# UNIVERSITA' DEGLI STUDI DI PAVIA

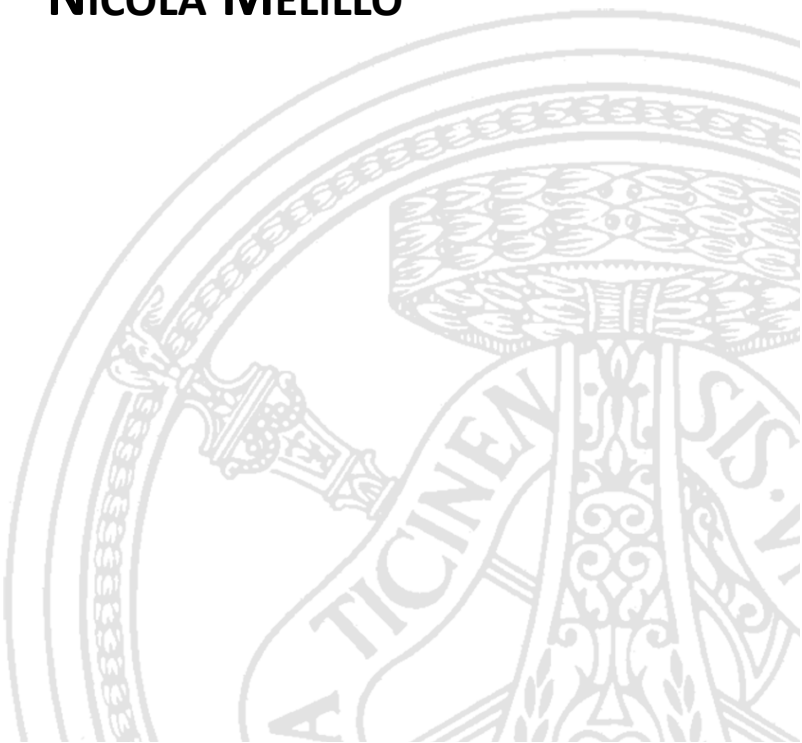
FACOLTA' DI INGEGNERIA
DIPARTIMENTO DI INGEGNERIA INDUSTRIALE E DELL'INFORMAZIONE



DOTTORATO DI RICERCA IN BIOINGEGNERIA E BIOINFORMATICA
XXXII CICLO - 2019


# UNCERTAINTY AND SENSITIVITY ANALYSIS FOR MECHANISTIC MODELS IN PHARMACOMETRICS


PhD Thesis by
**NICOLA MELILLO**


**Advisor:**
**Prof. Paolo Magni**

**PhD Program Chair:**
**Prof. Riccardo Bellazzi**

# Abstract (English)

Mechanistic models combine accepted laws to describe different parts of the system under analysis; their variables and parameters are typically related to specific elements of the system. *Physiologically based pharmacokinetics* (PBPK) models are a class of mechanistic models used in pharmacology. From a mathematical point of view, PBPK models are compartmental models in which each compartment corresponds to a specific organ or tissue. All the compartments are linked by flow rates representing the blood circulation. PBPK models nowadays are routinely applied by pharmaceutical companies from early discovery to late drug development. Moreover, they are used during regulatory review and for informing the drug labels.

However, the lack of appropriate parameter values hampers the PBPK models expansion. Along this line, there is a substantial problem that PBPK models share with almost all the mechanistic models: *the uncertainty in the parameter values*. Many PBPK parameters are derived from *in vitro* experiments, so, they are uncertain due to the measurement errors and to the difference between the *in vitro* and *in vivo* contexts. Other parameters can be predicted by using models, thus, they are uncertain due to the prediction errors. Other parameters instead can be calibrated on given sets of data, therefore, they are subject to the estimation errors. Moreover, when the PBPK models are used in a population context, their parameters are variable as well as uncertain. If the parameters are uncertain or variable,

then, the model outputs (e.g., plasma AUC, $C_{max}$) would be uncertain or variable too. As highlighted by regulatory agencies, there is the need of understanding the confidence on the PBPK modelling results. This could be done with *uncertainty* and *sensitivity analysis*. Uncertainty analysis refers to the quantification of the model output variation. Instead, sensitivity analysis refers to the act of apportioning the output variation to the sources of uncertainty or variability in the model inputs. To cope with this issue, in this work we used *global sensitivity analysis* (GSA) techniques. GSA, with respect to other types of sensitivity analyses, is model independent, detects interaction effects and assess each parameter impact while all the other parameters are allowed to vary as well.

In this thesis, some of the possible applications of uncertainty analysis and GSA for mechanistic models used in the field of pharmacology and, in particular, for PBPK models, are shown. With the example of a PBPK model for the anticancer pro-drug gemcitabine, we showed how GSA can be used to understand what parameters mainly drive the variability of some metrics of interest in a population. Moreover, GSA was applied to absorption models for orally administered compounds. In this case, GSA was used to gain insight into the structure of the model and to guide the choice of parameters that can safely be assumed, estimated or require data generation to allow informed model prediction. We applied uncertainty analysis and GSA also during the development of a model describing the pharmacokinetics (PK) of inhaled compounds. Here we identified two ways of performing GSA, that we called *inter-compounds* and *intra-compound*. Inter-compounds GSA was found to be particularly useful during the phase of model building, in fact it helps in understanding if the model behaves as expected and, if not, it gives useful information in identifying the reasons. Instead, intra-compound GSA is applied when the PBPK model is used for a particular drug. Here, it helps in selecting what parameters should be known with lower degree of uncertainty in order to give more precise predictions. Finally, with the case study of mechanistic models including the correlation between the expression of different enzymes and transporters, we showed how uncertainty analysis and GSA can be used to inform experimental design.

As highlighted by regulatory agencies and practitioners from multiple disciplines, sensitivity analysis is crucial for the quality assessment of model based inference. In this context, the aim of this thesis is to show the utility of uncertainty analysis and GSA in PBPK modelling and simulation, with the view of seeing these techniques routinely applied in model development and use.

# Abstract (Italian)

I modelli meccanicistici sono composti da leggi che descrivono diverse parti del sistema in analisi. Tutte le loro variabili e i loro parametri hanno una diretta corrispondenza con elementi specifici del sistema. I modelli di farmacocinetica basati sulla fisiologia (PBPK, dall'inglese *physiologically based pharmacokinetics*) sono una tipologia di modelli meccanicistici usati in farmacologia. Da un punto di vista matematico, essi sono modelli compartimentali, in cui ogni compartimento rappresenta uno specifico organo o tessuto. I compartimenti di questi modelli sono connessi da flussi rappresentanti la circolazione sanguigna. I modelli PBPK sono comunemente utilizzati durante tutto il processo di ricerca e sviluppo dei farmaci, inoltre, essi sono utilizzati anche da agenzie regolatorie.

Nonostante ciò, la diffusione dei modelli PBPK è fortemente limitata dalla mancanza di valori appropriati per i parametri. In questo senso, c'è una caratteristica sostanziale che i modelli PBPK condividono con la quasi totalità dei modelli meccanicistici: *l'incertezza nei valori dei parametri*. Molti parametri di questi modelli sono derivati da esperimenti *in vitro*, quindi, sono incerti a causa dell'errore sperimentale e della diversità tra il contesto *in vitro* con quello *in vivo*. Alcuni parametri possono essere predetti mediante altri modelli e quindi sono affetti dall'errore di predizione. Altri parametri vengono invece calibrati e quindi, sono soggetti agli errori di stima. Inoltre, se i modelli PBPK sono utilizzati in un contesto

di popolazione, i parametri risultano essere variabili oltre che incerti. Se i parametri in ingresso al modello sono incerti o variabili, allora, le uscite del modello (che possono essere metriche di interesse come AUC e $C_{max}$ plasmatiche) risultano essere incerte o variabili a loro volta. Come sottolineato dalle agenzie regolatorie, c'è il bisogno di stabilire la confidenza che si può attribuire ai risultati dei modelli PBPK. Questo può essere fatto mediante *analisi di incertezza e sensitività*. L'obbiettivo delle analisi di incertezza è quantificare la variazione nelle uscite del modello, mentre quello delle analisi di sensitività è di attribuire la variazione dell'uscita alle varie fonti di incertezza o variabilità negli ingressi. In questo lavoro, abbiamo utilizzato tecniche di analisi di sensitività globale (GSA, dall'inglese *global sensitivity analysis*). Le tecniche globali, a differenza di altri metodi per l'analisi di sensitività, sono modello-indipendenti, permettono la quantificazione degli effetti di interazione e determinano l'impatto di un ingresso sull'uscita del modello considerando tutti gli altri ingressi variabili a loro volta.

In questa tesi vengono mostrati alcuni dei possibili utilizzi dell'analisi dell'incertezza e della GSA per i modelli meccanicistici usati nel campo della farmacologia e, in particolare, per i modelli PBPK. Con l'esempio di un modello PBPK costruito per descrivere la farmacocinetica del profarmaco gemcitabina, abbiamo mostrato l'utilizzo della GSA per l'individuazione dei parametri che maggiormente causano la variabilità di certe metriche di interesse in una data popolazione. Inoltre, la GSA è stata applicata a modelli di assorbimento per farmaci somministrati oralmente. In questo caso, essa è stata utilizzata per aumentare la conoscenza della struttura del modello e per guidare la scelta dei parametri che possono essere fissati ad un certo valore oppure di quelli che necessitano di essere conosciuti con minore incertezza, con l'obbiettivo di aumentare la precisione nelle predizioni del modello. Abbiamo applicato l'analisi dell'incertezza e la GSA anche durante lo sviluppo di un modello PBPK per composti somministrati per via inalatoria. In questo contesto sono stati proposti due diversi approcci alla GSA, che abbiamo nominato *inter-composti* e *intra-composto*. La prima tipologia di analisi risulta essere particolarmente utile durante la fase di costruzione del modello, mentre la seconda quando il modello (o piattaforma) PBPK viene utilizzato per un particolare composto. Infine, abbiamo

mostrato come l'analisi dell'incertezza e la GSA possono essere utilizzate per informare il *design* sperimentale. Questo è stato fatto con l'esempio dello studio di modelli meccanicistici che includono correlazioni tra diversi enzimi e trasportatori

Come sottolineato da agenzie regolatorie e da professionisti di diversi settori, l'analisi di sensitività è uno strumento cruciale per valutare la qualità delle decisioni basate sull'utilizzo di modelli. L'obbiettivo di questa tesi è mostrare l'utilità dell'analisi dell'incertezza e della GSA per i modelli PBPK, con la prospettiva di vedere queste tecniche comunemente utilizzate durante lo sviluppo e l'uso di questi modelli.

# List of Abbreviations

**ADME** Absorption, Distribution, Metabolism and Excretion

**ANOVA** Analysis Of Variance

**AUC** Area Under the Curve

**BCS** Biopharmaceutics Classification System

**CAT** Compartmental Absorption and Transit

**CDA** cytidine deaminase

**CL** Clearance

$C_{max}$ Maximum Concentration

**COPD** Chronic Obstructive Pulmonary Disease

**CYP** Cytochrome P450

**dCK** deoxycytidine kinase

**DDI** Drug-Drug Interaction

**dFdC** 2',2'-difluorodeoxycytidine

**dFdCDP** dFdC diphosphate

**dFdCMP** dFdC monophosphate

**dFdCTP** dFdC triphosphate

**dFdU** 2',2'-difluorodeoxyuridine

$E$ Expected value

**ECOG** Eastern Cooperative Oncology Group

**EFPIA** European Federation of Pharmaceutical Industries and Associations

**EMA** European Medicines Agency

**FAST** Fourier Amplitude Sensitivity Test

**FDA** Food and Drug Administration

**GET** Gastric Emptying Time

**GIGO** Garbage In, Garbage Out

**GSA** Global Sensitivity Analysis

**HDMR** High Dimensional Model Representation

**hCNT1** human Concentrative Nucleoside Transporter 1

**hENT1** human Equilibrative Nucleoside Transporter 1

**iiv** Inter Individual Variability

**IMI** Innovative Medicines Initiative

**iov** Inter Occasion Variability

**IV** Intravenous

**IVIVE** In Vitro to In Vivo Extrapolation

**Lasso** Least Absolute Shrinkage and Selection Operator

**MID3** Model Informed Drug Discovery and Development

**MPPGL** Microsomal Protein Per Gram of Liver

**MT** Mixing Tank

**M&S** Modelling and Simulation

**NDAs** New Drug Applications

**NIH** National Institute of Health

**NME** New Molecular Entity

**OAT** One At a Time

**OATP** Organic-Anion-Transporting Polypeptide

**ODE** Ordinary Differential Equation

**OrBiTo** Oral Biopharmaceutics Tools

**PBPK** Physiologically Based Pharmacokinetic

**PCC** Partial Correlation Coefficient

**PD** Pharmacodynamics

**pdf** Probability Density Function

**PDUFA** Prescription Drug User Fee Act

**PhRMA** Pharmaceutical Research and Manufacturer of America

**PK** Pharmacokinetics

**PRCC** Partial Rank Correlation Coefficient

**QSAR** Quantitative Structure Activity Relationship

**QSP** Quantitative Systems Pharmacology

**R&D** Research and Development

**SITT** Small Intestinal Transit Time

**SRC** Standardized Regression Coefficient

**SRRC** Standardized Rank Regression Coefficient

**UGT** UDP-glucuronosyltransferases

$V$ Variance

# Contents

# List of Figures

# List of Tables

Chapter 1

# Introduction

## 1.1 Drug development

Drug development is a lengthy, complex, costly and risky process. From the discovery of a new molecular entity (NME) to the drug commercialization, it takes approximately 13.5 years [1] and the capitalized research and development (R&D) cost is more than $2.5 billions [2]. Drug development can be divided in several phases, including *drug discovery*, *preclinical phase* (also known as *phase zero*) and *clinical phase*, that is further divided into *phase 1*, *phase 2*, *phase 3* and *phase 4* [3, 4].

In drug discovery, after the understanding of the nature of the disease, new compounds are designed with the aim to reverse or stop the disease effect. During this phase there are thousands of potential compounds. However, after early testing, only few among them will proceed [4]. In the preclinical phase the number of candidate compounds is further reduced and the *lead compounds* that most likely have desirable *pharmacokinetics* (PK), *pharmacodynamics* (PD) and clinical properties in humans are identified[1]. This is done, for example, by using *in vitro* experiments and

[1]Pharmacokinetics can be defined as the kinetics of the processes of drug absorption,

by understanding an appropriate clinical formulation. Moreover, animal models and biomarkers are used to provide early information on efficacy and toxicity. All the information gathered during this phase can be integrated and used for decision-making and the design of the early clinical investigations. This information can be useful for understanding the drug mechanism of action, for the design of new animal experiments and for translating the exposure-response relationships from animals to humans[2] [3].

Clinical phase 1 studies are conducted in a small number of healthy volunteers (from 20 to 200) [4]. The aims of this phase are to inform on tolerability and safety, to find the maximal tolerated drug dose and concentration, to understand metabolism and elimination routes and to provide initial information on population variability [3]. Phase 1 usually lasts several months and the probability of a compound to move to the next phase is approximately 70% [4]. The phase 2 studies are conducted on several hundreds of people with the disease/condition under analysis [4]. The aims of these studies are to confirm the drug efficacy, the acute tolerability, the maximal dose and plasma concentration and the absence of acute safety issues. Moreover, an exploration of dosages regimen is performed. Finally, in order to optimize the dose regimen for patients subgroups, patient and external factors influencing the exposure-response relationship are identified [3]. The length of phase 2 studies goes from several months up to two years and the probability for a compound to move further is roughly 33% [4]. Phase 3 is a confirmatory phase and, with respect to the previous clinical phases, involves a larger number of patients (from 300 to 3000) [4]. Here the aims are to document the clinical safety and efficacy, to characterize the adverse reactions and to understand the sources of variability in the exposure-response. Moreover, dosing regimens can be identified for special populations (e.g., paediatric and elderly patients) [3]. The length of these

distribution and elimination or, more generally, as 'what the body does to the drug'; pharmacodynamics instead includes all the pharmacological actions, beneficial or adverse, of the drug on the body and can be defined as 'what the drug does to the body' [5, 6].

[2]Exposure generally refers to the dose or metrics related with drug concentration, such as AUC or $C_{max}$. Response refers to the drug effects, beneficial or adverse [6].

studies varies approximately from one to four years and the probability to move to the next phase is 25-30% [4]. Phase 4 is after the drug approval and commercialization. Here, the drug safety and efficacy are monitored in a population of several thousands of patients [4].

In each phase there is a risk of interrupting the drug development process (attrition). The majority of failures are reported to be due to efficacy or safety issues in both phase 2 and 3 [7, 8, 9]. In fact, the likelihood of being approved for a drug entering the clinical phase is approximately 11.8% [2].

## 1.2 Mechanistic models in drug development: the need for sensitivity analysis

The use of modelling and simulation (M&S) during drug discovery and development was reported to significantly impact the delivery of new therapies to patients by increasing the decision-making confidence, reducing cycle times and eliminating costs [10, 11, 12, 13, 14, 15]. In fact, many pharmaceutical companies and regulatory agencies are increasing the use of M&S for solving critical problems [13, 14]. In 2016, the term *model informed drug discovery and development* (MID3) was defined by the European Federation of Pharmaceutical Industries and Associations (EFPIA) as a *"quantitative framework for prediction and extrapolation, centered on knowledge and inference generated from integrated models of compound, mechanism and disease level data and aimed at improving the quality and efficiency and cost effectiveness of decision making"* [11]. A recent survey reported that the focus of most pharmaceutical companies in MID3 strategic plans is on PK [12]. Moreover, MID3 was officially recognized to enable an effective and efficient drug development and it was included in the *Prescription Drug User Fee act* (PDUFA) VI of the US Food and Drug Administration (FDA) [16, 17].

M&S is a term that includes a variety of different quantitative approaches. EFPIA highlighted a possible subdivision of M&S in drug development into two classes: *pharmacometrics* and (quantitative) *systems*

*pharmacology* (QSP). A definition of pharmacometrics can be *"the science of developing and applying mathematical and statistical methods to (a) characterize, understand, and predict a drug's pharmacokinetic and pharmacodynamic behavior; (b) quantify uncertainty of information about that behavior; and (c) rationalize data-driven decision making in the drug development process and pharmacotherapy"* [6]. The history of pharmacometrics is long, its first definition was given in 1982, however, some authors claim that its story begins with PK in 1847 [6, 18]. The term QSP was defined in 2011 by the National Institute of Health (NIH) QSP working group as *"an approach to translational medicine that combines computational and experimental methods to elucidate, validate and apply new pharmacological concepts to the development and use of small molecule and biologic drugs determining mechanisms of action of new and existing drugs in preclinical and animal models and patients"* [19]. Other definitions of pharmacometrics and QSP can be found [11, 18, 20].

One might argue that the definitions above are quite general and that there could exist models classified in both the classes. One case could be for example a model describing the human PK by 'translating' *in vitro* and *in vivo* information from the preclinical phase.

A crisp classification that looks more at the technical characteristics of the model, rather than at the application, is the one between *data driven*, or *top-down*, and *mechanistic*, or *law driven*, models (even if in this case the distinction can be blurry sometimes too) [21]. Data driven models have typically a standard structure and they are built to be parsimonious, so, their complexity is 'just enough' to describe the data [22]. The parameters of these models are generally identified from a given set of data (e.g., with least squares techniques). Pharmacometrics is typically a top-down (or data driven) discipline where PK and PD are described by using simple compartmental models and the parameters variability in a given population is described by using statistical models [22, 23]. The usefulness of these empirical approaches was proven in a variety of cases, for example, in supporting the approval for dosing regimens that were not tested, in supporting the dose response in paediatric patients and in dose recommendation for renally impaired patients [11, 17]. However *"no data-driven system model,*

*conceptual or mathematical, will have the level of detail needed to anticipate all the potential consequences of altering the system via pharmacological intervention"* [3] [22]. In fact, their behaviour tend to adhere to the one of the data used for the parameters identification [21].

Mechanistic models combine accepted laws to describe different parts of the system under analysis and all their variables and parameters are typically related to specific elements of the system. For example, the *Noyes Whitney* model [25] can be used to describe the drug dissolution in the intestinal lumen, while the *Michaelis Menten* [26] equation can be used to model the phosphorylation of a drug catalysed by a given kinase in a certain cell. Mechanistic models give higher importance to plausible representation of the system rather than model parsimony. For this reason, they are typically over-parametrized and they contain many more laws than the ones that could be supported by the data. However, this characteristic leads to a greater capacity of mechanistic models to describe the system behaviour in situations not already tested or observed [21] (e.g., PK in special population or in case of interactions between compounds).

When searching for mechanistic models in drug discovery and development, one often find three terms: *systems biology*, QSP and *physiologically based pharmacokinetic* (PBPK) models. Systems biology typically refers to an approach aimed at understanding the biological processes as a whole integrated system rather than isolated parts [27]. QSP can be included in the former definition as well [22]. However, the difference between systems biology and QSP is generally intended as follows: systems biology describe events at the very foundamental biological scale (e.g., metabolic pathways) while QSP aims to provide links between this low order scale and higher order behaviours, such as phenotypes or clinical outcomes. Moreover, QSP aims to fit to all the drug discovery and development stages [22].

PBPK models are (for some authors arguably [28]) a branch of QSP

[3] One might argue that the validity of this sentence is not only restricted to *data driven* models. In fact, *all* the models do not have the level of detail needed to anticipate *all* the potential consequences of altering the system. This because is not theoretically possible for a model to accurately represent *all* the processes happening in a real (open) system [24].

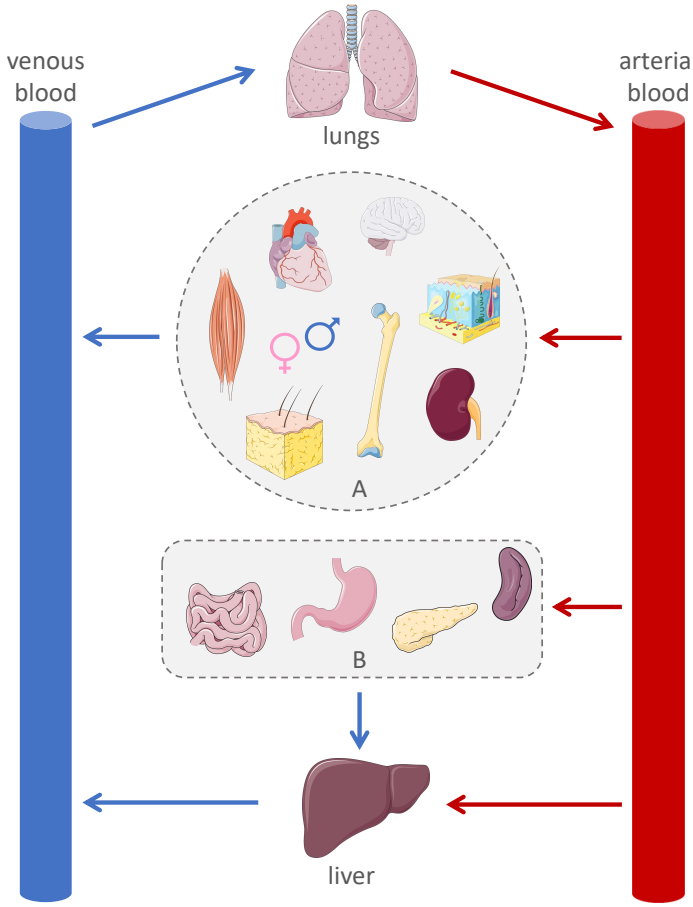



Figure 1.1: General structure of a PBPK model. Each organ correspond to a specific compartment and the red and blue arrows represent arterial and venous blood flows, respectively. The organs and tissues represented in boxes $A$ and $B$ are 'in parallel' with respect to the blood flow, this means that they have separate blood inflows and outflows. Box $A$: adipose tissue, bone, brain, gonads, heart, kidney, muscle and skin. Box $B$: gut, pancreas, spleen and stomach.

[29, 30]. From the mathematical point of view, PBPK models are compartmental models in which each compartment corresponds to a specific organ or tissue. All these compartments are linked by flow rates representing the blood circulation [31, 32]. The generic structure of a PBPK model is shown in figure 1.1. Given the mechanistic nature of PBPK models, all their parameters correspond to specific system properties. Usually, these parameters are classified as *system-specific* and *drug-specific* [29, 31]. Examples of systems-specific parameters are the organ volumes, blood flows and the ones relative to tissue composition (e.g., in terms of water and phospholipids content). Examples of drug-specific parameters can be the tissue to plasma partition coefficients, the intrinsic clearance, the gut-wall permeability or the intrinsic solubility [31].

Although PBPK models are classified as QSP, they have a longer history. Some authors suggest that the origin of PBPK modelling was in 1937 with the work of *Teorell* [31, 33]. He recognize that events occurring in one part of the body are influenced by, and in turn influence, events that occur in other parts. So, the body behaves as an integrated system [31, 33]. From 1960s to the early 1970s PBPK models progresses significantly, although the specific term 'PBPK' was not yet used [31]. As it can be seen from figure 1.2, from 1974 to nowadays there was a continuous increase of publications regarding/mentioning PBPK models. Approximately until the beginning of 2000, the use of these models was mostly restricted to environmental pollutants and toxicants, where strong constraints for data generation in humans exist [34]. Until then, the interest of pharmaceutical companies and regulatory agencies on PBPK modelling was low and most of the applications for pharmaceuticals was performed in academia [31]. However, from the early 2000s the situation changed and an exponential growth of their use in drug research and development occurred [34, 35]. Now, PBPK modelling is routinely applied from early discovery to late drug development [36, 37]. This situation was helped by several factors [29, 31, 34, 35, 38], including: 1) the presence of adequate *in vitro* systems informing the *in vivo* processes of absorption, distribution, metabolism and excretion (ADME); the availability of models predicting plasma-tissue partition coefficients, thus allowing an estimation of drug

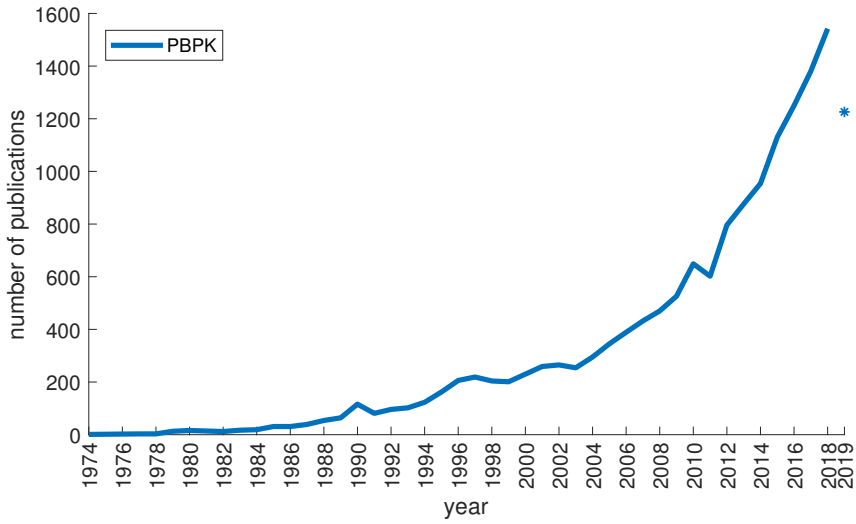



Figure 1.2: Trend in PBPK publication from 1974 to nowadays. The following research was performed on Scopus on 5 September 2019: `ALL ( "PBPK" OR "physiologically based pharmacokinetics")`.

tissue distribution without the need of performing expensive experiments [39, 40, 41, 42]; 3) the development of easy to use commercial tools for PBPK modelling, such as *Simcyp* (`https://www.certara.com/software/physiologically-based-pharmacokinetic-modeling-and-simulation/simcyp-simulator`), *Gastroplus* (`https://www.simulations-plus.com/software/gastroplus/`) and *PK-Sim* (`http://www.systems-biology.com/products/pk-sim/`).

PBPK models are used during the process of drug discovery and development in a variety of different cases, for example to perform *in vitro* to *in vivo* extrapolation (IVIVE) and to predict drug PK in preclinical species only by using *in vitro* data. They are used to predict the human PK before performing first in human studies, to predict the effect of drug-drug interactions (DDI), the food effect for orally administered compounds and the effect of age (especially for children younger than two years old). Moreover, they can be used in a Monte Carlo framework to explore the PK variability

in a population much numerous than the one that can be observed and thus, they can predict the characteristics leading to an extreme PK. Comprehensive reviews of PBPK applications in pharmaceutical companies can be found in [11, 31, 36, 37, 38, 43]. During early drug discovery, PBPK are typically used in a pure *bottom-up* fashion. In this case, the parameters are estimated by using *in silico* (e.g., quantitative structure activity relationship, QSAR [44]) and *in vitro* models and PBPK models are used to predict the PK in humans or preclinical species [37]. PBPK models are used also with a *middle-out* approach, where some of the model parameters are calibrated on certain data (e.g., clinical), with the aim of predicting the system behaviour in unobserved situations [30]. Here 'reduced' version of the models, with 'lumped' compartments can be used [31, 45, 46].

PBPK modelling has shown its utility during regulatory review [47, 48] and for informing the drug labels [37, 35]. In these contexts, PBPK models are used in decision making and to address questions such as how extrinsic factors (e.g., DDI) or intrinsic factors (e.g., age, genetics) could influence the drug exposure-response [47]. From 2008 to 2017 the FDA's Office of Clinical Pharmacology received 94 new drug applications (NDAs) including the use of PBPK models. Here, the aim of PBPK analyses was mainly to address DDI, followed by the application for paediatric population [49], as shown in figure 1.3. Of recent significance, EMA and FDA published guidelines for PBPK modelling and simulations reporting [50, 51]. Concerning QSP models (excluding PBPK), they are generally present as support of large evidentiary packages [28].

Referring to PBPK models, Poggesi and co-workers in 2014 posed the following question: *"is it all a success story?"*. In these few years the answer is probably not changed: *"no, it is not"*.

A study inspired by the Pharmaceutical Research and Manufacturer of America (PhRMA) initiative reported that the accuracy of PBPK models to predict the intra-venous (IV) PK was relatively good, while lower performances were observed for oral administration [52]. Similar results were obtained in the Oral Biopharmaceutics Tools (OrBiTo) project, where a large scale evaluation of PBPK models for oral drug absorption showed high variability in the predictive performances [53, 54, 55]. Moreover, many systems-

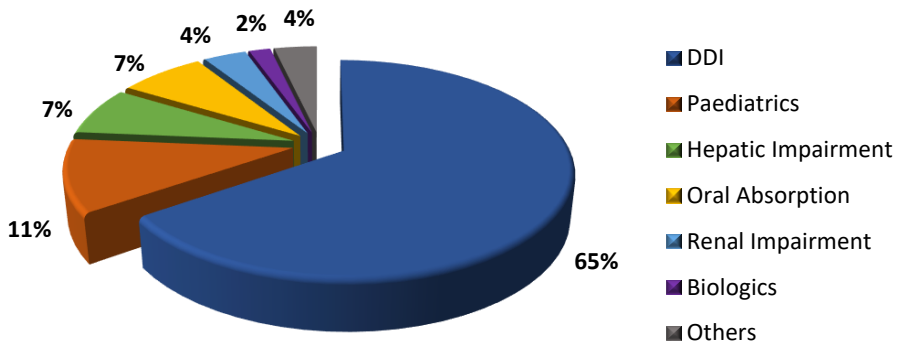



Figure 1.3: Aims of PBPK analyses in the 94 NDAs submitted to the FDA's Office of Clinical Pharmacology from 2008 to 2017 [49].

related parameters are still lacking, such as enzymes and transporters abundances in specific tissues, their population variability, inter-correlation and their relation with demographics [29, 32, 56]. Jamei reported that the lack of appropriate systems parameters is probably the biggest challenge in the PBPK models expansion [35].

Along this line, there is a substantial problem that PBPK models share with almost all the mechanistic models (or, in general, almost all the mathematical models): *the uncertainty in the parameter values* [21].

Many drug-dependent parameters are derived from *in vitro* experiments, so, they have uncertainties associated with the measurement errors and to the difference between the *in vitro* and *in vivo* contexts. Other parameters can be predicted by using models, thus, they are uncertain due to the prediction errors [57]. Other parameters again can be identified from a given set of data, therefore, they are subject to the estimation errors. For example, the human jejunal permeability is a parameter used in PBPK models to estimate the intestinal drug absorption. This value can be obtained through a linear regression by using as independent variable the apparent permeability coefficient ($P_{app}$), that in turn is derived by *in vitro* experi-

ments involving the Caco-2 cell line and the PAMPA assay [43, 58]. These measurements can have a significant degree of uncertainty and a substantial inter-laboratories variability was observed for the $P_{app}$ values derived from Caco-2 experiments [43, 59]. Therefore, the human jejunal permeability is an uncertain parameter and its uncertainty is driven by both measurement and prediction errors. A similar reasoning could be done for many other parameters. Moreover, when the PBPK models are used in a population context, their parameters are variable as well as uncertain.

If the model parameters are uncertain or variable, then, the model outputs (e.g., the plasma concentration time curve) would be uncertain or variable too. As highlighted in the recent EMA guideline, there is the need of understanding the confidence on the PBPK modelling results [50]. For this reason, the effect of uncertain and variable parameters should be assessed by understanding *how much* is the extent of the model outputs variation and *how much* this variation is apportioned to the various sources of uncertainty and variability in the model input parameters [21, 50]. This could be done with *uncertainty* and *sensitivity analyses*.

## 1.3 Introduction to global sensitivity analysis (GSA)

Sensitivity analysis can be defined as *"the study of how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input"* [60]. Uncertainty analysis refers instead to the quantification of the model output uncertainty, without specifying the relation with the input variation and ideally, it precedes sensitivity analysis [21]. For the purposes of these analyses, the 'model outputs' are *scalar* quantities of interest obtained after the model evaluation, while the 'model inputs' are everything that could produce a variation in the model outputs [21, 61]. Examples of model outputs in the field of pharmacology are plasma $AUC$ and $C_{max}$ of a certain compound, the drug concentration in the target tissue, the probability of survival after a given treatment, the nadir value of neutrophil concentration and so on. Examples of model inputs are the model parameters, such as the clearance, the

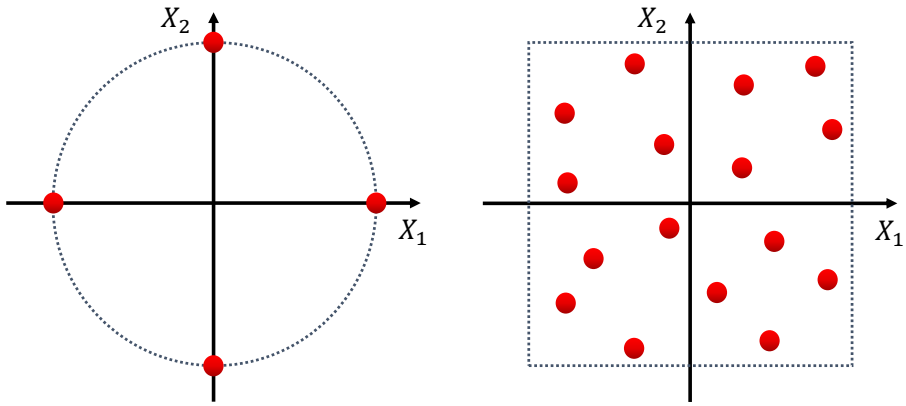



Figure 1.4: Univariate (OAT) versus multivariate sampling design. With the hypothesis of all the factors uniformly distributed between 0 and 1 (in this example, two factors, $X_1$ and $X_2$), it is easy to observe that all the points of an OAT design lay on the surface of a k-dimensional hypersphere of radius equal to 1/2 (a circle in two dimension), while all the points of a multivariate design better explore the whole space. This image was inspired by [63].

hepatic enzymes and transporters abundance, the jejunal permeability or the radius of the particle size of the formulation. Model inputs can be also the initial conditions, such as the dose, for example for inhaled compounds. Finally, the parameters relative to the numerical solution of the models can be considered as inputs as well [21, 62].

In the standard sensitivity analysis setup, the model is seen as an input-output map, as in equation 1.1.

$$Y = f(X_1, X_2, ..., X_k) \tag{1.1}$$

$Y$ is the model output, $X_i$, $i = 1, ..., k$, are the $k$ model inputs (that are also called factors) and $f$ is the model. In general, it is possible to group sensitivity analysis methods into two classes: *local* or *global*. Local sensitivity analysis is performed when output uncertainty is obtained thorough small variation around a nominal value. Instead, global sensitivity analysis (GSA) deals with the presence of uncertain input factors. In this case, a probability distribution is assigned to each model input and a multivariate variation of the parameters is performed. Here, $Y$ is obtained by model

evaluation after sampling from the joint probability distribution of the inputs. [21, 61]. For simplicity sake, from now on we would consider the input factors uniformly distributed in an unit $k$-dimensional hypercube (as it is explained in chapter 2, this is not a limitation).

A first mathematical definition of sensitivity is the derivative of $Y$ with respect to $X_i$, computed at a given point $\mathbf{x}^*$ of the input space, as in equation 1.3.

$$S^p_{X_i} = \frac{\partial Y}{\partial X_i}\bigg|_{\mathbf{X}=\mathbf{x}^*} \tag{1.2}$$

One characteristic of this method is that it is informative only at the point $\mathbf{x}^*$ in which the derivative is computed. For this reason, this method can be seen as a local method. This could be a limitation if the model inputs are uncertain and thus could vary within a predefined range [21].

Most of the published sensitivity analyses use te so called *one at a time* (OAT) approaches [64]. With these methods, the sensitivity ranking is obtained by increasing (or decreasing) each input factor of a given percentage and then, by quantifying the model output change. One possible index can be the following, known as *sensitivity index* [65].

$$SI_{X_i} = \frac{Y_{max,X_i} - Y_{min,X_i}}{Y_{max,X_i}} \tag{1.3}$$

$Y_{max,X_i}$ and $Y_{min,X_i}$ are the maximum and minimum output values obtained by varying $X_i$ over its range, respectively. Although the OAT are the most commonly used methods to perform a sensitivity analysis, they have two strong limitations [66]: 1) inefficient exploration of the parameter space; 2) impossibility to detect interactions among factors. These two limitations are strictly related to each other and they are both caused by the univariate variation of the input factors.
Concerning the first limitation, from figure 1.4 it is clear that an OAT design does not efficiently explore the parameter space, conversely to a multivariate variation. In 2010 Saltelli and Annoni introduced a geometric proof of the OAT inefficiency [66]. All the points of an OAT design are *by construction* included in a $k$-dimensional hypersphere (although they can

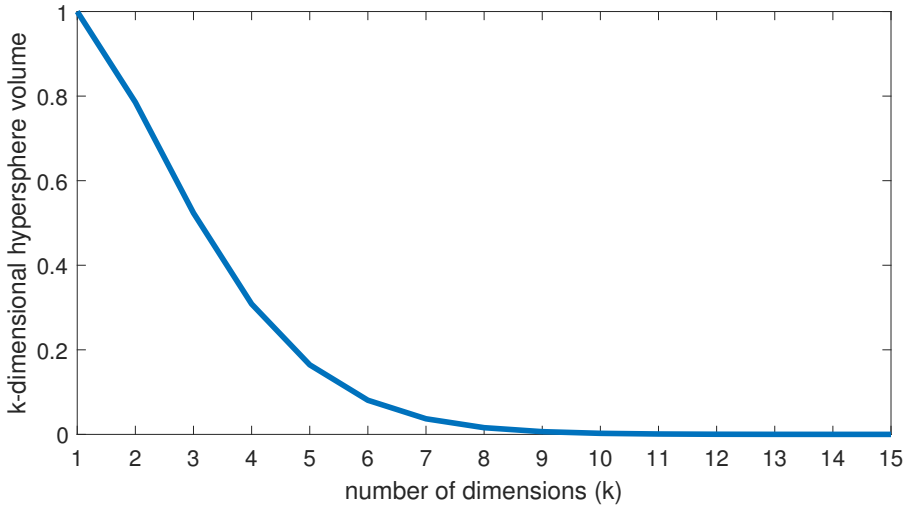



Figure 1.5: Geometric proof of OAT inefficiency in the exploration of the factor space. All the points of an OAT design are internal to a k-dimensional hypersphere (in particular they lay on the surface), while all the points of a multivariate design are internal to a k-dimensional hypercube, of volume equal to 1. The volume of a k-dimensional hypersphere is equal to $V(k) = \frac{\pi^{k/2}}{\Gamma(0.5 \cdot k+1)} (0.5)^k$, with $\Gamma$ the 'gamma' function (`gamma` command in MATLAB) and $k$ the number of dimensions. It is possible to observe that the volume of the hypersphere rapidly drops to zero as $k$ increases.

be included in a much smaller volume as well). As it can be seen from figure 1.5, as the number of dimensions increases, the ratio of the hypersphere volume with the whole input factors space volume (i.e., the volume of an unit $k$-dimensional hypercube, that is equal to 1), drops rapidly to zero. Thus, in a high dimensional space, OAT is non-exploratory and so, it is equivalent to a local method.

Concerning the second limitation, interaction effects happen when the impact of some parameter on a given model output depends on the value of other parameters, as exemplified in figure 1.6. In sensitivity analysis it is crucial to detect the interaction effects, because they could be the most important in determining the model output variation [67]. By performing an univariate variation, for OAT methods is impossible to detect these effects

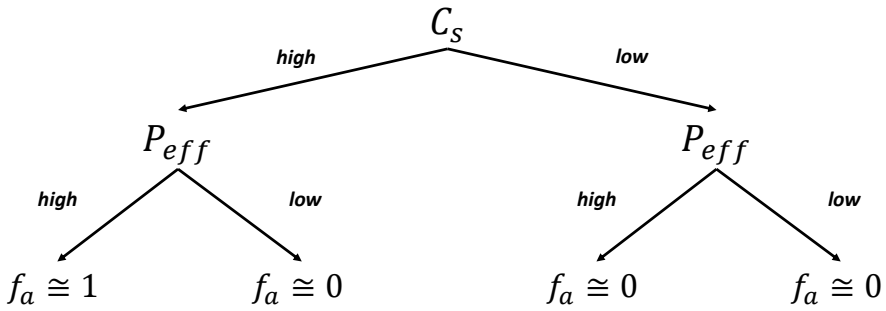



Figure 1.6: Example of interactions among factors. Let us consider a simple model describing the fraction absorbed ($f_a$) of a given compound after oral administration. This model has two parameters, drug solubility ($C_S$) and permeability across the gut wall layer ($P_{eff}$). If $C_S$ is supposed to be high, then the variation of $P_{eff}$ (high or low), will control the extent of $f_a$. For high permeabilities we would have almost a complete absorption, while no absorption occurs for low permeabilities. If $C_S$ is instead supposed to be (very) low, then no matter how is the value of $P_{eff}$, $f_a$ would always be low. Thus, the impact of $P_{eff}$ variation on $f_a$ *depends* on $C_S$ values. Interaction effects happen when the impact of the variation of some parameters ($P_{eff}$) on a given model output ($f_a$) depends on the value of other parameters ($C_S$).

