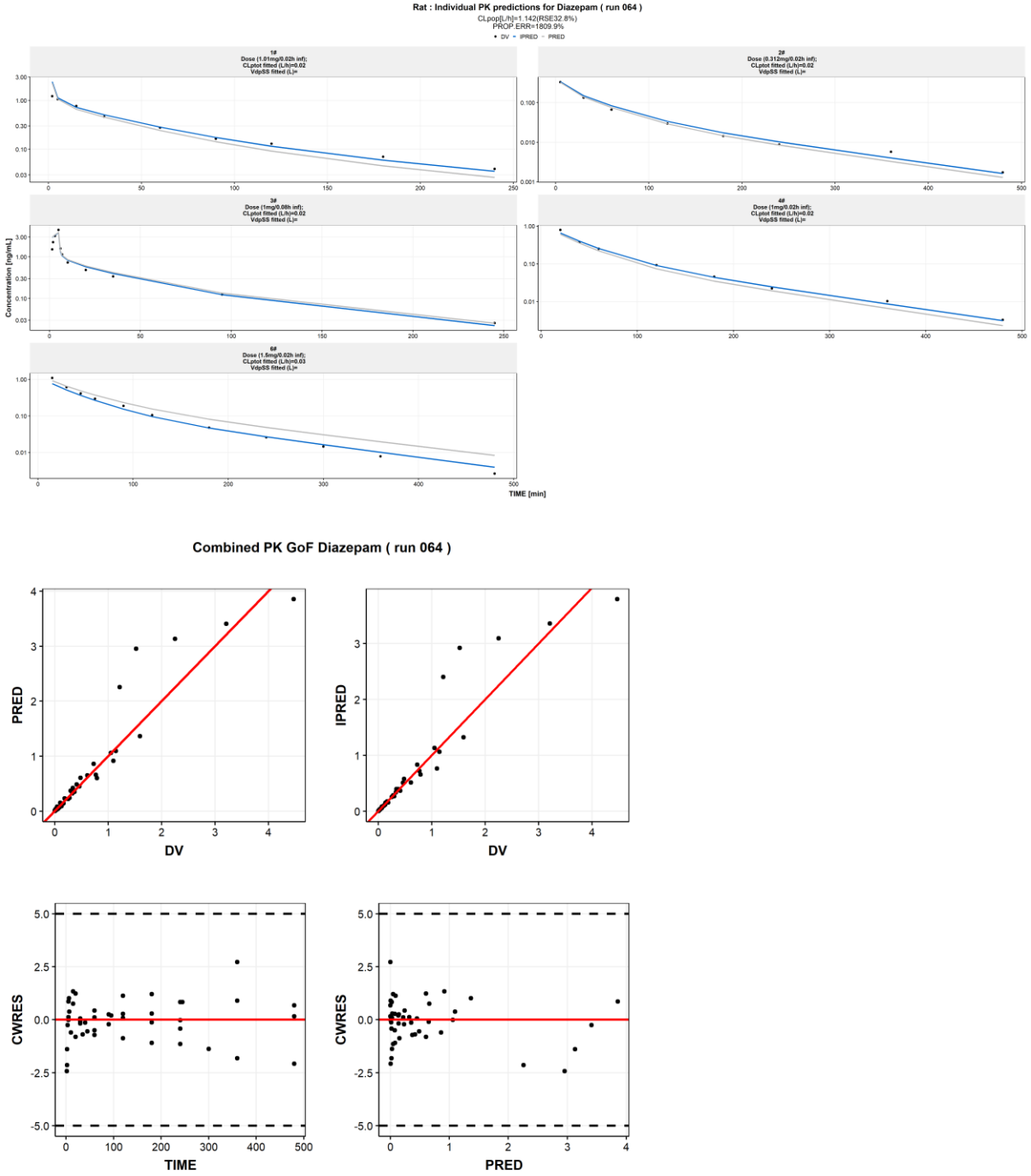
Fitting of the best simplified models in preclinical species

Figure A3.5: Fits of PBPK model with 3 common Kpu scalars model (K-means clustering) in rat