[66].
Moreover, OAT methods are inadequate for the uncertainty analysis as well. In fact, for all the reasons explained above, they are not able to appropriately identify the output distributions [66].

There is the need of methods for sensitivity analysis that consider the input factors in their whole range of uncertainty or variability and that perform a multivariate variation. These are the characteristics of GSA [21]. A variety of methods exists to perform GSA and some of them are reported in chapter 2. Probably, the most simple and *qualitative* method is to look at the *scatter plots* of the model output versus each input factor [21]. Although its simplicity, this method is really informative. In fact, it is possible to appreciate the type of dependency of $Y$ on $X_i$ and the presence or absence of interaction effects.

Concerning *quantitative* methods, the properties of good GSA indices are listed below [68].

- *Cope with the influence of scale and shape.* The effect of the input factors on the model output should depend on their range of variation and on the shape of their probability density function (pdf).

- *Include multidimensional averaging.* This property refers to the fact that the effect of each factor should be evaluated when all the other inputs are allowed to vary as well. By doing this, it is possible to detect the interaction effects.

- *Be model independent.* The GSA method should be suitable for all the models, regardless of their linearity or additivity.

- *Be able to treat grouped factors as if they were single factor.* This property refers to the simplification of the sensitivity analysis results, especially in presence of a large number of factors.

In the works reported in this thesis, we generally used the *variance based* method [21, 69]. In variance based GSA two sensitivity indices, which are called *main effect* ($S_i$) and *total effect* ($S_{Ti}$), are derived from the decomposition of the variance of $Y$ ($V(Y)$). $S_i$ is the portion of $V(Y)$ explained by the factor $X_i$ taken singularly, while $S_{Ti}$ is equal to $S_i$ plus all the interaction effects involving $X_i$.

Dependently on the aims of GSA, it is possible to identify several settings for the analysis [21].

- *Factor prioritization.* Here the aim is to detect and rank the factors that if fixed cause the greatest reduction in the output variance. This setting could be useful to understand what parameters need to be better known to reduce the uncertainty in the model predictions.

- *Factor fixing.* Here the aim is to identify the factors that if left free to vary do not impact the model output variation. This setting could be useful to simplify the models.

- *Variance cutting.* This setting is used when one wants to reduce the model output variance below a certain threshold. This could

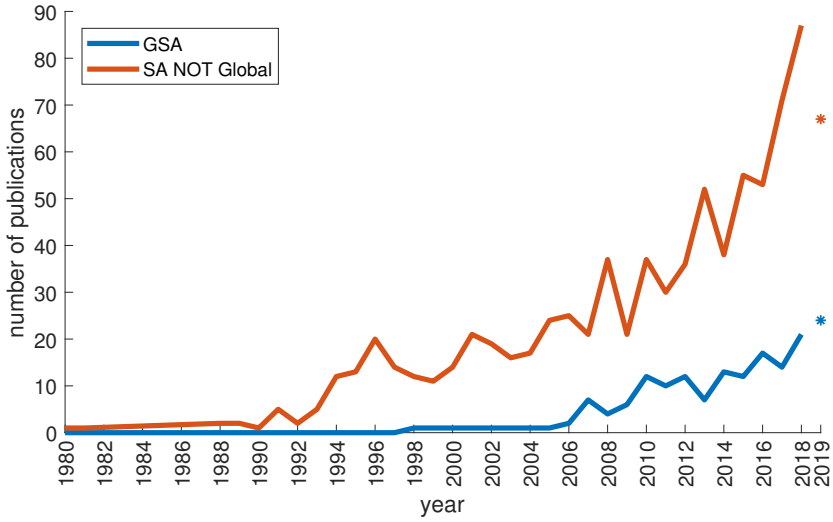



Figure 1.7: Trend in GSA publication for PBPK models from 1980 to nowadays. The following researches were performed on Scopus on 3 September 2019: `ALL ( ( "PBPK" OR "physiologically based pharmacokinetics" ) AND "global sensitivity analysis" )`; `ALL ( ( "PBPK" OR "physiologically based pharmacokinetics" ) AND "sensitivity analysis" AND NOT "global sensitivity analysis" )`. In the figure label, 'SA' stands for sensitivity analysis.

be useful to select the characteristics of a population in order to be homogeneous in terms, for example, of drug exposure.

- *Factor mapping.* This setting is used to identify what factors cause the model output to be or not to be in a certain group, for example, over or below a threshold. This could be useful to understand what parameters or covariates lead the plasma concentration to be above the toxicity level. For this setting, Monte Carlo filtering methods can be used [70].

Many authors advocated the use of sensitivity analysis for PBPK models [31, 37, 43, 48, 50, 51] and some authors advocated and used global approaches [46, 71, 72, 73, 74, 75, 76, 77]. However, as it can be seen in figure 1.7, the number of PBPK publications mentioning 'global sensitiv-

ity analysis' is still much lower than the ones mentioning just 'sensitivity analysis' without the word 'global'. Moreover, by comparing figure 1.1 and figure 1.7 it is possible to observe that only few publication regarding PBPK models include the words 'sensitivity analysis'.

## 1.4 Thesis outline

During these three years of PhD studentship I applied uncertainty and sensitivity analyses in a variety of different situations, mainly involving PBPK modelling and simulation. Through these examples, the aim of the thesis is to show the utility of uncertainty analysis and GSA for mechanistic models in the field of pharmacology.

In chapter 2, I reported a brief review of the main techniques used for GSA. Then, I applied them on two simple benchmark models.

In chapter 3, I described my first application of GSA on PBPK models, that was done in the context of a collaboration with the *Pharmacometrics and Systems Pharmacology* group of the *Universidad de Navarra*. The objective of that work was to build a mechanistic model describing the PK of the anticancer pro-drug gemcitabine and to predict the *in vivo* formation of its active metabolite in patients tumour tissue. Here GSA was used to understand what are the physiological and genetic characteristics that lead to different active metabolite exposures in the target tissue and thus, that drive the treatment outcome.

In chapter 4, I described one of the activities that I performed during my abroad period at the *University of Manchester*. Here the context is the one of the outcomes of the OrBiTo project. One of the various objectives of OrBiTo was to perform a large scale evaluation of PBPK models for oral drug absorption, to identify strengths and weaknesses of these models. However, the results of the analysis showed high variability in the performances. Here we used GSA to improve the understanding of PBPK absorption models by identifying what are the parameters that mainly drive the 'between-drugs' variability of the model predictions. Later, we called this type of analysis '*inter-compounds* GSA'. With the GSA results, it is possible to guide the

choice of the parameters that can safely be assumed or that require data generation in order to allow informed model predictions.

In chapter 5, I reported the application of GSA during the development of a PBPK model for inhaled compounds in rats. This was done in the context of a collaboration with the company *Chiesi Farmaceutici*. Here we identified two way of performing GSA, that we called *inter-compounds* and *intra-compound*. *Inter-compounds* GSA considers the between-drug parameters variability and it was found to be particularly useful during the model building phase. *Intra-compound* GSA mostly considers the uncertainty associated with the parameters relative to a particular drug, thus, it was found to be useful when the PBPK model is used to describe the PK for a given compound.

In chapter 6, I described another activity that I performed at the *University of Manchester*. Here we used uncertainty analyses to characterize the effect of plausible correlations between enzymes and transporters involved in drug metabolism and disposition on the outputs of mechanistic models. Moreover, we quantified the potential impact of these correlations on GSA. Here uncertainty and sensitivity analyses helped in highlighting what correlations are of potential interest and therefore, these analyses can be useful for informing experimental design.

Finally, the overall conclusions are reported in chapter 7.

Chapter 2

# Methods for Global Sensitivity Analysis

In this chapter we are firstly going to explain some of the techniques used for GSA, in particular the methods based on linear models, the variance based method and one example of moment-independent methods. For a more complete review see for example [21, 61, 65, 78, 79]. Then, we report the application of these techniques on two benchmark models: one simple linear model and a widely used model in the pharmacometrics filed, that is the one describing the course of neutropenia in patients treated with paclitaxel [80, 81, 82].

## 2.1 GSA methods - theory

All the GSA techniques here investigated share, to a certain extent, a similar framework. The model is generally considered as a black box:

$$Y = f(\mathbf{X}) = f(X_1, X_2, ..., X_k), \tag{2.1}$$

where $Y$ is a *scalar* model output, $X_i$, $i = 1, ..., k$, are the $k$ *scalar* input parameters (that, in GSA literature, are called factors), $\mathbf{X}$ is a vector con-

Table 2.1: Notation

| **Name** | Description |
|---|---|
| $\mathbf{X}$ | vector of model input factors |
| $X_i$ | the $i$-th element of $\mathbf{X}$ |
| $\mathbf{X}_{\sim i}$ | all the elements of $\mathbf{X}$ except $X_i$ |
| $\mathbf{x}_i$ | a vector of $n$ samples extracted from the pdf of $X_i$ |
| $x_i^j$ | the $j$-th element of $\mathbf{x}_i$ |
| $Y$ | model output |
| $\mathbf{y}$ | vector of outputs obtained after model evaluation |
| $y^j$ | the $j$-th element of $\mathbf{y}$ |
| $E(z)$ | expected value of the random variable $z$ |
| $E_\xi(z)$ | expected value of $z$ over $\xi$ |
| $V(z)$ | variance of $z$ |
| $V_\xi(z)$ | variance of $z$ over $\xi$ |
| $\Omega_z$ | support of $z$ |

taining all the $X_i$ and $f$ is the input-output relationship, that would be our model [21]. The notation used in this chapter is shown in table 2.1.

GSA methods deal with the presence of uncertainty in the model input factors, so, both $X_i$, $i = 1, ..., k$, and $Y$ are considered random variables. All the $X_i$ are generally considered independent (although, in certain cases, this hypothesis can be relaxed) and uniformly distributed between 0 and 1. So, the support of $\mathbf{X}$ would be

$$\Omega_{\mathbf{X}} = \{\mathbf{X} \,|\, 0 \leq X_i \leq 1;\, i = 1, ..., k\}. \tag{2.2}$$

To perform GSA, first samples must be extracted from the distributions of the parameters. The extractions can be designed by using pseudo-random strategies (e.g., `rand` function in MATLAB), *latin hypercube sampling method* or quasi-random strategies [21, 83]. The latter two methods would explore the parameters space better with respect to the former. All these methods extract samples from the unit hypercube, so, the inverse cumulative distribution function can be used to convert each sample from the uniform distribution to the desired probability density function (pdf). Then, for each of the parameters sets, the model is evaluated and the outputs of interest are computed. Finally, sensitivity indices are calculated for each of the considered model output.

### 2.1.1 Methods based on linear models

Some global sensitivity measures can be derived from the fitting of linear models, considering the factors ($\mathbf{X}$) as the independent variables and the model output ($Y$) as the dependent variable. Sensitivity indices are calculated on a dataset of $n$ samples of the model input parameters, $\mathbf{x}_i$, $i = 1, ..., k$, and the relative model output $\mathbf{y}$. For a review of these methods see, for example, [65, 78].
The main indices are:

- *Pearson correlation coefficient*:

$$\rho(\mathbf{x}_i, \mathbf{y}) = \frac{\sum_{j=1}^{n}(x_i^j - E(\mathbf{x}_i))(y^j - E(\mathbf{y}))}{\sigma_{\mathbf{x}_i}\,\sigma_{\mathbf{y}}} \tag{2.3}$$

  $\sigma_{\mathbf{x}_i}$ and $\sigma_{\mathbf{y}}$ are the standard deviation of $\mathbf{x}_i$ and $\mathbf{y}$. This index ranges from -1 to 1 and it is equal to 0 when no *linear* correlation exists between $\mathbf{x}_i$ and $\mathbf{y}$.

- *Partial correlation coefficient* (PCC):

$$PCC_i = \rho(\mathbf{x}_i - \hat{\mathbf{x}}_i, \mathbf{y} - \hat{\mathbf{y}}_{\sim i}) \tag{2.4}$$

  $\rho(a, b)$ is the Pearson correlation coefficient of $a$ and $b$. $\hat{\mathbf{x}}_i$ is the value of $X_i$ predicted by using a linear model with all the other factors ($\mathbf{x}_{\sim i}$) as independent variables. Similarly, $\hat{\mathbf{y}}_{\sim i}$ is the model output predicted by using a linear model with $\mathbf{x}_{\sim i}$ as independent variables. $PCC$ eliminates linear correlations that may exist between $\mathbf{x}_i$ and $\mathbf{x}_{\sim i}$ and between $\mathbf{y}$ and $\mathbf{x}_{\sim i}$. So, this index can be useful in case of correlation between the factors. In case of no correlation, the $PCC$ sensitivity ranking would not differ from the Pearson correlation coefficient ranking [65].

- *Standardized regression coefficient* (SRC):

$$SRC_i = \beta_i \sqrt{\frac{V(X_i)}{V(Y)}} \tag{2.5}$$

$\beta_i$ is the linear regression coefficient of $X_i$. $SRC_i^2$ is the portion of the variance of $Y$, $V(Y)$, explained by the variation of $X_i$, if the model is linear [21, 78]. $SRC$ can be viewed as a kind of variance decomposition method.

The strength of the linear assumption can be tested with standard statistical techniques such as the $R^2$. If the relationship between the factors and the output is not linear, but it is monotonic, these indices can still be used with the ranked transformed data, obtaining: *Spearman correlation coefficient*, *partial rank correlation coefficient* (PRCC) and *standardized rank regression coefficient* (SRRC).

If the objective of the GSA is the factor fixing, then, linear regression techniques with regularization, such as the *lasso* (least absolute shrinkage and selection operator) [84] can be used. The lasso minimize the sum of residual squares with a $L_1$ penality. This can be written as in equation 2.6, or, in the *Lagrangian form*, as in equation 2.7 [85],

$$\begin{aligned} \hat{\beta}^{lasso} = \arg\min_{\beta} \sum_{j=1}^{n} \left( y^j - \beta_0 - \sum_{i=1}^{k} x_i^j \, \beta_i \right)^2 \\ \text{subject to} \sum_{i=1}^{k} |\beta_i| \leq t \end{aligned} \tag{2.6}$$

$$\hat{\beta}^{lasso} = \arg\min_{\beta} \left\{ \frac{1}{2} \sum_{j=1}^{n} \left( y^j - \beta_0 - \sum_{i=1}^{k} x_i^j \, \beta_i \right)^2 + \lambda \sum_{i=1}^{k} |\beta_i| \right\} \tag{2.7}$$

where $t$ and $\lambda$ are the regularization parameters. One characteristics of the $L_1$ penality it is that, with $t$ sufficiently small (or $\lambda$ sufficiently big), it causes some of the $\beta$s to be exactly equal to 0. Thus, lasso makes a kind of continuous feature selection [85]. The best $\lambda$ (or $t$), can be chosen such as it minimize the *cross validation* error. Another commonly used rule is the one called 'one-standard error', that choose $\lambda$ such as its error is one-standard deviation higher than the best model (in the direction of higher regularization) [85].

All these methods based on linear models are easy to use (they are all implemented in MATLAB and in libraries of R and Python) and the sensitivity indices are easy to interpret. However, the linearity (or monotonicity) assumption sometimes can be too strong. Moreover, to estimate the interaction effects all the possible combination of parameters (e.g., $X_1 \cdot X_2$ or $\prod_{i=1}^{k} X_i$) should be included as regressors in the linear model. This can be challenging, especially with a high number of factors and with non-linear or non-monotonic dependencies between $Y$ and $\mathbf{X}$.

### 2.1.2 Variance based method

#### Theory

Variance based sensitivity indices can be derived from the functional decomposition, known as high dimensional model representation (HDMR), presented by Sobol in 1993 [69]. Let $Y = f(\mathbf{X})$ be a function defined in a $k$ dimensional unit hypercube $\Omega_{\mathbf{X}}$ (equation 2.2), the HDMR of $f(\mathbf{X})$ is defined as follows:

$$f(\mathbf{X}) = f_0 + \sum_{i=1}^{k} f_i(X_i) + \sum_{i} \sum_{j>i} f_{ij}(X_i, X_j) + ... + f_{1,...,k}(X_1, ..., X_k), \quad (2.8)$$

where $f_i(X_i)$ are first order functions, $f_{ij}(X_i, X_j)$ are second order functions and so on. If $f_0$ is constant and for each of the summands $f_{i_1,...,i_s}$, $1 \leq s \leq k$, is valid

$$\int_{\Omega_{X_j}} f_{i_1,...,i_s} dX_j = 0, \ 1 \leq j \leq s, \quad (2.9)$$

that is, the integral of each summand taken over one of 'its' variables is equal to zero, then, the HDMR has the following properties [69]:

- all the HDMR summands are orthogonal,

$$\int_{\Omega_{\mathbf{X}}} f_{i_1,...,i_s} f_{j_1,...,j_t} d\mathbf{X} = 0, \ (i_1, ..., i_s) \not\equiv (j_1, ..., j_t); \quad (2.10)$$

- the HDMR decomposition is unique for any function $f(\mathbf{X})$ integrable in $\Omega_{\mathbf{X}}$.

From equation 2.8, considering the summands orthogonality (equation 2.10) and considering that all the $X_i$ are uniformly distributed random variables (thus, their pdf $p(X_i)$ is equal to one in $\Omega_{X_i}$), it is possible to uniquely express all the summands $f_{i_1,\dots,i_s}$ with different integrals of $f$.
$f_0$ is obtained as in equation 2.11 by integrating $f(\mathbf{X})$ over $\mathbf{X}$ and it results equal to the expected value of $Y$, $E(Y)$ [69].

$$f_0 = \int_{\Omega_{\mathbf{X}}} f(\mathbf{X}) d\mathbf{X} = E(Y) \tag{2.11}$$

It is possible to obtain the definition of the first order term $f_i$ by integrating $f$ over $\mathbf{X}_{\sim i}$, as in equation 2.12.

$$\int_{\Omega_{\mathbf{X}_{\sim i}}} f(\mathbf{X}) d\mathbf{X}_{\sim i} = f_0 + f_i$$

$$f_i = \int_{\Omega_{\mathbf{X}_{\sim i}}} f(\mathbf{X}) d\mathbf{X}_{\sim i} - f_0 = E_{\mathbf{X}_{\sim \mathbf{i}}}(Y \mid X_i) - f_0 \tag{2.12}$$

Similarly, it is possible to define the second order terms (equation 2.13) and so on.

$$\int_{\Omega_{\mathbf{X}_{\sim ij}}} f(\mathbf{X}) d\mathbf{X}_{\sim ij} = f_0 + f_i + f_j + f_{ij}$$

$$f_{ij} = \int_{\Omega_{\mathbf{X}_{\sim ij}}} f(\mathbf{X}) d\mathbf{X}_{\sim ij} - f_0 - f_i - f_j = E_{\mathbf{X}_{\sim \mathbf{ij}}}(Y \mid X_i, X_j) - f_0 - f_i - f_j \tag{2.13}$$

If $f(\mathbf{X}) \in \mathbb{L}_2$, we subtract $f_0$ from both the sides of equation 2.8, we

square them and we integrate over $\mathbf{X}$, we can obtain the following relation.

$$\int_{\Omega_{\mathbf{X}}} \big(f(\mathbf{X}) - f_0\big)^2 d\mathbf{X} = \int_{\Omega_{\mathbf{X}}} \Big(\sum_{i=1}^{k} f_i(X_i) + \sum_i \sum_{j>i} f_{ij}(X_i, X_j) + ... + f_{1,...,k}(X_1, ..., X_k)\Big)^2 d\mathbf{X} \tag{2.14}$$

The left term of equation 2.14 is equal to the variance of $Y$, $V(Y)$. Thanks to the summands orthogonality, we can obtain the relation in equation 2.15 [69].

$$V(Y) = \sum_{i=1}^{k} V_i + \sum_i \sum_{j>i} V_{ij} + ... + V_{1,...,k} \tag{2.15}$$

The functional decomposition of the variance presented in equation 2.15 is also known as *functional ANOVA* [67, 86]. $V_i = V_{X_i}(E_{\mathbf{X}_{\sim \mathbf{i}}}(Y \mid X_i))$ is the first order term and it is the portion of $V(Y)$ explained by the variation of each $X_i$ taken singularly [87]. $V_{ij} = V_{X_i,X_j}(E_{\mathbf{X}_{\sim \mathbf{ij}}}(Y \mid X_i, X_j)) - V_i - V_j$ is the second order term and it is the portion of $V(Y)$ explained by the interactions between $X_i$ and $X_j$. Similarly, it is possible to define all the higher order interaction terms.
*Variance based* or *Sobol's sensitivity indices* can be defined from 2.15 as in equation 2.16 [69].

$$S_{i_1,...,i_s} = \frac{V_{i_1,...,i_s}}{V(Y)} \tag{2.16}$$

There are $2^k - 1$ indices and they are always between 0 and 1. It is possible to define the indices also for group of factors. From equations 2.16 and 2.15 it is possible to find that the sum of all the indices is always equal to 1 (2.17).

$$1 = \sum_{i=1}^{k} S_i + \sum_i \sum_{j>i} S_{ij} + ... + S_{1,...,k} \tag{2.17}$$

The most important among all the Sobol's indices is the one related to the first order terms, known as *main effect* or *first order sensitivity index*,

in equation 2.18 [87].

$$S_i = \frac{V_i}{V(Y)} = \frac{V_{X_i}(E_{\mathbf{X}_{\sim \mathbf{i}}}(Y \mid X_i))}{V(Y)} \tag{2.18}$$

The main effect is generally interpreted as the expected reduction in $V(Y)$ if $X_i$ could be fixed [88]. $S_i$ has a long history, in fact it was originally defined by Pearson in 1905 and it was known as *correlation ratio* [89].
To perform a 'complete' GSA, one should compute all the terms for all the orders, but this could be computationally demanding and it would increase the difficulty in the analysis interpretation (this because there would be too many indices to look at) [21, 78]. One could, for example, limit the analysis to the main effect. However, by doing this all the interactions between the parameters would be ignored. This fact could lead to an underestimation of the factors importance in explaining $V(Y)$. In fact, the higher order terms can be the most important ones and so, their importance must be assessed [1] [67].

To overcome this limitation, Homma & Saltelli in 1996 [90] and Wagner in 1995 [91] introduced the so-called *total effect*. The total effect is a sensitivity index that considers the impact of each parameter taken alone, plus all the interactions of all the orders in which that parameter is involved. To account for the interactions, let us consider the value $V_{\mathbf{X}_{\sim i}}(E_{X_i}(Y \mid \mathbf{X}_{\sim i}))$. This term includes the effect of any order, for any factor but $X_i$ [21]. In fact, the dependency on $X_i$ is removed by the expected value $E_{X_i}$ and the variance is computed over all the $\mathbf{X}$ but $X_i$. So, the larger $V_{\mathbf{X}_{\sim i}}(E_{X_i}(Y \mid \mathbf{X}_{\sim i}))$ becomes, the smaller the overall effect of $X_i$ (first order plus all the interactions) is. Considering the relation in equation 2.19, the total effect for a given factor $X_i$ can be defined as in equation 2.20 [21, 87].

[1] Only in the case of a linear model, such as $Y = \sum_{i=1}^{k} X_i$, we know in advance that the interaction terms are all equal to 0. Let us consider instead the model $Y = X_1 \cdot X_2$, with $X_1$ distributed normally with mean equal to 1 and variance equal to 1 and $X_2$ distributed normally with mean equal to 0 and variance equal to 1. The main effect of $X_1$ is equal to 0, because $X_2$ has mean 0. Thus, by limiting the analysis on the main effect, one may conclude that $X_1$ has no impact on $V(Y)$. Intuitively, this conclusion is wrong. In fact, $X_1$ impact can be observed if $X_2$ is allowed to vary from its mean value. Thus, $X_1$ impact on $V(Y)$ is due to the interaction effect with $X_2$.

$$V(Y) = V_{\mathbf{X}_{\sim i}}(E_{X_i}(Y \mid \mathbf{X}_{\sim i})) + E_{\mathbf{X}_{\sim i}}(V_{X_i}(Y \mid \mathbf{X}_{\sim i})) \tag{2.19}$$

$$S_{Ti} = 1 - \frac{V_{\mathbf{X}_{\sim i}}(E_{X_i}(Y \mid \mathbf{X}_{\sim i}))}{V(Y)} = \frac{E_{\mathbf{X}_{\sim i}}(V_{X_i}(Y \mid \mathbf{X}_{\sim i}))}{V(Y)} \tag{2.20}$$

As previously written, $S_{Ti}$ contains any term of any order in 2.15 that include the factor $X_i$, therefore $S_{Ti} \geq S_i$.

Variance based GSA substantially consists in the computation of the main and the total effect for all the model parameters that are considered variable. The larger $S_{Ti}$ is, the more important $X_i$ is in explaining $V(Y)$; $S_{Ti} = 0$ is a necessary and sufficient condition for the factor $X_i$ to be considered non-influential. The difference between $S_{Ti}$ and $S_i$ gives information about the extent of the interactions involving $X_i$. The results of a variance based GSA can be presented, for example, by using a barplot, as shown in figure 2.1.

The characteristic of variance based GSA is that the importance of each factor is related with the portion of output variance that it explains with its variation. Thus, with this method, the variance is used to represent the uncertainty of $Y$. This could be a problem, because the variance is a sufficient measure of the variability only under certain assumptions (e.g., normality) [61]. Moreover, it is difficult to robustly estimate the variance of fat-tailed or skewed distributions. This fact could lead to instability in the estimation of variance based indices from samples to samples. One way to overcome this problem could be to use a transformation of $Y$ as output variable, such as $\log Y$. However, results in log-scale (or, generally, in other scales) do not easily translate back to a linear scale [92].

### Implementation

It is possible to estimate the variance based sensitivity indices by using Monte Carlo based methods [21, 69, 87] or by using the Fourier Amplitude Sensitivity Test (FAST) [93, 94]. It was shown that for a number of input factors higher that 10, FAST is biased, unstable and costly [78, 95]. So,

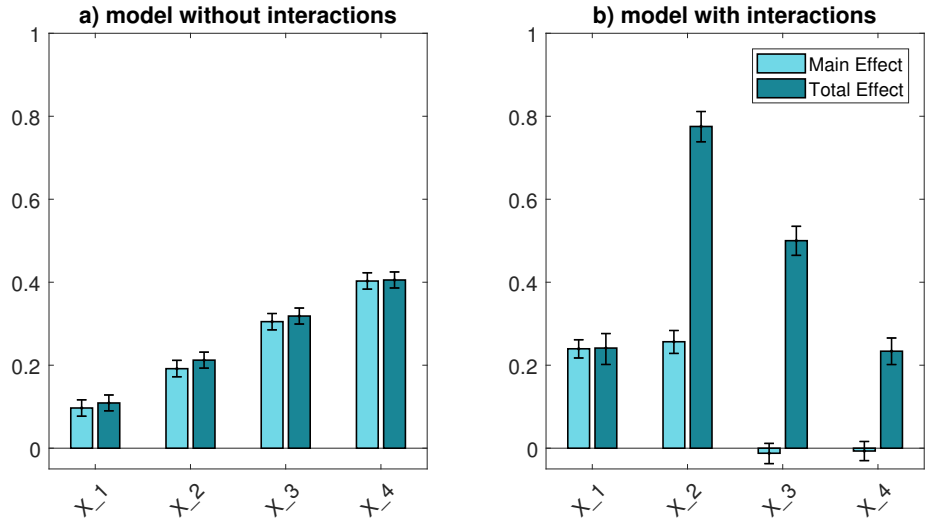



Figure 2.1: Example of GSA on models with and without interactions, with 4 factors all normally distributed with mean 0 and variance 1. (a) $Y = X_1 + \sqrt{2}X_2 + \sqrt{3}X_3 + \sqrt{4}X_4$, it is possible to observe that $S_i \backsimeq S_{Ti}$ and that the factors explain around 10%, 20%, 30% and 40% of $V(Y)$. (b) $Y = X_1 + X_2 + X_2X_3 + X_2X_3X_4$, it is possible to observe that there are interaction effects, in particular, $X_3$ and $X_4$ impacts are only related to terms with order higher than 1. All the indices were calculated by using the method presented in this section, with 10000 samples. Error bars represent the 95% confidence interval of the sensitivity indices numerical estimation. Negative values for indices whose value is close to zero are due to numerical errors [21].

we used the Monte Carlo approach presented in the book *Global Sensitivity Analysis: The Primer* [21], that refers to Saltelli's work in 2002 [87], that in turn is an optimization of the method proposed by Sobol [69] and Homma & Saltelli [90].

The Monte Carlo method that we used consists in four steps:

1. Generate two $(n, k)$ matrices $\mathbf{A}$ and $\mathbf{B}$, that contain $n$ samples drawn from the $k$ factors pdf.

$$\begin{aligned} \mathbf{A} &= [\mathbf{x}_{A,1}, \mathbf{x}_{A,2}, ..., \mathbf{x}_{A,k}] \\ \mathbf{B} &= [\mathbf{x}_{B,1}, \mathbf{x}_{B,2}, ..., \mathbf{x}_{B,k}] \end{aligned} \tag{2.21}$$

with $\mathbf{x}_{A,i}$ and $\mathbf{x}_{B,i}$ vectors containing $n$ samples of $X_i$ in $\mathbf{A}$ and $\mathbf{B}$,

respectively.

2. Create $k$ matrices $\mathbf{C}_i$ that are equal to $\mathbf{B}$ except the $i$-th column that would be equal to $\mathbf{x}_{A,i}$.

$$\mathbf{C}_i = [\mathbf{x}_{B,1}, \mathbf{x}_{B,2}, ..., \mathbf{x}_{A,i}, ..., \mathbf{x}_{B,k}] \tag{2.22}$$

3. Compute the model output for all the $n$ rows of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}_i$, in order to obtain $k+2$ model output vectors, $\mathbf{y}_A$, $\mathbf{y}_B$ and $\mathbf{y}_{C_i}$.

4. Compute the main ($S_i$) and the total effect ($S_{Ti}$) as follows.

$$\begin{aligned} S_i &= \frac{\frac{1}{n}\sum_j y_A^j y_{C_i}^j - \frac{1}{n^2}\sum_j y_A^j \sum_j y_B^j}{\frac{1}{n}\sum_j (y_A^j)^2 - \hat{f}_0^{\,2}} \\ S_{Ti} &= \frac{\frac{1}{n}\sum_j y_B^j y_{C_i}^j - \hat{f}_0^{\,2}}{\frac{1}{n}\sum_j (y_A^j)^2 - \hat{f}_0^{\,2}} \\ \hat{f}_0^{\,2} &= \Big(\frac{1}{2n}\sum_j y_{AB}^j\Big)^2 \end{aligned} \tag{2.23}$$

$\mathbf{y}_{AB}$ is a vector obtained concatenating $\mathbf{y}_A$ and $\mathbf{y}_B$. This is done to allow a more robust estimation of $\hat{f}_0$.

With this approach, the model would be evaluated $n(k+2)$ times. To calculate the confidence intervals of the sensitivity indices estimates, a bootstrap approach can be used [67, 83].

### 2.1.3 Moment independent methods

GSA methods that consider the whole output distribution, rather than a singular moment (such as the variance based method), are called *moment independent*. Here we describe only one particular index, known as $\delta$ *sensitivity measure*, proposed by Borgonovo in 2006 [96]

Let $p_Y$ be the probability density function of the model output $Y$. If we fix one factor to a particular value, $X_i = x_i^*$, we can obtain the conditional

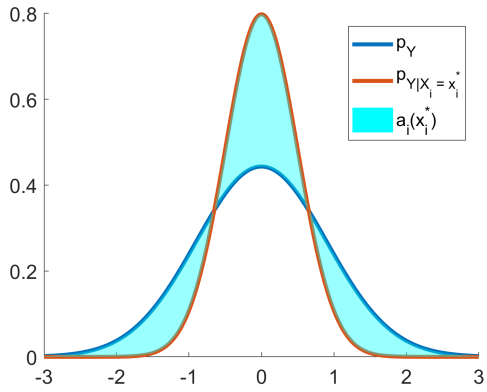



Figure 2.2: Unconditional and conditional densities. Blue area represent the measure of the distances between the two pdf.

density of $Y$, $p_{Y|X_i=x_i^*}$. To measure the impact of fixing $X_i$ to $x_i^*$ on $p_Y$, one can introduce a distance measure between $p_Y$ and $p_{Y|X_i=x_i^*}$ (blue area in figure 2.2), such as:

$$a_i(x_i^*) = \int_{\Omega_Y} |p_Y - p_{Y|X_i=x_i^*}| dY. \tag{2.24}$$

However, $X_i$ is a random variable and can assume other values in its support. Thus, a sensitivity measure can be derived by computing the expected value of $a_i$ over all the possible $X_i$ values. By doing this, we can obtain the $\delta$ sensitivity index for $X_i$, shown in equation 2.25.

$$\delta_i = \frac{1}{2} E_{X_i}\left[\int_{\Omega_Y} |p_Y - p_{Y|X_i=x_i^*}| dY\right] \tag{2.25}$$

$\delta$ index has several interesting properties (see [61, 96]), such as:

- $0 \leq \delta_i \leq 1$;
- $\delta_i = 0$ if and only if the $Y$ is independent from $X_i$;

- the sensitivity measure of a group containing all the factors is equal to 1, $\delta_{1,...,k} = 1$;
- $\delta_i$ is monotonic invariant.

The last property refers to the fact that $\delta_i$ for $Y$ is equal to $\delta_i$ for any monotonic transformation of $Y$ (e.g., $\log Y$).

Concerning the computation of the indices, it seems to be not trivial. In fact, to calculate the sensitivity indices, the samples of each $X_i$ have to be grouped in $M$ classes and then the $p_Y$ and $p_{Y|X_i=x_i^j}$, $j = 1, ..., M$, are fitted using the *kernel smoothing* method [97]. So, one should carefully asses how the choice of the design parameters (such as the number of classes $M$ and the kernel functions) impacts the values of $\delta_i$.

Another popular sensitivity measure belonging to the class of moment independent methods, is the one called PAWN [98]. The idea here is to use the cumulative density function rather than the pdf and then to characterize the distance between $p_Y$ and $p_{Y|X_i=x_i^*}$ by using the Kolmogorov-Smirnov statistic. This method is implemented in the SAFE toolbox (`https://www.safetoolbox.info/`). However, a recent work showed that the sensitivity indices obtained by using PAWN may be sensitive to the design parameters (`arXiv:1904.04488 [stat.AP]`).

## 2.2 Application to a model describing neutropenia time course

### 2.2.1 Model description

This benchmark model was taken from [80, 81], that in turn is derived from the widely known *Friberg* model [82]. The model describes the time course of neutropenia in response to a treatment with paclitaxel, a cytotoxic drug. The model is composed by two parts, one describing the paclitaxel PK and the other one the neutrophils level in response to the treatment (PD).

The PK model has three compartments and it is reported in equation system 2.26.

$$\begin{aligned}
\frac{dm_1}{dt} &= -\left(\frac{V_{M,el}}{K_{M,el} + m_1/V_1} + \frac{V_{M,tr}}{K_{M,tr} + m_1/V_1} + Q\right)\frac{m_1}{V_1} + k_{21}\, m_2 + Q\,\frac{m_3}{V_3} \\
\frac{dm_2}{dt} &= \frac{V_{M,tr}}{K_{M,tr} + m_1/V_1}\,\frac{m_1}{V_1} - k_{21}\, m_2 \\
\frac{dm_3}{dt} &= Q\left(\frac{m_1}{V_1} - \frac{m_3}{V_3}\right)
\end{aligned} \tag{2.26}$$

$m_1$, $m_2$ and $m_3$ are the drug masses (in $\mu mol$) in the three compartments, $V_1$ and $V_3$ are compartment 1 and 3 volumes. $V_{M,el}$, $K_{M,el}$ and $V_{M,tr}$, $K_{M,tr}$ are the Michaelis-Menten parameters relative to drug elimination from compartment 1 and drug transfer from the first to the second compartment, respectively. $k_{21}$ is the rate transfer from compartment 2 to compartment 1 and $Q$ is the flow between compartment 1 and 3.

The PD model is reported in equation system 2.27.

$$\begin{aligned}
\frac{dP}{dt} &= k_{prol}\, P\, F_B\, (1 - E_{drug}) - k_{tr}\, P \\
\frac{dT_1}{dt} &= k_{tr}\, (P - T_1) \\
\frac{dT_2}{dt} &= k_{tr}\, (T_1 - T_2) \\
\frac{dT_3}{dt} &= k_{tr}\, (T_2 - T_3) \\
\frac{dC}{dt} &= k_{tr}\, T_3 - k_{el}\, C \\
F_B &= \left(\frac{C_0}{C}\right)^{\gamma} \\
E_{drug} &= SL\, \frac{m_1}{V_1}
\end{aligned} \tag{2.27}$$

$P$, $T_1$, $T_2$, $T_3$ and $C$ are the proliferative cells, transit compartment 1, 2, 3 and circulating cells concentration (in $10^9 cells/L$). $k_{prol}$ is the time constant of cellular proliferation, $k_{tr}$ is the transit time between the compartments and $k_{el}$ is the time constant associated with the elimination of the circulating cells. $k_{tr}$ was set equal to $4/MMT$, where $MMT$ is the mean maturation time. Moreover, $k_{tr} = k_{prol} = k_{el}$ [80, 81]. $F_B$ is the feedback term and depends on $C$, $C_0$ (baseline concentration of circulating cells) and $\gamma$. $E_{drug}$ is the drug effect on cells proliferation and it is a linear function of drug concentration in compartment 1, with $SL$ the 'slope' parameter.

Both *inter-individual* and *inter-occasion* variabilities (iiv and iov, respectively) are present for some of the model parameters. An exponential model was used to describe the iiv, as in equation 2.28.

$$\theta_i = \theta_{pop}\, e^{\eta_\theta} \tag{2.28}$$

$\theta_i$ is the subject parameter, $\theta_{pop}$ is the population mean and $\eta_\theta \sim \mathcal{N}(0, \omega^2_{\theta,\eta})$ is the inter individual variability parameter, with $\omega^2_{\theta,\eta}$ the variance. The

Table 2.2: Paclitaxel model parameters

| **Parameters**[a] | mean | iiv (CV [%][b]) | iov (CV [%][b]) | units |
|---|---|---|---|---|
| | | **PK** | | |
| $V_1$ | 10.8 | | 37.3 | $L$ |
| $V_3$ | 275 | 46.2 | | $L$ |
| $K_{M,el}$ | 0.576 | | | $\mu M$ |
| $V_{M,el}$ | 35.8 | 17.8 | 15.2 | $\mu mol/h$ |
| $K_{M,tr}$ | 1.43 | 69.8 | | $\mu M$ |
| $V_{M,tr}$ | 177 | 28.7 | | $\mu mol/h$ |
| $k_{21}$ | 1.11 | 9.31 | | $h^{-1}$ |
| $Q$ | 15.6 | 45.8 | | $L/h$ |
| | | **PD** | | |
| $MMT$ | 141 | 27 | | $h$ |
| $SL$ | 2.6 | 44.9 | | $L/\mu mol$ |
| $\gamma$ | 0.2 | | | |
| $C_0$ | 6.48 | 31.6 | | $10^9 cells/L$ |
| | | **residual variabilities** | | |
| $RES^b_{PK}$ | 18.2 | | | % |
| $RES^b_{PD}$ | 31.6 | | | % |

[a] All the parameters were taken from [81].
[b] CV in natural scale is considered approximately equal to the standard deviation in logarithmic scale [99].

iov is also modelled by using 2.28. In case a parameter has both iiv and iov, equation 2.29 was used.

$$\theta_i = \theta_{pop}\, e^{\eta_\theta + \kappa_\theta} \tag{2.29}$$

$\kappa_\theta \sim \mathcal{N}(0, \omega^2_{\theta,\kappa})$ is the inter occasion variability parameter, with $\omega^2_{\theta,\kappa}$ the variance. In this model, the occasion is the chemotherapy cycle.
All the parameters values are reported in table 2.2.

### 2.2.2 GSA results

The model was evaluated on 10000 samples extracted from the input parameters pdf by using a Latin hypercube sampling strategy. In figure 2.3 the paclitaxel plasma concentration ($m_1/V_1$) and the circulating neutrophils

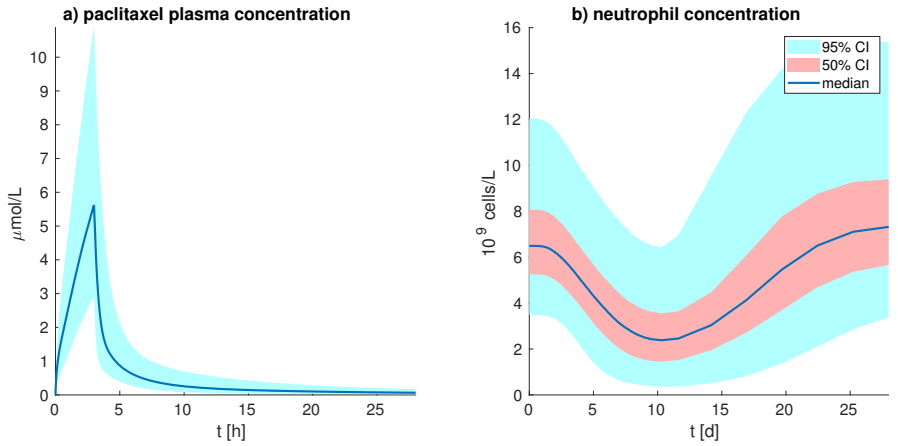



Figure 2.3: Output of paclitaxel PKPD model. Residual variability on the observations was not included in the simulated time curves.

level ($C$) are reported, respectively. For GSA, nadir value of neutrophil concentration ($C_{nadir}$) and its time of occurrence ($t_{nadir}$) were used as output references. All the model input parameters samples and the relative outputs were used to perform the GSA with the methods presented in section 2.1, except for the variance based method. For the latter, the algorithm presented in 2.1.2 was used, with a number of samples for the construction of **A** and **B** matrices equal to 10000.

Before doing a GSA, it is appropriate to look at the model output distributions (figure 2.4) and at the scatter plot of the model output versus the model input parameters (figures 2.5 and 2.6). From figure 2.5 it is possible to observe that $Q$, $MMT$, $SL$, $C_0$ and $RES_{PD}$ are the parameters with the highest correlation with $C_{nadir}$. Moreover, it could be seen that interaction effects occur for $SL$, $C_0$ and $RES_{PD}$. This happens because the variation of these parameters causes a modification in the confidence interval width of $C_{nadir}$. From figure 2.6 it is clear that the most important parameter in explaining $t_{nadir}$ variation is $MMT$.

All the GSA results obtained by using the Pearson correlation coefficient, PCC, SRC, lasso, the variance based method and the $\delta$ sensitivity

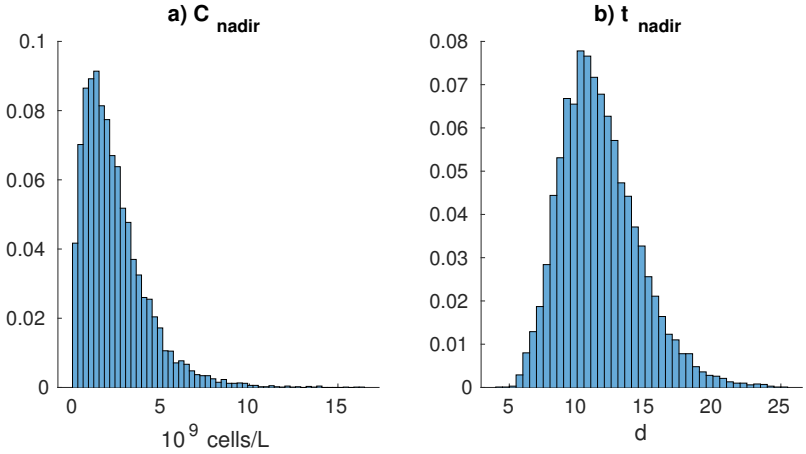



Figure 2.4: Histograms of model outputs. Residual variability in the observation was included in the histograms.

indices are reported in table 2.3 for $C_{nadir}$ and in table 2.4 for $t_{nadir}$. Results of variance based GSA are also shown in figure 2.7.

Concerning $C_{nadir}$, it is possible to observe that all the methods give, more or less, the same factor ranking (at least for the first five sensitive parameters). This probably happens because the input-output relationship is almost linear. In fact, $R^2$ of the linear regression is equal to 0.8. Moreover, from figure 2.7 it can be seen that the interaction effects are not strong enough to change the factor ranking done considering only the first order terms. However, one difference between the variance based, the $\delta$ sensitivity indices and the other methods is in the factor fixing setting. For all the methods based on linear models, the sensitivity indices relative to $V_{M,el}$ variabilities have approximately the same order of magnitude of $SL$. For Sobol and $\delta$ methods, instead, the $V_{M,el}$ sensitivity indices are one order of magnitude lower than the one of $SL$.

Concerning $t_{nadir}$, as for $C_{nadir}$, all the methods give approximately the same factor ranking too, at least for the two most important parameters. In this case, $R^2$ of the linear regression is close to 0.9.
Similarly to the results for $C_{nadir}$, one difference between Sobol and $\delta$ in-

Table 2.3: $C_{nadir}$ GSA results

| **Parameters** | Pearson | PCC | SRC | lasso | $S_T$ | $\delta$ |
|---|---|---|---|---|---|---|
| | **sensitivity indices** | | | | | |
| $V_1$ | 0.04 | 0.08 | 0.03 | 0.01 | -0.01 | 0.02 |
| $V_3$ | 0.05 | 0.14 | 0.06 | 0.04 | -0.01 | 0.02 |
| $V_{M,el}$ iiv | 0.17 | 0.37 | 0.17 | 0.15 | 0.04 | 0.05 |
| $V_{M,el}$ iov | 0.15 | 0.33 | 0.15 | 0.13 | 0.01 | 0.05 |
| $K_{M,tr}$ | -0.08 | -0.18 | -0.08 | -0.05 | -0.01 | 0.03 |
| $V_{M,tr}$ | 0.09 | 0.2 | 0.09 | 0.06 | 0 | 0.03 |
| $k_{21}$ | -0.04 | -0.06 | -0.03 | 0.004 | -0.01 | 0.02 |
| $Q$ | 0.21 | 0.42 | 0.21 | 0.18 | 0.05 | 0.07 |
| $MMT$ | 0.31 | 0.58 | 0.31 | 0.29 | 0.1 | 0.11 |
| $SL$ | -0.52 | -0.77 | -0.53 | -0.5 | 0.36 | 0.24 |
| $C_0$ | 0.4 | 0.67 | 0.4 | 0.37 | 0.26 | 0.16 |
| $RES_{PK}$ | 0.01 | -0.01 | 0 | 0 | -0.01 | 0.01 |
| $RES_{PD}$ | 0.4 | 0.67 | 0.4 | 0.37 | 0.25 | 0.15 |
| | **factor ranking** | | | | | |
| 1st | $SL$ | $SL$ | $SL$ | $SL$ | $SL$ | $SL$ |
| 2nd | $C_0$ | $RES_{PD}$ | $RES_{PD}$ | $RES_{PD}$ | $C_0$ | $C_0$ |
| 3rd | $RES_{PD}$ | $C_0$ | $C_0$ | $C_0$ | $RES_{PD}$ | $RES_{PD}$ |
| 4th | $MMT$ | $MMT$ | $MMT$ | $MMT$ | $MMT$ | $MMT$ |
| 5th | $Q$ | $Q$ | $Q$ | $Q$ | $Q$ | $Q$ |
| 6th | $V_{M,el}$ iiv | $V_{M,el}$ iiv | $V_{M,el}$ iiv | $V_{M,el}$ iiv | $V_{M,el}$ iiv | $V_{M,el}$ iov |
| 7th | $V_{M,el}$ iov | $V_{M,el}$ iov | $V_{M,el}$ iov | $V_{M,el}$ iov | $RES_{PK}$ | $V_{M,el}$ iiv |
| 8th | $V_{M,tr}$ | $V_{M,tr}$ | $V_{M,tr}$ | $V_{M,tr}$ | $k_{21}$ | $V_{M,tr}$ |
| 9th | $K_{M,tr}$ | $K_{M,tr}$ | $K_{M,tr}$ | $K_{M,tr}$ | $V_{M,el}$ iov | $K_{M,tr}$ |
| 10th | $V_3$ | $V_3$ | $V_3$ | $V_3$ | $V_1$ | $V_3$ |
| 11th | $V_1$ | $V_1$ | $V_1$ | $V_1$ | $V_3$ | $V_1$ |
| 12th | $k_{21}$ | $k_{21}$ | $k_{21}$ | $k_{21}$ | $K_{M,tr}$ | $k_{21}$ |
| 13th | $RES_{PK}$ | $RES_{PK}$ | $RES_{PK}$ | $RES_{PK}$ | $V_{M,tr}$ | $RES_{PK}$ |

dices, with respect to all the other methods, is still in the factor fixing setting. In fact, for the methods based on linear models, other parameters (such as $SL$, $Q$, $V_{M,el}$) have the sensitivity indices with the same order of magnitude of $MMT$. Instead, Sobol and $\delta$ methods tend to discriminate better the importance of $MMT$.

Table 2.4: $t_{nadir}$ GSA results

| **Parameters** | Pearson | PCC | SRC | lasso | $S_T$ | $\delta$ |
|---|---|---|---|---|---|---|
| | | | **sensitivity indices** | | | |
| $V_1$ | -0.02 | -0.05 | -0.02 | 0 | 0.01 | 0 |
| $V_3$ | 0.01 | -0.01 | 0 | 0 | 0.01 | 0 |
| $V_{M,el}$ iiv | -0.09 | -0.23 | -0.08 | -0.05 | 0.02 | 0.02 |
| $V_{M,el}$ iov | -0.07 | -0.21 | -0.07 | -0.04 | 0.02 | 0.01 |
| $K_{M,tr}$ | 0 | 0.07 | 0.02 | 0 | 0.01 | 0 |
| $V_{M,tr}$ | -0.03 | -0.09 | -0.03 | 0.002 | 0.01 | 0 |
| $k_{21}$ | 0.02 | 0.04 | 0.01 | 0 | 0.01 | 0 |
| $Q$ | -0.08 | -0.23 | -0.08 | -0.05 | 0.02 | 0.01 |
| $MMT$ | 0.92 | 0.94 | 0.92 | 0.89 | 0.93 | 0.65 |
| $SL$ | 0.2 | 0.54 | 0.2 | 0.18 | 0.06 | 0.06 |
| $C_0$ | 0 | 0 | 0 | 0 | 0 | 0 |
| $RES_{PK}$ | 0 | 0.02 | 0.01 | 0 | 0 | 0.01 |
| $RES_{PD}$ | -0.01 | 0.01 | 0 | 0 | 0 | 0.01 |
| | | | **factor ranking** | | | |
| 1st | $MMT$ | $MMT$ | $MMT$ | $MMT$ | $MMT$ | $MMT$ |
| 2nd | $SL$ | $SL$ | $SL$ | $SL$ | $SL$ | $SL$ |
| 3rd | $V_{M,el}$ iiv | $V_{M,el}$ iiv | $V_{M,el}$ iiv | $V_{M,el}$ iiv | $Q$ | $V_{M,el}$ iiv |
| 4th | $Q$ | $Q$ | $Q$ | $Q$ | $V_{M,el}$ iiv | $Q$ |
| 5th | $V_{M,el}$ iov | $V_{M,el}$ iov | $V_{M,el}$ iov | $V_{M,el}$ iov | $V_{M,el}$ iov | $RES_{PK}$ |
| 6th | $V_{M,tr}$ | $V_{M,tr}$ | $V_{M,tr}$ | $V_{M,tr}$ | $V_{M,tr}$ | $V_{M,el}$ iov |
| 7th | $k_{21}$ | $K_{M,tr}$ | $K_{M,tr}$ | $V_1$ | $V_3$ | $RES_{PD}$ |
| 8th | $V_1$ | $V_1$ | $V_1$ | $V_3$ | $K_{M,tr}$ | $K_{M,tr}$ |
| 9th | $V_3$ | $k_{21}$ | $k_{21}$ | $K_{M,tr}$ | $V_1$ | $V_{M,tr}$ |
| 10th | $RES_{PD}$ | $RES_{PK}$ | $RES_{PK}$ | $k_{21}$ | $k_{21}$ | $V_1$ |
| 11th | $RES_{PK}$ | $RES_{PD}$ | $RES_{PD}$ | $C_0$ | $RES_{PK}$ | $C_0$ |
| 12th | $K_{M,tr}$ | $V_3$ | $V_3$ | $RES_{PK}$ | $RES_{PD}$ | $V_3$ |
| 13th | $C_0$ | $C_0$ | $C_0$ | $RES_{PD}$ | $C_0$ | $k_{21}$ |

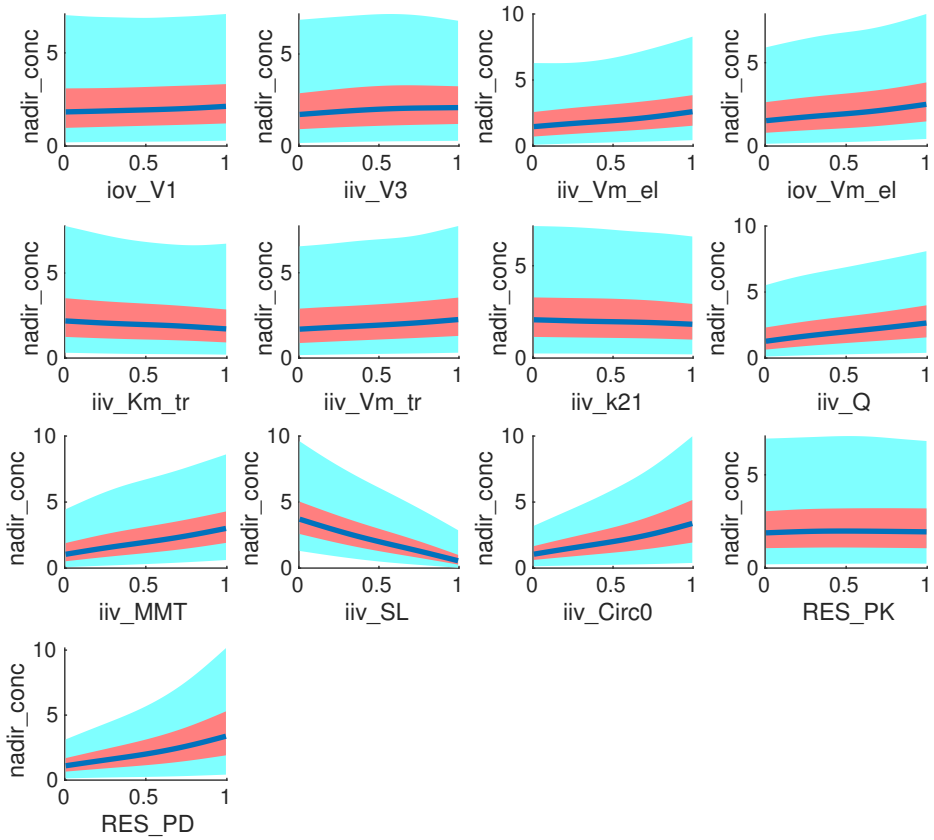



Figure 2.5: Scatter plot nadir concentration ($10^9 cells/L$) versus input parameters (normalized between 0 and 1). The blue line is the median, the red and blue shaded area are the 50% and 95% confidence interval of the data.

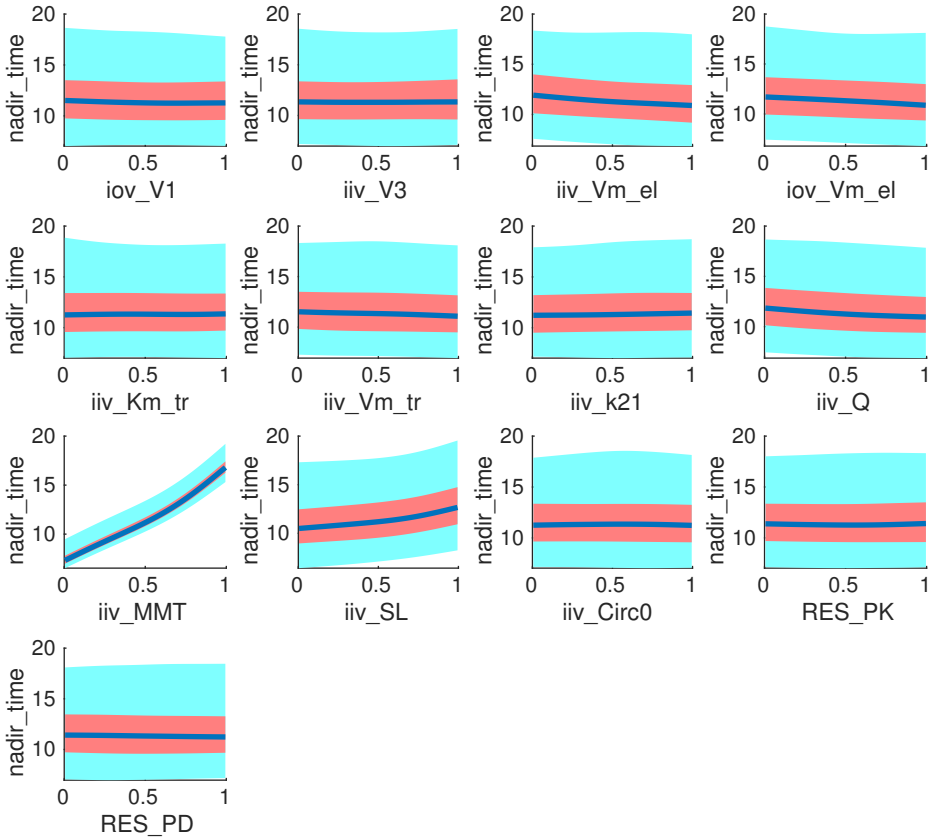



Figure 2.6: Scatter plot time to nadir (*days*) versus input parameters (normalized between 0 and 1). The blue line is the median, the red and blue shaded area are the 50% and 95% confidence interval of the data.

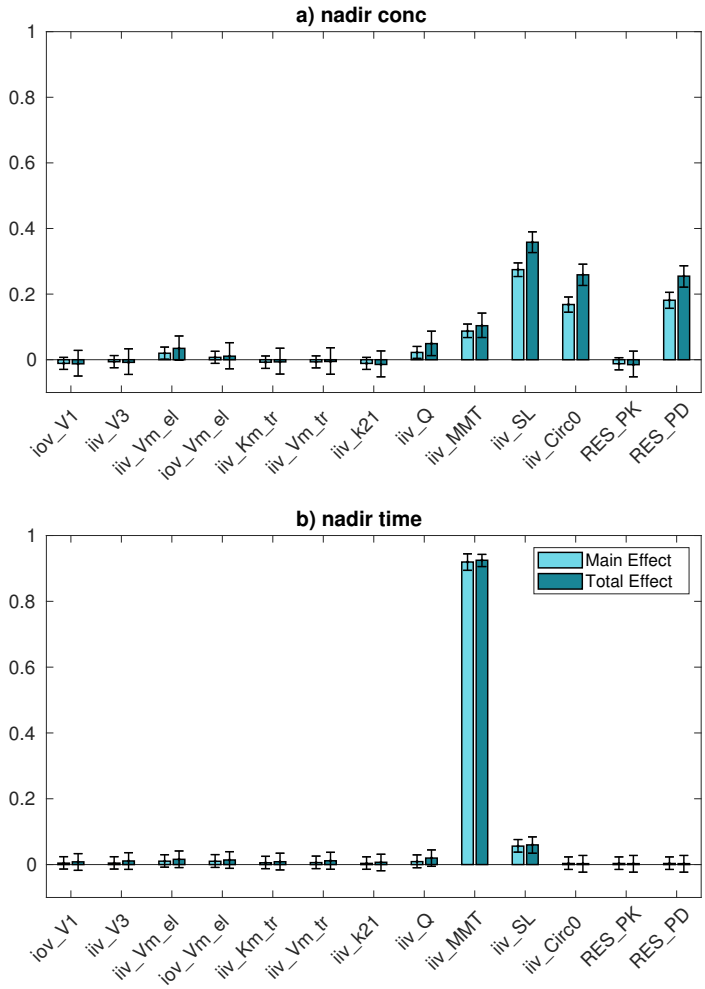



Figure 2.7: Variance based indices of $C_{nadir}$ and $t_{nadir}$.

## 2.3 Application to a linear model

In the neutropenia model presented in section 2.2, the parameter ranking is approximately the same for all the GSA methods. This happened because the interaction effects were not strong enough to change the ranking done with the first order terms, as it can be seen from the results in figure 2.7. To make the point of the advantage in using the variance based GSA, we applied all the presented methods to a simple linear model,

$$Y = X_1 + X_2 + X_2\,X_3 + X_2\,X_3\,X_4, \tag{2.30}$$

where $Y$ is the model output and $X_i \sim \mathcal{N}(0,1)$, $i = 1, ..., 4$, are the model input factors. Similarly to what was done for the neutropenia model, 10000 samples were extracted with a Latin hypercube sampling strategy from the input factor distribution and for each sample, the output was computed. All the samples were used to perform the GSA with all the methods except the variance based one. For the latter, the algorithm in section 2.1.2 was used, with a number of samples for the construction of $\mathbf{A}$ and $\mathbf{B}$ matrices equal to 10000.

In figures 2.8 and 2.9 the histogram of the model output $Y$ and the scatterplots of $Y$ against the input factors are reported, respectively. From the scatterplots it is possible to observe that $X_3$ and $X_4$ change the distribution width of $Y$, but not its central tendency. This is a clear signal that the impact of $X_3$ and $X_4$ on $V(Y)$ is mainly due to interaction effects.

In table 2.5 the GSA results are reported. It is possible to observe that all the linear methods identify $X_1$ and $X_2$ as the most important factors, while $X_3$ and $X_4$ as non-influential factors. For the variance based method all the factors have an impact on the model output variability and the most important parameter is $X_2$. In figure 2.1 (b) it can be seen that the importance of $X_2$, $X_3$ and $X_4$ in explaining $V(Y)$ is mainly related with interaction terms. Concerning the $\delta$ method, it recognize $X_3$ as an important factor, however, with respect to the variance based method, it underestimates the importance of $X_4$.

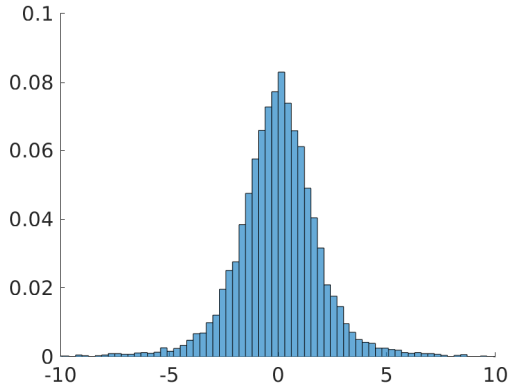



Figure 2.8: Histogram of the linear model output.

Table 2.5: GSA results for the linear model

| **Parameters** | Pearson | PCC | SRC | lasso | $S_T$ | $\delta$ |
|---|---|---|---|---|---|---|
| | **sensitivity indices** | | | | | |
| $X_1$ | 0.49 | 0.57 | 0.5 | 0.39 | 0.25 | 0.28 |
| $X_2$ | 0.5 | 0.58 | 0.51 | 0.4 | 0.74 | 0.25 |
| $X_3$ | 0 | 0.01 | 0 | 0 | 0.52 | 0.1 |
| $X_4$ | 0.01 | 0 | 0 | 0 | 0.26 | 0.04 |
| | **factor ranking** | | | | | |
| $1^{\text{st}}$ | $X_2$ | $X_2$ | $X_2$ | $X_2$ | $X_2$ | $X_1$ |
| $2^{\text{nd}}$ | $X_1$ | $X_1$ | $X_1$ | $X_1$ | $X_3$ | $X_2$ |
| $3^{\text{rd}}$ | $X_4$ | $X_3$ | $X_3$ | $X_3$ | $X_4$ | $X_3$ |
| $4^{\text{th}}$ | $X_3$ | $X_4$ | $X_4$ | $X_4$ | $X_1$ | $X_4$ |

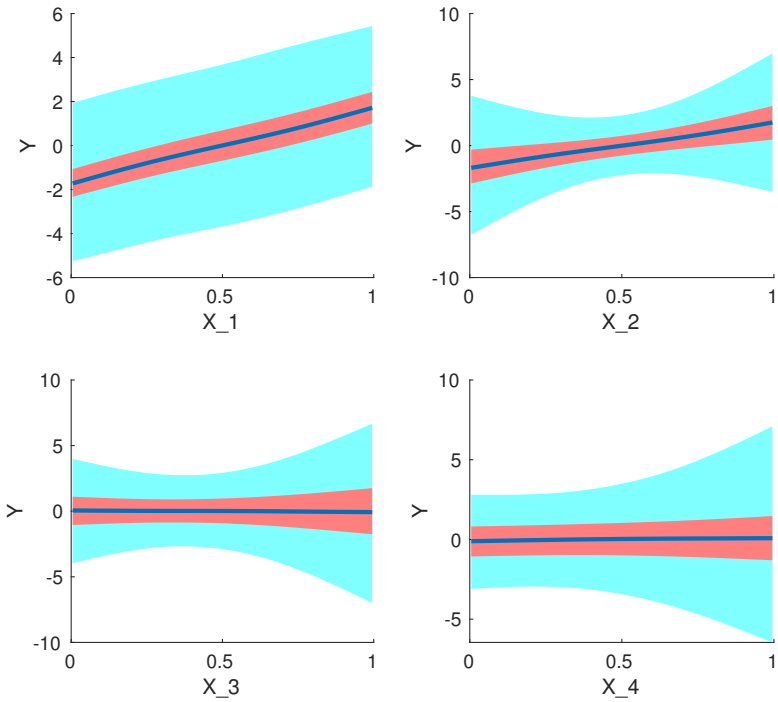



Figure 2.9: Scatter plot of the linear model output against the input parameters (normalized between 0 and 1). The blue line is the median, the red and blue shaded area are the 50% and 95% confidence interval of the data.

## 2.4 Discussion

In this chapter we presented some of the methods to perform GSA. The methods based on linear models are commonly implemented in data analysis software, are easy to use and the sensitivity indices are generally easy to interpret. However, their use is challenging in presence of non-linear or non-monotonic dependencies between the model output and the input factors. Moreover, by using them it is difficult to appropriately estimate the interaction effects. The variance based method is easy to implement and the sensitivity indices are easy to interpret. Moreover, this method allows to appreciate how much is the impact of each factor, taken singularly, on the model output variation and how much is the extent of its interaction effects. However, the variance based method takes $V(Y)$ as a proxy of model output variation, with all the limitation reported in section 2.1.2. Finally, the $\delta$ method overcomes some of the limitation of the variance based GSA and it can be used also for skewed distribution of $Y$. However, with this method it is impossible to detect interaction effects and the design parameters have to be chosen appropriately.

The focus of the next chapters is to apply GSA to different models, considering several levels of uncertainty and variability for each one of them. The method that we choose to use for performing GSA is the variance based method, that is considered to be the gold standard by the sensitivity analysis community. We believe that conducting all the sensitivity analyses with only one method would be advantageous. This because it would simplify the results interpretation and presentation, it would uniform the works and put the focus on the applications rather than on the comparison between different methods.

Chapter 3

# Mechanistic model for the anticancer pro-drug Gemcitabine: use of GSA to understand the parameters impact on population variability [1]

## 3.1 Introduction

*Gemcitabine* (2',2'-difluorodeoxycytidine, dFdC) is a nucleoside antimetabolite pro-drug effective against several solid tumours [100, 101, 102, 103]. Treatment with dFdC represents the first line therapy of pancreatic cancer,

[1] This work was published in "M. García-Cremades, N. Melillo, I.F. Tróconiz, P. Magni. Mechanistic multi-scale pharmacokinetic model for the anticancer drug gemcitabine in pancreatic cancer. *Clinical and Translational Science*, 2020".

that constitutes one of the most aggressive and lethal oncology diseases, with an overall 5-year survival rate of less than 5% [104]. In this case, dFdC is either given in combination with nab-paclitaxel for patients with ECOG (eastern cooperative oncology group) performance status 0-1, or as a single agent for advanced patients (ECOG>1), and for those patients who cannot receive combination treatments [105].

As a pro-drug, dFdC has to be intracellularly metabolized to its active metabolite, dFdC triphosphate (dFdCTP), to exert its cytotoxic action [106]. Firstly, dFdC is taken into the cell by active transporters (hENTs, hCNTs) [107] and then it is phosphorylated by deoxycytidine kinase (dCK) to its monophosphate form, dFdCMP. dFdCMP is subsequently metabolized by nucleoside kinases to dFdC diphosphate (dFdCDP) and then to dFdCTP that binds to the DNA promoting apoptosis [108]. dFdC also suffers inactivation by cytidine deaminase (CDA), leading to inactive metabolite 2',2'-difluorodeoxyuridine (dFdU), which is excreted in urine [109].

One of the biggest complications associated to treatment with dFdC is the *variability* in responses, ranging from lack of efficacy to severe toxicity [110]. These different rates of responses to dFdC could be in part explained by individual genetic factors affecting its metabolic pathway, leading to different dFdCTP intracellular tumour concentrations [111]. As an example, a high activity of CDA enzyme is related with a higher depletion of dFdC and so, lower dFdCTP concentrations [112]. It is also stated that treatment efficacy may be explained by a non-functional transport of the pro-drug into the cell [107]. Moreover, cells with low dCK levels are associated with resistance to dFdC [113, 114]. In addition, some clinical studies in patients with pancreatic cancer treated with dFdC, associated different expressions of the transporters or the target enzymes activity with a high or low survival probabilities [113, 115, 116].

dFdC effects on pancreatic cancer have been described previously by using PKPD, mechanistical and semi-mechanistical models in *in vitro* [117, 118, 119, 120], preclinical *in vivo* [121, 122] and clinical stages [123, 124]. However, to the best of our knowledge, the models developed in clinical stages do not consider the inter-subject variability in dFdC metabolism (e.g., individual concentrations of enzymes involved in dFdC metabolism).

Information on the systemic and cellular pharmacokinetics of dFdC could be used to develop a quantitative model that describes mechanistically processes such as drug distribution, metabolism and active metabolite formation. A model like this could help in improving the knowledge of the system, for example by understanding what subject characteristics determine mostly the predicted active metabolite exposure in the site of action.

Systems pharmacology is an approach that aims to develop multi-scale mechanistic models that *"span the divide between cell-level biochemical models and organism-level PK/PD models"* [19]. These models integrate the knowledge from various sources (e.g., *in vitro* experiments, physiological data) [15]. PBPK modelling approach provides a framework for integrating drug specific parameters and *in vitro* measurements with physiological system-specific parameters [32, 125]. This type of models can integrate inter-individual variabilities in the concentrations of enzymes involved in drug metabolism and allows the simulation of drug concentration in specific tissues (e.g., pancreatic tumour) [126].

In this context, we built a mechanistic systems pharmacology model to describe dFdC pharmacokinetics and dFdCTP tumour concentrations, in a population of pancreatic cancer patients. The developed model was built using data from the literature, including genetic and physiological inter-subject variabilities. In summary, our aims were to: 1) propose a translational multi-scale system pharmacokinetic modelling approach for gemcitabine able to describe different concentrations of dFdC metabolites; 2) to show the capabilities and the limitations of this kind of modelling strategy starting from a case study; 3) to understand what information is needed and what can be found or not in the literature; 4) to understand what are the parameters that mostly drive the dFdC and dFdCTP exposure variability in a population of patients. The latter point was performed with uncertainty analysis and GSA.

## 3.2 Development of the mechanistic model and GSA

The work was performed in three different steps. First, we developed a model (the so-called metabolic network) to describe dFdC metabolism *in vitro* in two pancreatic cancer cell lines (PK9 and RPK9) [114]. Then, we developed a PBPK model to describe the dFdC pharmacokinetics in a population of pancreatic cancer patients. The *in vitro* derived metabolic network was coupled with the PBPK model in a compartment representing the pancreatic tumour after an appropriate rescaling of the network parameters. This was done to describe the dFdC metabolism and predict dFdCTP concentrations in the site of action. Finally, we performed GSA on the developed model to identify what are the characteristics that mostly drive the dFdC and dFdCTP exposure variability in a population of patients. The analyses were performed in MATLAB R2019a [127]. Parameters were estimated by using the covariance matrix adaptation evolution strategy (CMA-ES) [128].

### 3.2.1 Metabolic network

An extensive literature review was performed looking for knowledge and *in vitro* data to build a mathematical model of the dFdC metabolic pathway. The structure of this pathway, which has been defined over the years [106, 109, 108], is schematized in figure 3.1 (c). For this, the mathematical model was built by assuming that enzymatic reactions were described by first order rate constants, except for those catalysed by CDA, dCK and hENT1 enzymes, which were described by a *Michaelis Menten* model. The

metabolic network equations are reported in equation system 3.1.

$$
\begin{aligned}
\frac{d\,dFdC_{ext}}{dt} &= -\frac{V_{max,hENT1}\,dFdC_{ext}}{K_{M,hENT1}+dFdC_{ext}} - \frac{V_{max,CDA,ext}\,dFdC_{ext}}{K_{M,CDA,ext}+dFdC_{ext}} \\
\frac{d\,dFdC_{int}}{dt} &= \frac{V_{max,hENT1}\,dFdC_{ext}}{K_{M,hENT1}+dFdC_{ext}} - \frac{V_{max,CDA,int}\,dFdC_{int}}{K_{M,CDA,int}+dFdC_{int}} \\
&\quad - \frac{V_{max,dCK}\,dFdC_{int}}{K_{M,dCK}+dFdC_{int}} + K_{MPC}\,dFdCMP_{int} \\
\frac{d\,dFdCMP_{int}}{dt} &= \frac{V_{max,dCK}\,dFdC_{int}}{K_{M,dCK}+dFdC_{int}} - (K_{MPC}+K_{NMPK})\;dFdCMP_{int} \\
&\quad - \frac{K_{CMPD}\,dFdCMP_{int}}{1+INH\,dFdCTP_{int}} + K_{DPMP}\,dFdCDP_{int} \\
\frac{d\,dFdCDP_{int}}{dt} &= K_{NMPK}\,dFdCMP_{int} - (K_{NDPK}+K_{DPMP})\;dFdCDP_{int} \\
&\quad + K_{TPDP}\,dFdCTP_{int} \\
\frac{d\,dFdCTP_{int}}{dt} &= K_{NDPK}\,dFdCDP_{int} - (K_{TPDP}+K_{DNA})\;dFdCTP_{int}
\end{aligned}
\tag{3.1}
$$

$dFdC_{ext}$, $dFdC_{int}$, $dFdCMP_{int}$ and $dFdCDP_{int}$ are the dFdC extracellular, dFdC, dFdCMP, dFdCDP and dFdCTP intracellular amounts. $V_{max,x}$ and $K_{M,x}$ are the parameters of the *Michaelis Menten* equation relative to the protein $x$. $K_x$ is the time constant relative to the reaction $x$ and $INH$ is the inhibition constant of $dFdCTP_{int}$ with respect to $K_{CMPD}$.

Experimental data used to identify network parameters were taken from [114]. *In vitro* concentrations of dFdC metabolites (extracellular dFdC and dFdU, intracellular dFdC, dFdCMP, dFdCDP, dFdCTP, dFdU and dFdUMP) for two pancreatic cancer cell lines (i.e., PK9 and its resistant version to dFdC, RPK9) were available [114, 129]. Parameters were jointly estimated on both cell lines data by including the ratio of the three target enzymes (CDA, dCK and hENT1) concentrations between the two cell lines as covariates of the model . $COV_{CDA}$, $COV_{dCK}$ and $COV_{hENT1}$, the covariates, were set equal to 1 for PK9 and equal to 1.64, 0 and 1.35 for RPK9, respectively [114]. However, with the available data, it was only

possible to identify the parameters of a reduced metabolic network involving dFdC, dFdCMP, dFdCDP, dFdCTP but not dFdU and dFdUMP. In addition, to fit data, an extracellular dFdC deamination due to the activity of an extracellularly secreted CDA was added. The hypothesis is supported by some observations reported in the literature for other cancer cell lines [130].

To allow writing the mass balance equations, the metabolite profiles were transformed from concentrations ($pmol/mg\ prot$), as reported in the original publication [114], to amount ($pmol$). Information regarding the *in vitro* experiments needed to establish the correction factor was obtained from the original publication and from personal communication by the first author: cells used in the *in vitro* experiment were seeded onto non-coated tissue culture dishes at the concentration of $1.5 \cdot 10^4\ cells/cm^2$ [114] using a 6-well plate (personal communication). So, the well area was set equal to a standard value for a 6-well plate[2] ($9.6\ cm^2$). The cell number per protein amount of PK9 and RPK9 was fixed equal to $10^6$ cells per $mg$ of proteins (personal communication). Original *in vitro* data were then transformed from concentration to mass units by multiplying their values for a correction factor. This correction factor was calculated dividing the total amount of cells (well area times cell density) for the cell number per $mg$ of proteins. The medium volume was calculated dividing the initial amount of extracellular dFdC ($1.63\ nmol/well$) for the solution molarity ($1\ \mu M$). Intracellular volume was calculated multiplying the number of cells in the well (culture area times cell density) for the volume of a pancreatic ductal cell ($200\ fL$ [131]).

### 3.2.2 PBPK model

A PBPK model was developed to describe dFdC distribution and metabolism in the body. Drug specific parameters used in the model are listed in table 3.1. The model consists on fourteen organs and tissues and it is represented in figure 3.1 (a). Each organ and tissue in the PBPK (excluding arterial and

[2]Standard area for a 6-well plate was found in `https://www.thermofisher.com/it/en/home.html`.

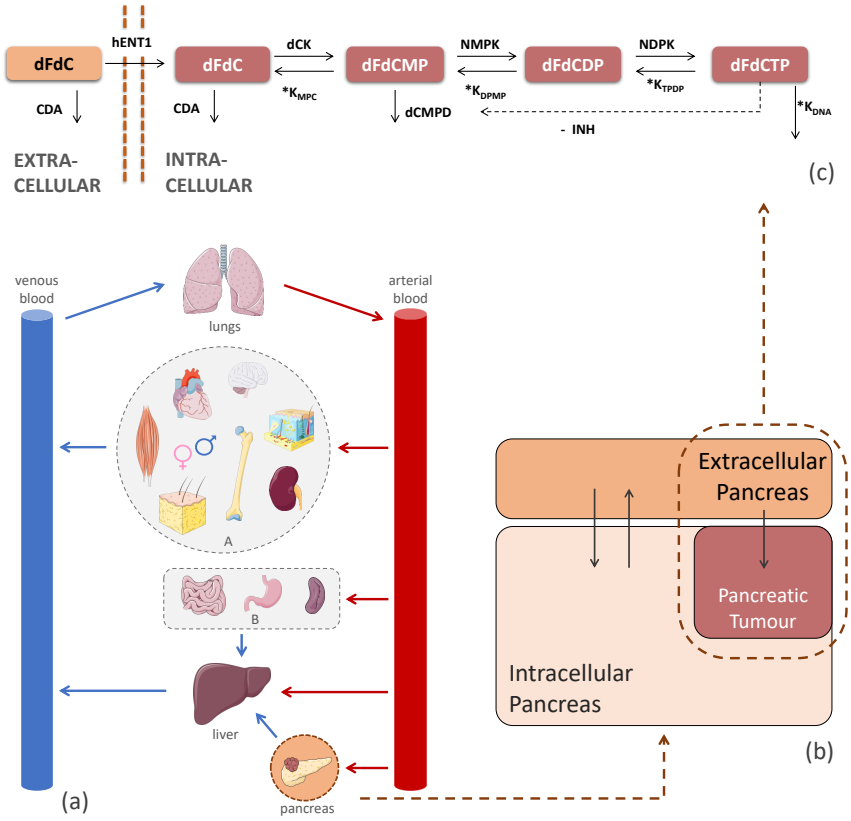



Figure 3.1: Schematic structure of the whole body PBPK model coupled with the metabolic network representing dFdC metabolism in the pancreatic tumour tissue. (a) PBPK model structure. Red arrows represent arterial blood flows, while blue arrows represent venous blood flows. The organs and tissues represented in boxes $A$ and $B$ are 'in parallel' with respect to the blood flow, this means that they have separate blood inflows and outflows. Box $A$: adipose tissue, bone, brain, gonads, heart, kidney, muscle and skin. Box $B$: gut, spleen and stomach. (b) Model structure of the pancreas. Pancreatic tumour and intracellular space share the same extracellular environment. (c) Schematic representation of dFdC metabolism network, including metabolites (dFdC, dFdCMP, dFdCDP, dFdCTP, dFdU and dFdUMP), transporters (hENT1) and target enzymes responsible of driving the metabolism reactions (dCK, NMPK, NDPK, CDA and dCMPD). Reactions catalysed by unknown enzymes were named as $MPC$, $DPMP$ and $TPDP$.

venous blood) was described by using a permeability limited model [132]. This choice was supported by the hydrophilic nature of dFdC hampering distributions into the cells [107, 133].

dFdC is transported inside the cell by concentrative nucleoside transporters (mainly hCNT1) and equilibrative nucleoside transporters (mainly hENT1) proteins [107, 134]. The activity of both hCNT1 and hENT1 was modelled as a first order reaction, namely $R_{hCNT1}$ and $R_{hENT1}$ (equations 3.2 and 3.3). hCNT1 mediates a unidirectional flux from the extracellular to the intracellular space, while hENT1 was considered as a bidirectional transporter. Once inside each organ intracellular space, the drug was supposed to be metabolized by CDA. This process was modelled with first order reaction too ($R_{CDA}$), as in equation 3.4.

$$R_{hCNT1,t} = V_{ext,t}\, e_{hCNT1,t}\, k_{hCNT1}\, \frac{V_{int,t}}{V_{ext,t}}\, Cu_{ext,t} \tag{3.2}$$

$$R_{hENT1,t} = e_{hENT1,t} \Big( V_{ext,t}\, k_{hENT1,in}\, \frac{V_{int,t}}{V_{ext,t}}\, Cu_{ext,t} - V_{int,t}\, k_{hENT1,out}\, Cu_{int,t} \Big) \tag{3.3}$$

$$R_{CDA,t} = V_{int,t}\, e_{CDA,t}\, CL_{CDA}\, Cu_{int,t} \tag{3.4}$$

$e_{x,t}$ is the relative expression of enzyme or transporter $x$ in the tissue $t$. They were taken from the *Open Systems Pharmacology Suite* version 7.1[3] (PK-sim) and are numbers always between 0 and 1; their values are reported in table A.2. $k_x$ is the time constant relative to the protein $x$ activity, assumed to be equal in all the organs. $Cu_{ext,t}$ and $Cu_{int,t}$ are the unbound extracellular and intracellular dFdC concentrations, respectively. $V_{ext,t}$ and $V_{int,t}$ are the extracellular and intracellular volumes of the tissue $t$. They are calculated by multiplying the tissue volume $V_t$ for the extracellular and intracellular water fractions ($f_{ew}$ and $f_{iw}$, respectively): $V_{ext,t} = f_{ew} \cdot V_t$ and $V_{int,t} = f_{iw} \cdot V_t$. Unbound fraction of dFdC was considered equal to 1 [135].