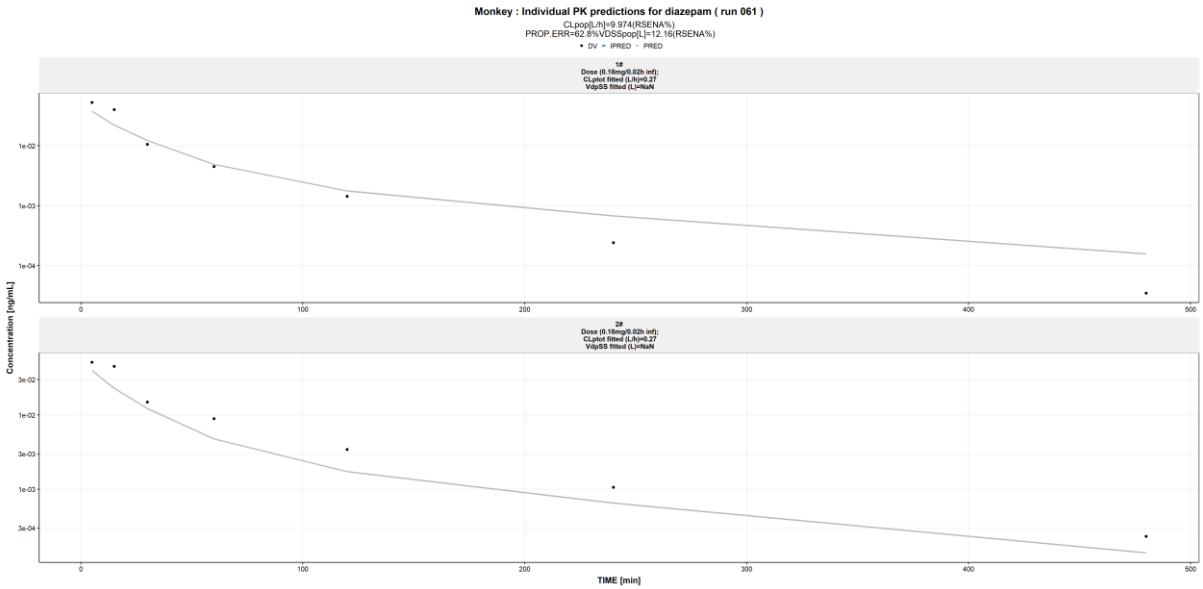
DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

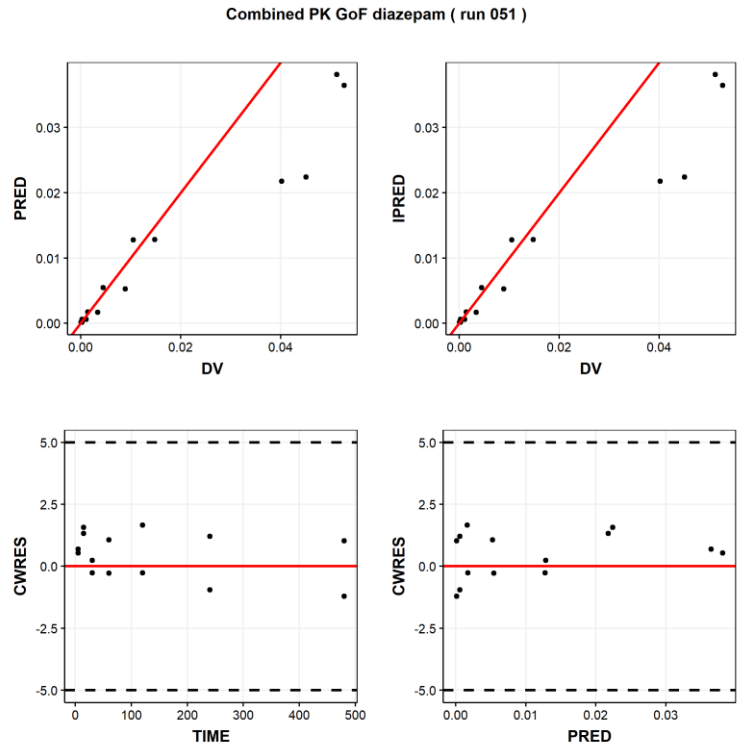
Figure A3.6: Fits of PBPK model with 4 common Kpu scalars model (K-means clustering) in rat



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

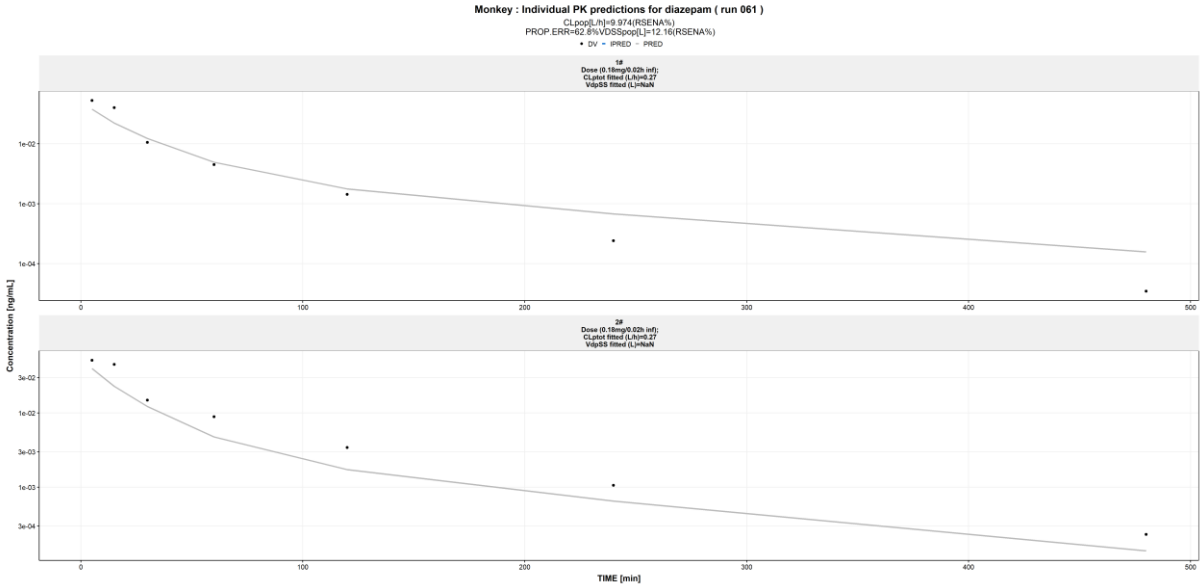
Figure A3.7: Fits of PBPK model with 3 common Kpu scalars model (K-means clustering) in monkey