[3] `https://github.com/open-systems-pharmacology`

To account for the different transporters and enzymes expressions on each organ, $k_x$ was multiplied for $e_{x,t}$, as done in [126]. The main hypothesis here is that the intracellular enzymes and transporters concentrations are proportional between the different organs. In equations 3.2 and 3.3 the linear time constants relative to the transport from the extracellular to the intracellular compartment are multiplied for $V_{int,t}/V_{ext,t}$. The time constant $k_x$ could be written as $V_{max}/K_M$, where $V_{max}$ and $K_M$ are the parameters of the *Michaelis Menten* equation. So, the following relationship is valid.

$$k_{x,ext} = \frac{V_{max,ext}}{K_M} = \frac{k_{cat}\,[x]_{ext}}{K_M} = \frac{k_{cat}\,[x]_{int}}{K_M}\,\frac{V_{int,t}}{V_{ext,t}} = k_x \frac{V_{int,t}}{V_{ext,t}} \tag{3.5}$$

$[x]_{int}$ and $[x]_{ext}$ are the intracellular and extracellular transporter concentration and $k_{cat}$ is the turnover number. Equation 3.5 is valid for both hCNT1 and hENT1.

Each organ modelled in the PBPK was described by using two compartments, representing intracellular and extracellular spaces. The generic dFdC tissue extracellular and intracellular unbound concentration dynamics are represented in equation system 3.6.

$$\begin{aligned} V_{ext,t}\,\frac{dCu_{ext,t}}{dt} &= Q_t\left(C_{art} - \frac{Cu_{ext,t}}{P_{t:p}/B:P}\right) - R_{hCNT1,t} - R_{hENT1,t} \\ V_{int,t}\,\frac{dCu_{int,t}}{dt} &= R_{hCNT1,t} + R_{hENT1,t} - R_{CDA,t} \end{aligned} \tag{3.6}$$

$Q_t$ is the tissue blood flow, $C_{art}$ corresponds to the arterial dFdC concentration, $P_{t:p}$ is the tissue to plasma partition coefficient, calculated as in [132] and $B:P$ is the blood to plasma partition coefficient. These equations are valid for all the tissues except arterial and venous blood, lungs and pancreas. After appropriate parameters rescaling, the metabolic network describing dFdC metabolism was included into a compartment representing the pancreatic tumour, as explained in section 3.2.3. All the model equations and parameter values are reported in appendix A.

The four time constants associated with enzymes and transporters activities in the PBPK model ($k_{hCNT1}$, $k_{hENT1,in}$, $k_{hENT1,out}$ and $k_{CDA}$, in

equations 3.2, 3.3 and 3.4) were identified for a male mean subject (height 175 $cm$, weight 73 $kg$ and age 30 years) using the data simulated with a pharmacokinetic model taken from the literature (*Zhang model*, reported in section A.2), for 30 minutes infusion of 3.34 $mmol/m^2$ of dFdC [136].

Finally, by adding variability to the physiological parameters, a population model was obtained. The variabilities in organ volumes and blood flows were modelled following *Willmann et al.* [137]. Briefly, in this model: 1) the sex and age of the subjects are extracted; 2) for each subject the height is extracted from a distribution given the particular sex and age; 3) mean organs weight and blood flows are generated given the mean subject characteristics and 4) a residual variability is added. $k_{hCNT1}$, $k_{hENT1,in}$, $k_{hENT1,out}$ and $k_{CDA}$ were considered variable too in order to account for the different protein expression in a population. They were supposed log-normally distributed with mean equal to the estimated values and coefficient of variation (CV) equal to that of the enzyme concentration in the population [138]. To our knowledge, no information regarding CDA variability in tissues is present in the literature; thus, we decided to fix the CDA CV equal to the one of the pancreatic cancer, obtained from [139]. Distribution parameters are reported in table 3.1.

### 3.2.3 Inclusion of the metabolic network in the PBPK model

A compartment representing the pancreatic tumour cells was included into the PBPK, as shown in figure 3.1 (b). The main hypothesis is that the tumour and the pancreatic intracellular space share the same extracellular environment and they compete for drug uptake. The network was included into the PBPK as follows: *in vitro* intracellular compartment corresponds to the PBPK tumour compartment while the *in vitro* medium corresponds to the pancreatic extracellular space.

The parameters were appropriately rescaled considering the different volumes and enzymatic concentrations between the *in vitro* and *in vivo* situations. In the publication where we took the *in vitro* data, the concentrations of some of the enzymes and transporters involved into dFdC metabolism, like CDA, dCK and hENT1, were reported [114]. In another

publication, the concentration of these enzymes in pancreatic tumour samples from ten different subjects was measured [139]. That information was used for the *in vitro* to *in vivo* rescaling.

In the *in vitro* metabolic network, the metabolites were considered in units of mass ($pmol$). So, the units of the estimated $V_{max}$ is $pmol/h$, whereas that of $K_M$ is $pmol$ and of the time constants is $1/h$.

$V_{max,\ dCK}$ can be expressed as:

$$V_{max,dCK} = k_{cat,dCK}\ [dCK]_{int,vitro}\ V_{int,vitro} \tag{3.7}$$

where $k_{cat,dCK}$ is the turnover number, $[dCK]_{int,vitro}$ is the enzymatic intracellular concentration and $V_{int,vitro}$ is the *in vitro* culture volume. In order to rescale $V_{max,dCK}$ (*in vitro*) to $\widetilde{V}_{max,dCK}$ (*in vivo*), one has to consider that the *in vitro* and *in vivo* situations have different volumes and enzymatic concentrations. Thus, $\widetilde{V}_{max,dCK}$ could be obtained from $V_{max,dCK}$ as follows.

$$\begin{aligned} \widetilde{V}_{max,dCK} &= k_{cat,dCK}\ [dCK]_{int,vivo}\ V_{int,vivo} \\ &= V_{max,dCK}\ \frac{[dCK]_{int,vivo}}{[dCK]_{int,vitro}}\ \frac{V_{int,vivo}}{V_{int,vitro}} \end{aligned} \tag{3.8}$$

However, enzymes concentrations were available as $pmol/mg\ prot$ [114, 139], and not in $pmol/mL$, as required in the previous equations. Thus, we made the hypothesis that the *in vivo*/*in vitro* enzymatic concentration ratio expressed in $pmol/mg\ prot$ is equal to the one expressed in $pmol/mL$.

$K_{M,dCK}$ is expressed in the metabolic network in units of mass, thus, it is valid the following relationship.

$$K_{M,dCK} = K_{M,dCK,conc}\, V_{int,vitro} \tag{3.9}$$

$K_{M,dCK,conc}$ is the $K_M$ expressed in concentration, as generally it is. $\widetilde{K}_{M,dCK}$ could be obtained from $K_{M,dCK}$ as follows.

$$\widetilde{K}_{M,dCK} = K_{M,dCK}\ \frac{V_{int,vivo}}{V_{int,vitro}} \tag{3.10}$$

The same rationale can be followed for the reaction catalysed by intracellular CDA. $V_{max,hENT1}$ was rescaled like $V_{max,dCK}$. This was done

because, even if the substrate is present in the extracellular compartment, the transporter abundance depends on the intracellular volume. In fact, is valid the following equation:

$$\begin{aligned} V_{max,hENT1} &= k_{cat,hENT1}\ [hENT1]_{ext,vitro}\ V_{ext,vitro} \\ &= k_{cat,hENT1}\ [hENT1]_{int,vitro}\ V_{int,vitro} \end{aligned} \tag{3.11}$$

where $[hENT1]_{ext,vitro}\, V_{ext,vitro} = [hENT1]_{int,vitro}\, V_{int,vitro}$ is the total amount of transporter in the system. Considering that the substrate of hENT1 is in the extracellular environment, $K_{M,hENT1}$ was rescaled as follows.

$$\widetilde{K}_{M,hENT1} = K_{M,hENT1}\, \frac{V_{ext,vivo}}{V_{ext,vitro}} \tag{3.12}$$

The extracellular CDA concentration was supposed equal in both *in vitro* and *in vivo* situation, thus, the $V_{max,CDA}$ was rescaled only by using the ratio of the extracellular volumes: $\widetilde{V}_{max,CDA} = V_{max,CDA}\, V_{ext,vivo}/V_{ext,vitro}$. $K_{M,CDA,ext}$ was rescaled as $K_{M,hENT1}$.

Finally, the concentration of the enzymes catalysing all the other reactions were supposed to be equal between the *in vitro* and the *in vivo* situations. Thus, all the linear constants were not rescaled between the two systems. This was done because, with the hypothesis of equal concentration of the enzymes catalysing the reaction in the two situations, the following relationship is valid.

$$K_r = \frac{V_{max}}{K_M} = \frac{\widetilde{V}_{max}}{\widetilde{K}_M} = \widetilde{K}_r \tag{3.13}$$

Concerning $INH$, we made the hypothesis that the inhibition depends on the metabolite concentration. Thus, from the *in vitro* to the *in vivo* situation the parameter was corrected for a factor equal to $V_{int,vivo}/V_{int,vitro}$.

All the *in vivo* enzymatic concentrations were supposed log-normally distributed with the mean and CV derived from the pancreatic tumour samples [139]. The tumour volume was supposed uniformly distributed between 32.3 $mL$ and 224.3 $mL$ [140].

### 3.2.4 GSA for the mechanistic Gemcitabine model

A variance-based GSA was performed on the gemcitabine PBPK model coupled with the metabolic network. AUC of plasma dFdC and tumour dFdCTP concentrations were considered as outputs of interest. AUC was calculated from time 0 (dose administration) to 7 days. The parameters that were considered variables in the population are: the sex, age and height, the residual variability of the organs volumes and blood flows, the dFdC blood to plasma ratio, all the estimated time constants associated with drug transport and elimination in the PBPK, the tumour volume and the tumour concentrations of the enzymes involved in dFdC metabolism. Given that $k_{hENT1,in}$ and $k_{hENT1,out}$ are related to the activity of the same enzyme, in the GSA they were jointly considered (grouped): they shared the same variability and so they were considered as a unique parameter ($k_{hENT1}$). For readability purposes, all the organ volumes and blood flows residual variabilities were grouped too [21]. The distributions of all the parameters are reported in table 3.1.

The number of samples, $n$, extracted in the GSA was set to 5000. The uncertainty of the sensitivity indices were calculated using 10000 bootstrap samples [67].

Table 3.1: Drug related parameters and parameters distributions for GSA

| **Parameters** | distribution parameters | distribution type | units |
|---|---|---|---|
| $pKa$ [135] | 3.6 | fixed | |
| $B:P$ [141] | 1.94 | fixed | |
| $f_{u,p}$ [135] | 1 | fixed | |
| molecular weight [135] | 299.66 | fixed | g/mol |
| $logP_{ow}$ [135] | -1.4 | fixed | |
| sex[c,d] | 0, 1 | uniform[a] | |
| age[d] | 20, 65 | uniform[a] | $years$ |
| $E:P$[e] | 1, 5 | uniform[a] | |
| tumour volume | 32.3, 224.3 | uniform[a] | $mL$ |
| $k_{hCNT1}$ | 920.17 (33%) | log-normal[b] | $1/min$ |
| $k_{hENT1,in}$ | 20.17 (24.3%) | log-normal[b] | $1/min$ |
| $k_{hENT1,out}$ | 25.66 (24.3%) | log-normal[b] | $1/min$ |
| $k_{CDA}$ | 0.33 (109.6%) | log-normal[b] | $1/min$ |
| $[dCK]_{int,vivo}$ | 0.45 (20%) | log-normal[b] | $pmol/mg\ prot$ |
| $[hENT1]_{int,vivo}$ | 3.08 (53.4%) | log-normal[b] | $pmol/mg\ prot$ |
| $[CDA]_{int,vivo}$ | 0.67 (109.6%) | log-normal[b] | $pmol/mg\ prot$ |

[a] For distribution parameters, *minimum*, *maximum* of the parameter.
[b] For distribution parameters, *mean (CV)* of the log-normal variable.
[c] If the extracted value is <0.5 the subject is female (0), otherwise male (1).
[d] Height, organ volumes and blood flows were generated by using the Willmann model [137] and are function of sex and age.
[e] $B:P$ calculated from $E:P$ values [141], as $E:P = 1/H \cdot (B:P - 1 + H)$ [39].

Table 3.2: Metabolic network parameters

| **Parameters** | value | units |
|---|---|---|
| $V_{max,dCK}$ (x$COV_{dCK}$) | $1.45{\cdot}10^5$ | $pmol/h$ |
| $K^a_{M,dCK}$ | 4.6 | $\mu M$ |
| $V_{max,CDA,int}$ (x$COV_{CDA}$) | $1.1{\cdot}10^5$ | $pmol/h$ |
| $K^a_{M,CDA,int}$ | $0.43{\cdot}10^5$ | $pmol$ |
| $V_{max,CDA,ext}$ (x$COV_{CDA}$) | $0.84{\cdot}10^5$ | $pmol/h$ |
| $K^a_{M,CDA,ext}$ | $4.2{\cdot}10^5$ | $pmol$ |
| $V_{max,hENT1}$ (x$COV_{hENT1}$) | 5.46 | $pmol/h$ |
| $K^a_{M,hENT1}$ | 4.7 | $pmol$ |
| $K_{NMPK}$ | $0.11{\cdot}10^5$ | $1/h$ |
| $K_{DCMPD}$ | $0.027{\cdot}10^5$ | $1/h$ |
| $K_{DPMP}$ | $0.14{\cdot}10^5$ | $1/h$ |
| $K_{DNA}$ | $1{\cdot}10^{-7}$ | $1/h$ |
| $K_{INH}$ | $0.8{\cdot}10^5$ | $1/pmol$ |

[a] Value taken from [142]. To include it into the model it was multiplied for the intracellular volume.
[b] Given that its value is significantly less than the other rate constants, it was set to 0 without an impact on the simulations.

## 3.3 Results

### 3.3.1 *In vitro* metabolic network

The parameters of the metabolic network were identified on the *in vitro* data from [114]. The estimated parameter values are listed in table 3.2, where it is also indicated for which parameter the covariates were included. In order to reduce the number of parameters to identify, $K_{NMPK}$, $K_{NDPK}$ were considered equal and $K_{DPMP}$, $K_{TPDP}$, $K_{MPC}$ were considered equal too. In figure 3.2, the results of the fitting process for each metabolite profile and each pancreatic cell line are shown. In the PK9 cell line, $R^2$ values for extracellular dFdC, intracellular dFdC, dFdCMP, dFdCDP and dFdCTP are equal to 0.98, 0.1, 0.99, 0.61 and 0.87, respectively. The metabolic network was believed sufficiently capable of describing the time course of extracellular and intracellular concentrations of dFdC and its phosphorylated metabolites (dFdCMP, dFdCDP and dFdCTP), for the two different cell lines.

### 3.3.2 PBPK model coupled with the *in vivo* metabolic network

We fitted the PBPK model coupled with the reduced metabolic network against a typical plasma profile of dFdC, given a single dose of 3.34 $mmol/m^2$ infused in 30 minutes (standard administration in the clinical setting), simulated with the Zhang model [136]. The identified parameters are reported in table 3.1 (mean values of the time constants) and the fitting results are shown in figure 3.3, panels a and b. It is possible to observe that the PBPK model well reproduces the typical subject profiles provided by the Zhang model. From the logarithmic scale it can be appreciated that the elimination rate is well captured. The value of $R^2$ is equal to 0.94.

Once the model parameters were identified on the typical profile, the PK profiles of a population of 500 individuals were simulated. In figure 3.3, panels c, d, e and f the plasmatic dFdC concentration profiles and the dFdC and its metabolites pancreatic tumour profiles are shown.

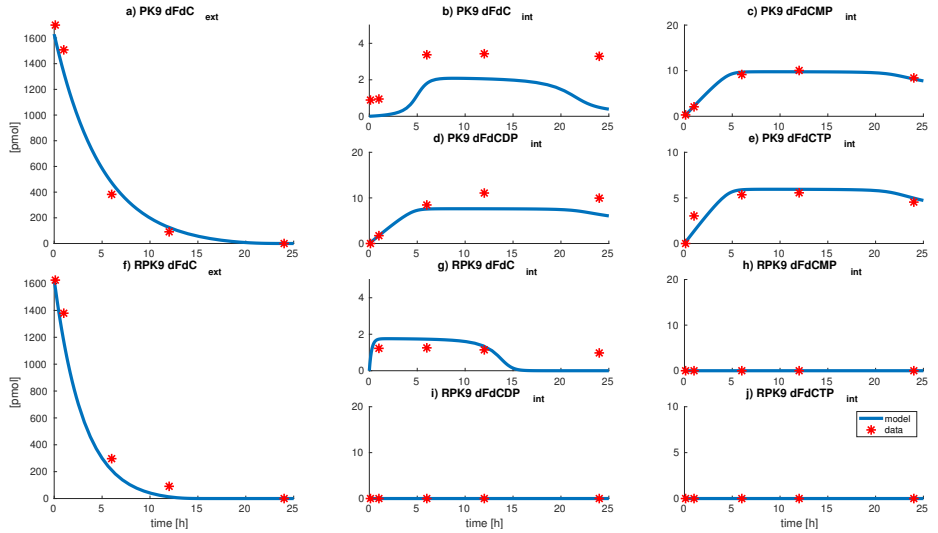



Figure 3.2: Fitting results of the *in vitro* metabolic network for PK9 and RPK9 pancreatic cancer cell lines. Blu lines are the model predictions and red stars the data from [114]. In panel a, the results for dFdC medium amount are reported and in panels b, c, d and e the results for dFdC, dFdCMP, dFdCDP and dFdCTP intracellular amounts are reported for the PK9 cell line. In panel f, the results for dFdC medium amount are reported and in panels g, h, i and j the results for dFdC, dFdCMP, dFdCDP and dFdCTP intracellular amounts are reported for the RPK9 cell line.

In figure 3.3, it is possible to observe that while the dFdC plasma concentration drops to zero for at least the 95% of the subjects in almost 24 hours, the metabolite concentrations in tumour decreases much slowly. These results are qualitatively in agreement with the simulations obtained by the Zhang model, that predicts a drop to zero in 70 hours for the typical value of the dFdCTP concentration in the white blood cells (WBC), used as a surrogate of the intracellular dFdCTP concentration.

In table 3.3 the metrics obtained with the PBPK model for plasmatic dFdC and tumour dFdCTP concentrations, together with those obtained with the Zhang model for plasmatic dFdC and WBC dFdCTP concentrations, are reported. Concerning the dFdC plasmatic AUC, the results of both the models show good agreement. However, the dFdCTP tu-

mour AUC predicted with the PBPK results slightly lower with respect to dFdCTP WBC intracellular concentration AUC predicted with the Zhang model. It is possible to observe that the PBPK model overestimated the population variability of all the metrics (especially the dFdC and dFdCTP AUC).

As a further simulation exercise, 3.34 $mmol/m^2$ of dFdC were infused weekly for 20 weeks; results are shown in section A.3. In figure A.1, it is possible to see that there is no accumulation of dFdC and dFdCTP, in agreement with the observations presented in [143].

### 3.3.3 GSA results

We performed a variance-based GSA on the PBPK model coupled with the metabolic network, with the aim of understanding what are the most important parameters in explaining plasma dFdC and tumour dFdCTP concentrations AUC variability in the population. Results are shown in figure 3.4.

The parameter that mainly explains the dFdC plasmatic concentration AUC is the time constant relative to the dFdC elimination in tissues, $k_{CDA}$. A critical aspect related with this parameter is that we have considered its CV equal to the one found in pancreatic tumour tissue samples. This was done because, to our knowledge, a value relative to the other tissues was not present in the literature. Given that the variation of $k_{CDA}$ explains almost the totality of the dFdC AUC variance, a better characterization of its variability in the population is needed for a more reliable prediction of the AUC variability.

dFdCTP tumour concentration AUC variability is mainly due to dCK and CDA tumour concentrations. Between dCK and CDA tumour concentrations, the most important one in determining the dFdCTP AUC is the former. This is in agreement with the fact that the resistance against dFdC in some cell lines is obtained by reducing the dCK levels [114].

It is interesting to notice that hENT1 tumour expression is not important in determining dFdCTP AUC. This could be due to the fact that hENT1 concentration was quite homogeneous in the population that we

used to estimate its variability [139]. Thus, it is possible that with these data, the hENT1 population variability was underestimated.

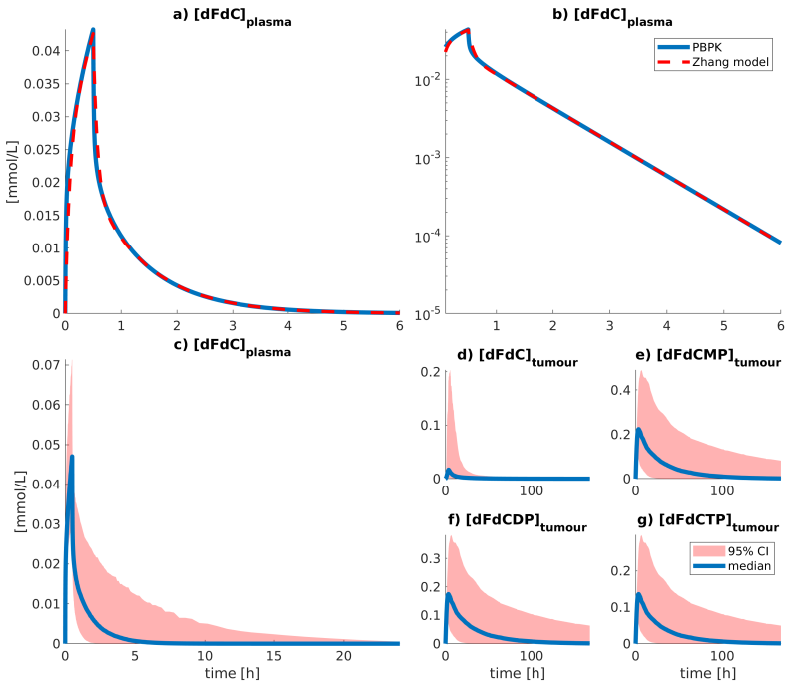



Figure 3.3: Fitting and population simulation results of the PBPK model. In panel a and b the results of the fitting process are shown, in natural and semi-logarithmic scale, respectively. Continuous blue line represent the dFdC plasma concentration simulated by using the PBPK model, while dashed red line represent the dFdC plasma concentration simulated with the Zhang model [136]. Panels c, d, e and f and g show the results of the simulation of dFdC and its metabolites pharmacokinetics, obtained by using the PBPK model coupled with the metabolic network. In panel c the dFdC plasma concentrations are reported. In panels d, e, f and g the tumour dFdC, dFdCMP, dFdCDP and dFdCTP concentrations are reported, respectively. In this case, blue line represents the median of the compound concentrations in the population, while red shaded area represents the 95% confidence interval.

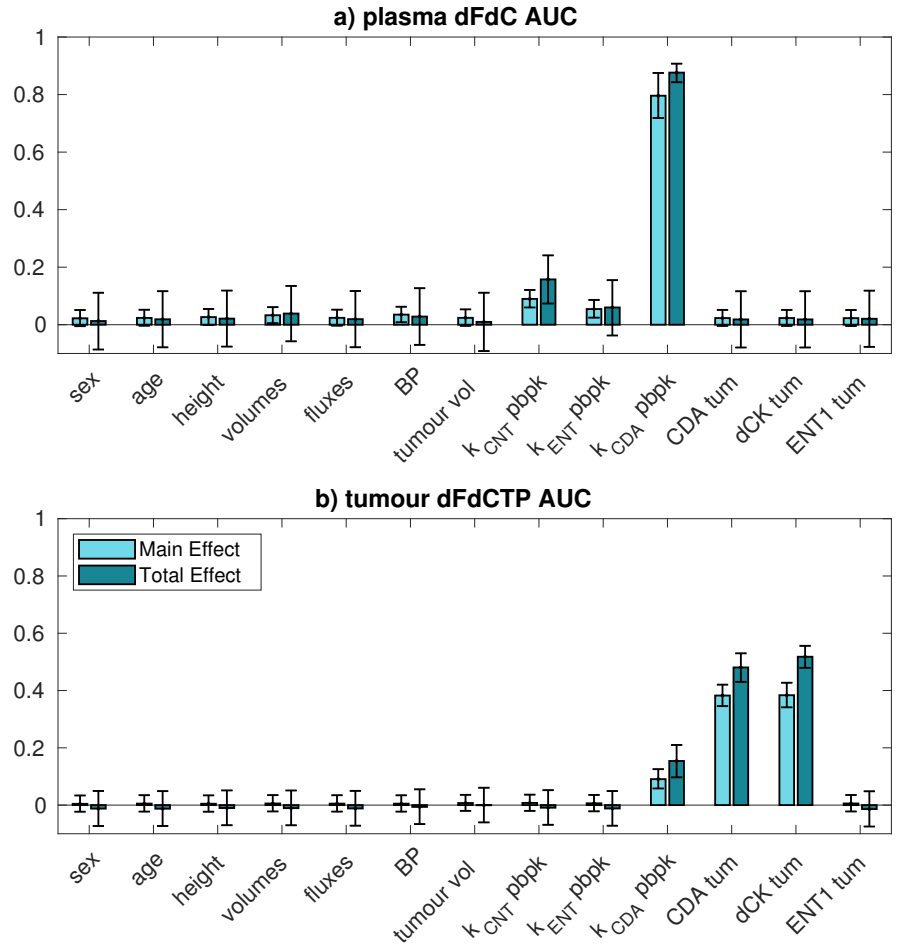



Figure 3.4: GSA results performed for: a) plasma dFdC AUC; b) tumour dFdCTP AUC. The parameters that are considered variables are: sex, age, residual variability on height, organ volumes and blood flows (height, volumes and fluxes), blood to plasma ratio (BP), tumour volume (tumour vol), all the estimated linear constants relative to the transport and elimination of dFdC in the PBPK model ($k_{CNT}$ pbpk, $k_{hENT1}$ pbpk and $k_{CDA}$ pbpk) and the enzymes tumour concentrations (CDA tum, dCK tum and hENT1 tum). Error bars represent the 95% confidence interval of the sensitivity indices calculated with 10000 bootstrap samples.

Table 3.3: PBPK and *Zhang* model metrics

| **Metrics** | mean subject | mean | median | 2.5 perc | 97.5 perc | CV (%) | units |
|---|---|---|---|---|---|---|---|
| | | | | PBPK | | | |
| Plasma dFdC AUC | 2.19 | 3.57 | 2.81 | 1.14 | 11 | 78.06 | $mmol \cdot min/L$ |
| Plasma dFdC $C_{max}$ | 0.04 | 0.048 | 0.047 | 0.03 | 0.07 | 22.99 | $mmol/L$ |
| Tumour dFdCTP AUC | 201.28 | 343.86 | 252.85 | 31.18 | 1127.3 | 94.36 | $mmol \cdot min/L$ |
| Tumour dFdCTP $C_{max}$ | 0.16 | 0.15 | 0.14 | 0.07 | 0.3 | 42.99 | $mmol/L$ |
| | | | | *Zhang model*[a] | | | |
| Plasma dFdC AUC | 2.17 | 2.21 | 2.07 | 1.22 | 3.59 | 28.92 | $mmol \cdot min/L$ |
| Plasma dFdC $C_{max}$ | 0.043 | 0.041 | 0.041 | 0.024 | 0.06 | 22.14 | $mmol/L$ |
| WBC dFdCTP AUC[b] | 545.1 | 564.95 | 528.52 | 331.25 | 931.17 | 29.12 | $mmol \cdot min/L$ |
| WBC dFdCTP $C^b_{max}$ | 0.43 | 0.44 | 0.42 | 0.26 | 0.68 | 25.75 | $mmol/L$ |

[a] 500 subjects were simulated (250 male and 250 female, age 29) including inter-subject variability reported in Zhang's model. Mean subject metrics were calculated for a 29 year old man.
[b] Values of dFdCTP AUC and $C_{max}$ were simulated with the Zhang model in $pmol \cdot h/10^6 cells$ and $pmol/10^6 cells$, respectively. These values were converted to intracellular WBC concentration by dividing them for the mean volume of a neutrophil, equal to 299 $fL$ [144].

## 3.4 Discussion

We developed a multi-scale systems pharmacology model describing the dFdC metabolic pathway and predicting the levels of dFdCTP in the active site for a population of pancreatic cancer patients. This model was built by integrating different resources obtained from the literature: we used *in vitro* information regarding dFdC metabolism in pancreatic cancer cell lines [114], a compartmental model describing plasma dFdC concentrations in pancreatic cancer patients [136] and physiological and genetical information of a given population [137, 139]. Finally, we performed a GSA in order to understand what parameters explain the predicted population variation of plasma dFdC and tumour dFdCTP AUC.

Regarding the dFdC *in vitro* metabolism, data used for developing the network were obtained from *in vitro* experiments performed after a single dose exposure of gemcitabine, collecting a single profile per metabolite for each cell line. With the data available, the results of the fitting process are biased, as it can be appreciated by looking at figure 3.2. Despite of it, the metabolic network was believed sufficiently capable of describing the profiles of dFdC and its phosphorylated metabolites. The development of this model presented several difficulties. First, important information regarding the experimental setup were not present in the original publication of the *in vitro* data, like the *mg* of protein per number of cells and the area of the well in which the cells were cultured. Moreover, it was not possible to describe the profiles of dFdU and its metabolites. Given that the metabolic network structure depends, to a certain extent, on the data available, a kind of *structural uncertainty* of the model is present. This uncertainty could potentially impact the *in vivo* predictions of dFdCTP concentration once the network is included in the PBPK model.

To model dFdC distribution and metabolism in the body, a permeability limited PBPK model was developed accounting for the activity of plasmatic membrane transporters [107, 133]. One of the main advantages of the developed PBPK model is that the metabolic network was easily coupled with the model, leading to the possibility of describing the active metabolite concentration in the site of action. In order to do this, a

compartment representing the pancreatic tumour was introduced into the model and was supposed to share the same extracellular environment with the pancreas sane tissue (figure 3.1 c). By doing this, there was a direct correspondence between *in vitro* and *in vivo* intracellular and extracellular environments. In this context, information regarding the target enzymes concentration both in *in vitro* cancer cell lines and *in vivo* tumour samples, was found to be particularly useful for the parameters rescaling and thus, the inclusion of the metabolic network in the PBPK model.

With the model presented here, there was no need of estimating the blood flow directed to the tumour. This approach presents, however, some drawbacks. In fact, by using this model it would be difficult to describe processes such as the angiogenesis and the effect of a potential antiangiogenic compound. Moreover, the thick stroma surrounding the tumour cells that characterize the pancreatic cancer was not modelled [145] and this could potentially impact the predicted drug disposition in the tumour tissue.

Another advantage of the developed systems pharmacology model is that it includes the interpatient variability of parameters such as organ volumes, blood flows and abundances of enzymes and transporters. Then, by performing GSA, it is possible to understand what are the parameters that with their variation in the population mostly explain the inter-patient variability of some metrics of interest, such as AUC of plasma dFdC and tumour dFdCTP. The GSA results highlight that the tumour dFdCTP AUC variability is mainly explained by the variation of CDA and dCK tumour concentration. These results are in accordance with the fact that the dFdC clinical response is probably related to the patients genotype and to different expressions of the transporters or target enzymes [111].

The results of this modelling study suggest that individual genetic factors affecting gemcitabine metabolism would lead to different amounts of its metabolites and, consequently, different treatment responses, as dFdCTP exposure has been previously related to tumour response, and the later, to survival [124]. Apart from the genetic variability associated with gemcitabine's metabolism pathway highlighted in this work, different individual mutations affecting its mechanism of action regarding cell cycle progression, apoptosis and survival signalling pathways in pancreatic cancer cells

can also have an impact on treatment response. A recent multiscale network characterizes the effect of proteomics on gemcitabine mechanism of action and its signalling pathways, in combination with birinapant [120]. Future integration of their results with those present in this study could provide insights to better understand gemcitabine variability and drug effects.

In conclusion, further research should be done for characterizing *in vitro* different pancreatic cancer cell coming from patients receiving dFdC, measuring the target enzyme level expression and the degree of their polymorphisms. This would be key to assess and refine the current model.

Chapter 4

# GSA to gain insight into the structure of physiological intestinal absorption models for BCS I-IV drugs [1]

## 4.1 Introduction

The oral route is the preferred method of drug administration, mainly because of its convenience and minimal invasiveness. However, the bioavailability of drugs (i.e., the fraction that reaches the systemic circulation unchanged) is limited by several processes such as dissolution and absorption in the gut lumen, metabolism in the gut wall and liver [54].
In order to facilitate the development of oral formulations the *Biophar-*

[1] This work was published in "N. Melillo, L. Aarons, P. Magni, A.S. Darwich. Variance based global sensitivity analysis of physiologically based pharmacokinetic absorption models for BCS I-IV drugs. *Journal of Pharmacokinetics and Pharmacodynamics*, 46(1):27-42, February 2019" [146].

*maceutics Classification System* (BCS) was created [147]. The BCS uses physicochemical and physiological parameters to classify drugs into four different classes based on their permeability and solubility characteristics: class I (highly permeable, highly soluble); class II (highly permeable, lowly soluble); class III (lowly permeable, highly soluble); and class IV (lowly permeable, lowly soluble). The BCS is widely used by the *European Medicines Agency* (EMA) and United States *Food and Drug Administration* (FDA) for developing guidance on formulation development and by the pharmaceutical industry during drug discovery & development [148]. Although considered an oversimplification of complex drug and formulation characteristics, the BCS is useful for informing experimental and clinical design, especially for class I compounds [54, 149].

Considerable efforts have been carried out to combine *in silico* mathematical modelling with the design and evaluation of experimental studies to reduce the number of *in vivo* bioequivalence studies needed, therefore reducing time and cost of biopharmaceutical development [150]. Among various types of *in silico* modelling techniques, PBPK models have been used to investigate complex biopharmaceutical problems [54].

Several PBPK absorption models have been developed over the last decades and integrated into bespoke PBPK software (such as: *GastroPlus*, *PK-Sim* and *Simcyp Simulator*) or more general modelling platforms, such as MATLAB [151, 152, 153, 154, 155, 156]. In general, these represent drug transit through the small intestine, release from formulation, dissolution/precipitation and absorption in the gastrointestinal tract, gut wall metabolism and active efflux/uptake transport. PBPK absorption models are used from lead optimization through phase 2 studies. For example, during lead optimization physiological models can be used to predict absorption from in vitro data. Moreover, these models are used to predict drug absorption in humans in combination with animal data obtained during pre-clinical development. Such predictions are possible because of the incorporation of physiological and biochemical differences between species. During clinical development, physiological models can be used to mechanistically interpret clinical data, to explore hypotheses and to guide formulation development [43]. The use of PBPK models has the potential

to reduce the number of animal studies and replace or supplement clinical trials [35, 36].

The *OrBiTo* (Oral Biopharmaceutics Tools) project (Innovative Medicines Initiative, IMI), started in 2012 and aimed to address the gaps in gastrointestinal drug absorption knowledge and support a rational use of predictive tools for oral drug delivery. This was done by refining existing tools and defining new methodologies for oral drug delivery [157]. One of the various objectives of OrBiTo was to perform a large scale evaluation of PBPK models for oral drug absorption, to identify strengths and weaknesses of these models. The results of the analysis showed high variability in the performance [53, 55].

We believe that a better comprehension of the relationship between the model input parameters (e.g., drug/formulation-specific and physiological parameters) and outputs (e.g., drug exposure and secondary pharmacokinetic parameters) would be useful for the development and refinement of PBPK models. Performing a sensitivity analysis is useful for understanding how the uncertainty in input parameters translates to uncertainty in the outputs and, by this, identifying the most important parameters for a given output [21]. PBPK models have a complex structure and, usually, a significant variability in input parameters, for example, the variability that occurs in a given population, for parameters such as the gastric emptying time, the intestinal transit time and the enzymatic liver expression. Often there is a significant uncertainty in the estimation of some of these inputs, where the parameters are typically fixed to mean values or fitted to experimental data. Depending on the knowledge and information available, these parameters could vary within a certain defined range of values [74]. Thus, for these types of models, it is appropriate to perform GSA. Furthermore, there is currently a strong regulatory interest from EMA and FDA in the use of sensitivity analysis to evaluate PBPK models in pharmaceutical research & drug development and in regulatory submissions [50, 149].

In this context, the aim of our work was to give a demonstration of the GSA methodology, by applying it on compartmental PBPK models that describe drug absorption, dissolution and transit in the gastrointestinal tract. This was done in order to identify what are the most important

physiological and drug related parameters in determining the variability of the fraction absorbed ($f_a$) and bioavailability ($F_{oral}$) within each BCS class, for acidic, basic and neutral drugs, after an oral administration. Between various methods for GSA we choose the variance based method because it is model independent, considers each parameter in its full range of variation and allows estimation of the interaction effects between input parameters [21]. The analysis was firstly done for neutral compounds on a mixing tank derived model [158] and then on a compartmental absorption and transit (CAT) derived model [159] for acidic, basic and neutral compounds. The GSA was performed separately for each BCS class because we expected that the order of importance of the parameters (e.g., relative to dissolution and absorption) could vary among classes.

## 4.2 PBPK intestinal absorption models and characterization of parameters uncertainty

### 4.2.1 PBPK absorption models

Two different compartmental PBPK absorption models with different levels of detail were implemented in MATLAB, both aiming to describe the oral absorption process. One model was based on the mixing-tank model [158], describing drug dissolution and absorption in the gastrointestinal tract, where the small intestine was represented by one well-stirred luminal segment. The other model was based on the CAT model [159] and described drug transit, dissolution and absorption in the gastrointestinal tract. All the models parameters are presented in tables 4.1, 4.3 and 4.2.

#### Compartmental Absorption and Transit based model

In the CAT based model, represented in figure 4.1, the gastrointestinal tract is subdivided into eight different sections: the stomach, six small intestine segments (one for the duodenum, two for the jejunum, three for the ileum) and one for the large intestine. In the gut lumen, drug can be present in two states: solid and dissolved. It is supposed that absorption occurs only

of dissolved drug in the small intestine. Drug is absorbed from the small intestine into the enterocytes where it can be metabolised or transported to the liver via the blood flow. Once in the liver the drug can be metabolised or reach the systemic circulation via the hepatic vein. Equation system 4.1 represents the dissolution and the transit out of the stomach. Equation system 4.2 describes the processes of transit, dissolution, absorption and metabolism that occur in the small intestine, in the enterocytes and in the liver. The large intestine is modelled as a sink, receiving input from the third section of the ileum. Equation system 4.3 represents the dynamics of the system output.

$$\begin{aligned}
\frac{dA_{st,s}}{dt} &= -k_{t,0}\,A_{st,s} - K_{st}\,A_{st,s} \\
\frac{dA_{st,d}}{dt} &= K_{st}\,A_{st,s} - k_{t,0}\,A_{st,d} \\
K_{st} &= \frac{3D}{\rho hr}\left(C_{s,st} - \frac{A_{st,d}}{V_{st}}\right) \\
C_{s,st} &= C_s\frac{\alpha_{st}}{\alpha_{ref}}
\end{aligned} \tag{4.1}$$

$$\begin{aligned}
\frac{dA_{i,s}}{dt} &= k_{t,i-1}\,A_{i-1,s} - k_{t,i}\,A_{i,s} - K_i\,A_{i,s} \\
\frac{dA_{i,d}}{dt} &= k_{t,i-1}\,A_{i-1,d} - k_{t,i}\,A_{i,d} - k_{a,i}\,A_{i,d} + K_i\,A_{i,s} \\
\frac{dA_{i,ent}}{dt} &= k_{a,i}\,A_{i,d} - CL_{ent,i}\frac{A_{i,ent}}{V_{ent,i}} - Q_{ent,i}\frac{A_{i,ent}}{V_{ent,i}} \\
\frac{dA_{liv}}{dt} &= -CL_{liv}\frac{A_{liv}}{V_{liv}} - Q_{HV}\frac{A_{liv}}{V_{liv}} + \sum_{i=1}^{6} Q_{ent,i}\frac{A_{i,ent}}{V_{ent,i}} \\
K_i &= \frac{3D}{\rho hr}\left(C_{s,i} - \frac{A_{i,d}}{V_i}\right) \\
C_{s,i} &= C_s\frac{\alpha_i}{\alpha_{ref}} \\
i &= 1, ..., 6
\end{aligned} \tag{4.2}$$

$$\begin{aligned} \frac{dA_a}{dt} &= \sum_{i=1}^{6} k_{a,i}\, A_{i,d} \\ \frac{dA_{oral}}{dt} &= Q_{HV} \frac{A_{liv}}{V_{liv}} \end{aligned} \tag{4.3}$$

$A_{st,s}$, $A_{st,d}$, $A_{i,s}$, $A_{i,d}$ and $A_{i,ent}$ are the amount of solid and dissolved drug in stomach, in the $i$-th compartment of the small intestine and the amount of drug in the $i$-th enterocytic compartment ($i = 1 \dots 6$), respectively. $A_{liv}$, $A_a$ and $A_{oral}$ are the amount of drug in liver, the total amount of absorbed drug and the total amount of drug that reaches the systemic circulation. $V_{st}$, $V_i$, $V_{ent,i}$ and $V_{liv}$ represent the volume of the stomach, of the $i$-th compartment of the small intestine, of the $i$-th compartment of the enterocytes and of the liver. $k_{t,0}$ is the time constant for the drug output from the stomach and is calculated as the inverse of the gastric emptying time ($GET$). $k_{t,i}$ with $i = 1, ..., 6$ is the time constant for the $i$-th small intestine compartment and is calculated as $k_{t,i} = (SITT \cdot l_i / l_{tot})^{-1}$ where $SITT$ is the small intestinal transit time, $l_i$ is the small intestine segment length and $l_{tot}$ is the total length of the small intestine. $k_{a,i}$ is the absorption constant of the $i$-th compartment of the small intestine and is calculated from the effective jejunal permeability ($P_{eff}$) as $k_{a,i} = 2\, P_{eff} / R_i$ [157], where $R_i$ is the radius of the intestinal compartment. $Q_{HV}$ and $Q_{ent,i}$ are the hepatic vein and $i$-th enterocyte compartment blood flow, respectively. Linear metabolic clearance occurs in each enterocyte compartment and in the liver and is implemented as a function of regional cytochrome P450 3A4 (CYP3A4) abundance. The expression for the clearance in each enterocytes compartment ($CL_{ent,i}$) and in the liver ($CL_{liv}$) are represented in equation 4.4 and 4.5.

$$CL_{ent,i} = CL_{int} \cdot A_{3A4,ent_i} \tag{4.4}$$

$$CL_{liv} = CL_{int} \cdot C_{3A4,LM} \cdot MPPGL \cdot W_{liv} \tag{4.5}$$

$CL_{int}$ is the intrinsic clearance and $A_{3A4,ent_i}$ is the amount of CYP3A4 in the $i$-th enterocytes compartment, obtained by multiplying the CYP3A4

total amount in enterocytes for the proportion of CYP3A in each compartment. $C_{3A4,LM}$ is the concentration of CYP3A4 per $mg$ of microsomal proteins, $MPPGL$ is the amount of microsomal protein per gram of liver and $W_{liv}$ is the liver weight in grams [160].