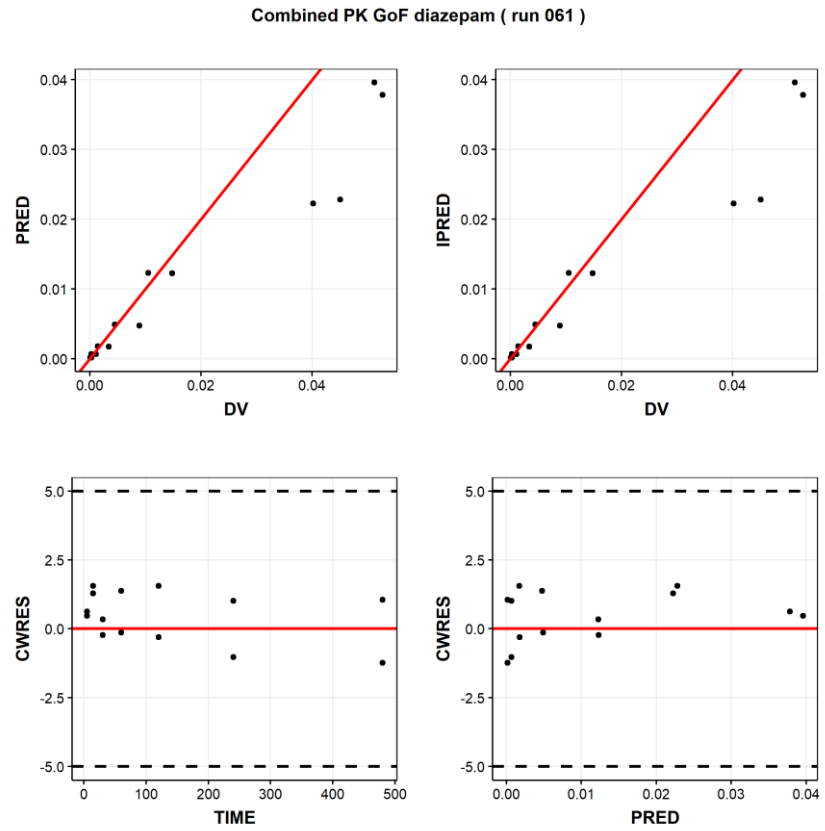




DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

Figure A3.8: Fits of PBPK model with 4 common Kpu scalars model (K-means clustering) in monkey





DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A3.6.2. Midazolam

Fitting of the investigated models in preclinical species

Table A3.11: Parameter estimates of the investigated model for midazolam in rat

Models	BIC	SF1 (RSE%)	SF2 (RSE%)	SF3 (RSE%)	SF4 (RSE%)	IIV CL as %CV (RSE%)	Prop err	Vss,b estimated [L]	Vss within 20% (YYY), 25% (YY), 30% (Y)
3A	-203.717	0.553 (34.9)	24.7 (25.8)	0.626 (18)	N/A	28 (15.7)	23.5 (11.7)	1.04	N/A
3B	-166.558	0.584 (32.3)	24.7 (25.2)	0.644 (13.3)	0.511 (46.3)	27.9 (15.5)	23.6 (12)	1.04	N/A
3C	-193.811	1.19 (17.2)	4.80 (22.7)	0.236 (92)	N/A	30.6 (15.2)	25.7 (17)	0.51	YYY
3D	-157.038	1.17 (11.5)	5.18 (41.7)	0.761 (27.8)	0.487 (44.9)	29.7 (15.4)	25.8 (17.3)	0.54	YYY
3E	-195.665	0.647 (14.6)	28.5 (28.6)	14.9 (66.4)	N/A	29.6 (14.6)	25.2 (16.1)	0.34	N/A
3F	-158.207	0.667 (14.4)	27.2 (27.1)	1.19 (69.1)	13.9 (62.1)	30 (14)	25.4 (15.6)	0.31	N/A

N/A: no value due to model failing to converge
*Estimates were reported with relative standard error (RSE) in brackets. Relative standard errors (RSEs) were calculated as: (standard error/estimate)*100. To obtain the RSEs in the domain of the reported original parameter instead of the log-transformed domain, normal/log-normal reverse algebra was applied (see A3.7). Intra-individual variability (IIV) was expressed as Coefficient of variation (% CV) which was calculated as:* $\sqrt{(e^{\omega^2} - 1)} \cdot 100$
Abbreviations for model are defined in Table 4.3.

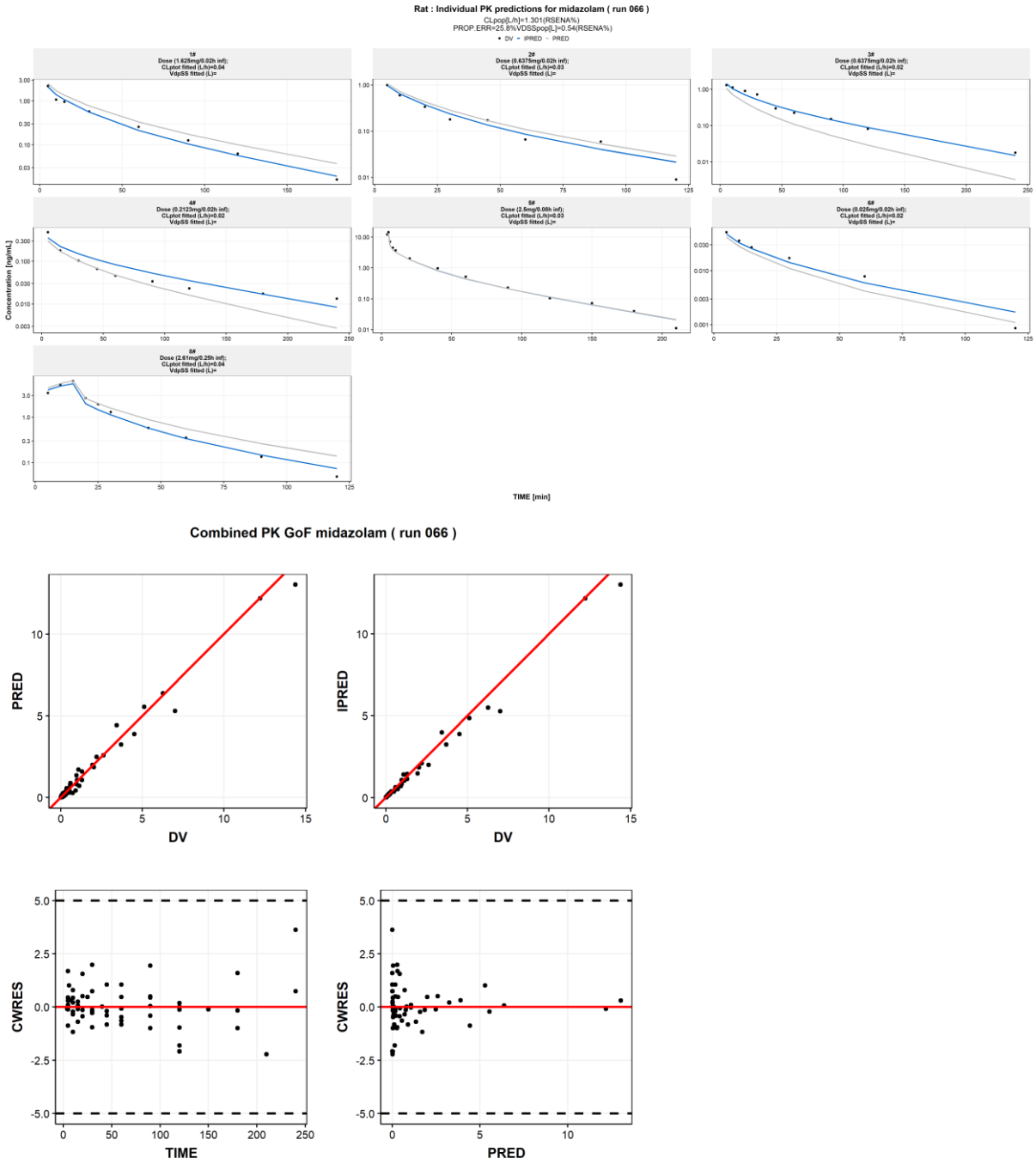
Table A3.12: Parameter estimates of the investigated model for midazolam in monkey

Models	BIC	SF1 (RSE%)	SF2 (RSE%)	SF3 (RSE%)	SF4 (RSE%)	IIV CL as %CV (RSE%)	Prop err	Vss,b estimated [L]	Vss within 20% (YYY), 25% (YY), 30% (Y)
3A	-120.577	0.367 (15.8)	6.50 (19.2)	0.225 (17.8)	N/A	27.2 (7.9)	25 (17.7)	18.5	-
3B	-114.154	0.409 (14.6)	6.55 (18.8)	0.201 (16.9)	0.153 (15.4)	26.9 (8.3)	24.9 (17.5)	18.6	-
3C	-117.811	0.323 (16.3)	11.565 (37.7)	1.481 (13)	N/A	28.1 (7.5)	28.5 (20.8)	11.43	Y
3D	-124.556	0.324 (16)	21.22 (9.6)	0.203 (14.7)	1.878 (13.2)	21.6 (3.3)	31.5 (21.5)	10.31	YYY
3E	-105.884	0.644 (8.8)	0.617 (40.2)	5.73 (91.8)	N/A	35.2 (10.9)	25.2 (27.9)	11.87	N/A
3F	-104.170	0.625 (7.0)	8.72 (26.3)	10.8 (64.4)	2.37 (98.7)	34.6 (11.4)	24.8 (27.6)	11.93	N/A

N/A: no value due to model failing to converge
*Estimates were reported with relative standard error (RSE) in brackets. Relative standard errors (RSEs) were calculated as: (standard error/estimate)*100. To obtain the RSEs in the domain of the reported original parameter instead of the log-transformed domain, normal/log-normal reverse algebra was applied (see A3.7). Intra-individual variability (IIV) was expressed as Coefficient of variation (% CV) which was calculated as:* $\sqrt{(e^{\omega^2} - 1)} \cdot 100$
Abbreviations for model are defined in Table 4.3.