$K_{st}$ and $K_i$ are the drug dissolution rates in the stomach and in the $i$-th section of the small intestine, described by the *Noyes-Whitney* model. $D$ is the drug diffusion coefficient and is calculated from the Stokes-Einstein equation:

$$D = \frac{k_b T}{6\pi\, \eta_w\, R_h}, \tag{4.6}$$

where $k_b$ is the *Boltzmann* constant, $T$ is the absolute temperature of the body in Kelvin, $\eta_w$ is the dynamic viscosity of water at 37°C and $R_h$ is the hydrodynamic radius of the diffusing drug. $R_h$ is calculated as in equation 4.7, assuming the drug molecule is spherical in shape [161].

$$R_h = \sqrt[3]{\frac{3\ mw}{4\pi\, N_A\, \rho}} \tag{4.7}$$

$\rho$ is the density of the drug particle, $mw$ the molecular weight and $N_A$ is Avogadro's number. In the Noyes-Whitney model $r$ is the particle radius of the formulation and $h$ the effective thickness of the hydrodynamic diffusion layer. $h$ is calculated from $r$ by the *Hintz and Johnson* model as in [162, 163]: $h = r$ if $r < 30\mu m$, otherwise $h = 30\mu m$. $C_s$ is the drug solubility, and $\alpha$ is defined using the *Henderson Hasselbalch* equation using the $pKa$ of the drugs and a pH equal to 6 for $\alpha_{ref}$ and equal to the pH of the $i$-th section of the gastrointestinal tract for $\alpha_i$. $\alpha$ for acids and bases are shown in equations 4.8 and 4.9.

$$\alpha_{acid} = 1 + 10^{pH-pKa} \tag{4.8}$$

$$\alpha_{base} = 1 + 10^{pKa-pH} \tag{4.9}$$

For basic compounds, precipitation was considered. Briefly, if the concentration of the dissolved drug in a given gastrointestinal compartment is

larger than $R_{ss}C_{s,i}$, where $R_{ss}$ is the supersaturation ratio, linear precipitation of the drug occurs, at a time constant equal to $k_p$, as detailed in equations system 4.10.

$$\begin{aligned}\frac{dA_{i,d}}{dt} &= k_{t,i-1}\,A_{i-1,d} - k_{t,i}\,A_{i,d} - k_{a,i}\,A_{i,d} - k_p\,A_{i,d} + K_i\,A_{i,s}\\ \frac{dA_{i,s}}{dt} &= k_{t,i-1}\,A_{i-1,s} - k_{t,i}\,A_{i,s} + k_p\,A_{i,d} - K_i\,A_{i,s}\end{aligned} \tag{4.10}$$

The model outputs $f_a$ and $F_{oral}$ are defined, respectively, as the value of $A_a$ and the value of $A_{oral}$, in equation 4.3, at steady state, both normalised with respect to the dose.

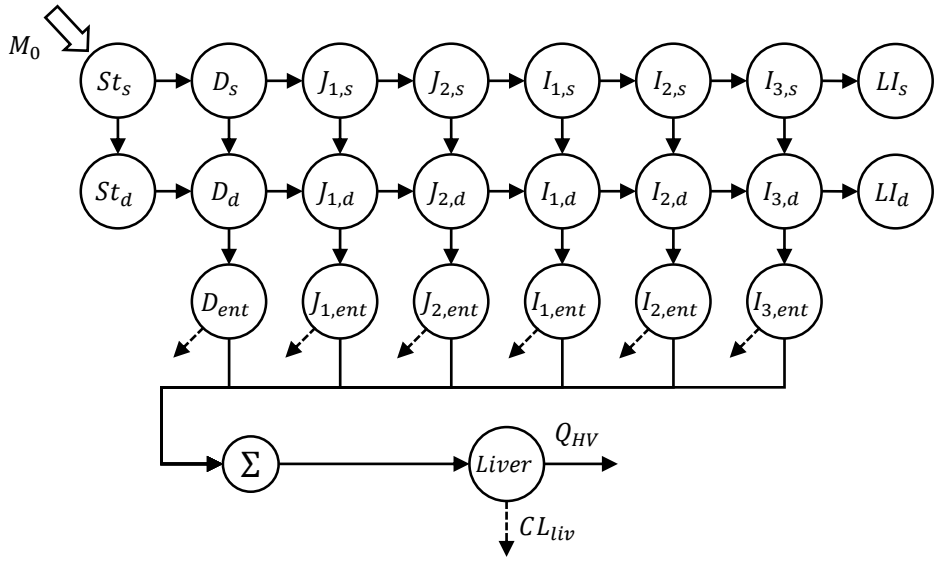



Figure 4.1: CAT derived model. *St* stands for stomach, *D* for duodenum, *J* for jejunum, *I* for ileum and *LI* for large intestine. Subscripts *s*, *d* and *ent* stand for solid, dissolved and enterocytes. Continuous and dashed arrows represent mass transfer and clearance processes, respectively. Drug is administered solid in the stomach compartment, then is subject to dissolution, transit, absorption in the small intestine and metabolism in gut wall and liver.

**Mixing-tank based model**

This model, represented in figure 4.2, is substantially similar to the CAT based model, where the gastrointestinal system is subdivided into three sections: the stomach, the small intestine and large intestine. For this reason, the only differences in the equations are in 4.2, where $i = 1$. So, $k_{t,1}$ is equal to the inverse of $SITT$ and $V_1$ and $V_{ent,1}$ represent the total volume of the small intestine lumen and the total volume of the enterocytes.

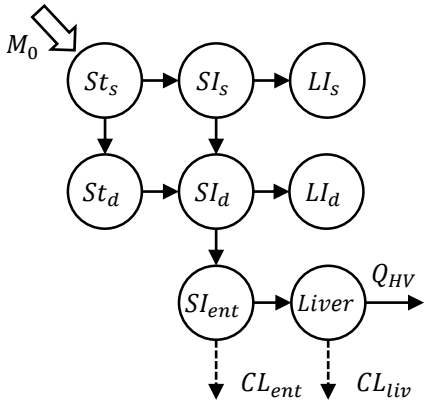



Figure 4.2: Mixing tank model derived model. *St* stands for stomach, *SI* for small intestine and *LI* for large intestine. Subscripts *s*, *d* and *ent* stand for solid, dissolved and enterocytes. Continuous and dashed arrows represent mass transfer and clearance processes, respectively.

### 4.2.2 Definition of the BCS classes and GSA

To perform the GSA, a probability distribution has to be defined for each input parameter of the model (see table 4.1). In order to simplify the analysis, a number of physiological parameters were fixed to their mean values for a fasted state, including: all volumes, luminal pH values, blood flows and each small intestine segment radii and length (see tables 4.3 and 4.2). The GSA algorithm extracts samples from the parameter spaces and, for each of them, evaluates the model and computes the outputs that, in our case, are $f_a$ and $F_{oral}$. Then, a drug is defined as a sample extracted

Table 4.1: Parameter distributions used for GSA

| **Parameters** | distribution parameters | distribution type | units | reference |
|---|---|---|---|---|
| $A_{3A4,ent}$: total enterocytes amount of CYP3A4 | 66.2 (60%) | Lognormal[a] | $nmol$ | [160] |
| $C_{3A4,LM}$: CYP3A4 concentration in liver microsomes | 137 (41%) | Lognormal[a] | $pmol/mg\ prot$ | [160] |
| $MPPGL$: microsomal protein per gram of liver | 39.79 (26.9%) | Lognormal[a] | $mg\ prot/g$ | [160] |
| $GET$: gastric emptying time | 0.25 (38%) | Lognormal[a] | $h$ | [155] |
| $\ln(CL_{int})$: intrinsic clearance | 2.0809, 2.4086 | Normal[b] | $mL/(h \cdot pmol)$ | [164] |
| $SITT$: small intestine transit time | 4.04, 2.92 | Weibull[c] | $h$ | [155] |
| $\rho$: density of the formulation | 1-1.8 | Uniform[d] | $g/cm^3$ | [163] |
| $pKa$: acid dissociation constant | Acid: 2.5 - 13.5<br>Base: 0.5 - 12.5 | Uniform[d] | | [165] |
| $r$: formulation radius of the particle | 0.5 - 500 | Uniform[e] | $\mu m$ | [163] |
| $k_p$: precipitation time constant | 0.4 - 40 | Uniform[e,f] | $h^{-1}$ | [160] |
| $R_{ss}$: supersaturation ratio | 1 - 100 | Uniform[e,f] | | [160] |
| $D_0$: dose number | BCS I & III: 0.01 - 1<br>BCS II & IV: 1 - 100 | Uniform[e,g] | | [148] |
| $P_eff$: effective permeability | BCS I & II: 1.5 - 8.70<br>BCS III & IV: 0.03 - 1.5 | Uniform | $10^{-4}\,cm/s$ | [157] |

[a] For distribution parameters, *mean (coefficient of variation)* of the lognormal random variable.
[b] For distribution parameters, *mean, standard deviation* of the natural logarithm of $CL_{int}$, estimated using the MATLAB *distribution fitter* toolbox.
[c] For distribution parameters, $A$, $B$ with $A$ scale parameter and $B$ shape parameter of the Weibull distribution (*WeibullDistribution* object of MATLAB). The distribution was truncated between 1.8 and 8 $h$ [166] by using the MATLAB function `truncate`.
[d] Uniform distribution between *minimum*, *maximum*.
[e] For distribution parameters, *minimum*, *maximum* of the parameter. A uniform distribution of the natural logarithm of the parameter between $\ln(minimum)$ and $\ln(maximum)$ was used.
[f] *minimum* and *maximum* are, respectively, 1/10 and 10 times the mean value in Simcyp.
[g] For doses of 100 mg and 1000 $mg$ of BCS class I and III, $D_0$ limits were set to [0.1 1] in order to avoid too high solubilities and so a too stiff system.

Table 4.2: Physiological gastrointestinal parameters

| **Comp name** | lumen volume $[mL]^a$ | length $[cm]^{b,d}$ | diameter $[cm]$ $^d$ | pH $^a$ | volume enterocytes $[L]$ $^{c,d}$ | fraction CO to enterocytes $^{c,d}$ | CYP3A proportion $^c$ |
|---|---|---|---|---|---|---|---|
| Stomach$^a$ | 48.92 (+250)$^e$ | | | 1.3 | | | |
| Duodenum | 44.57 | 21 | 4.75 | 6.0 | 0.0262 | 0.0038 | 0.1376 |
| Jejunum 1 | 166.6 | 105/2 | 3.25 | 6.20 | 0.119/2 | 0.0178/2 | 0.5448/2 |
| Jejunum 2 | 131.0 | 105/2 | 3.25 | 6.40 | 0.119/2 | 0.0178/2 | 0.5448/2 |
| Ileum 1 | 102.0 | 156/3 | 2.9 | 6.60 | 0.079/3 | 0.0264/3 | 0.3176/3 |
| Ileum 2 | 75.35 | 156/3 | 2.9 | 6.90 | 0.079/3 | 0.0264/3 | 0.3176/3 |
| Ileum 3 | 53.57 | 156/3 | 2.9 | 7.40 | 0.079/3 | 0.0264/3 | 0.3176/3 |

$^a$ [167].
$^b$ [168].
$^c$ [169].
$^d$ Measure relative to the total segment divided by the number of sections in which the segment is subdivided (for jejunum 2 and for ileum 3).
$^e$ Stomach volume (+ volume of water administered with the drug).
$^f$ [160].

from the joint space of the parameters (such as $P_{eff}$, $mw$...). The solubility relative to dose (dose number) and permeability were the only parameters that were assumed to differ between the BCS classes in this analysis (figure 4.3).

BCS classes I and II are characterised by high absorption, while classes III and IV by low absorption. The parameter that controls the absorption in equations system 4.2 is the absorption rate constant ($k_a$), defined as a function of the effective permeability ($P_{eff}$). The cut-off value for $P_{eff}$ that distinguish between high and low absorption was set to $1.5 \cdot 10^{-4}$ $cm/s$ [157] and the ranges of its variation were taken from the same publication.

The parameter that was used to distinguish between high and low solubility (between classes I and II and between III and IV) was the dose number [147, 148],

$$D_0 = \frac{M_0/V_{in}}{C_s}, \tag{4.11}$$

where $V_{in}$ is the volume of water taken with the drug (250 $ml$ [147]) and

Table 4.3: Constant parameters of the physiological intestinal absorption models

| **Parameters** | value | units | reference |
|---|---|---|---|
| $BW$: body weight | 70 | $kg$ | [170] |
| $CO$: cardiac output | 350.37 | $L/h$ | [169] |
| $W_{liv}$: liver weight (percentage of BW) | 5.53 (0.079) | $kg$ | [170] |
| $Q_{HV}$: hepatic vein blood flow (percentage of CO) | 89.34 (0.255) | $h^{-1}$ | [170] |
| $\rho_{liv}$: liver density[a] | 1.080 | $kg/l$ | [171] |
| $T$: absolute body temperature | 310.15 (37) | $K$ ($^oC$) | |
| $k_b$: Boltzmann constant | 1.3806504 | $10^{-23}$ $J/K$ | |
| $N_A$: Avogadro's number | 6.02214179 | $10^{23}$ $mol^{-1}$ | |

[a] Used to calculate $V_{liv}$ from $W_{liv}$: $V_{liv} = W_{liv}/\rho_{liv}$.

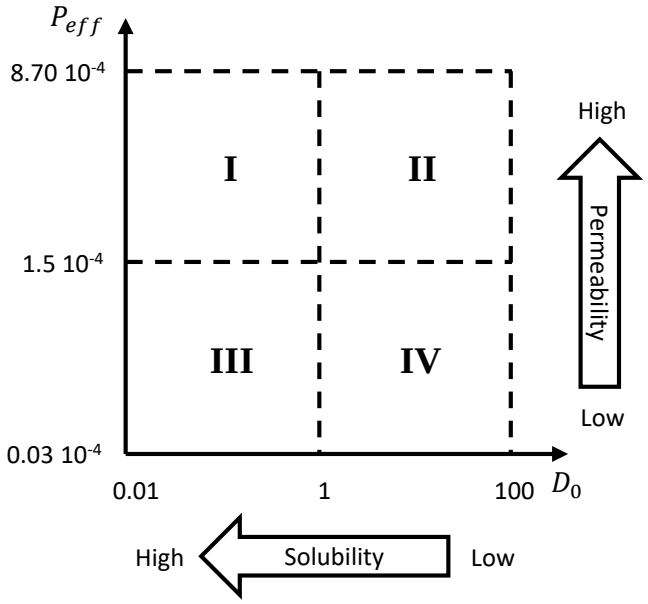



Figure 4.3: Model parameter space following to the definition of the BCS classes. A drug is defined as highly soluble if $D_0$ is between 0.01 and 1 and lowly soluble if it is between 1 and 100. A drug is defined highly permeable if $P_eff$ is between $1.5 \cdot 10^{-4}$ $cm/s$ and $8.7 \cdot 10^{-4}$ $cm/s$, meanwhile is lowly permeable if it is between $0.03 \cdot 10^{-4}$ $cm/s$ and $1.5 \cdot 10^{-4}$ $cm/s$. Roman numbers represent the BCS classes.

$M_0$ the drug dose. If $D_0 \leq 1$ a compound is highly soluble, while if $D_0 > 1$ it is solubility limited. The ranges for this parameter were arbitrarily set from $10^{-2}$ to 1 for classes I, III and from 1 to $10^2$ for classes II and IV. However, the solubility ($C_s$), and not $D_0$, is present in systems of equations 4.1 and 4.2. So, once the dose, $M_0$, is fixed the algorithm computing the sensitivity indices extract a value for $D_0$ and calculates $C_s$ (supposed for a pH equal to 6). By doing this, extracting $D_0$ was equivalent to extracting $C_s$ once dose was fixed. Then $C_s$ results to be depended on the dose. For this reason, different dose levels were tested (0.1 $mg$, 1 $mg$, 10 $mg$, 100 $mg$ and 1000 $mg$).

To perform GSA the number of samples, $n$, has to be chosen. Some au-

thors suggest to set n to 500 or 1000 [21], however, this may be insufficient. We decided to fix $n = 5000$ in order to have reasonable precise estimates, taking into account also the required computational time. The analysis was first carried out for the simple model, derived from the mixing tank model, for neutral drugs, then GSA was performed on the CAT derived model for acidic, basic and neutral compounds. For basic compounds the GSA was also performed in the presence of precipitation. Uncertainty of GSA results was estimated using 1000 bootstrap samples [67]. Coefficient of variation (CV) for the most sensitive parameter, given a certain BCS class and a certain dose, are shown in section B.2.

Differential equations were solved for a time span of 0 (dose administration) to 100 $h$, to assure of reaching the steady state, using the `ode23s` MATLAB solver. The analysis was performed using MATLAB R2017b on a 64-bit computer configured with Intel® Core™ i7-7700 3.60 GHz x 8 processor, running Ubuntu 16.04 LTS[2]. The computational time required to perform the sensitivity analysis for all the BCS classes and all the dosages of, for example, a neutral compound, was approximatively 18 hours.

[2] The codes used to perform the analysis are available at the following link: `http://aimed11.unipv.it/JPKPDMelillo18/`.

## 4.3 Results

A variance based GSA was performed on the two PBPK absorption models described above with the aim of identifying the relative importance of each parameter (both physiological and drug related), considered over its range of variation, in determining the variability of the predicted $f_a$ and $F_{oral}$. The analysis was performed for acidic, basic and neutral drugs from each BCS class. Figures 4.4 and 4.5 summarise the results of the analysis for $f_a$ and $F_{oral}$, respectively. The complete set of figures related to GSA results are presented in section B.1.

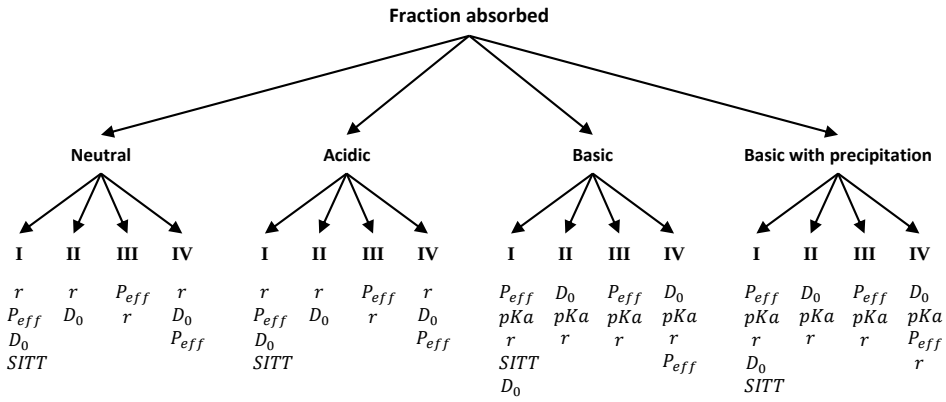



Figure 4.4: Summary of the CAT derived model results for $f_a$. This tree shows the parameters that mostly impact on the variance of $f_a$ for each BCS class, for neutral, acidic and basic compounds with and without the precipitation. The reported parameters have the total effect higher than 0.25. The parameters are written from up to down in descending order of their maximum total effect value through all the dose levels.

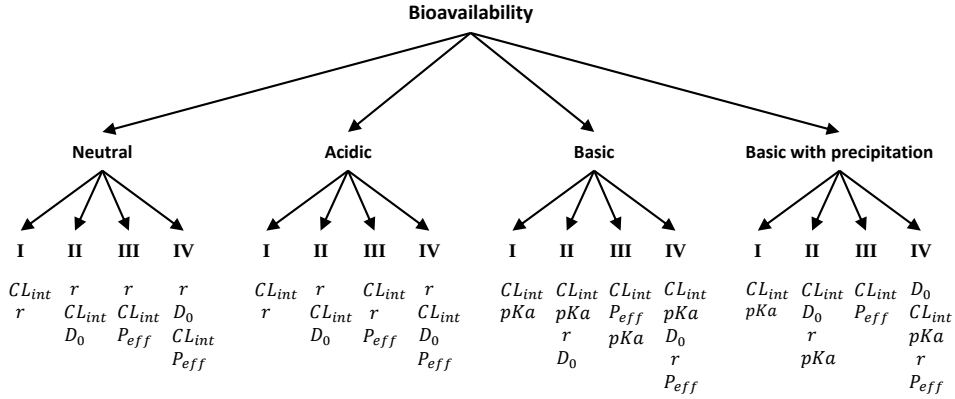



Figure 4.5: Summary of the CAT derived model results for $F_{oral}$. This tree shows the parameters that mostly impact on the variance of $F_{oral}$ for each BCS class, for neutral, acidic and basic compounds with and without the precipitation. The reported parameters have the total effect higher than 0.25. The parameters are written from up to down in descending order of their maximum total effect value through all the dose levels.

### 4.3.1 $f_a$, CAT based model

Main and total effect indices for $f_a$ values were calculated for neutral, acidic and basic drugs in each BCS class. Figures 4.6 and 4.7 represent the main and total effect for neutral compounds, figures B.9 and B.10 the main and total effect for acidic compounds and figures B.15 and B.16 represent the main and total effect for basic compounds. Each figure shows four heatmaps, one for each BCS class. Each heatmap shows the input parameters on the vertical axis and the different dose levels on the horizontal axis.

Considering the neutral case, in figures 4.6 and 4.7, for drugs of class I given at low doses, the most important parameter is the particle radius of the formulation, $r$. This means that the variance of $f_a$, among drug belonging form this class, is mainly explained by the variation of $r$. For drugs administered at higher doses the importance of $r$ is reduced and an increased importance of interactions can be seen, observable through the difference between the total and the main effect. For $r$ the difference

between the total and the main effect increases from 0.0640 at a dose of 0.1 $mg$, to 0.3278 for a dose of 100 $mg$. For doses of 10 $mg$ and 100 $mg$ the importance of $D_0$ increases and $D_0$ becomes the second most important parameter, this is mainly due to interaction effects. Thus, limiting the analysis to the main effect or using sensitivity analysis approaches that cannot detect the interactions (e.g., OAT methods), could lead to an underestimation of parameters influence on the output variance. At a dose of 1000 $mg$ there is an increase in the importance of the small intestinal transit time and the most important parameter becomes the effective permeability, $P_{eff}$. Within a given class, drugs administered at higher dose levels typically have higher solubility. This can be seen in equation 4.11, where if $D_0$ is fixed, higher values of $M_0$ imply higher values of $C_s$. Drugs of class I administered at 1000 $mg$ will most likely have high solubility values and therefore the dissolution process will generally become fast with respect to the absorption process, independent of the value of $r$, even if the drug is highly absorbed. Therefore, $P_{eff}$ becomes the rate-limiting parameter, and therefore the most important parameter in determining $V(Y)$. The main and total effect indices are normalised with respect to the total variance of the output $V(Y)$. So, for doses of 1000 $mg$, $P_{eff}$ will become the most important parameter, but the variability of $f_a$, as can be seen in figure 4.8, is lower with respect to the lower dose levels. As explained before, higher dose levels imply higher values of $C_s$ and a reduction in the influence of $r$ variation on $f_a$ variability. This causes a faster dissolution, resulting in an increase in $f_a$ and a reduction in $f_a$ variability for higher dose levels.

Moving from BCS class I to class II, there is an increase in the values of $D_0$ and therefore a reduction in the solubility for a given dose. Then, for class II compounds the most important parameters result to be $r$ and $D_0$, both related to the dissolution process. This is a consequence of dissolution rate being the limiting step of BCS class II drugs, in accordance with the definition of the class. When considering higher dose levels, and by as a consequence higher solubilities, $D_0$ becomes more sensitive than $r$. It can be seen that the interaction effect of $D_0$ decreases as the dose increases, in fact the difference between total and main effect is reduced from a value of 0.1997 for a dose of 0.1 $mg$ to 0.0680 for a dose of 1000 $mg$.

In BCS class III, we start to consider compounds with low absorption properties. For drugs administered at low doses a similar situation as class I can be seen (with slightly more importance on $P_{eff}$). This is most likely because $C_s$ is not high enough and therefore $r$ is the more influential parameter with regards to the dissolution process, making it the limiting step. When examining compounds administered at higher doses, a progressive reduction of importance of $r$ and $D_0$ is observable and an increase of importance of $k_a$ can be seen. This effect is due by an increase in $C_s$ and so the limiting step is no longer dissolution but absorption.
A more complex situation can be seen for BCS class IV compounds, where parameters related to both dissolution and absorption remain important across the simulated dose levels. This happens because in class IV both solubility and permeability are low, and therefore both could act as the limiting step.

Considering the case of acidic drugs, in figures B.9 and B.10, the results are similar to the case of a neutral drugs as in the stomach $\alpha_{st}$ is low compared to $\alpha_{ref}$ in equation system 4.1 due to the low pH in the gastric lumen ($pH_{stomach} = 1.3$), and so the drug dissolves to a lesser extent. The pH of the small intestine is around 6, which is the value used to calculate $\alpha_{ref}$, and so, in equations system 4.2, $\alpha_i \simeq \alpha_{ref}$ and therefore the solubility is similar to the neutral case. The fact that the drug dissolves less in the stomach does not change the importance of the variables with respect to the case of a neutral drug.

For basic compounds, in figures B.15 and B.16, results differ compared to the previous cases. For class I compounds, up till doses of 10 $mg$, the dissolution appears to be the limiting step, where $pKa$ is the most influential parameter. This is probably because $\alpha_{st}$ in the stomach could reach higher levels depending on the $pKa$ of the compound compared to $\alpha_{ref}$ and therefore the solubility in this compartment is enhanced. For compounds administered at higher doses the conclusions are similar to that of the neutral case. With respect to the neutral case, a stronger interaction effect can be seen, especially for $r$, $D_0$ and $pKa$ for doses of 10 $mg$ and 100 $mg$. For BCS class II compounds all parameters related to dissolution, including $pKa$, are of importance at all the simulated dose levels, similarly

to the neutral case. Also $r$ is associated with a strong interaction effect. Concerning class III, the $P_{eff}$ is the most important parameter at all the dosage levels. This is probably due to the enhancing of dissolution in the stomach, meaning that the drug is already dissolved when reaching the intestine and therefore the absorption process becomes the rate limiting step. As for neutral and acidic drugs, BCS class IV presents a more complicated situation, in fact, both parameters related to dissolution and absorption remain important throughout the simulations. Interaction effects can be observed, especially for $r$ and $pKa$ from doses of 1 $mg$ up to 100 $mg$.

For neutral and acidic compounds, the interaction effects seem to occur to a lesser extent for class I compounds administered at low dosages and for classes III at high dosages. For basic compounds, interactions occur to a lesser extent only for class III drugs administered at high dosages. This happens probably because these cases represent extreme situations, in which the variation of only one parameter seems to determine the variability of the $f_a$. In all the other cases, the variance of $f_a$ can be affected by the variation of multiple parameters, so, the effect of one factor may depend on the values of other factors and interaction effects may arise.

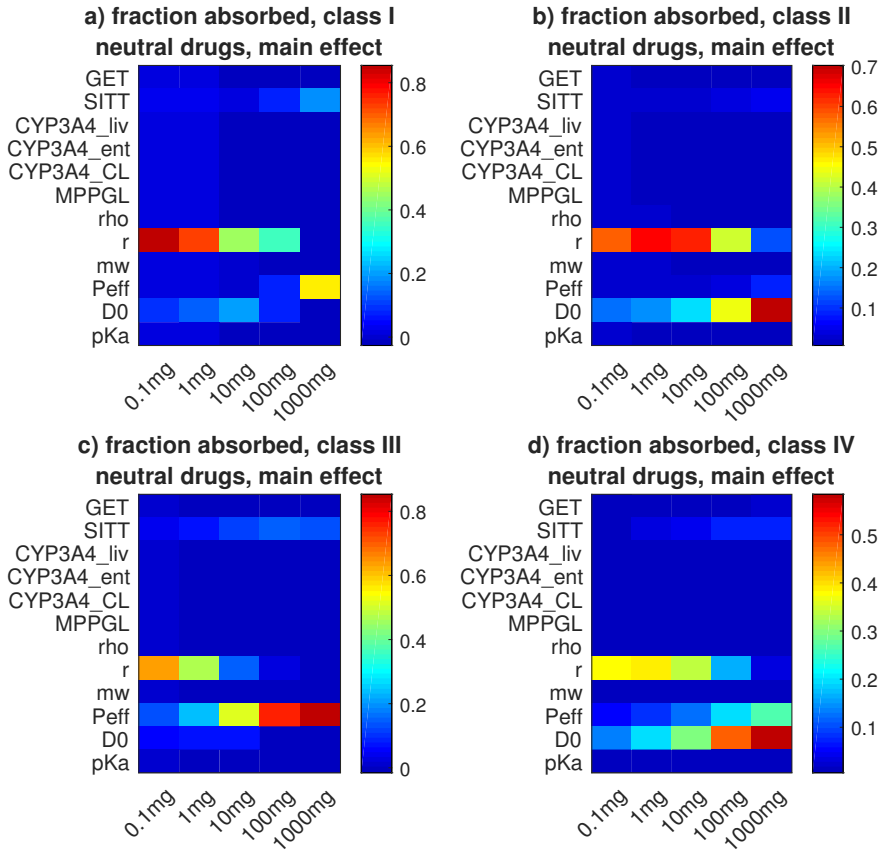



Figure 4.6: $f_a$ main effect of the CAT based model for neutral compounds. Panels A, B, C and D are relative to BCS class I, II, III and IV compounds. Each panel contains a heatmap that has the input parameters on the vertical axis and the different dose levels on the horizontal axis. Each heatmap cell contains the value of the main effect relative to a particular parameter and dose level. Colour legends are shown to the right of each heatmap. CYP3A4_liv, CYP3A4_ent and CYP3A4_CL stand to the microsomal concentration of CYP3A4 in the liver, the total amount of CYP3A4 in the enterocytes and the intrinsic clearance.

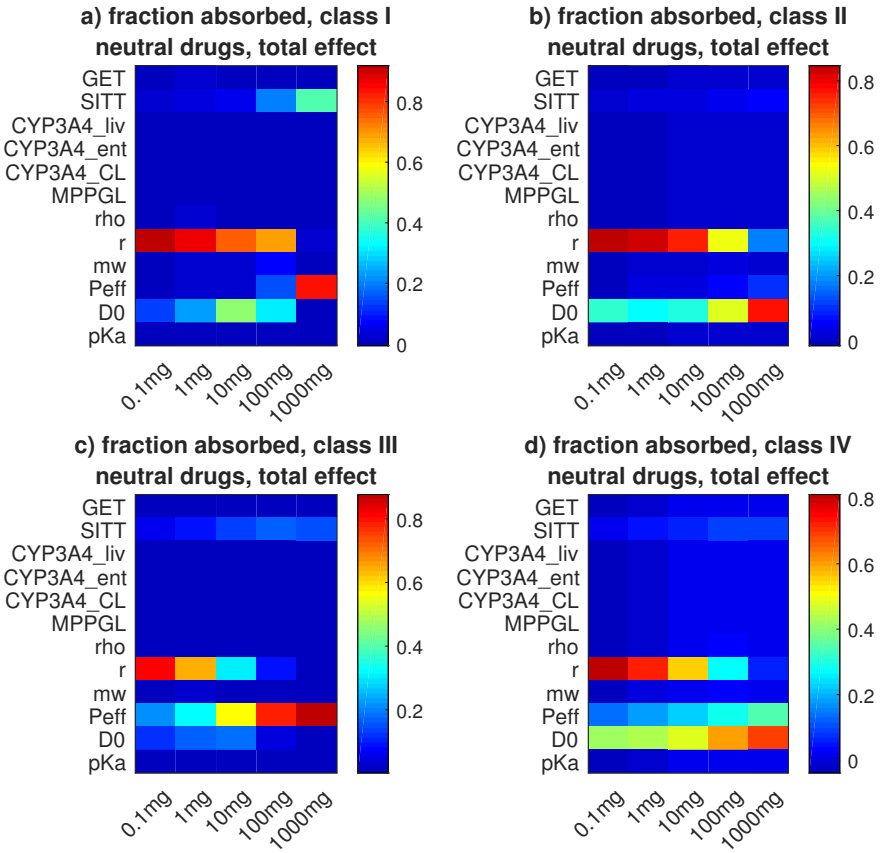



Figure 4.7: $f_a$ total effect of the CAT based model for neutral compounds. Panels A, B, C and D are relative to BCS class I, II, III and IV compounds. Each panel contains a heatmap that has the input parameters on the vertical axis and the different dose levels on the horizontal axis. Each heatmap cell contains the value of the total effect relative to a particular parameter and dose level.

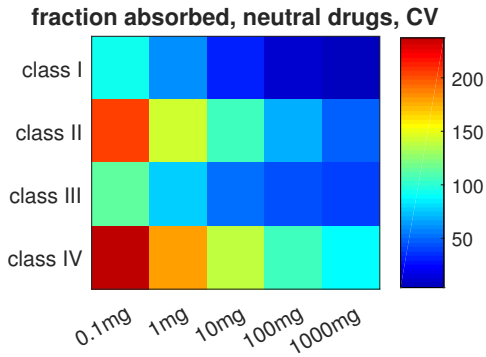



Figure 4.8: $f_a$ CV in percentage, predicted using the CAT based model for neutral compounds. The heatmap vertical axis represents the BCS classes and the horizontal axis represents the dose levels. Each cell contains the value of fraction absorbed CV for a specific BCS class and dose level. Each CV was calculated from the samples used to calculate the main and total effect of the variance based GSA.

### 4.3.2 $F_{oral}$, CAT based model

Main and total effect indices for $F_{oral}$ values were calculated for neutral, acidic and basic drugs in each BCS class. Figures 4.9 and 4.10 represent the main and total effect for neutral compounds, figures B.11 and B.12 the main and total effect for acidic compounds and figures B.17 and B.18 represent the main and total effect for basic compounds.

Considering the neutral case (figures 4.9 and 4.10) for class I compounds given at a low dose, the most important parameter is $r$, while for doses up to 10 $mg$ the intrinsic clearance $CL_{int}$ becomes the most influential parameter. This is most likely because, as explained for $f_a$, for low dosages $r$ is more important for determining dissolution as $C_s$ is not high enough to become as influential, by making it the limiting step. Moving towards higher doses an increasing importance of $C_s$ can be seen. Given the high permeability, the rate limiting step becomes the clearance. Amongst all the parameters involved in the clearance process (e.g., liver enzymatic concentration, $MPPGL$) the consistently most important parameter is $CL_{int}$, this is probably because the parameter was defined with a larger range of variation and because it appears at two sites in the model (gut wall and liver).

For compounds belonging to BCS class II the most important parameters are related to both dissolution and metabolism. At lower dose levels there is a higher importance of formulation related parameters, $r$, meanwhile moving towards higher dose levels, $CL_{int}$ and $D_0$ become the most important parameters. As seen for $f_a$, this is in accordance with the definition of the class properties. Notable interaction effects can be seen for $CL_{int}$, $r$ and $D_0$, especially for doses of 0.1 $mg$ and 1 $mg$.

Moving to class III, for low dosages, $r$ is the most important parameter in determining variation in $F_{oral}$ followed by $CL_{int}$. At higher dose levels clearance and absorption become the rate limiting steps. The reasoning around the differing importance of $r$ and $P_{eff}$ throughout the dose levels follows the same argument as for class III and $f_a$. $CL_{int}$ is more influential at higher doses as compared to $P_{eff}$, which is probably due to its higher range of variation.

For BCS class IV compounds, dissolution, absorption and clearance parameters remain important across the simulated dose levels, with a reduction of importance of $r$ and an increase of importance of $CL_{int}$, $P_{eff}$ and $D_0$ when moving towards higher doses. Notable interaction effects can be consistently observed across dose levels for parameters $r$, $CL_{int}$, $P_{eff}$ and $D_0$.

Considering the case of an acidic compound, as for $f_a$, the results are similar to the case of the neutral one. As explained for $f_a$ the fact that an acidic compound dissolves to a lesser extent in the stomach, does not change the importance of the variables found in the case of a neutral drug.

The situation for basic compounds (figures B.17 and B.18) is slightly different. For class I compounds the most important parameter for all the doses is $CL_{int}$. This is because for a base, as explained for $f_a$, the solubility could be highly enhanced in the stomach and so the drug could dissolve completely prior to reaching the small intestine. Given that the absorption is high, metabolic clearance becomes the rate limiting step. For class II compounds, as in the neutral case, dissolution and clearance are important determinants of variation in the output. Interaction can be mainly seen for the dissolution related parameters. For class III the most important parameters are $CL_{int}$ and $P_{eff}$ across all doses, with $pKa$ being relevant at a dose of 0.1 $mg$, but mainly due to interaction effects. Like in the previous cases, BCS class IV compounds present a more complicated situation, where dissolution, absorption and clearance parameters remain important at all the studied dose levels. Interaction effects can be observed especially for the dissolution related parameters.

For all the compounds, the interaction effects seem to occur to a lesser extent especially for class I drugs administered at dosages higher than 10 $mg$. Similarly to what was explained for $f_a$, these are situation in which the output variance can be addressed almost uniquely to the variation of one parameter and consequently, limited interaction effects arise.

It is possible to observe that there is an apparent discrepancy between $f_a$ and $F_{oral}$ results for BCS class I compounds administered at high dosages. In fact, for $f_a$ the only sensitive parameters are $P_{eff}$ and $SITT$, while for $F_{oral}$ it is $CL_{int}$. In this case, both solubility and permeability are high,

thus, practically all the drug gets absorbed. From figures B.1 and B.2, it can be observed that the variability of class I compounds administered at 1000 $mg$ is much higher for $F_{oral}$ than for $f_a$. So, the clearance processes explain almost all the $F_{oral}$ variability.

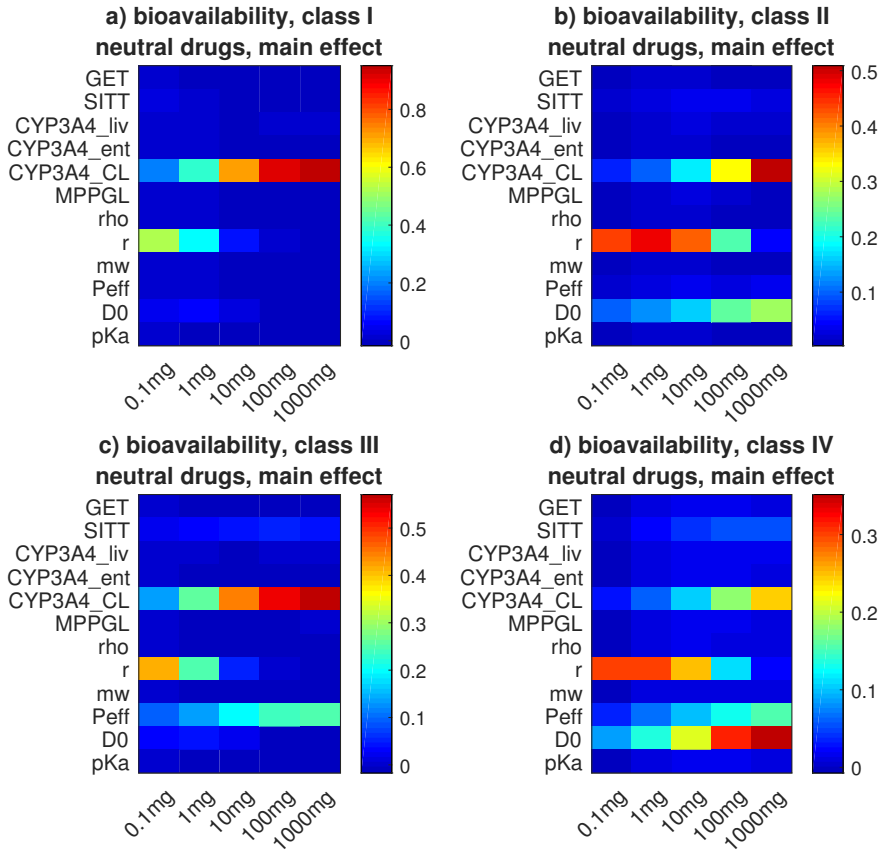



Figure 4.9: $F_{oral}$ main effect of the CAT based model for neutral compounds. Panels A, B, C and D are relative to BCS class I, II, III and IV compounds. Each panel contains a heatmap that has the input parameters on the vertical axis and the different dose levels on the horizontal axis. Each heatmap cell contains the value of the main effect relative to a particular parameter and dose level.

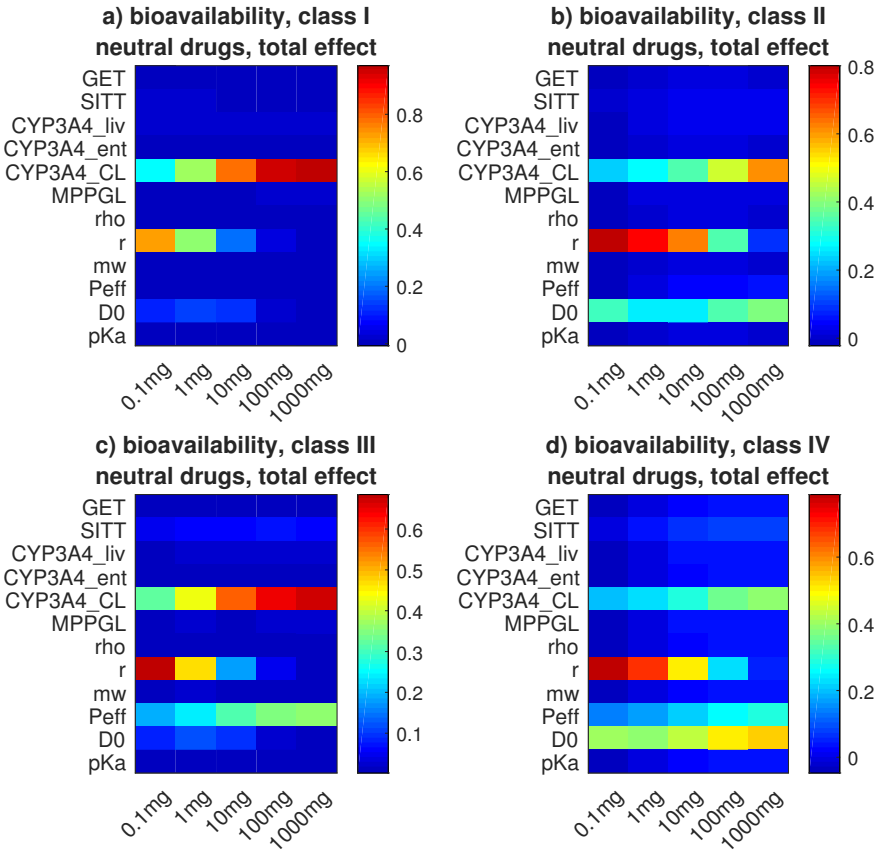



Figure 4.10: $F_{oral}$ total effect of the CAT based model for neutral compounds. Panels A, B, C and D are relative to BCS class I, II, III and IV compounds. Each panel contains a heatmap that has the input parameters on the vertical axis and the different dose levels on the horizontal axis. Each heatmap cell contains the value of the total effect relative to a particular parameter and dose level.

### 4.3.3 $f_a$ and $F_{oral}$, CAT based model, basic compound with precipitation

GSA was carried out for basic compounds using the CAT based model with the additional consideration of precipitation. Main and total effect for $f_a$, in figures B.21 and B.22, and for $F_{oral}$, in figures B.23 and B.24, were calculated.

Concerning both $f_a$ and $F_{oral}$, for compounds belonging to BCS classes I and III, the results in presence of precipitation resemble the case of absence of precipitation. It can be seen that the variability of the supersaturation ratio $R_{ss}$ and of the precipitation time constant $k_p$ are not important in determining the variance of the output. In the stomach, the maximal concentration that a drug could achieve is equal to $M_0/V_{st}$, with $M_0$ the drug dose and $V_{st}$ equal to 298.92 $mL$ (see table 4.2). For BCS classes I and III, the drug solubility, calculated by using equation 4.11, results higher or equal to $M_0/V_{in}$, with $V_{in}$ equal to 250 $mL$ (this happens because in this case $D_0$ was considered between $10^{-2}$ and 1). Moreover, in the stomach $\alpha_{st} > \alpha_{ref}$, so, the solubility is enhanced with respect to the neutral case. It follows that the maximal concentration that the drug could reach in the stomach is lower with respect to the solubility in that compartment. Consequently, the precipitation does not occur in the stomach. It can be observed that even if the precipitation occurs in the small intestine this does not make $R_{ss}$ and $k_p$ important in determining output variability.

Concerning BCS classes II and IV, for both $f_a$ and $F_{oral}$, in case of presence of precipitation, $D_0$ is slightly more important in determining output variance with respect to the case of absence of precipitation. With respect to BCS classes I and III, in this case the maximal concentration that a drug could reach in the stomach can be lower than the solubility in that compartment. This happens because $D_0$ was considered between 1 and $10^2$. Therefore, precipitation could occur in the stomach. Probably, in this case $D_0$ acquires importance because its value it is used to determine the threshold at which the precipitation starts to occur. Moreover, when a drug once dissolved could precipitate, an additional dissolution step is required to allow absorption and $D_0$ is a parameter involved in the process

of dissolution.

### 4.3.4 $f_a$ and $F_{oral}$, MT based model

Main and total effect for $f_a$, in figures B.27 and B.28, and for $F_{oral}$, in figures B.29 and B.30, were calculated for neutral compounds for each BCS class, using the mixing tank derived model. For both $f_a$ and $F_{oral}$ the results are consistent with the CAT derived model for neutral compounds. This fact does not mean that the outputs of the CAT derived model are similar to that of the mixing-tank derived model, instead in both the models the variability in the output is explained to a similar extent by the same parameters.

## 4.4 Discussion

We performed a variance based GSA on PBPK models describing drug dissolution, transit and absorption in the gastrointestinal tract, with the aim of finding the most important parameters that determine the variability of the predicted $f_a$ and $F_{oral}$ for acidic, basic and neutral drugs within each BCS class. In figures 4.4 and 4.5 the results of the analysis are summarised.

Performing a GSA could help in identifying limiting steps and bottlenecks, in different situations, and in understanding the behaviour of the model as a function of the variation of different parameters. This kind of information is difficult to obtain during performance evaluation exercises, such as OrBiTo [53, 54, 55], where the model predictions are affected by the quality of the data informing the values of compound specific parameters. In fact, in the OrBiTo compound database a high level of missingness for parameters such as particle size of solid formulation and solubility vs pH profiles was observed [54]. In this case, for example, following the analysis here reported, it is possible to conclude that a performance evaluation of PBPK absorption models where the radius of the formulation-specific particle size is fixed at an assumed or mean value could result in an incomplete interpretation, especially for compounds administered at low dose levels where the particle radius explains the majority of the output variation.

However, it must be considered that the validity of the GSA results is limited to the specific model and to the specific ranges of parameters investigated. This means that the level of importance of each parameter is relative to the model and to all parameter distributions, and thus it would be incorrect to attribute the results presented here to different absorption models or parameter distributions. Nevertheless, we could consider the theoretical case of a model that describes the absorption of a class III drug, administered at a dose of 100 $mg$, that exhibits uncertainty, or variability, in the $P_{eff}$ within a range that is wide with respect to the $P_{eff}$ range tested in this analysis[3]. Considering the results of the GSA presented in this study we could conclude that the variability of the $P_{eff}$ for that particular drug should be further investigated, otherwise, by fixing the value of $P_{eff}$ to the mean, it is possible that we ignore an important source of variability in estimating $f_a$ and $F_{oral}$. Anyway, if one wants to assess the impact of the variability of each parameter on the model outputs for a specific drug (and not to a class of compounds, as done in this study), one should perform a GSA adapting the parameters distributions to that particular case [72, 173].

One limitation of this analysis is that the classification between highly and lowly soluble drugs was defined using only $D_0$, as in [148], and does not take into account the effect of the formulation properties. Thus, it is possible that this will result in an overestimation of the importance of formulation related parameters (e.g., $r$) for BCS classes I and III. Another limitation is in the derivation of $F_{oral}$, we hypothesised that the metabolism in the gut wall and in the liver was due to CYP3A4 abundance and so, the results are limited to that particular case. Finally, we chose to use the variance based GSA method as per [87], but is also possible to use, for example, moment independent methods, or regression based methods [61, 75, 98].

[3]Atenolol is a BCS class III drug with a mean in vivo $P_{eff}$ of $0.15 \cdot 10^{-4}$ $cm/s$ and a standard deviation of $0.2 \cdot 10^{-4}$ $cm/s$ [157, 172]. Supposing that $P_{eff}$ is distributed log-normally, the $95^{\text{th}}$ percentile is equal to 0.47 and the $5^{\text{th}}$ percentile is below the inferior limit in table 4.1. So, for Atenolol the range of variability of $P_{eff}$, from the lower limit in table 4.1 to its $95^{\text{th}}$ percentile, represents around 30% of the whole range of variation considered for class III drugs.

In mathematical modelling sensitivity analysis should be performed to gain insight into the structure of the model and understand its behaviour in the parameter space of uncertainty and/or variability [21, 61]. In general, PBPK models include many parameters that are uncertain or variable at a population or individual level and whose impact on the outputs of interest is not trivial to predict. For example, when these models are used, there is a tendency to fix some uncertain parameters (e.g., radius of the particle, solubility) to an assumed value, mean value or to use *in silico* methods, such as quantitative structure-activity relationship (QSAR) models, to predict parameter values, without sufficiently exploring the impact on model development and predictions. Performing GSA could help in identifying the few key parameters amongst many [21] that are mainly responsible for the variation in output.

Typically, if a model is well constructed, understood and characterized, the results of a sensitivity analysis should reflect the qualitative expectations of the model behaviour and thus, may appear to be obvious. However, especially if a model is involved in regulatory decisions, a sensitivity analysis should be performed to objectively confirm these expectations, as highlighted by regulatory agencies [50, 174]. In fact, sensitivity analysis can quantitatively assess the impact of each parameter variation on the variability of some output metrics. Understanding *how much* the input parameters influence the model outputs (so, the magnitude of the sensitivity indices) is crucial information that helps to understand if a given parameter can be assumed, fixed or require further investigation in order to allow informed model prediction. That information would be difficult to obtain without performing a sensitivity analysis.

In conclusion, this work aimed to identify the importance of different parameters for different types of drugs, to improve our understanding of PBPK absorption models and guide the choice of parameters that can safely be assumed, estimated or require data generation to allow informed model prediction. Pharmaceutical regulators have identified the importance of sensitivity analysis in PBPK model qualification [50, 174]. Here we give a demonstration of the GSA methodology and highlight its utility by using a generalised example, spanning across a number of hypothetical compounds

and showing its importance in identifying the key parameters that may be targeted for further investigation during pharmaceutical research and drug development.

Chapter 5

# Inter & intra compound GSA for the development and use of a physiological pulmonary absorption model [1]

## 5.1 Introduction

Due to the opportunity of directly targeting the biophase of interest, the inhalation route has been considered a convenient way of drug administration for local treatment of lung-specific diseases, such as asthma and chronic obstructive pulmonary disease (COPD). This route allows the administration of drugs at lower dosages, minimizing potential side effects

[1] "N. Melillo, S. Grandoni, N. Cesari, G. Brogin, P. Puccini, P. Magni. Global sensitivity analysis of a physiological model for pulmonary absorption of inhaled compounds. *Manuscript in preparation*".

driven by high systemic exposures. Topically active compounds for lung diseases have normally an adequate, and generally sustained, lung residence time [175, 176, 177]. However, efforts have to be placed in the optimization of drug lung disposition looking for an optimal lung retention, since an increased residence time in the airways could potentially translate into the risk of drug removal from the lung due to mucociliary clearance or into the risk of unsafe drug accumulation in pulmonary tissues. For this reason, it is necessary to maintain an appropriate balance between lung retention and absorption by the modulation of the interplay of some key properties, such as solubility, permeability and lung tissue binding [175, 177].

Nowadays, administration by inhalation to rodents is still an important step in preclinical development of new drugs designed for the inhalation route [175]. A mathematical model able to predict compounds pharmacokinetics properties after inhalation in preclinical species could be extremely beneficial during early drug discovery for aims such as the compounds prioritization before animal experiments and for preclinical to clinical translation. For these reasons, in a previous work we developed a PBPK model for inhaled drugs [178]. The model was used to predict the compounds lung disposition in preclinical species (e.g., rodents) starting from physiological and in vitro parameters, such as mucociliary clearance rate, drug solubility and permeability.

In this context, the aim of our work was to adequately characterize the developed physiological absorption model, by understanding how much the uncertainties and variabilities in the input parameters drive the predictions uncertainties and variabilities. This was done by performing a variance based GSA [21, 69].

We performed two types of GSA with two different aims: *inter-compounds* and *intra-compound* GSA. The inter-compounds GSA resembles the analysis done in [146] for intestinal absorption models. Each of the drug-related model parameters was considered variable in a range given by the minimum and the maximum value in the considered set of compounds. Thus, inter-compound GSA mainly focus on the 'between-drugs' parameters variability and would be useful to understand the main model behaviour in the space of all the considered compounds. The aim of this analysis is to understand

what are the key parameters that mostly explain the differences between drugs. On the other hand, the intra-compound GSA, focalized on the parameter uncertainties related to a specific compound, has the purpose to find the most important parameters that, with their uncertainty, mostly causes the output uncertainty. In this work, intra-compound GSA was performed on three representative compounds (namely A, B and C). For each compound, the drug-related parameters were considered variable in a range representative of the parameters uncertainty.

## 5.2 Physiological pulmonary absorption model and inter & intra compound GSA

### 5.2.1 Pulmonary absorption model

The considered physiologically based model was originally presented in [178] and was inspired by the work of Boger *et al.* [179]. The model is composed of three parts describing the pulmonary absorption, the intestinal absorption and the systemic disposition. The pulmonary absorption model was built to take into account the principal PK processes occurring when a drug is inhaled: deposition, mucociliary clearance, dissolution, absorption in lung tissue and absorption in blood circulation [180]. In the model, the lung was divided in two parts, the central region and the peripheral region. The central region roughly corresponds to the tracheobronchial region, while the peripheral region to the alveolar region. Both regions were further divided in four compartments: the undissolved drug, the dissolved drug, the extravascular and vascular lung tissue. The central region was considered perfused by the systemic circulation, while the peripheral region by the pulmonary circulation. The model structure is shown in figure 5.1.

In this work, drugs are intra-tracheally administered to rats. Only a fraction of the drug amount administered to the animal actually reaches the lungs ($F_{inh}$) whereas the rest is deposited in the oropharyngeal region and gets swallowed ($F_{swa}$). Of the fraction delivered to the lungs, a part is deposed in the central region ($F_C$) and a part reaches the peripheral

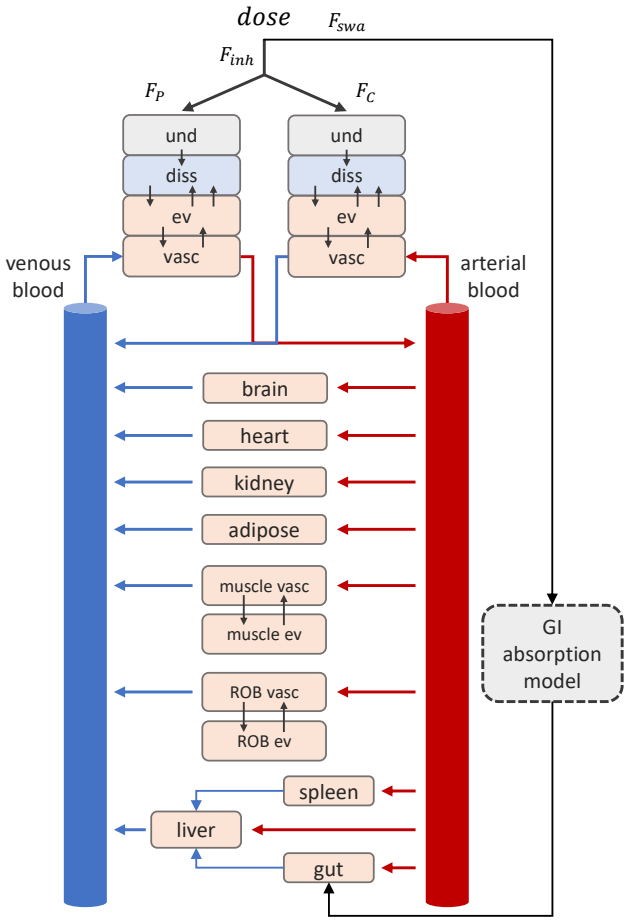



Figure 5.1: Physiologically based model structure for inhaled compounds. The model is composed of three parts: pulmonary absorption model, intestinal absorption model and systemic PBPK. *und*, *diss*, *ev* and *vasc* stand for undissolved, dissolved, extravascular and vascular. *ROB* stands for rest of the body.

region ($F_P$). $F_C$ and $F_P$ values were calculated from formulation properties, as explained in appendix C. Once deposited, in both central and

peripheral regions the drug dissolves in the physiological fluids and then is supposed to be passively absorbed in the tissues. Here, the drug can diffuse to the vascular compartment or back to the dissolved drug compartment. A monodirectional transport from the tissue to the dissolved drug compartment was included to account for the possible action of efflux transporters, such as the P-glyco proteins. The mucociliary clearance mechanism has been considered acting only on the undissolved drug compartment of the central region, since this mechanism should be negligible in the alveoli [180].

The model equations are shown in 5.1.

$$
\begin{aligned}
\frac{da_{u,C}}{dt} &= -(k_{d,C} + k_{MC})\, a_{u,C} \\
\frac{da_{d,C}}{dt} &= k_{d,C}\, a_{u,C} - P_p\, S_C\, \frac{a_{d,C}\, f_{u,elf}}{V_{elf,C}} - (P_p + P_a)\; S_C\, \frac{a_{ev,C}\, f_{u,l}}{V_{ev,C}} \\
\frac{da_{ev,C}}{dt} &= P_p\, S_C\, \frac{a_{d,C}\, f_{u,elf}}{V_{elf,C}} + P_p\, S_C\, \frac{a_{v,C}\, f_{u,b}}{V_{v,C}} - (2P_p + P_a)\; S_C\, \frac{a_{ev,C}\, f_{u,l}}{V_{ev,C}} \\
\frac{da_{v,C}}{dt} &= Q_C \left( c_{ven} - \frac{a_{v,C}}{V_{v,C}} \right) + P_p\, S_C\, \frac{a_{ev,C}\, f_{u,l}}{V_{ev,C}} - P_p\, S_C\, \frac{a_{v,C}\, f_{u,b}}{V_{v,C}} \\
\frac{da_{u,P}}{dt} &= -k_{d,P}\, a_{u,P} \\
\frac{da_{d,P}}{dt} &= k_{d,P}\, a_{u,P} - \alpha\, P_p\, S_P\, \frac{a_{d,P}\, f_{u,elf}}{V_{elf,P}} - (\alpha P_p + P_a)\; S_P\, \frac{a_{ev,P}\, f_{u,l}}{V_{ev,P}} \\
\frac{da_{ev,P}}{dt} &= \alpha\, P_p\, S_P\, \frac{a_{d,P}\, f_{u,elf}}{V_{elf,P}} + \alpha\, P_p\, S_P\, \frac{a_{v,P}\, f_{u,b}}{V_{v,P}} - (2\alpha P_p + P_a)\; S_P\, \frac{a_{ev,P}\, f_{u,l}}{V_{ev,P}} \\
\frac{da_{v,P}}{dt} &= Q_P \left( c_{art} - \frac{a_{v,P}}{V_{v,P}} \right) + \alpha\, P_p\, S_P\, \frac{a_{ev,P}\, f_{u,l}}{V_{ev,P}} - \alpha\, P_p\, S_P\, \frac{a_{v,P}\, f_{u,b}}{V_{v,P}}
\end{aligned}
\tag{5.1}
$$

Subscripts $C$ and $P$ stand for central and peripheral lung regions, respectively; $a_u$, $a_d$, $a_{ev}$ and $a_v$ are the drug amounts in undissolved, dissolved, extravascular and vascular compartments, respectively; $c_{ven}$ and $c_{art}$ are the drug concentrations in venous and arterial compartments; $f_{u,elf}$, $f_{u,l}$ and $f_{u,b}$ are the drug fraction unbound in the epithelial lining fluids, lung

tissues and blood; $V_{elf}$, $V_{ev}$ and $V_v$ are the volumes of the epithelial lining fluids and the extravascular and vascular compartments, respectively; $Q_P$ and $Q_C$ are the blood flows directed to the peripheral and central lung regions; $P_p$ and $P_a$ are the passive and active permeabilities, calculated from experiments with Calu3 cells as explained in appendix C, which have characteristics similar to the tracheobronchial region; $S_P$ and $S_C$ are the surfaces of peripheral and central regions; $k_{MC}$ is the time constant relative to mucociliary elimination, supposed to happen only in peripheral region, while $k_d$ it is the dissolution time constant, modelled with the *Noyes-Whitney* model, as explained in section 4.2.1.

In equation system 5.1, $\alpha$ is a scalar constant used to model the higher passive permeability in peripheral region, with respect to the central one, due to the minor thickness of alveoli epithelium. $\alpha$ was calculated as in equation 5.2.

$$\alpha = \frac{BT}{ALT} \tag{5.2}$$

$BT$ and $ALT$ are the thickness of the bronchial and alveolar wall, respectively. To describe the systemic drug disposition and the intestinal absorption, the pulmonary absorption model was coupled with the whole body PBPK model presented in [181], as shown in figure 5.1. The physiological lung related parameters are reported in table 5.1. The ranges of variability of physiological parameters are reported in table 5.2. All the remaining physiological parameter values (e.g., organ volumes and blood flows) of the whole body PBPK model for a mean rat of 250 $g$ are reported in the supplementary material of [181].

### 5.2.2 Inter-compounds & intra-compound GSA

Inter-compounds and intra-compound are two ways of performing a GSA that differ in the aims and then in the considered parameter variability, as mentioned in the introduction. Figure 5.2 didactically shows the difference.

We performed the inter-compounds GSA on the physiological pulmonary absorption model decoupled from the distribution PBPK and the intestinal

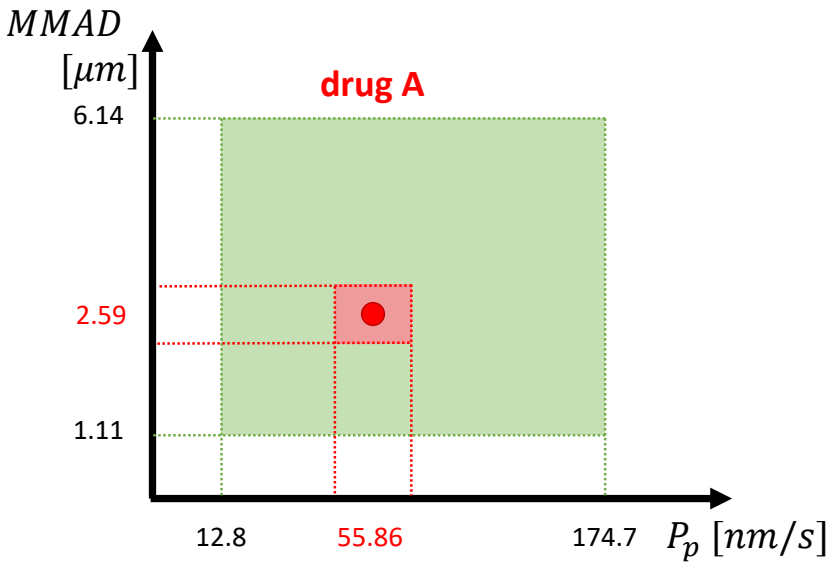



Figure 5.2: Difference in the parameters space used for inter-compounds and intra-compound GSA. Inter-compounds GSA (green area) considers the 'between-drugs' variability, while intra-compound GSA (red area) considers the uncertainty associated to the parameters relatively to a specific compound. As shown, variability in general is wider than uncertainty.

absorption models. This was done to characterize the absorption process and simplifying the model and the results understanding. To decouple the model, blood inflow and outflow values of both peripheral and central lung vascular compartments were set equal to zero. Moreover, the fluxes due to passive permeability from the vascular to the extravascular compartments were set equal to zero too, in both the lung regions. Thus, lung vascular compartments behave like wells (integrators). The outputs considered in this analysis are: the fraction absorbed ($f_a$), the AUC and MRT of the drug concentration in the whole lung. Whole lung concentration was obtained by the sum of the solid and dissolved amounts in the epithelial lining fluids with the ones in the central and peripheral extravascular compartments, all divided for the total volume. $f_a$ was obtained as $f_a = 1 - f_{CL}$, where $f_{CL}$

is the fraction of the drug eliminated by the mucociliary clearance at the steady state.

In the inter-compounds GSA, we considered each of the drug-related parameters varying in a range given by the minimum and the maximum values of the set of the nine considered *Chiesi*'s compounds (table 5.2). Their distributions were considered uniform in these ranges. The inter-compounds GSA was performed separately for highly and poorly soluble compounds. The criterion used for the compounds classification in the two groups was inspired to the one adopted for oral administered compounds [147, 148, 182]. A dose number for inhaled compounds was defined as in equation 5.3.

$$D_{0,inh} = \frac{dose/V_{elf}}{C_{S,pH_{elf}}} \tag{5.3}$$

The drug dose was considered fixed to 10 $\mu g$. $V_{elf} = V_{elf,P} + V_{elf,C}$ is the lung epithelial lining fluid volume and $C_{S,pH_{elf}}$ is the drug solubility measured in simulated lung fluid at pH 6.9 [183, 184]. A compound was classified as highly soluble if $D_{0,inh} \leq 1$ or poorly soluble if $D_{0,inh} > 1$. As done in chapter 2, during GSA we first extracted the values of $D_{0,inh}$ and then we computed the $C_{S,pH_{elf}}$ using equation 5.3.

The intra-compound GSA was performed on the whole body PBPK model for three representative compounds, characterized by different properties. The outputs that were considered are the drug whole lung and plasma concentration AUC and MRT. Here, whole lung concentration was calculated as the sum of the drug amount in all the pulmonary absorption model compartments, divided by the lung total volume. All the drug-specific model parameters were considered uniformly distributed between the ranges reported in tables 5.4, 5.5 and 5.6, except for the dose, that was considered normally distributed with a CV equal to 15% [185]. When no experimental data supporting the variability range definition were available, arbitrary ranges reflecting the perceived parameter uncertainties were used. To account for the population variability of rat weight, all the volumes and blood flows were multiplied for $(w_{subj}/w_{mean})$ and $(w_{subj}/w_{mean})^{0.75}$, respectively. $w_{subj}$ is the extracted value of rat weight and $w_{mean}$ is the mean

rat weight equal to 250 $g$, as used in [181].

All the analyses were performed using MATLAB R2019a [127] on a 64-bit computer configured with Intel® Core™ i7-7700 3.60 GHz x 8 processor, running Ubuntu 16.04 LTS. The systems of differential equations were solved by using the `ode15s` MATLAB solver, for a time span ranging from 0 to 400 $h$. A program to perform variance based GSA was developed. To perform the GSA in both inter and intra compound cases, we used a number of samples ($n$) equal to 20000. Uncertainty on the calculation of the sensitivity indices was estimated by using 10000 bootstrap samples [67].

Table 5.1: Lung physiological parameters

| **Parameters** | central lung | peripheral lung | units | reference |
|---|---|---|---|---|
| surface area | 276.4 | 3.27 | $dm^2/kg$ | [179] |
| lining fluid volumes | 163.6 | 193.5 | $\mu L/kg$ | [179] |
| tissue volumes[a] | 1.01 | 0.2438 | $mL$ | [171, 179] |
| blood flows[b] | 89.61 | 1.88 | $mL/min$ | [171, 179] |
| proportion of ev tissue[c] | 0.55 | 0.55 | | [181, 186] |

[a] The tissue volumes were obtained multiplying the total lung volume, from [171], for the proportions reported in [179].
[b] The blood flows were obtained multiplying the CO, from [171], for the proportions reported in [179].
[c] The proportions of vascular and extravascular tissue in the central and peripheral lung regions were obtained from [186], as shown in [181]. Both the values were assumed to be the same.

Table 5.2: Physiological and drug related parameters used for both inter and intra compound GSA

| **Parameters** | baseline | min value | max value | units | reference |
|---|---|---|---|---|---|
| $F_{inh}$: inhaled fraction | 0.9 | 0.9 | 1 | | internal data |
| $k_{MC}$: mucociliary clearance | 0.5545 | 0.4621 | 0.6931 | $1/h$ | [187] |
| $w$: rat weight | 0.25 | 0.26 | 0.35 | $kg$ | internal data |
| $\alpha$: correction factor[a] | 15.6 | 9.2 | 103.57 | | |
| $GFR$: glomerular fraction rate | 1.62 | 1.134 (-30%) | 2.106 (+30%) | $mL/min$ | [188] |
| $\rho$: drug true density | 1 | 0.5 | 1.5 | $mg/mL$ | [189] |

[a] permeability correction factors were calculated from bronchial and alveolar cell layer thickness taken from [190, 191, 192, 193, 194, 195, 196].

Table 5.3: Drug related parameters used for inter-compounds GSA

| **Parameters**[a] | min value | max value | units |
|---|---|---|---|
| $mw$: molecular weight | 334.4 | 769.2 | $g/mol$ |
| $MMAD$: mean aerodynamic diameter | 1.106 | 6.136 | $\mu m$ |
| $GSD$: geometric standard deviation of the diameters | 0.84 | 3.432 | $\mu m$ |
| $D_0^b$: dose number | 0.045 (1) | 1 (160.9) | |
| $f_{u,l}$: fraction unbound lung tissue | 0.001 | 0.264 | |
| $f_{u,elf}$: fraction unbound fluid | 0.1 | 1 | |
| $P_p$ CALU3: fitted passive permeability | 12.85 | 174.7 | $nm/s$ |
| $P_a$ CALU3: fitted active permeability | $4 \cdot 10^{-6}$ | 60600 | $nm/s$ |

[a] all the parameters were internally available.
[b] for min and max values: *high solubility* (*poor solubility*).

Table 5.4: Drug A parameters for intra-compound GSA

| **Parameters**[a] | baseline | min value | max value | units |
|---|---|---|---|---|
| $dose^b$ | 10 | | | $\mu g$ |
| $logP^c$ | 3.87 | 3.72 (-30%) | 3.98 (+30%) | |
| $pKa$ | 8.7 | 8.6 | 8.8 | |
| $P_p$ CALU3: fitted passive permeability | 55.86 | 16.76 (-70%) | 94.96 (+70%) | $nm/s$ |
| $P_a$ CALU3: fitted active permeability | $4 \cdot 10^{-6}$ | $1.2 \cdot 10^{-6}$ (-70%) | $6.8 \cdot 10^{-6}$ (+70%) | $nm/s$ |
| $Caco2_{AB}$: gut wall permeability | 4.7 | 1.41 (-70%) | 7.99 (+70%) | $nm/s$ |
| $BP$: blood to plasma partition coefficient | 0.8 | 0.72 (-10%) | 0.88 (+10%) | |
| $E_r$: extraction ratio | 0.8 | 0.56 (-30%) | 1 | |
| $C_s$: drug solubility | 696 | 487.2 (-30%) | 904.8 (+30%) | $ng/mL$ |
| $MMAD^d$: mean aerodynamic diameter | 2.59 | 2.2 (-30%) | 4.9 (+30%) | $\mu m$ |
| $GSD^d$: geometric standard deviation | 2.1 | 1.12 (-30%) | 2.51 (+30%) | $\mu m$ |
| $f_{u,elf}$: fraction unbound fluid | 0.16 | 0.112 (-30%) | 0.208 (+30%) | |
| $f_{u,l}$: fraction unbound lung tissue | 0.0015 | 0.001 (-30%) | 0.002 (+30%) | |
| $f_{u,p}$: fraction unbound plasma | 0.032 | 0.0224 (-30%) | 0.0416 (+30%) | |

[a] minimum or maximum range limit (difference with respect to the baseline value, in percentage).
[b] the dose was considered normally distributed with a CV equal to 15% [185].
[c] the ranges were calculated as ±30% of the natural value.
[d] ranges were set equal to -30% the minimum and +30% the maximum of multiple measurements.

Table 5.5: Drug B parameters for intra-compound GSA

| **Parameters**[a] | baseline | min value | max value | units |
|---|---|---|---|---|
| $dose^b$ | 15 | | | $\mu g$ |
| $logP^c$ | 1.99 | 1.84 (-30%) | 2.1 (+30%) | |
| $pKa$ | 9.81 | 9.71 | 9.91 | |
| $P_p$ CALU3: fitted passive permeability | 16.06 | 4.81 (-70%) | 27.3 (+70%) | $nm/s$ |
| $P_a$ CALU3: fitted active permeability | 68.82 | 20.65 (-70%) | 117 (+70%) | $nm/s$ |
| $Caco2_{AB}$: gut wall permeability | 49.2 | 14.76 (-70%) | 83.64 (+70%) | $nm/s$ |
| $BP$: blood to plasma partition coefficient | 1.6 | 1.44 (-10%) | 1.76 (+10%) | |
| $E_r$: extraction ratio | 0.95 | 0.67 (-30%) | 1 | |
| $C_s$: drug solubility | 360000 | 252000 (-30%) | 468000 (+30%) | $ng/mL$ |
| $MMAD^d$: mean aerodynamic diameter | 3.2 | 2.18 (-30%) | 4.34 (+30%) | $\mu m$ |
| $GSD^d$: geometric standard deviation | 1.67 | 1.08 (-30%) | 2.3 (+30%) | $\mu m$ |
| $f_{u,elf}$: fraction unbound fluid | 1 | 0.7 (-30%) | 1 | |
| $f_{u,l}$: fraction unbound lung tissue | 0.26 | 0.18 (-30%) | 0.34 (+30%) | |
| $f_{u,p}$: fraction unbound plasma | 0.82 | 0.58 (-30%) | 1 | |

[a] minimum or maximum range limit (difference with respect to the baseline value, in percentage).
[b] the dose was considered normally distributed with a CV equal to 15% [185].
[c] the ranges were calculated as ±30% of the natural value.
[d] ranges were set equal to -30% the minimum and +30% the maximum of multiple measurements.

Table 5.6: Drug C parameters for intra-compound GSA

| **Parameters**[a] | baseline | min value | max value | units |
|---|---|---|---|---|
| $dose^b$ | 23 | | | $\mu g$ |
| $logP^c$ | 5.4 | 5.25 (-30%) | 5.51 (+30%) | |
| $pKa$ | 8.5 | 8.4 | 8.6 | |
| $P_p$ CALU3: fitted passive permeability | 20.54 | 6.16 (-70%) | 34.92 (+70%) | $nm/s$ |
| $P_a$ CALU3: fitted active permeability | 598.74 | 179.62 (-70%) | 1017.9 (+70%) | $nm/s$ |
| $Caco2_{AB}$: gut wall permeability | 0.3 | 0.09 (-70%) | 0.51 (+70%) | $nm/s$ |
| $BP$: blood to plasma partition coefficient | 1 | 0.9 (-10%) | 1.1 (+10%) | |
| $E_r$: extraction ratio | 0.95 | 0.67 (-30%) | 1 | |
| $C_s$: drug solubility | 14300 | 10010 (-30%) | 18590 (+30%) | $ng/mL$ |
| $MMAD^d$: mean aerodynamic diameter | 1.71 | 1.11 (-30%) | 2.54 (+30%) | $\mu m$ |
| $GSD^d$: geometric standard deviation | 2.33 | 1.51 (-30%) | 3.39 (+30%) | $\mu m$ |
| $f_{u,elf}$: fraction unbound fluid | 0.1 | 0.07 (-30%) | 0.13 (+30%) | |
| $f_{u,l}$: fraction unbound lung tissue | 0.001 | 0.0007 (-30%) | 0.0013 (+30%) | |
| $f_{u,p}$: fraction unbound plasma | 0.0015 | 0.001 (-30%) | 0.002 (+30%) | |

[a] minimum or maximum range limit (difference with respect to the baseline value, in percentage).
[b] the dose was considered normally distributed with a CV equal to 15% [185].
[c] the ranges were calculated as ±30% of the natural value.
[d] ranges were set equal to -30% the minimum and +30% the maximum of multiple measurements.

## 5.3 Results

### 5.3.1 Inter-compounds GSA results

Here the results of the inter-compounds GSA on the pulmonary absorption model decoupled from the whole-body PBPK, for both highly and poorly soluble compounds are reported. In figure 5.3, the distributions of the selected model outputs are showed. In figure 5.4 the sensitivity indices for $f_a$ and for the logarithms of AUC and MRT of the drug concentration in the lungs are reported. We choose the log scale for AUC and MRT to avoid possible imprecisions in the variance based sensitivity indices estimation due to the skewness of AUC and MRT distributions in natural scale [92].

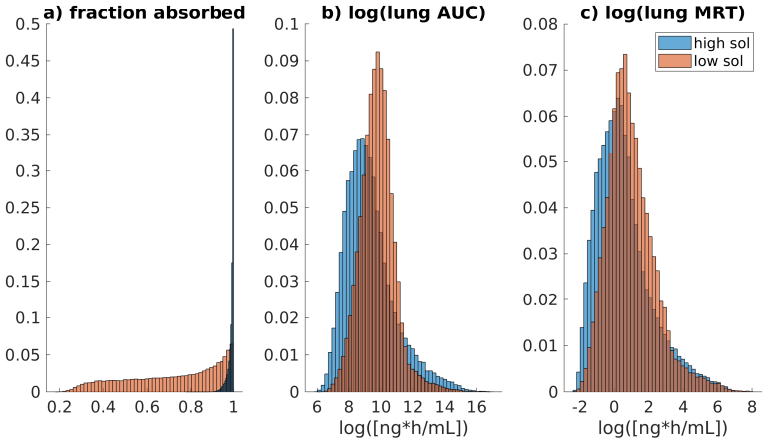



Figure 5.3: Inter-compounds variability for highly and poorly soluble compounds of: a) fraction absorbed; b) logarithm of whole lung AUC; c) logarithm of whole lung MRT.

For highly soluble compounds, the parameters that mostly explain the $f_a$ variability are $D_{0,inh}$ and $MMAD$. Concerning $D_{0,inh}$, it is probably important because it controls the solubility, and so the dissolution rate. The higher the dissolution rate is, the faster the drug is removed from the solid compartment in the central region. That is in fact the region in which

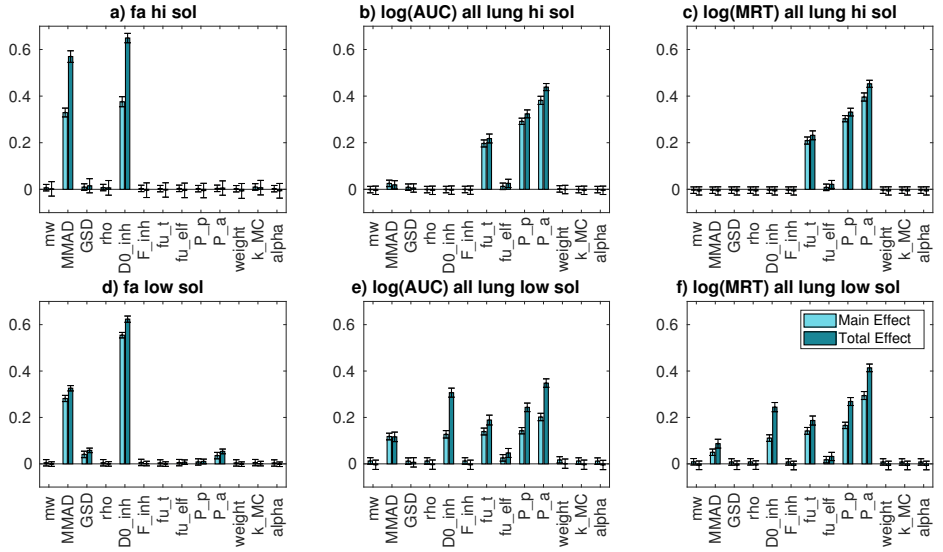



Figure 5.4: Inter-compounds GSA results for the output of interests: a) fraction absorbed for highly soluble compounds; b) logarithm of whole lung AUC for highly soluble compounds; c) logarithm of whole lung MRT for highly soluble compounds; d) fraction absorbed for poorly soluble compounds; e) logarithm of whole lung AUC for poorly soluble compounds; f) logarithm of whole lung MRT for poorly soluble compounds.

the drug could be eliminated via mucociliary clearance. However, the $f_a$ variability is quite low, as shown in figure 5.3. So, $D_{0,inh}$ and $MMAD$ are the most important parameters, but actually the variation of the model output is quite limited.

For poorly soluble compounds, both $MMAD$ and $D_{0,inh}$, even if with a minor contribution of $GSD$ and $P_a$, still impact the $f_a$ variability. Concerning $D_{0,inh}$, the reasons of its importance are probably the same of highly soluble compounds. $MMAD$ could be important in determining $f_a$ variability mainly for two reasons. First, the $MMAD$ value could impact the dissolution rate when the solubility is low. Second, it determines $F_C$ and, mainly for poorly soluble compounds, $f_a$ could be sensitive to the repartition between central and peripheral regions. In fact, in peripheral region the mucociliary clearance does not occur. The main difference with respect

to the high solubility case is that $f_a$ variability is quite high. Thus, $MMAD$ and $D_{0,inh}$ are responsible for a great variation of the model output.

Concerning lung AUC, the parameters that mostly explain its variability for highly soluble compounds are $f_{u,l}$, $P_p$ and $P_a$. This probably happens because $f_{u,l}$ and the permeabilities are parameters that determine the drug retention into the lungs and thus they control the AUC. For poorly soluble compounds, in addition to parameters that control the drug retention into the lungs, $D_{0,inh}$ and $MMAD$ are also important. This happens because, as explained before, they could impact the $f_a$ and, thus, the lung AUC. Concerning lung MRT, the parameters that mostly explain the variability for highly soluble compounds are still $f_{u,l}$ and both passive and active permeabilities. The reasons of their importance are probably similar to the one for the AUC: these are the parameters responsible to the drug retention into the lungs. For poorly soluble compounds, the most important parameters are similar to the ones for highly soluble compounds, with the addition of $D_{0,inh}$ and $MMAD$.

### 5.3.2 Intra-compound GSA results

Here the results of the intra-compound GSA for three representative compounds belonging to the *Chiesi* portfolio, namely A, B and C, are reported. With respect to all the other *Chiesi* compounds, compound A is characterized by a lower solubility, a higher permeability and a low $f_{u,l}$. Compound B has a higher solubility, a lower permeability and a higher $f_{u,l}$. Finally, compound C has a mean solubility and permeability and a low $f_{u,l}$. The parameters values and associated uncertainty or variability are reported in tables 5.4, 5.5 and 5.6.