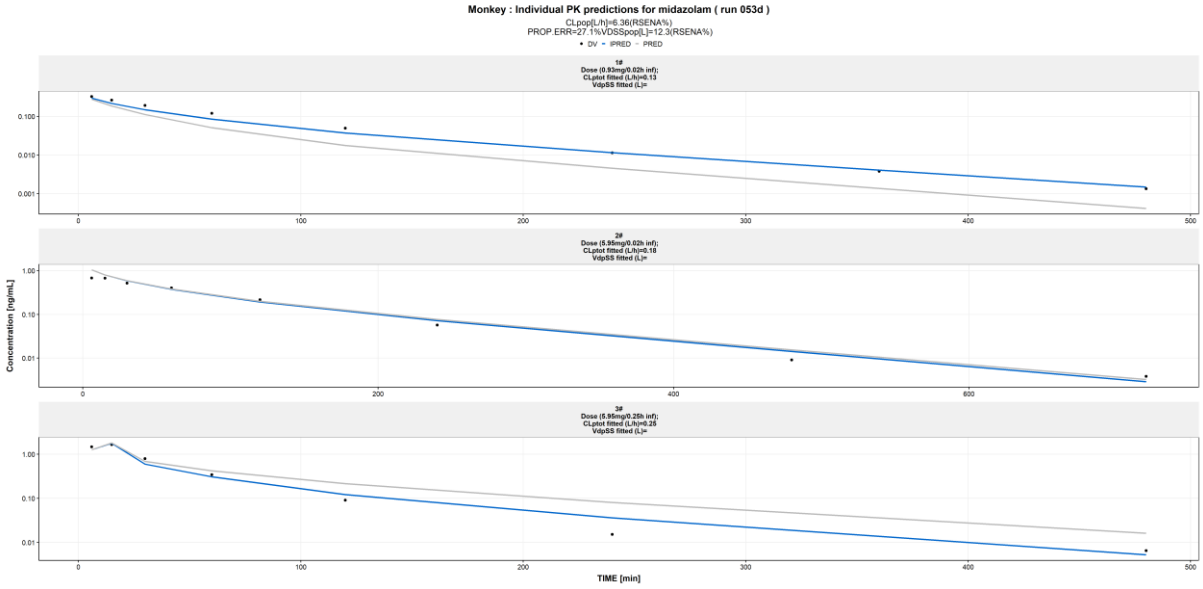
Fitting of the best simplified models in preclinical species

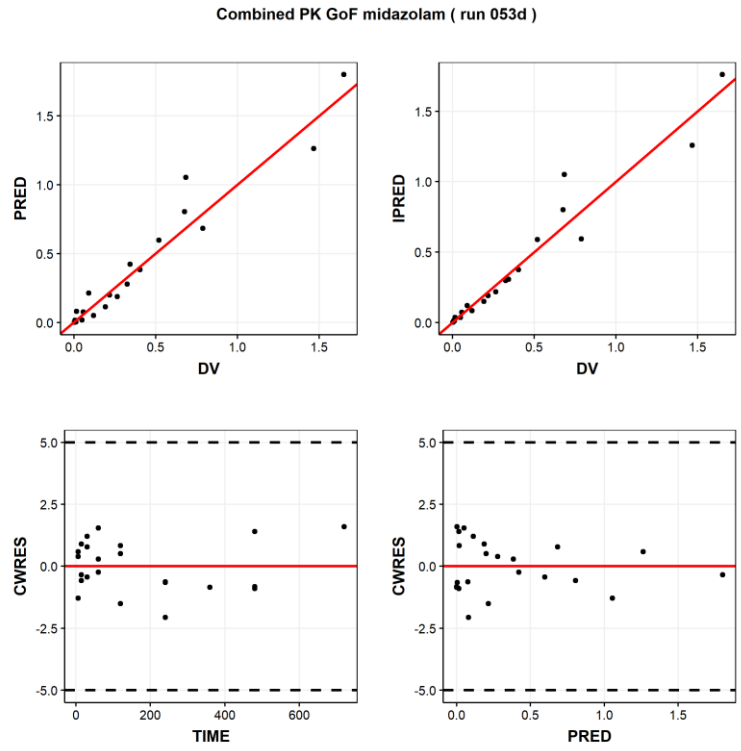
Figure A3.9: Fits of PBPK model with 4 common Kpu scalars model (K-means clustering) in rat



DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

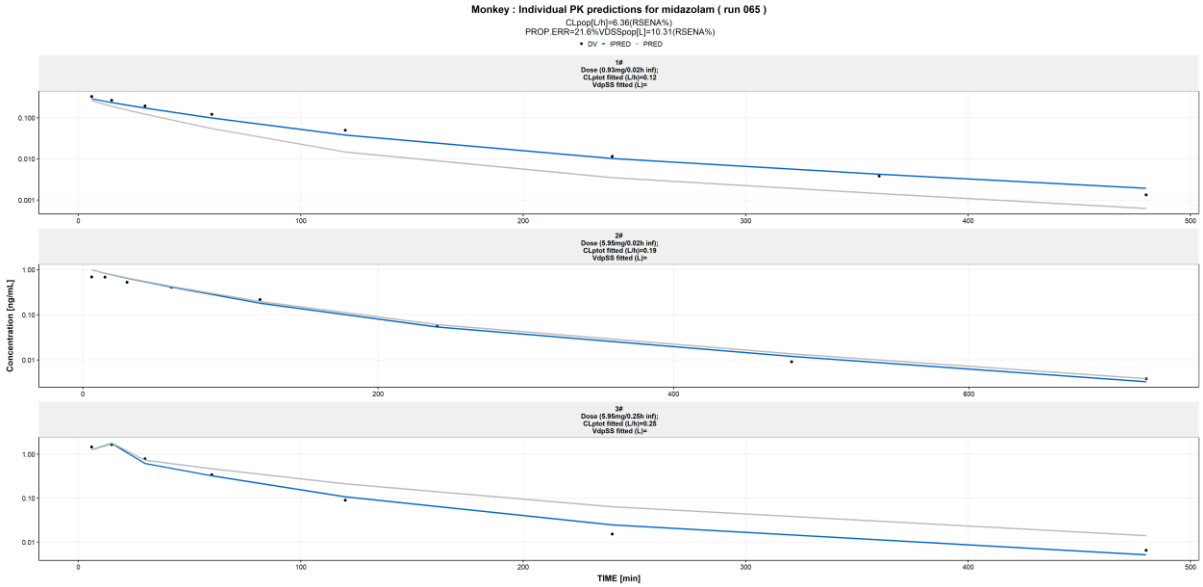
Figure A3.10: Fits of PBPK model with 3 common Kpu scalars model (K-means clustering) in monkey

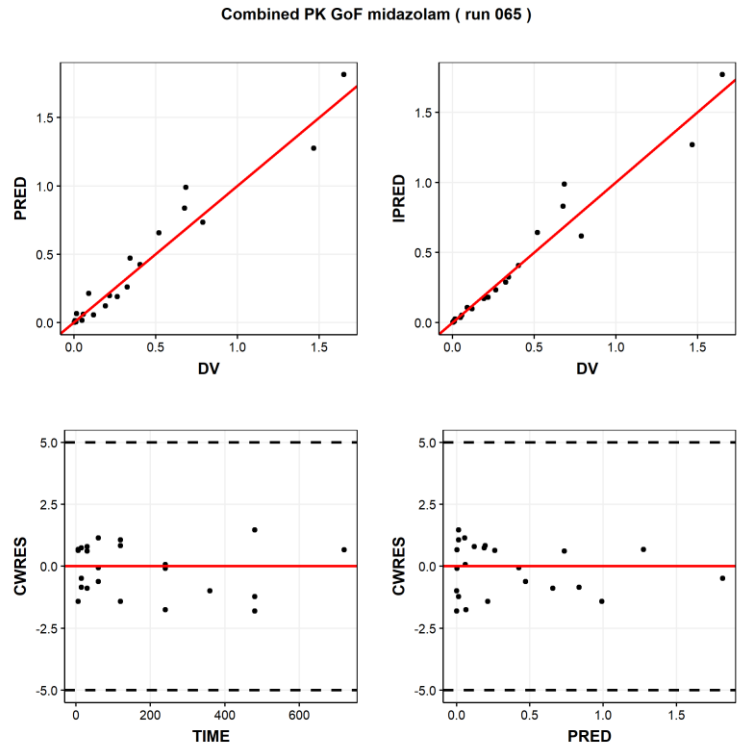




DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

Figure A3.11: Fits of PBPK model with 4 common Kpu scalars model (K-means clustering) in monkey





DV: observed; PRED: population predicted; IPRED: individual predicted concentration (ng/ml); CWRES: conditional weighted residuals: TIME: time (h)

A3.6.3. Basmisanil

Fitting of the investigated models in preclinical species

Table A3.13: Parameter estimates of the investigated model for basmisanil in rat

Models	BIC	SF1 (RSE%)	SF2 (RSE%)	SF3 (RSE%)	SF4 (RSE%)	Prop err	Vss,b estimated [L]	Vss within 20% (YYY), 25% (YY), 30% (Y)
3A	163.831	1.68 (2.5)	339 (8.9)	2292 (7.7)	N/A	43.6% (6.7)	111.02	N/A
3B	190.657	2.344	357	0.313	8344	48.4% (8.7)	153	N/A
3C	175.292	0.273 (240)	119 (31)	577 (3.2)	N/A	46.1% (1.8)	30.84	N/A
3D	181.136	2.84 (38.3)	1.56 (30.4)	553 (1.8)	0.033 (66.1)	46.3% (2.2)	29.59	N/A
3E	160	5.23 (19.2)	15.1 (9.6)	693 (10.5)	N/A	30.6% (9.2)	2.22	N/A
3F	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