#### Compound A

The distribution of drug plasma and whole lung AUC and MRT are reported in figure 5.5, while the sensitivity analysis results are reported in figure 5.6. Whole lung AUC variance is mainly explained by the passive permeability variation, together with a minor contribution of the dose, $f_{u,t}$,

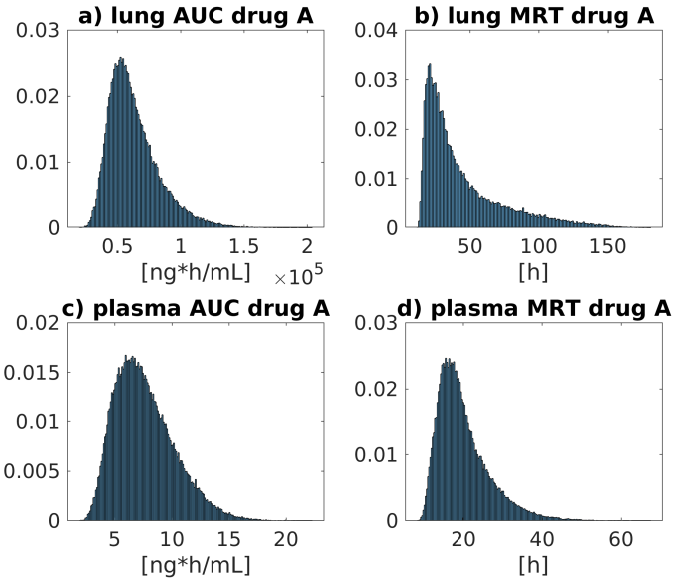



Figure 5.5: Intra-compound uncertainty for compound A: a) whole lung AUC; b) whole lung MRT; c) plasma AUC; d) plasma MRT.

rat weight and $k_{MC}$ variabilities. Even if the drug has a low solubility, it seems that the highest impact on the AUC is attributed to the passive permeability. This probably happens because compound A has very low fraction unbound in tissue, thus, even if the permeability parameter of the free compound is high, the overall drug permeability results to be the limiting step. In addition to that, it should be noted that the uncertainty associated to the passive permeability is higher with respect to those associated to other parameters. The passive permeability is still the most important parameter when whole lung MRT is considered. The reasons seem to be similar to those discussed for the AUC. Plasma AUC variability is mainly explained by the extraction ratio and, to a minor extent, by the dose, $MMAD$ and $GSD$ variabilities. These results highlight that the elimination process plays a major role in determining the plasma AUC vari-

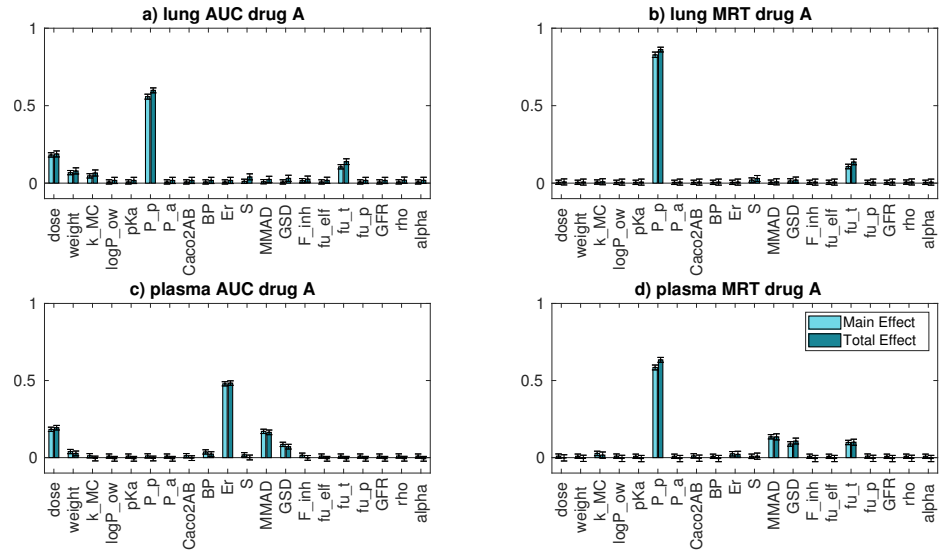



Figure 5.6: Intra-compound GSA results for compound A: a) whole lung AUC; b) whole lung MRT; c) plasma AUC; d) plasma MRT.

ability. Concerning the plasma MRT, the most important parameter is the passive permeability. This probably happens because the drug is slowly absorbed from the lungs into the systemic circulation.

### Compound B

The distribution of drug plasma and whole lung AUC and MRT are reported in figure 5.7, while the results of the sensitivity analysis are reported in figure 5.8. The parameter that mostly explains compound B whole lung AUC and MRT variation is the passive permeability. This probably happens because compound B has a lower permeability and higher solubility with respect to the other compounds of interest, thus, the absorption rate is probably permeability limited. Moreover, as explained for compound A, with respect to all the other parameters, the permeabilities have associated a greater uncertainty, thus, it is more likely that they have a relevant im-

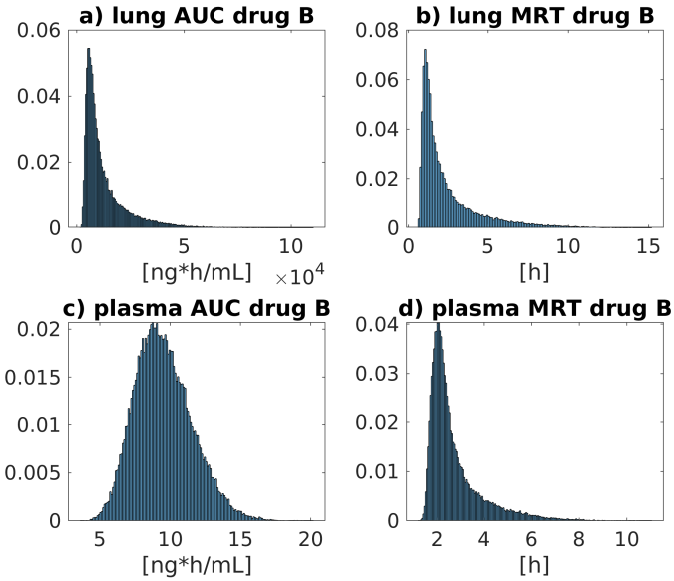



Figure 5.7: Intra-compound uncertainty for compound B: a) whole lung AUC; b) whole lung MRT; c) plasma AUC; d) plasma MRT.

pact in explaining the AUC variation. Concerning the plasma AUC, the most important parameters are the dose and the extraction ratio, followed by rat weight and $BP$. As for compound A, the elimination process is more important than the distribution or absorption processes in determining the AUC variation. The most important parameter in explaining the MRT variance is the passive permeability. As for compound A, this probably happens because the drug is slowly absorbed from the lungs into the systemic circulation.

### Compound C

The distribution of plasma and whole lung AUC and MRT are reported in figure 5.9, while the results of the sensitivity analysis are reported in

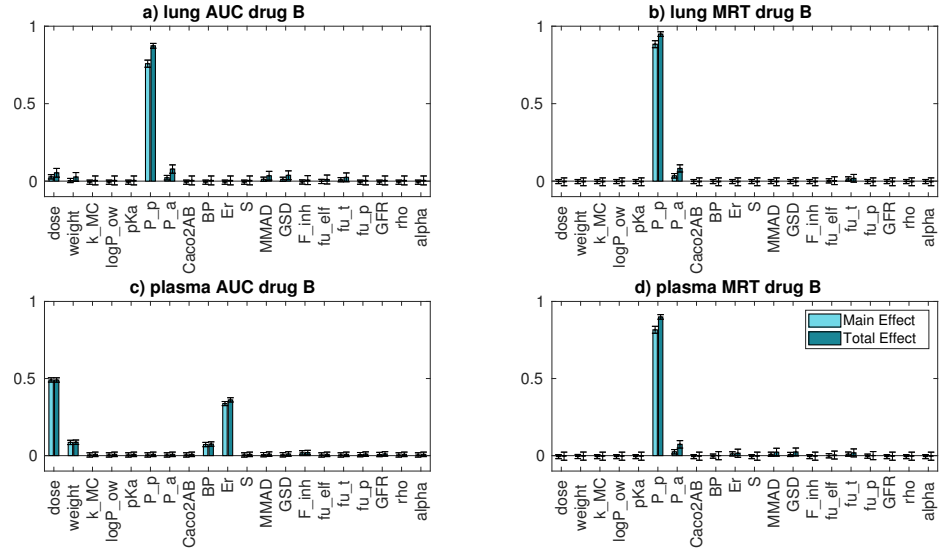



Figure 5.8: Intra-compound GSA results for compound B: a) whole lung AUC; b) whole lung MRT; c) plasma AUC; d) plasma MRT.

figure 5.10. For both whole lung AUC and MRT, the parameter that with its variation mainly explains their variability is the passive clearance. The reasons are probably similar to those reported for compounds A and B. Moreover, compound C has a higher solubility, but lower permeability than compound A. This, together with a low fraction unbound, could explain the slightly higher importance of the passive permeability with respect to compound A. Concerning the plasma AUC, the parameters that mainly explain its variation are the dose and the extraction ratio, followed by the passive permeability, rat weight and $B : P$. These results highlight that probably, for this drug, the elimination process has a greater role in determining the AUC variability with respect to the distribution or absorption. Concerning the plasma MRT, the situation resembles the one of compound A.

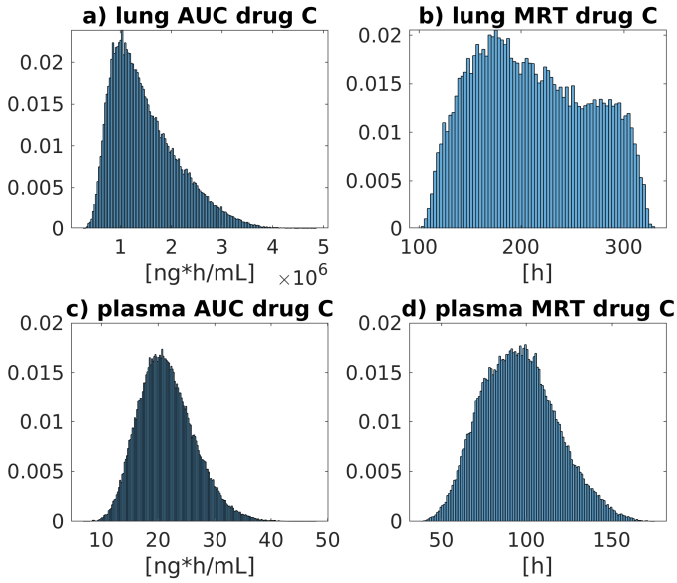



Figure 5.9: Intra-compound uncertainty for compound C: a) whole lung AUC; b) whole lung MRT; c) plasma AUC; d) plasma MRT.

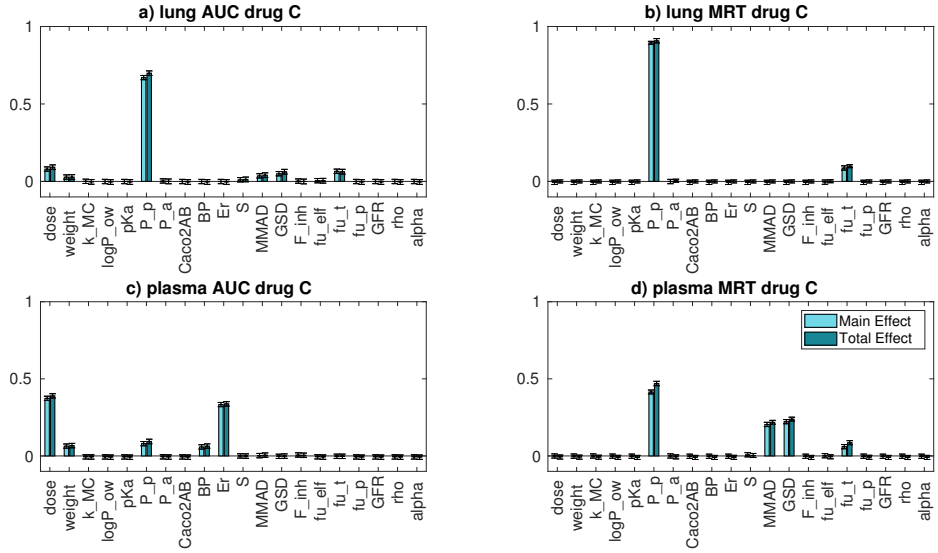



Figure 5.10: Intra-compound GSA results for compound C: a) whole lung AUC; b) whole lung MRT; c) plasma AUC; d) plasma MRT.

## 5.4 Discussion

In this manuscript, we showed how GSA techniques were used to assess model behaviours and support the development of a mechanistic model describing pulmonary absorption for orally inhaled compounds. We identified two ways of performing a GSA, that differ in the aims and thus, in the considered parameters variability: inter-compounds and intra-compound. Both the approaches helped in understanding different model aspects.

The inter-compounds GSA was performed for the absorption model decoupled from the distribution PBPK. In particular, this analysis can be performed in more 'homogeneous' sub-spaces of the whole parameter space, as we have done distinguishing highly soluble from poorly soluble compounds. Looking at the model output distributions gives the possibility of assessing the extent of the inter-compounds variability of the metrics of interest and then its relevance. Then, GSA helps in understanding what are the parameters that mostly determine the observed variation of the output predictions between different compounds. For example, from figure 5.3 it is possible to appreciate the difference in the model behaviour and in the impact of the model parameter variability between highly and poorly soluble compounds. We found that inter-compounds GSA is particularly useful during the process of model development. In fact, this analysis can help in understanding if the model behaves as expected and, in case of discrepancies between the expected and the actual model behaviours, GSA gives useful information that helps in identifying the reasons. Theoretically, if the model structure and the physiological and inter-compounds parameter variabilities are correctly identified and fixed, this analysis can be performed just once (e.g., when the model is firstly presented or at the platform release).

The intra-compound GSA was instead performed for three representative compounds on the whole body PBPK model. The parameters variation was defined to represent the uncertainties associated to their values for a specific compound. With this analysis it is possible to know how much the model output variation is apportioned to the uncertainty of the parameters. When doing this analysis, it is useful to look at the output distribution, to determine if it is narrow enough to be considered acceptable. If not, GSA

helps in selecting what parameters should be known with less degree of uncertainty in order to give a more accurate prediction. For example, we believe that the uncertainty associated to compound A lung MRT is too high. So, if one is interested in using this model for lung MRT predictions (e.g., for different dosages or species), from the GSA results we know that a better characterisation of the passive permeability is needed in order to reduce uncertainty of the considered metric. This situation probably does not happen for compound A lung AUC, given that the uncertainty associated with this metrics can be considered low. Differently from inter-compounds GSA, intra-compound GSA should be performed each time the model (or the PBPK platform) is used for a specific drug.

In our experience, GSA showed some criticisms. For example, due to the lack of available data, in certain situations determining the uncertainty or variability ranges was not an easy task, in particular for the intra-compound GSA. In these cases, expert opinion has to be used to fill the gap.

In conclusion, we suggest the use of GSA during the model development and evaluation, especially for the development of (complex) mechanistic models. GSA increases the knowledge of the model, it helps in finding errors, sensitive assumptions and it identifies the parameters that must be known with higher confidence if one is interested in reducing the model prediction uncertainties.

Chapter 6

# Accounting for inter-correlation between enzyme abundance: assessing implications on model predictions with uncertainty and sensitivity analysis [1]

[1]This work was published in "N. Melillo, A.S. Darwich, P. Magni, A. Rostami-Hodjegan. Accounting for inter-correlation between enzyme abundance: a simulation study to assess implications on global sensitivity analysis within physiologically-based pharmacokinetics. *Journal of Pharmacokinetics and Pharmacodynamics*, 46(2):137-154, April 2019" [56].

## 6.1 Introduction

In the recent EMA draft guidance on the qualification and reporting of PBPK models, the need to assess the biological plausibility and validity of model parameter values was highlighted. Further, EMA encouraged PBPK platforms to evolve with new published data [50]. In order to simulate the variability in pharmacokinetics of a given drug in a given population, a set of individuals are generated by sampling model parameter values within a predefined space (e.g., liver volume and blood flow) [197, 137]. In PBPK simulations for large populations of virtual patients, if during the process of assigning system parameters, the correlations between various anatomical, physiological and biological attributes are ignored, it is possible to generate *implausible combinations of parameters*. This could lead to an erroneous estimation of the interindividual variability of the main pharmacokinetic parameters, such as clearance [30, 198].

Accordingly, most PBPK models implement, for example, the known correlation between organ volumes and blood flows [137, 199]. However, there are other physiological and biological components of the system with known degrees of correlation which are not considered in the development of virtual populations. These include enzyme expressions in various organs and tissues. Recently, novel techniques in quantitative proteomics has made possible the reliable quantification of multiple enzyme and transporter expressions in the same experiment and sample. This allows robust measures of correlations between proteins [200, 201]. For example, it has been shown that various CYPs enzymes and *UDP-glucuronosyltransferases* (UGT) are correlated [202]. Recent work has demonstrated that including the correlation between hepatic CYP3A4 and CYP2C8 in PBPK modelling of repaglinide, leads to improved performance in predicting inter-individual variability in drug clearance and the magnitude of metabolic drug-drug interactions (DDIs), as compared to the same PBPK predictions without correlation [201]. While the correlation between hepatic CYPs and UGTs has been robustly assessed, that between CYP3A4 in liver and gut wall remains controversial: some work suggests its absence [203], while other work suggests its presence [204]. Moreover, to our knowledge, evidence

in the literature regarding any correlation between enzyme and transporter expressions and their relevance are still insufficient [201], however, this type of correlation might be possible (internal unpublished data).

In this context, the *uncertainty analysis* of mathematical models may be informative in highlighting possible correlations that may impact the PK of drugs and, by extension, may inform experimental design. The aim of our work was to perform an uncertainty analysis in order to theoretically assess the impact of the correlation between the expressions of two enzymes, under different conditions, on the PK of drugs that are substrates of both enzymes. For this purpose, we considered three simple compartmental semi-physiological models representing the cases of: 1) intravenously (IV) administered drugs metabolised by two enzymes expressed in the liver; 2) orally administered drugs metabolised by CYP3A4 expressed in the gut wall and liver; 3) IV administered drugs that are substrates of CYP3A4 and of the transporter *organic-anion-transporting polypeptide* (OATP) 1B1, both expressed in the liver. For each of these models we tested the impact of correlation between the expression of the two enzymes with respect to the case of absence of correlation. We choose to use simple compartmental models in order to highlight the correlation effect (if present), while limiting the effect of confounding variables and thus, simplifying the interpretation of the results.

Finally, the EMA report highlighted the importance of performing a sensitivity analysis *'for all the parameters that are likely to markedly influence the outcome of simulated pharmacokinetics and/or the model application'* [205]. Hence, we assessed how the correlation between enzymatic abundances impacts the results and interpretation of sensitivity analysis [30]. For this purpose, we performed GSA [60] on a semi-physiological model describing *repaglinide* PK after IV infusion. Repaglinide is a substrate of CYP3A4 and CYP2C8, where the abundances of these two enzymes have been shown to be correlated in the liver [202]. Different degrees of correlation between the two enzymes were tested in terms of their impact on the GSA results. As suggested in Saltelli *et al.*, [21], we choose to express the correlation between the two enzymes by using an explicit relationship (e.g., linear regression) with the addition of an independent noise term and

then performed a GSA on the uncorrelated parameters. The technique that we choose to use to perform the GSA was the variance based method [21, 87, 69].

## 6.2 Simulation study: uncertainty and sensitivity analysis

A simulation study was performed to assess the effect of the correlation between enzymatic expressions, in different situations, on the PK of drugs with different intrinsic clearances. Then, the impact of correlation between hepatic CYP3A4 and CYP2C8 expressions on the results of a variance based GSA of a semi-physiological model describing repaglinide PK after IV infusion was assessed. The model parameter values that are considered fixed are reported in table 6.1, the distributions of the enzyme concentrations are reported in table 6.2 and the parameters related to repaglinide metabolism and distribution are given in 6.3. The analysis was performed using MATLAB R2017b on a 64-bit computer configured with Intel® Core™ i7-7700 3.60 GHz x 8 processor, running Ubuntu 16.04 LTS.

### 6.2.1 Model for two CYPs in the liver

A semi-physiological two-compartmental model was developed to simulate IV administration at an arbitrary dose of 1 $mg$ of drug. The model, represented in figure 6.1 (a), is composed of one central compartment and one compartment representing the liver. Drug metabolism occurs in the liver and depends on the activity of two different enzymes, namely enzyme 1 ($E_1$) and enzyme 2 ($E_2$), with equal intrinsic clearances. Model equations are shown in equation system 6.1.

$$\begin{aligned} \frac{dx_c}{dt} &= \frac{Q_{liv,art}}{V_{liv}} x_{liv} - \frac{Q_{liv,art}}{V_c} x_c \\ \frac{dx_{liv}}{dt} &= \frac{Q_{liv,art}}{V_c} x_c - \frac{Q_{liv,art}}{V_{liv}} x_{liv} - \frac{CL_1 + CL_2 I}{V_{liv}} x_{liv} \end{aligned} \tag{6.1}$$

$x_c$ and $x_{liv}$ are the drug amounts in the central and liver compartment, respectively. $V_c$ and $V_{liv}$ are the central and liver compartment volumes and $Q_{liv,art}$ is the arterial liver blood flow. The value of $V_c$ was taken equal to the median of drug distribution volumes in [206] minus $V_{liv}$. $CL_1$ and

Table 6.1: Model parameters values

| **Parameters** | value | units | reference |
|---|---|---|---|
| $BW$: body weight | 70 | $kg$ | [170] |
| $CO$: cardiac output | 350.37 | $L/h$ | [169] |
| $W_{liv}$: liver weight (fraction of BW) | 1.81 (0.026) | $kg$ | [170] |
| $V_c$: central compartment volume (for models in Figures 1a and 1c)[a] | 115.91 | $L$ | [206] |
| $V_{ent}$: volume of the enterocytes | 0.2242 | $L$ | [169] |
| $Q_{liv,art}$: liver blood flow (percentage of CO) | 68.32 (0.195) | $L/h$ | [170] |
| $Q_{liv,ven}$: hepatic vein blood flow (percentage of CO) | 89.34 (0.255) | $L/h$ | [170] |
| $Q_{ent}$: blood flow out of the enterocytes | 16.82 | $L/h$ | [169] |
| $k_t$: time constant relative to the transit in the intestine | 0.3 | $h^{-1}$ | Calculated[d] |
| $P_{eff}$: effective permeability across the gut wall[b] | 8.70 | $10^{-4} cm/s$ | [157] |
| $\rho_{liv}$: liver density[c] | 1.080 | $kg/L$ | [171] |
| $f_{ew,liv}$: fraction of extracellular water in liver | 0.161 | | [40] |
| $f_{iw,liv}$: fraction of intracellular water in liver | 0.573 | | [40] |
| $MPPGL$: $mg$ of microsomal proteins per $g$ of liver | 39.79 | $mg\ prot/g\ liver$ | [160] |
| $HPGL$: hepatocellularity per $g$ of liver | 117.52 | $10^6$ $hepatocytes/g\ liver$ | [160] |

[a] Calculated as the median of the values in [206] times $BW$ (70 $kg$) minus the liver volume.
[b] Taken equal to the maximum value in [157].
[c] Used to calculate $V_{liv}$ from $W_{liv}$: $V_{liv} = W_{liv}/\rho_{liv}$.
[d] Calculated dividing the minimum of the mean flow rate in the jejunum (0.5 $mL/min$) by the minimum volume (100 $mL$) reported in [158].

Table 6.2: Parameters of the enzymatic distributions

| **Parameters** | distribution parameters | distribution type | units | reference |
|---|---|---|---|---|
| $[CYP3A4]_{mic}$: CYP3A4 microsomal concentration | 137 (41%) | Lognormal[a] | $pmol/mg\ prot$ | [160] |
| $CYP3A4_{ent}$: total amount of CYP3A4 in small intestine | 66.2 (60%) | Lognormal[a] | $nmol$ | [160] |
| $[CYP2C8]_{mic}$: CYP2C8 microsomal concentration | 24 (81%) | Lognormal[a] | $pmol/mg\ prot$ | [160] |
| $[OATP1B1]_{cells}$: OATP1B1 concentration per $10^6$ cells | 4.28 (74%) | Lognormal[a] | $pmol/10^6\ hepatocytes$ | [160] |

[a] For distribution parameters, *mean (coefficient of variation)* of the lognormal random variable.

Table 6.3: Repaglinide drug-specific parameters

| **Parameters** | value | units | reference |
|---|---|---|---|
| $mw$: molecular weight | 452.6 | $g/mol$ | [201] |
| $V_{max,\ 3A4}$[a] | 958.2 | $pmol/min/mg\ prot$ | [207] |
| $K_{M,3A4}$ | 13.2 | $\mu M$ | [207] |
| $V_{max,2C8}$[a] | 300.8 | $pmol/min/mg\ prot$ | [207] |
| $K_{M,2C8}$ | 2.3 | $\mu M$ | [207] |
| $CL_{int,OATP}$: intrinsic clearance of OATP | 246 | $\mu L/min/10^6\ hepatocytes$ | [201] |
| $CL_{PS}$: passive clearance | 0.089 | $mL/min/10^6\ hepatocytes$ | [201] |
| $V_{ss}$: volume of distribution at the steady state | 0.24 | $L/kg$ | [201] |
| $fu_t$: fraction unbound in liver tissue | 0.072 | | [208, 209] |

[a] HLM parameters in [201].

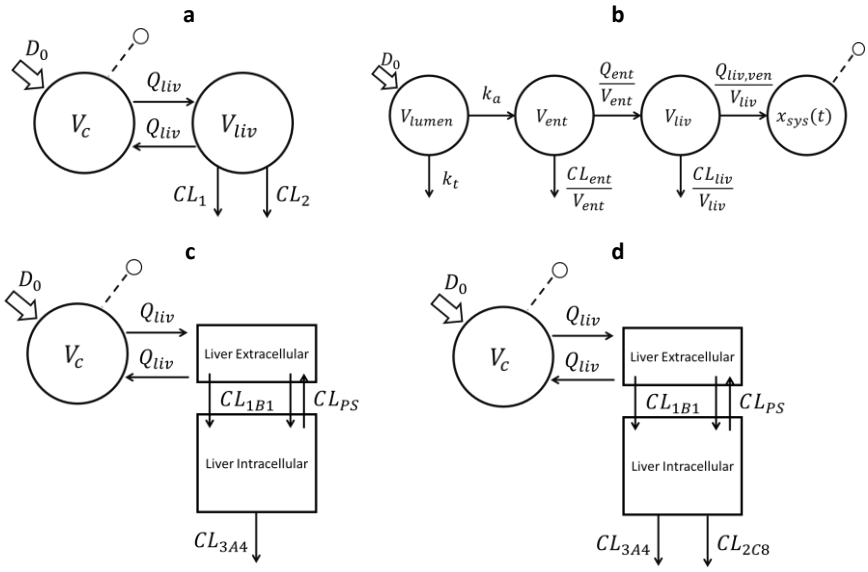



Figure 6.1: Different models used in this study: a) intravenously administered drugs metabolised by two enzymes expressed in the liver; b) orally administered drugs metabolised by CYP3A4 expressed in the liver and gut wall; c) intravenously administered drugs metabolised by CYP3A4 and OATP1B1, both expressed in the liver; d) intravenous administration of repaglinide, a drug substrate of CYP3A4, CYP2C8 and OATP1B1.

$CL_2$ are the clearances of enzyme 1 and enzyme 2, respectively, and $I$ is an inhibition constant. The clearance expression is given in equation 6.2.

$$CL = CL_{int} \cdot [E]_{mic} \cdot MPPGL \cdot M_{liv} \tag{6.2}$$

$CL_{int}$ is the intrinsic clearance, $[E]_{mic}$ is the enzymatic concentration per $mg$ of microsomal proteins, $MPPGL$ is the $mg$ of microsomal protein per $g$ of liver and $M_{liv}$ the liver mass in $g$ [199].

As an exercise, the distributions of microsomal concentrations of both $E_1$ and $E_2$ were taken to equal the one of CYP3A4. We choose to set the intrinsic clearance and the distributions of $E_1$ and $E_2$ to equal values. This was done to give equal importance to the two enzymes on the drug

metabolism (for an average individual) and thus highlight the impact of correlation effect where present. All the model parameters were considered fixed to their mean value, shown in table 6.1, except $CL_{int}$, $I$ and the concentrations of the two enzymes. As for CYP3A4, both concentrations of enzyme 1 and 2 were assumed to be log-normally distributed [160] as shown in 6.3.

$$\begin{bmatrix} \ln([E]_1) \\ \ln([E]_2) \end{bmatrix} \sim \mathcal{N}\left( \begin{bmatrix} l\mu_1 \\ l\mu_1 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \rho\,\sigma_1\,\sigma_2 \\ \rho\,\sigma_1\,\sigma_2 & \sigma_2^2 \end{bmatrix} \right) \tag{6.3}$$

where $l\mu_i$ and $\sigma_i$ are the mean and standard deviation of the natural logarithm of the concentration of enzyme $i$. Both presence and absence of correlation between the two enzymes were considered. In particular, when the correlation is present it was assumed that $\rho = 0.7$, close to the physiological correlation between the logarithms of CYP3A4 and CYP2C8 microsomal concentrations[2]. $CL_{int}$ was considered variable within the ranges of 0.005 to 2949 $\mu L/min/mg\ microsomal\ protein$ as reported in [164]. Given that in [164] the values of $CL_{int}$ are expressed in $\mu L/min/mg\ microsomal\ protein$, in order to express the clearance in $\mu L/min/pmol\ CYP$, the original values were divided by the mean concentration of CYP3A4 per $mg$ of microsomal protein. Values of $I$ were considered between 0 and 2, which corresponds to complete inhibition and two-fold induction of the $E_2$, respectively.

A total of 5000 samples were extracted from the enzyme concentration distribution in both cases of presence and absence of correlation. The considered timespan ranged from 0 (time of dose administration) to 100 $h$. For the given values of $CL_{int}$ and $I$, the AUC from 0 to 100 $h$ of the central compartment following an IV bolus of 1 $mg$ was calculated for all the paired enzyme concentration samples. Thus, for each $CL_{int}$ and $I$ combination, two AUC distributions were obtained: one in presence and the other one in absence of correlation between the enzyme concentrations. The index $R_P$

[2]Data on CYP3A4 and CYP2C8 paired microsomal concentration on 23 Caucasian subjects were internally available [202]. The Pearson linear coefficient was calculated by using the MATLAB function `corr`, on the logarithms of the CYPs concentrations, resulting in $\rho = 0.7436$.

was defined as in equation 6.4, with $AUC_{P,corr}$ and $AUC_{P,not\ corr}$ given percentiles relative to the correlated and not correlated distribution of the AUC, respectively.

$$R_P = \frac{AUC_{P,corr} - AUC_{P,\ not\ corr}}{AUC_{P,not\ corr}} \tag{6.4}$$

$R_P$ represents the increase in the $P^{\text{th}}$ AUC percentile, $AUC_P$, due to the presence of correlation, relative to absence of any correlation. An $R_P > 0$ for the upper limit of the 95% confidence interval, means that this limit is increased to $(R_P \cdot 100)\%$ due to the presence of correlation. Instead, an $R_P < 0$ for the lower limit of the 95% confidence interval, means that this limit is reduced to $(R_P \cdot 100)\%$. The reasoning could be easily extended for $R_P < 0$ for the upper limit and $R_P > 0$ for the lower limit. $R_P$ was calculated for the $5^{\text{th}}$, $50^{\text{th}}$ and $95^{\text{th}}$ percentiles. Given that $R_P$ depends on $CL_{int}$ and $I$, its value was calculated for a grid of values of these two parameters. This grid was obtained by the combination of 20 values for both $CL_{int}$ and $I$, equally spaced between the ranges given before. Given that $CL_{int}$ spans in several orders of magnitude, a logarithmic spacing with lower and upper bounds set as the logarithms of minimum and maximum value in [164], respectively, was used for values of $CL_{int}$. To establish the values of $I$, a linear spacing with a lower bound equal to 0 and an upper bound equal to 2 was used. Finally, for each value of $CL_{int}$, the ratio ($R_{inh}$) between central compartment AUC when $I = 0$ (complete inhibition of $E_2$) and when $I = 1$ (absence of inhibition), was calculated for both the presence and absence of correlation.

### 6.2.2 Model for CYP3A4 in gut wall and in liver

A semi-physiological model was set up to describe the absorption and first pass effect for a generic drug. The model, shown in figure 6.1 (b), is composed of four compartments, representing intestinal lumen, the enterocytes, the liver and the cumulative amount of drug reaching the systemic

circulation. Model equations are shown in equation system 6.5.

$$\begin{aligned}
\frac{dx_{lum}}{dt} &= -\left(k_a + k_t\right) x_{lum} \\
\frac{dx_{ent}}{dt} &= k_a x_{lum} - \left(\frac{CL_{ent} I}{V_{ent}} + \frac{Q_{ent}}{V_{ent}}\right) x_{ent} \\
\frac{dx_{liv}}{dt} &= \frac{Q_{ent}}{V_{ent}} x_{ent} - \left(\frac{CL_{liv} I}{V_{liv}} + \frac{Q_{liv,ven}}{V_{liv}}\right) x_{liv} \\
\frac{dx_{sys}}{dt} &= \frac{Q_{liv,ven}}{V_{liv}} x_{liv}
\end{aligned} \tag{6.5}$$

$x_{lum}$, $x_{ent}$, $x_{liv}$ and $x_{sys}$ represent drug amount in gastrointestinal lumen, enterocytes, liver and cumulative amount that reaches the systemic circulation, respectively. $k_a = 2\,P_{eff}/r_i$ is the absorption constant, with $r_i = 1.75$ $cm$ the mean intestine lumen radius and $P_{eff}$ the effective permeability [157]. $P_{eff}$ was taken to equal to the maximum value in [157] ($8.70 \cdot 10^{-4}$ $cm/s$) to allow maximal drug absorption. $k_t$ is the time constant relative to small intestine transit. $V_{ent}$ and $V_{liv}$ are the enterocyte and liver volumes and $Q_{ent}$ and $Q_{liv,ven}$ are the enterocyte and venous liver blood flows, respectively. $CL_{ent}$ and $CL_{liv}$ are the clearances due to CYP3A4 activity in the gut wall (equation 6.6) and the liver (equation 6.2) and $I$ is the inhibition constant.

$$CL_{ent} = CL_{int,3A4} \cdot CYP3A4_{ent} \tag{6.6}$$

$CL_{int,3A4}$ is the intrinsic clearance and $CYP3A4_{ent}$ is the CYP3A4 amount in enterocytes.

$F_{oral}$ is equal to $\lim_{t\to+\infty}(x_{sys}/M_0)$. Its analytical expression was derived and is shown in equation 6.7, where $M_0$ is the drug dose. Derivation is shown in appendix D.

$$F_{oral} = \frac{k_a \cdot (Q_{ent}/V_{ent}) \cdot (Q_{liv,ven}/V_{liv})}{(k_a + k_t)\left(\frac{Q_{ent}}{V_{ent}} + \frac{CL_{ent}}{V_{ent}} \cdot I\right)\left(\frac{Q_{liv,ven}}{V_{liv}} + \frac{CL_{liv}}{V_{liv}} \cdot I\right)} \tag{6.7}$$

All the parameters were considered fixed, except $CL_{int,3A4}$, $I$, $CYP3A4_{ent}$ and the microsomal CYP3A4 concentration $[CYP3A4]_{mic}$. $CL_{int,3A4}$ and $I$

were considered as in the section 6.2.1. Both $CYP3A4_{ent}$ and $[CYP3A4]_{mic}$ were considered log-normally distributed [210, 160]. As an exercise, both presence and absence of correlation between CYP3A4 amount in the enterocytes and CYP3A4 concentration in liver microsomes, were considered. In the former case, the correlation coefficient $\rho$ was set equal to 0.7.

As in the previous section, 5000 samples were extracted from the joint distribution of CYP3A4 in gut wall and liver in the case of presence and absence of correlation. Then, the index $R_P$ was calculated for $F_{oral}$, for each combination of $CL_{int}$ and $I$. $R_P$ was calculated for the $5^{\text{th}}$, $50^{\text{th}}$ and $95^{\text{th}}$ percentiles. Finally, for each value of $CL_{int}$ the ratio ($R_{inh}$) was calculated between bioavailability when $I = 0$ (complete inhibition of both CYP3A4 in gut wall and liver) and when $I = 1$ (absence of inhibition). The analytical expression of $R_{inh}$ is reported in equation 6.8.

$$R_{inh} = \frac{\left(\frac{Q_{ent}}{V_{ent}} + \frac{CL_{ent}}{V_{ent}}\right)\left(\frac{Q_{liv,ven}}{V_{liv}} + \frac{CL_{liv}}{V_{liv}}\right)}{\frac{Q_{ent}}{V_{ent}} \cdot \frac{Q_{liv,ven}}{V_{liv}}} \tag{6.8}$$

### 6.2.3 Model for CYP3A4 and OATP1B1 in liver

A semi-physiological model was developed to describe IV administration of 1 $mg$ of a generic substrate of CYP3A4 and OATP1B1. The model, represented in figure 6.1 (c), is composed of one central compartment and two compartments representing the extracellular and intracellular space of

the liver. Model equations are shown in equation system 6.9.

$$\begin{aligned}
\frac{dx_c}{dt} &= \frac{Q_{liv,art}}{V_{liv,ext}} x_{liv,ext} - \frac{Q_{liv,art}}{V_c} x_c \\
\frac{dx_{liv,ext}}{dt} &= \frac{Q_{liv,art}}{V_c} x_c - \left( \frac{Q_{liv,art}}{V_{liv,ext}} + \frac{CL_{OATP}}{V_{liv,ext}} + \frac{CL_{PS}}{V_{liv,ext}} \right) x_{liv,ext} \\
&\qquad + \frac{CL_{PS}}{V_{liv,int}} x_{liv,int} \qquad (6.9) \\
\frac{dx_{liv,int}}{dt} &= \left( \frac{CL_{OATP}}{V_{liv,ext}} + \frac{CL_{PS}}{V_{liv,ext}} \right) x_{liv,ext} \\
&\qquad - \left( \frac{CL_{CYP}}{V_{liv,int}} + \frac{CL_{PS}}{V_{liv,int}} \right) x_{liv,int}
\end{aligned}$$

$x_c$, $x_{liv,ext}$ and $x_{liv,int}$ are drug amount in central, extracellular and intracellular liver compartments, respectively. $V_c$, $V_{liv,ext}$ and $V_{liv,int}$ are central, extracellular and intracellular liver volumes, respectively. $Q_{liv,art}$ is the liver arterial blood flow. $CL_{CYP}$ represents the CYP3A4 mediated clearance, as shown in the previous sections. $CL_{OATP}$ is the active transport clearance from the extracellular to the intracellular liver compartment and $CL_{PS}$ is the passive clearance across the hepatocytes' plasma membrane. $CL_{OATP}$ is defined in equation 6.10.

$$CL_{OATP} = CL_{int,OATP} \cdot [OATP1B1]_{cells} \cdot HPGL \cdot M_{liv} \qquad (6.10)$$

$CL_{int,OATP}$ is the intrinsic clearance, $[OATP1B1]_{cells}$ is the OATP1B1 concentration per $10^6$ hepatocytes, $HPGL$ the hepatocellularity per gram of liver and $M_{liv}$ the liver mass in $g$.

All the model parameters were fixed, except $CL_{int,3A4}$, $CL_{int,OATP}$, $[OATP1B1]_{cells}$ and $[CYP3A4]_{mic}$. Both $[OATP1B1]_{cells}$ and $[CYP3A4]_{mic}$ were considered log-normally distributed and, as an exercise, both the presence and absence of correlation between their expression were considered. In the former case the correlation coefficient $\rho$ was set equal to 0.7.
The ranges of $CL_{OATP}$ were taken from the uptake parameters in human hepatocytes [211]. With the hypothesis that active transport was mainly

due to OATP1B1 activity, clearance values in [211] were converted from $\mu L/min/10^6$ *hepatocytes* to $\mu L/min/pmol\ OATP1B1$, by dividing them by the mean OATP1B1 concentration per $10^6$ hepatocytes. $CL_{PS}$ was defined as 21.65% of the total transport clearance at a mean OATP1B1 concentration[3].

Similarly to what was done in the previous sections, 5000 samples from $[OATP1B1]_{cells}$ and $[CYP3A4]_{mic}$ distributions were extracted both in presence and absence of correlation. Then, the index $R_P$ was calculated for the AUC in the central compartment, for each combination of $CL_{int,OATP}$ and $CL_{int,3A4}$. $R_P$ was calculated for the $5^{\text{th}}$, $50^{\text{th}}$ and $95^{\text{th}}$ percentiles. The time span that was considered ranged from 0 (dose administration) to 100 $h$.