*Estimates were reported with relative standard error (RSE) in brackets. Relative standard errors (RSEs) were calculated as: (standard error/estimate)*100. To obtain the RSEs in the domain of the reported original parameter instead of the log-transformed domain, normal/log-normal reverse algebra was applied (see A3.7). Intra-individual variability (IIV) was expressed as Coefficient of variation (% CV) which was calculated as:* $\sqrt{(e^{\omega^2} - 1)} \cdot 100$

Abbreviations for model are defined in Table 4.3.

Table A3.14: Parameter estimates of the investigated model for basmisanil in monkey

Models	BIC	SF1 (RSE%)	SF2 (RSE%)	SF3 (RSE%)	SF4 (RSE%)	IIV CL as %CV (RSE%)	Prop err	Vss,b est [L/kg]	Vss within 20% (YYY), 25% (YY), 30% (Y)
3A	74.52	1.7 (22.3)	$3.6x10^{-4}$ (221)	0.23 (195)	N/A	341 (40.9)	73% (19.1)	5.49	N/A
3B	75.34	1.84 (96.6)	$5.7x10^{-5}$ (N/A)	0.24 (71.8)	0.6 (41)	341 (40.8)	72% (19.3)	5.67	N/A
3C	69.04	1.99 (27.9)	25.4 (62.1)	0.19 (94.8)	N/A	348 (40.8)	69% (18.2)	8.63	N/A
3D	69.89	2.15 (33.8)	26.1 (68)	0.19 (53.3)	0.6 (88.8)	348 (41)	68% (18.4)	8.87	N/A
3E	75.16	1.41 (16.2)	$7.8x10^{-5}$ ($1.9x10^{3}$)	$9.8x10^{-5}$ ($1.6x10^{-4}$)	N/A	340 (40.8)	73% (18.5)	6.30	N/A
3F	70.32	1.68 (23)	24.6 (64.6)	520 (585)	$4.9x10^{-6}$ ($6.4x10^{-4}$)	348 (41)	69% (18.8)	42.7	N/A

*Estimates were reported with relative standard error (RSE) in brackets. Relative standard errors (RSEs) were calculated as: (standard error/estimate)*100. To obtain the RSEs in the domain of the reported original parameter instead of the log-transformed domain, normal/log-normal reverse algebra was applied (see A3.7). Intra-individual variability (IIV) was expressed as Coefficient of variation (% CV) which was calculated as:* $\sqrt{(e^{\omega^2} - 1)} \cdot 100$

Abbreviations for model are defined in Table 4.3.

A3.7. Notes on the calculation of RSE%

The parameters estimated in the simplified models were log-transformed when fitting the simplified models to the data. In order to calculate the relative standard errors (RSEs) in the domain of the reported original parameter instead of the log-transformed domain normal/log-normal reverse algebra needed to be applied.

Considering a variable following a lognormal distribution with mean (M) and variance (V), then the mean (μ) and standard-deviation (σ) of the respective normally untransformed variable is calculated as followed:

$$\mu = \exp\left(M + \frac{V}{2}\right) \quad \textbf{Eq. A3.3}$$

$$\sigma^2 = \exp(2 \cdot M + V) \cdot (\exp(V) - 1) \quad \textbf{Eq. A3.4}$$

And finally, $RSE = \frac{\sigma}{\mu} \cdot 100$

A3.8. References

[1] Zamek-Gliszczynski MJ, Ruterbories KJ, Ajamie RT, Wickremsinhe ER, Pothuri L, Rao MV, et al. Validation of 96-well equilibrium dialysis with non-radiolabeled drug for definitive measurement of protein binding and application to clinical development of highly-bound drugs. J Pharm Sci. 2011;100(6):2498-507.

[2] Banker MJ, Clark TH, Williams JA. Development and validation of a 96-well equilibrium dialysis apparatus for measuring plasma protein binding. J Pharm Sci. 2003;92(5):967-74.

[3] Poulin P, Theil FP. Development of a novel method for predicting human volume of distribution at steady-state of basic drugs and comparative assessment with existing methods. J Pharm Sci. 2009;98(12):4941-61.

[4] Berezhkovskiy LM. Volume of distribution at steady state for a linear pharmacokinetic system with peripheral elimination. Journal of Pharmaceutical Sciences. 2004;93(6):1628-40.

[5] Poulin P, Krishnan K. An algorithm for predicting tissue: blood partition coefficients of organic chemicals from n-octanol: water partition coefficient data. J Toxicol Environ Health. 1995;46(1):117-29.

[6] Poulin P, Schoenlein K, Theil FP. Prediction of adipose tissue: Plasma partition coefficients for structurally unrelated drugs. Journal of Pharmaceutical Sciences. 2001;90(4):436-47.

[7] Poulin P, Theil FP. A Priori Prediction of Tissue:Plasma Partition Coefficients of Drugs to Facilitate the Use of Physiologically-Based Pharmacokinetic Models in Drug Discovery. Journal of Pharmaceutical Sciences. 2000;89(1):16-35.

[8] Rodgers T, Leahy D, Rowland M. Physiologically based pharmacokinetic modeling 1: predicting the tissue distribution of moderate-to-strong bases. J Pharm Sci. 2005;94(6):1259-76.

[9] Rodgers T, Rowland M. Physiologically based pharmacokinetic modelling 2: predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. J Pharm Sci. 2006;95(6):1238-57.

[10] Assmus F, Houston JB, Galetin A. Incorporation of lysosomal sequestration in the mechanistic model for prediction of tissue distribution of basic drugs. Eur J Pharm Sci. 2017;109:419-30.

[11] Graham H, Walker M, Jones O, Yates J, Galetin A, Aarons L. Comparison of in-vivo and in-silico methods used for prediction of tissue: plasma partition coefficients in rat. J Pharm Pharmacol. 2012;64(3):383-96.

[12] Yau E, Olivares-Morales A, Gertz M, Parrott N, Darwich AS, Aarons L, et al. Global Sensitivity Analysis of the Rodgers and Rowland Model for Prediction of Tissue: Plasma Partitioning Coefficients: Assessment of the Key Physiological and Physicochemical Factors That Determine Small-Molecule Tissue Distribution. AAPS J. 2020;22(2):41.

[13] Rodgers T, Rowland M. Mechanistic approaches to volume of distribution predictions: understanding the processes. Pharm Res. 2007;24(5):918-33.

[14] Ruark CD, Hack CE, Robinson PJ, Mahle DA, Gearhart JM. Predicting passive and active tissue:plasma partition coefficients: interindividual and interspecies variability. J Pharm Sci. 2014;103(7):2189-98.

[15] Poulin P, Jones RD, Jones HM, Gibson CR, Rowland M, Chien JY, et al. PHRMA CPCDC initiative on predictive models of human pharmacokinetics, part 5: prediction of plasma concentration-time profiles in human by using the physiologically-based pharmacokinetic modeling approach. J Pharm Sci. 2011;100(10):4127-57.

[16] Kuwahira I, Moue Y, Ohta Y, Mori H, Gonzalez NC. Distribution of pulmonary blood flow in conscious resting rats. Respir Physiol. 1994;97(3):309-21.

[17] Brown RP, Delp MD, Lindstedt SL, Rhomberg LR, Beliles RP. Physiological parameter values for physiologically based pharmacokinetic models. Toxicol Ind Health. 1997;13(4):407-84.

[18] Valentin J. Basic anatomical and physiological data for use in radiological protection: reference values. ICRP Publication 89. Ann ICRP. 2002;32(3-4):1-277.

[19] Williams LR, Leggett RW. Reference values for resting blood flow to organs of man. Clin Phys Physiol Meas. 1989;10(3):187-217.

[20] Nestorov I. Modelling and simulation of variability and uncertainty in toxicokinetics and pharmacokinetics. Toxicol Lett. 2001;120(1-3):411-20.

[21] Davies B, Morris T. Physiological parameters in laboratory animals and humans. Pharm Res. 1993;10(7):1093-5.

[22] Shah DK, Betts AM. Towards a platform PBPK model to characterize the plasma and tissue disposition of monoclonal antibodies in preclinical species and human. J Pharmacokinet Pharmacodyn. 2012;39(1):67-86.

[23] Forsyth RP. Hypothalamic control of the distribution of cardiac output in the unanesthetized rhesus monkey. Circ Res. 1970;26(6):783-94.

[24] Igari Y, Sugiyama Y, Sawada Y, Iga T, Hanano M. Prediction of diazepam disposition in the rat and man by a physiologically based pharmacokinetic model. J Pharmacokinet Biopharm. 1983;11(6):577-93.

[25] Gueorguieva I, Aarons L, Rowland M. Diazepam pharamacokinetics from preclinical to phase I using a Bayesian population physiologically based pharmacokinetic model with informative prior distributions in WinBUGS. J Pharmacokinet Pharmacodyn. 2006;33(5):571-94.

[26] Gueorguieva I, Nestorov IA, Murby S, Gisbert S, Collins B, Dickens K, et al. Development of a whole body physiologically based model to characterise the pharmacokinetics of benzodiazepines. 1: Estimation of rat tissue-plasma partition ratios. J Pharmacokinet Pharmacodyn. 2004;31(4):269-98.

[27] Igari Y, Sugiyama Y, Sawada Y, Iga T, Hanano M. Tissue distribution of 14C-diazepam and its metabolites in rats. Drug Metab Dispos. 1982;10(6):676-9.

[28] Bjorkman S, Fyge A, Qi Z. Determination of the steady state tissue distribution of midazolam in the rat. J Pharm Sci. 1996;85(8):887-9.