### 6.2.4 GSA in case of parameters correlation: the repaglinide example

A semi-physiological model, represented in figure 6.1 (d), was developed to describe repaglinide PK after an IV infusion of 2 $mg$ over 15 minutes. The model is composed of one central compartment and two compartments representing the extracellular and intracellular space of the liver. Active transport due to OATP1B1 and metabolism due to CYP3A4 and CYP2C8

[3]In [211] the percentages of active clearance with respect to the total transport clearance are reported. The fixed percentage of passive clearance was calculated as 1 minus the median of the active clearance percentages reported in the article. Total clearance was obtained dividing the active clearance for the fraction of active clearance.

are included. Model equations are shown in equation system 6.11.

$$\begin{aligned}
\frac{dx_c}{dt} &= \frac{Q_{liv,art}}{V_{liv,ext}} x_{liv,ext} - \frac{Q_{liv,art}}{V_c} x_c \\
\frac{dx_{liv,ext}}{dt} &= \frac{Q_{liv,art}}{V_c} x_c - \left( \frac{Q_{liv,art}}{V_{liv,ext}} + \frac{CL_{OATP} \cdot fu_t}{V_{liv,ext}} + \frac{CL_{PS} \cdot fu_t}{V_{liv,ext}} \right) x_{liv,ext} \\
&\quad + \frac{CL_{PS} \cdot fu_t}{V_{liv,int}} x_{liv,int} \\
\frac{dx_{liv,int}}{dt} &= \left( \frac{CL_{OATP} \cdot fu_t}{V_{liv,ext}} + \frac{CL_{PS} \cdot fu_t}{V_{liv,ext}} \right) x_{liv,ext} \\
&\quad - \left( met_{3A4} + met_{2C8} + \frac{CL_{PS} \cdot fu_t}{V_{liv,int}} \right) x_{liv,int} \\
met_{3A4} &= \frac{V_{max,3A4} CYP3A4_{liv}}{K_{M,3A4} + x_{liv,int} \cdot fu_t / V_{liv,int}} \cdot \frac{fu_t}{V_{liv,int}} \\
met_{2C8} &= \frac{V_{max,2C8} CYP2C8_{liv}}{K_{M,2C8} + x_{liv,int} \cdot fu_t / V_{liv,int}} \cdot \frac{fu_t}{V_{liv,int}}
\end{aligned} \tag{6.11}$$

$x_c$, $x_{liv,ext}$ and $x_{liv,int}$ are the drug amounts in central compartment, extracellular and intracellular liver compartments, respectively. $V_c$, $V_{liv,ext}$ and $V_{liv,int}$ are the central compartment, extracellular and intracellular liver volumes, respectively. $Q_{liv,art}$ is the liver arterial blood flow and $fu_t$ is the fraction unbound drug in the liver tissue [208, 209]. $CL_{OATP}$ is the transporter clearance, defined as in equation 6.10, and $CL_{PS}$ is the passive clearance. $V_{max,3A4}$, $K_{M,3A4}$, $V_{max,2C8}$ and $K_{M,2C8}$ are the Michaelis-Menten parameters of CYP3A4 and CYP2C8 catalysed reactions, respectively [207]. $V_{max}$ values in [207] were converted from *pmol/min/mg microsomal protein* to *pmol/min/pmol CYP* by dividing them for the mean microsomal CYP concentration. $CYP3A4_{liv}$ and $CYP2C8_{liv}$ are the amounts of CYP3A4 and CYP2C8 in the liver. Enzyme amounts are calculated as the enzymatic concentration per $mg$ of protein, times $MPPGL$, times liver mass in grams [199]. All the parameters were considered fixed, except OATP1B1, CYP3A4 and CYP2C8 abundances which were considered log-normally distributed.

A variance based GSA was performed on the predicted central com-

partment AUC, considering different levels of correlation between CYP3A4 and CYP2C8 microsomal concentrations. The joint distribution of the two CYPs' microsomal concentration is shown in equation 6.3. We considered different linear correlation coefficients ($\rho$) between the natural logarithm of the two cytochromes' concentrations: 0 (absence of correlation), 0.1, 0.3, 0.5, 0.7, 0.9 and a physiological $\rho$ equal to 0.7436 [202]. OATP1B1 abundance was considered to be independent from CYP3A4 and CYP2C8 since there information on correlations between transporters and enzymes are currently sparse and starting to emerge [212].

To express the correlation between CYP3A4 and CYP2C8 a linear regression with the addition of a noise term was used, as suggested by [21]. For each of the correlation coefficient, 1000 samples were extracted from the joint probability distribution of the natural logarithms of CYP3A4 and CYP2C8 concentrations. On these samples, parameters of a linear regression ($\beta_0$ and $\beta_1$) were identified considering CYP2C8 concentrations as a function of CYP3A4 concentration, a constant standard deviation for the residual error. CYP2C8 concentrations were then expressed as the values predicted by using the regression with the addition of a noise term $\varepsilon$, as reported in equation 6.12.

$$log([CYP2C8]_{mic}) = \beta_0 + \beta_1 \log([CYP3A4]_{mic}) + \varepsilon \tag{6.12}$$

$\varepsilon$ was considered normally distributed with a mean equal to 0 and variance equal to the one estimated from the residuals of the linear regression.

Then, for each correlation level, a variance based GSA was performed. $n$ samples extracted from the $k$-dimensional unit hypercube, with $k = 3$ (number of parameters), were manipulated to obtain the OATP1B1 and CYP3A4 concentration distributions and the distribution of $\varepsilon$. $n$ was set equal to 7000. The analysis was repeated five times to assess the variability of the sensitivity indices estimation. In conclusion, for each correlation level we calculated the main and the total effect for all the variable model parameters (i.e., CYP3A4, CYP2C8 and OATP1B1 concentrations). Finally, the same analysis was performed again, considering CYP3A4 expressed as a function of CYP2C8, with the purpose to assess the impact of the arbitrary

decision of what variable is the independent one in the linear regression model.

## 6.3 Results

### 6.3.1 Results for two CYPs in the liver

For the model represented in figure 6.1 (a) and described in equation system 6.1, the distribution of central compartment AUC was simulated for each combination of $CL_{int}$ and $I$, both in case of presence and absence of correlation between the liver expression of the two enzymes metabolising the drug. Then, for each of these AUC distribution pairs, the index $R_P$, in equation 6.4, was calculated for the 5$^{\text{th}}$, 50$^{\text{th}}$ and 95$^{\text{th}}$ percentiles. Results are shown in figure 6.2.

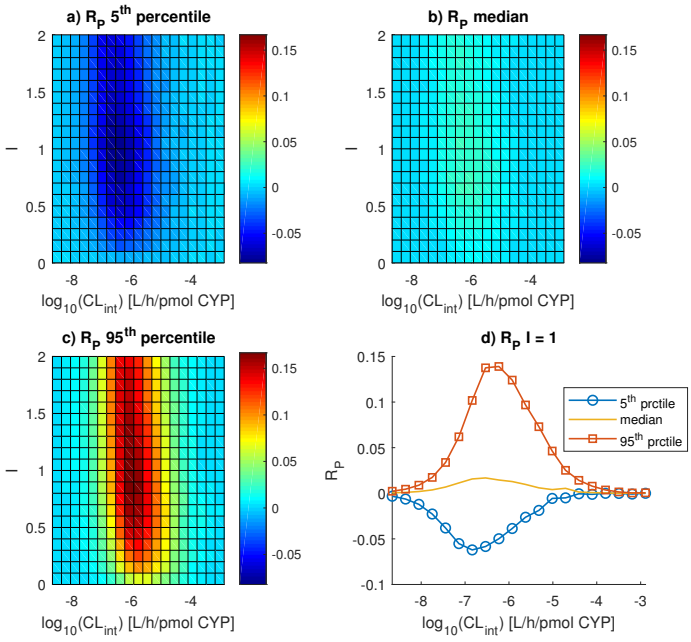



Figure 6.2: $R_P$ indices for the model in figure 6.1 (a), computed for different inhibition levels and $CL_{int}$ values: a) $R_P$ indices for the 5$^{\text{th}}$ percentile; b) $R_P$ indices for the median; c) $R_P$ indices for the 95$^{\text{th}}$ percentile; d) $R_P$ indices for 5$^{\text{th}}$ percentile, median and 95$^{\text{th}}$ percentile in absence of inhibition. For central $CL_{int}$ values the AUC distribution is wider ($R_{P,5^{th}} < 0$ and $R_{P,95^{th}} > 0$) in case of correlation with respect to the absence of that.

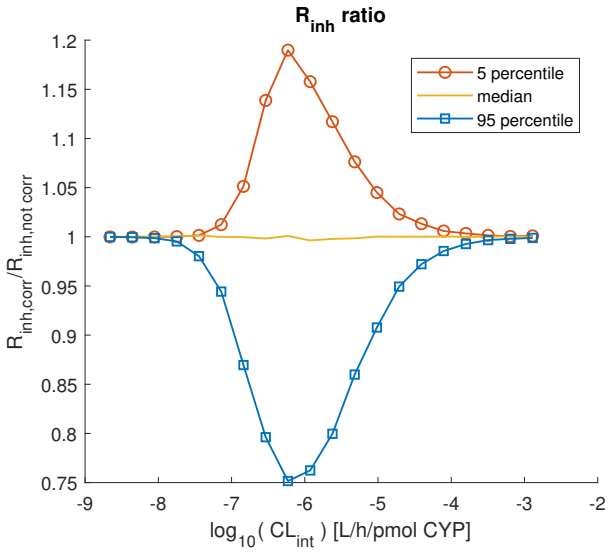



Figure 6.3: Ratio between $R_{inh,corr}$ and $R_{inh,not\,corr}$. It is possible to observe that in case of correlation the distribution of $R_{inh,corr}$ is narrower with respect to the one of $R_{inh,not\,corr}$ ($95^{\text{th}}$ percentile $< 0$ and $5^{\text{th}}$ percentile $> 0$). Moreover, it can be seen that the median of the ratios in case of correlation is equal to the one in case of absence of that.

Due to the presence of correlation between $E_1$ and $E_2$, the $5^{\text{th}}$ percentile of the central compartment AUC is reduced by more than 5% and the $95^{\text{th}}$ percentile increases by more than 15%, while the median remains almost stable. In the presence of correlation, when $E_1$ is highly expressed, $E_2$ is likely to be highly expressed as well. Instead, in case of absence of correlation, if $E_1$ is highly expressed then $E_2$ is equally likely to have any expression level according to its distribution. Therefore, in presence of correlation, the metabolism is likely to be higher, and thus the AUC is lower. For the same reason, in case of low expression of one of the two enzymes, the metabolism is likely to be lower, and so the AUC is higher if the enzymes are correlated, as compared to the case of uncorrelated expressions. For central values of $CL_{int}$ a higher correlation effect can be observed while a lower effect can be observed for extreme $CL_{int}$ values. For higher $CL_{int}$

values, this is probably a consequence of the metabolism becoming flow-limited and therefore less dependent on enzymatic abundances. For lower $CL_{int}$ values the clearance is so low that even inhibiting one enzyme does not change the overall absolute metabolic clearance.

For each value of $CL_{int}$, the ratio ($R_{inh}$) between the central compartment AUC when $I = 0$ and when $I = 1$ was calculated in case of presence ($R_{inh,corr}$) and absence ($R_{inh,not\,corr}$) of correlation. Then, the ratio $R_{inh,corr}/R_{inh,not\ corr}$ was computed and the results are shown in figure 6.3. When the correlation between enzymatic expression is present, the distribution of $R_{inh}$ is narrower in comparison to the absence of correlation. This effect is more apparent for central $CL_{int}$ values and it is weaker for extreme $CL_{int}$ values. In presence of correlation, when $E_1$ has high expression, $E_2$ is likely to have high expression and so when $E_2$ is completely inhibited, the capacity for drug metabolism via $E_1$ is still high. In this case, the change in AUC values would be limited compared to the case of absence of correlation, in which most of the metabolism could be due to the activity of one enzyme alone. Instead, if $E_1$ has low expression, $E_2$ is likely to have low expression, thus the metabolism would still be low in case of complete inhibition of $E_2$, therefore the AUC values would still be high. In the case of absence of correlation, the complete inhibition of one enzyme could dramatically change the AUC values. For these reasons, the distribution of $R_{inh}$ is narrower in the presence of correlation. The correlation effect is higher for central $CL_{int}$ values and lower for extreme $CL_{int}$ values, probably for the same reasons as it is for $R_P$: for higher $CL_{int}$ values the metabolism becomes flow-limited and so less dependent on enzymatic expression, while for lower $CL_{int}$ values the clearance is so low that even inhibiting one enzyme does not change the overall absolute metabolic clearance.

### 6.3.2 Results for CYP3A4 in gut wall and in liver

For the model represented in figure 6.1 (b) and described by equation system 6.5, the $F_{oral}$ distribution was calculated for each combination of $CL_{int}$ and $I$, both in the case of presence and absence of correlation between CYP3A4 expression in the gut wall and liver. Then, for each of these AUC

distribution pairs, the index $R_P$ was calculated for the $5^{\text{th}}$, $50^{\text{th}}$ and $90^{\text{th}}$ percentiles. Results are shown in figure 6.4.

In presence of the correlation, the $5^{\text{th}}$ $F_{oral}$ percentile decreases of more than 20% and the $95^{\text{th}}$ percentile increases of more than 35%, while the median remains stable. This effect mainly occurs for high $CL_{int}$ values. In the case of correlation, for high liver CYP3A4 expression, it is likely to be high also the gut wall CYP3A4 expression. Instead, in absence of correlation, for high values of liver CYP3A4 expression, the gut wall expression could equally likely be any value, according to its distribution. In such cases, when the correlation is present, the metabolism is likely to be higher, thus the $F_{oral}$ lower. For the same reasons, in case of low CYP3A4 expression in one of the two sites, the metabolism is likely to be lower, and so the $F_{oral}$ to be higher, if the enzymes are correlated, with respect to the case of absence of correlation. The correlation effect could mainly be observed for high $CL_{int}$ values, as compared to mean or low $CL_{int}$ values, in which little or no effect is observed. Therefore, the effect of the correlation could only be seen for low values of oral bioavailability. In figure 6.5, for central or lower $CL_{int}$ values, the fraction of the drug metabolised in the liver is much higher with respect to the one metabolised in the gut wall. The two fractions start to be comparable only for high $CL_{int}$ levels, and so this is probably why the effect of the correlation is stronger in these cases.

For each value of $CL_{int}$, the ratio ($R_{inh}$) of $F_{oral}$ in case of complete inhibition of CYP3A4 in both gut wall and liver ($I = 0$) and in case of absence of inhibition ($I = 1$) was calculated. Results are reported in figure 6.6. It can be seen that the distribution of the ratio is slightly wider in the case of presence of correlation as compared to no correlation. It is possible to explain this behaviour by looking at the $R_{inh}$ analytical expression in equation 6.8. In fact, if the expressions of CYP3A4 in gut wall and liver are correlated, it is more likely that low values of clearance in the gut wall correspond to low values of clearance in the liver, leading to higher $F_{oral}$ in comparison to the case of absence of correlation. Similarly, in the presence of correlation, high values of clearance in the gut wall are likely to correspond to high values of clearance in the liver, leading to a lower $F_{oral}$.

In case of complete inhibition of CYP3A4 metabolism in only one of the

two sites, for example the gut wall (as is the case for grapefruit juice interaction studies [213]), the ratio of $F_{oral}$ in the inhibited state over $F_{oral}$ in absence of inhibition, is equal to $\left(\frac{Q_{ent}}{V_{ent}} + \frac{CL_{ent}}{Vent} I\right) / \left(\frac{Q_{ent}}{V_{ent}} + \frac{CL_{ent}}{Vent}\right)$. Thus, in these cases the ratio does not depend on the correlation between CYP3A4 in the gut wall and liver.

### 6.3.3 Results for CYP3A4 and OATP1B1 in liver

For the model represented in figure 6.1 (c) and described in equation system 6.9, the distribution of central compartment AUC was simulated for each combination of $CL_{int,CYP}$ and $CL_{int,OATP}$, in both the cases of presence and absence of correlation between the liver expression of the two enzymes metabolising the drug. Then, for each of these AUC distribution pairs, the index $R_P$ was calculated for $5^{\text{th}}$, $50^{\text{th}}$ and $95^{\text{th}}$ percentiles. Results are shown in figure 6.7.

Due to the presence of correlation between OATP1B1 and CYP3A4, the central compartment AUC $5^{\text{th}}$ percentile is reduced by more than 20%, the $95^{\text{th}}$ percentile increases by more than 25% and the median remains stable. In case of correlation, high OATP1B1 concentrations are likely to correspond to high CYP3A4 concentrations. Therefore, the drug is quickly transported into the intracellular liver compartment and is then rapidly metabolised. This leads to lower central compartment AUC values compared to the absence of correlation, where high OATP1B1 concentrations could equally likely correspond to high or low CYP3A4 concentrations. Similarly, in the case of correlation low OATP1B1 concentrations are likely to correspond to low CYP3A4 concentrations and therefore leads to higher central compartment AUC values.

### 6.3.4 GSA results for the repaglinide model

Variance based global sensitivity analysis was performed on central compartment AUC, as predicted using the reduced repaglinide PBPK model represented in figure 6.1 (d) and described by equation system 6.11. The main and total effects were calculated for CYP3A4, CYP2C8 and OATP1B1

concentrations. The correlation between the two cytochromes microsomal concentrations was modelled through a linear regression with CYP3A4 and CYP2C8 being the independent and dependent variable, respectively. Results are shown in figures 6.8 and 6.9.

In case of absence of correlation ($\rho = 0$), the parameter that mostly explains the central compartment AUC variance is the OATP1B1 concentration, followed by CYP2C8 and CYP3A4 concentrations, with total effects equal to 0.5456, 0.434 and 0.1617, respectively. Despite the lower mean microsomal concentration and lower $V_{max}$ in the case of repaglinide metabolism, variability in the abundance of CYP2C8 is more important than variability in CYP3A4 when it comes to explaining the variance in AUC. This could be explained by the higher variability of microsomal CYP2C8 and by the lower value of $K_M$ for repaglinide, as compared to that of CYP3A4.
Moving towards higher correlation coefficients, the importance of OATP1B1 and CYP2C8 concentrations is reduced and the importance of CYP3A4 concentrations is increased. In the case of physiological correlation ($\rho = 0.7436$), the total effects of OATP1B1, CYP2C8 and CYP3A4 concentrations are equal to 0.4499, 0.122 and 0.5174, respectively. The increased importance of CYP3A4 and reduced importance of CYP2C8 as correlation increases are probably mainly a consequence of CYP2C8 being dependent on CYP3A4. This because CYP2C8 concentrations are derived through a linear regression where CYP3A4 is the independent variable. Moreover, the variability of $\varepsilon$, the residual variability or error, decreases as $\rho$ increases. In fact, the variance of $\varepsilon$ is reduced from 1.8629 when $\rho = 0$, to 0.8329 when $\rho = 0.7436$. Instead, in equation system 6.11, the parameters related with OATP1B1 concentration distribution do not change through different correlation levels between the two CYPs. So, the reduction of OATP1B1 main and total effect when $\rho$ increases, is explained by the increase of the level of correlation between CYP3A4 and CYP2C8. Thus, the more CYP3A4 and CYP2C8 are correlated, the less OATP1B1 is important on central compartment AUC variance.

The results of the analysis performed considering CYP3A4 as a function of CYP2C8 are shown in figure 6.10. In summary, the observations are: 1)

when $\rho$ is equal to 0 the main and total effect of all the three parameters are the same to the ones shown in figure 6.8; 2) main and total effect of OATP1B1 do not change across the different correlation levels with respect to the ones shown in figure 6.8. The differences seen in the results of the analysis shown in figure 6.8 are that in this case CYP2C8 main and total effect increase and CYP3A4 main and total effect are reduced, as $\rho$ increases. This because, similarly to the results explained above, CYP3A4 concentrations depend on CYP2C8 concentrations.

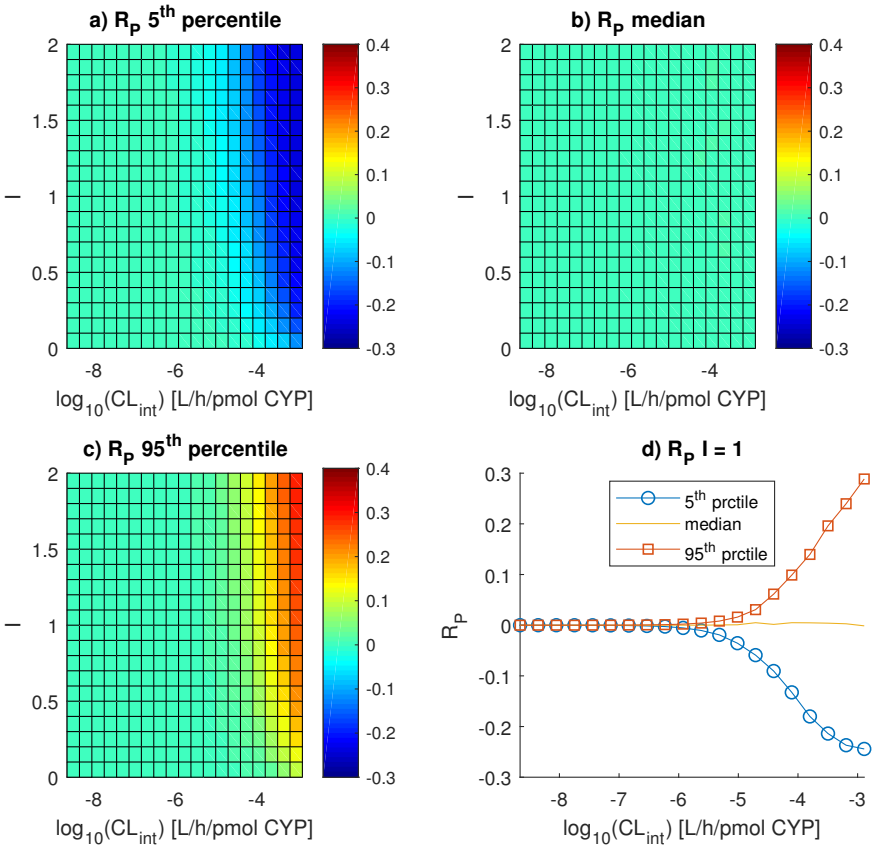



Figure 6.4: $R_P$ indices for the model in figure 6.1 (b), computed for different inhibition levels and $CL_{int}$ values: a) $R_P$ indices for the $5^{\text{th}}$ percentile; b) $R_P$ indices for the median; c) $R_P$ indices for the $95^{\text{th}}$ percentile; d) $R_P$ indices for $5^{\text{th}}$ percentile, median and $95^{\text{th}}$ percentile in absence of inhibition. For higher $CL_{int}$ values the AUC distribution is wider ($R_{P,5^{th}} < 0$ and $R_{P,95^{th}} > 0$) in case of correlation with respect to the absence of that.

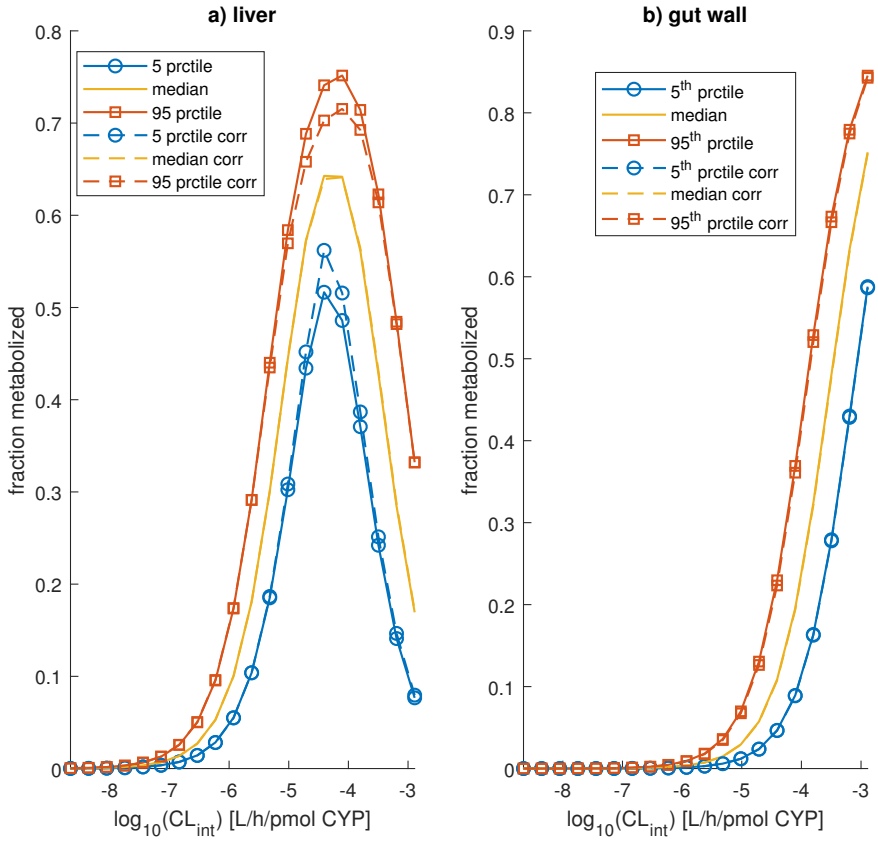



Figure 6.5: $5^{\text{th}}$ percentile, median and $95^{\text{th}}$ percentile of the fraction metabolized by the liver (panel a) and the gut wall (panel b) distributions for different values of $CL_{int}$, in case of absence of inhibition, predicted by using the model in figure 6.1 (b). It is possible to observe that the percentiles of the fraction metabolized by the gut wall distributions are equal in case of presence and absence of correlation. This result is trivial because the fraction metabolized by the gut wall depends only on the CYP3A4 expression in that site. Moreover, it is possible to observe that the two yellow median lines overlap in both panel a and b.

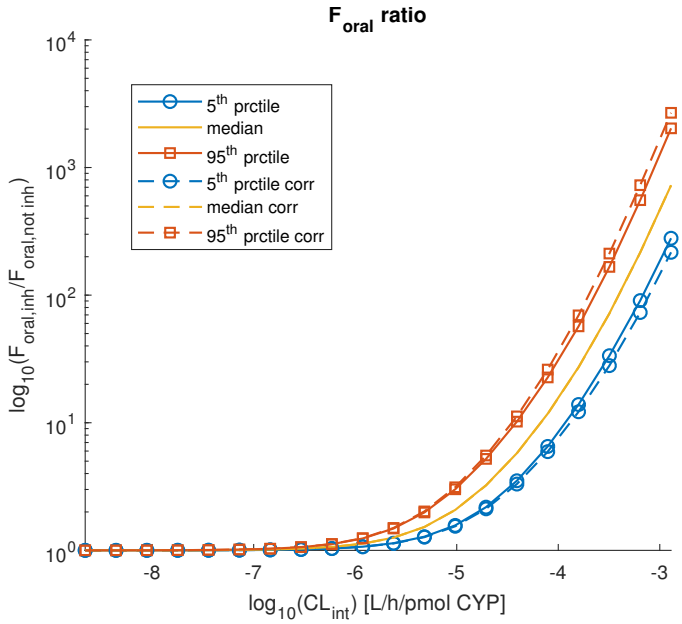



Figure 6.6: Ratio between the bioavailability in presence and absence of inhibition of both CYP3A4 gut wall and liver clearance ($I = 0$ and $I = 1$, respectively), predicted by using the model in 6.1 (b) and calculated for different values of $CL_{int}$. Continuous and dashed lines represent the various percentiles ($5^{\text{th}}$, $50^{\text{th}}$ and $95^{\text{th}}$) in absence and presence of correlation between CYP3A4 in gut wall and in liver.

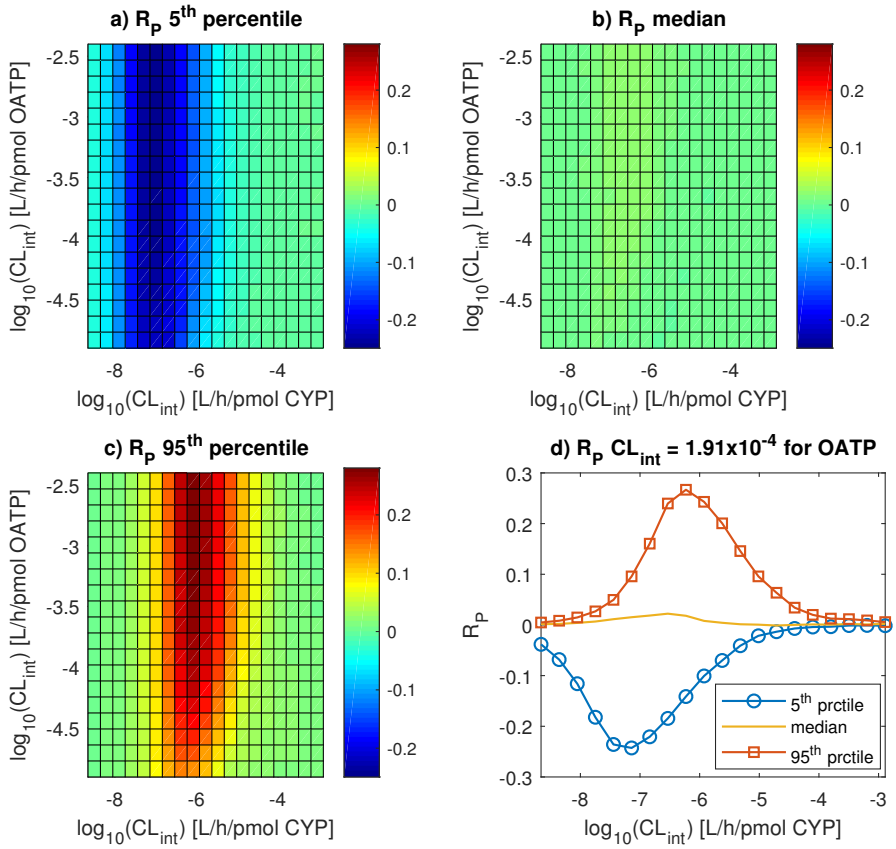



Figure 6.7: $R_P$ indices for the model in figure 6.1 (c), computed for different levels of CYP3A4 and OATP1B1 intrinsic clearances: a) $R_p$ indices for the $5^{\text{th}}$ percentile; b) $R_P$ indices for the median; c) $R_P$ indices for the $95^{\text{th}}$ percentile; d) $R_P$ indices in function of CYP3A4 intrinsic clearance for the $10^{\text{th}}$ entry (out of 20) of OATP1B1 intrinsic clearance vector. For central $CL_{int}$ values the AUC distribution is wider ($R_{P,5^{th}} < 0$ and $R_{P,95^{th}} > 0$) in case of correlation with respect to the absence of that.

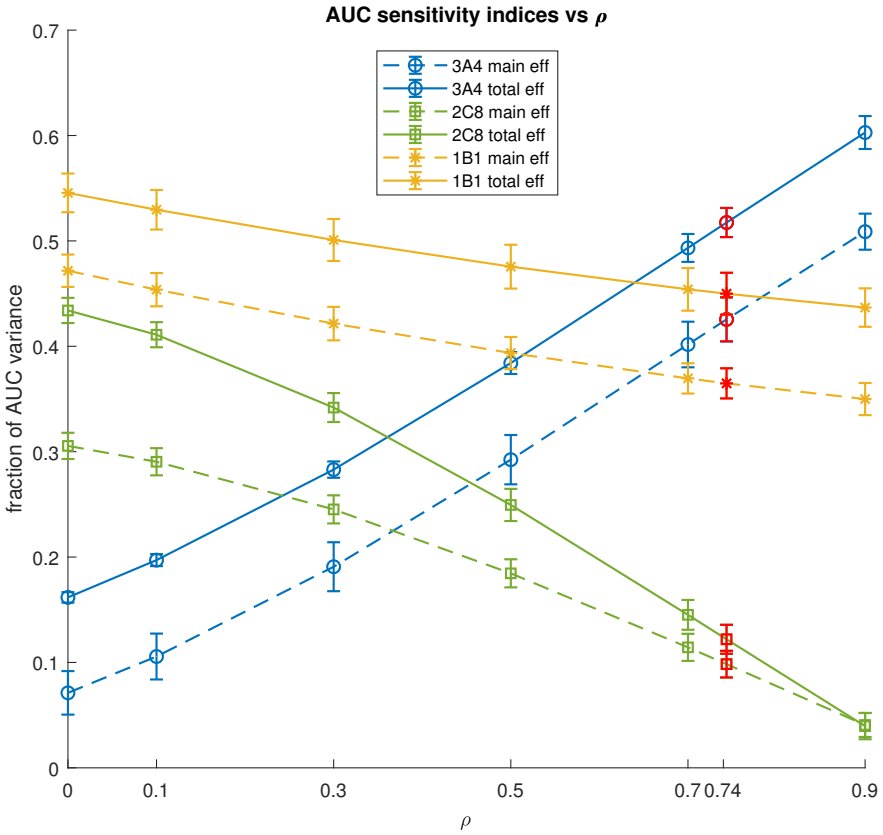



Figure 6.8: Main and total effect of CYP3A4, CYP2C8 and OATP1B1 concentrations, for different correlation levels between CYP3A4 and CYP2C8, for the model in figure 6.1 (d). Indices corresponding to the physiological $\rho$ are highlighted in red. To deal with the correlation between the two enzymatic concentrations, here CYP2C8 was expressed as a function of CYP3A4 by using a linear regression plus the addition of a noise. Error bars represent the standard deviation of the estimated sensitivity indices.

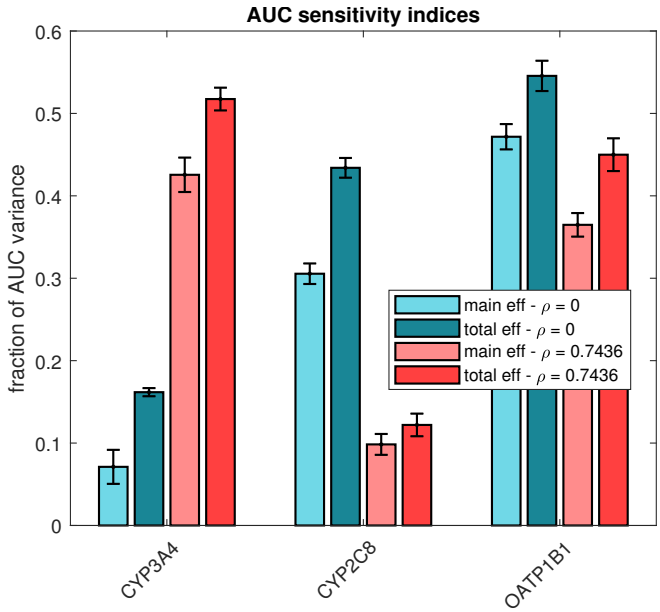



Figure 6.9: Main and total effect of CYP3A4, CYP2C8 and OATP1B1 concentrations, in case of absence and presence of a physiological correlation between the expression of CYP3A4 and CYP2C8, for the model in figure 6.1 (d). To deal with the correlation between the two enzymatic concentrations, here CYP2C8 was expressed as a function of CYP3A4 by using a linear regression plus the addition of a noise. Error bars represent the standard deviation of the estimated sensitivity indices.

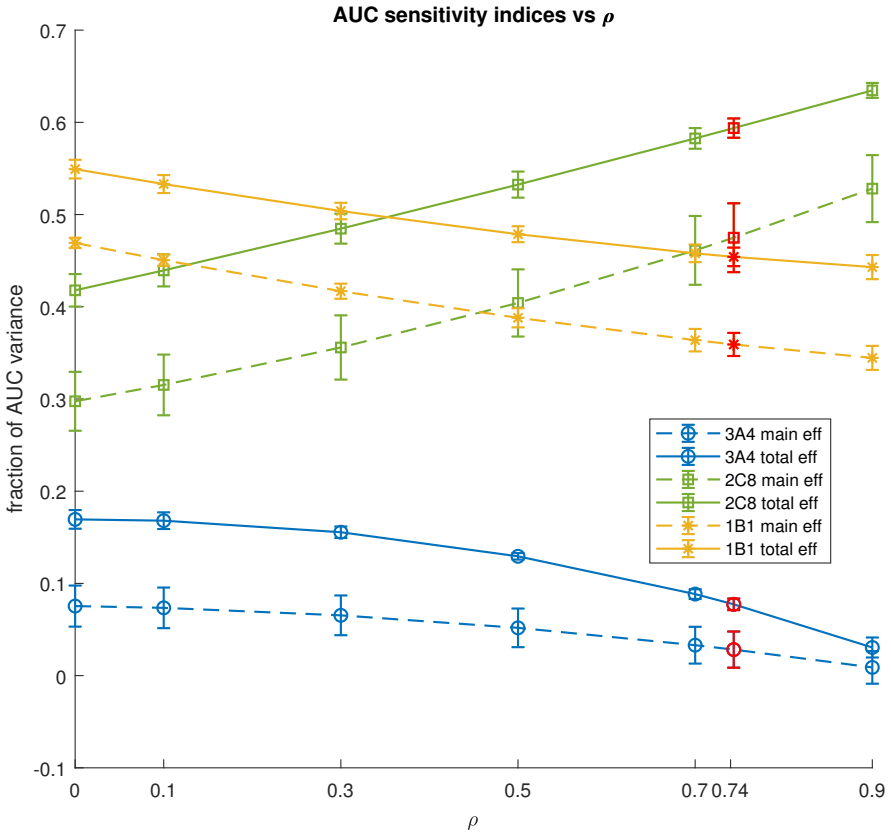



Figure 6.10: Main and total effect of CYP3A4, CYP2C8 and OATP1B1 concentrations, for different correlation levels between CYP3A4 and CYP2C8, for the model in figure 6.1 (d). Indices corresponding to the physiological $\rho$ are highlighted in red. To deal with the correlation between the two enzymatic concentrations, here CYP3A4 was expressed as a function of CYP2C8 by using a linear regression plus the addition of a noise. Error bars represent the standard deviation of the estimated sensitivity indices.

## 6.4 Discussion

The aims of this work were, first of all, to perform an uncertainty analysis to assess the effect of the correlation between enzymatic expressions at different sites in the body on the PK of different types of drugs and second, to understand how the correlation between enzymatic abundances impacts the results of global sensitivity analysis. In this context, three semi-physiological models were considered to describe the PK of generic drugs metabolised by two enzymes with correlated expressions, in order to analyse the following cases: 1) impact of the correlation between the expression of two enzymes present in the liver on the plasma AUC of drugs that are metabolized by both these enzymes; 2) impact of the correlation between gut wall and liver CYP3A4 expression levels on the bioavailability of CYP3A4 metabolised drugs; 3) impact of the correlation of OATP1B1 and CYP3A4 expressed in the liver on the plasma AUC of drugs that are substrates of both proteins.

Concerning the first case, the correlation acts to enlarge the plasma AUC distribution ($R_P < -0.05$ for the $5^{\text{th}}$ percentile, $R_P > 0.1$ for the $95^{\text{th}}$ percentile) and results in a more narrow distribution of the ratio between plasma AUC in presence and in absence of inhibition, especially for mean values of $CL_{int}$ (figures 6.2 and 6.3). These results are in accordance with the observations reported in [201], in fact in that study it was shown that a PBPK model ignoring correlation between two drug-metabolising enzymes may overestimate the effect of the DDI when one of the two enzymes is inhibited. Concerning the second case, the effect on the bioavailability seems to be relevant ($R_P < -0.2$ for the $5^{\text{th}}$ percentile, $R_P > 0.2$ for the $95^{\text{th}}$ percentile) only for high levels of intrinsic clearance. However, it must be considered that in this analysis we did not included the dissolution process and we considered high absorption. So, the results are limited to the above situation. Finally, concerning the third case, the correlation shows its effect on plasma AUC especially for mean levels of $CL_{int}$ ($R_P < -0.2$ for the $5^{\text{th}}$ percentile, $R_P > 0.2$ for the $95^{\text{th}}$ percentile). The correlation between OATP1B1 and CYP3A4 expressions may be considered a theoretical exercise alone given the fact that, to our knowledge, there are no data available

to support (or reject) this hypothesis.

Further, we have to consider the fact that in this analysis we used simple semi-physiological models, with a limited number of compartments and limited sources of variability (practically exclusively limited to the parameters that are considered correlated), and not complex whole body PBPK models (such as: *Simcyp Simulator*, *GastroPlus* and more). This was done to highlight the effect of the correlation, if present, and to allow an easier interpretation of the results. The problem of using a whole body PBPK is that there are too many variables to control, consequently, the parameters space would become too large to analyse and so, one is forced to reduce it considering, for example, a specific case drug (e.g., repaglinide [201]). With simple compartmental models the parameters space is reduced, and it becomes possible to investigate more generally in what parameter space interaction effects occur and their extent in function of set of parameters (e.g., the intrinsic clearance). This would not directly correspond to a specific drug or class of drugs. However, our study could be a useful guide to indicate in what situations the correlation may have a potential effect.

The results of this analysis show the necessity of considering the correlation between enzymatic expressions in physiological model when confidence in biological evidence for such correlations is high. Otherwise, there is a risk of underestimating the population variability or overestimate the effect of DDI (though this is shown for direct correlations of enzymes and the reverse could be true if the enzyme abundances were inversely related). To remain true to the knowledge of the system, PBPK platforms should not ignore known correlations of any of the model parameters. However, one has to consider that in a whole body PBPK model (unlike the minimal model used in our investigation), it is likely that some of our findings might be mitigated due lack of dominance of parameters that we intentionally selected in this study. Our study considered some conditions (e.g., equal intrinsic clearance between two correlated enzymes) which were in favour of propagating the inter-correlation effect to make the point and these may not be the case at the presence of multiple parameters variabilities and correlations.

This analysis could be also useful for informing experimental design, for

instance to assess the impact of hypothetical correlations prior to the generation of the data. For example, in order to asses if a correlation exists between CYP3A4 in gut wall and in liver, an appropriate number of paired samples in the same subjects should be collected. This could be challenging in terms of sample collection and expensive. Knowing that this correlation, even if present, has little or no impact except for high intrinsic clearance drugs, could be useful information when choosing what experiments to prioritise and their design.

Given the recent interest of the EMA in sensitivity analysis [50], we wanted to investigate the impact of the correlation between enzymatic abundances when performing GSA. A variance based GSA was performed on a reduced PBPK model for repaglinide, a substrate drug for CYP3A4 and CYP2C8, these two enzymes have been shown to be correlated in the liver [202]. One initial problem was to find an appropriate methodology to conduct a GSA in the case of correlated parameters. Naïvely, an easy way could be to extract correlated samples from the space of parameters (e.g., enzymatic expression), compute the model outputs (e.g., plasma AUC) for each one of these parameters sets and then perform a GSA using the linear correlation coefficient between the outputs and the input parameters. In case of correlated parameters, the results of this analysis could be misleading. For example, let us consider the theoretical case of a model describing the PK of a drug metabolised by CYP3A4 in liver. In a population simulation, plasma AUC would probably be correlated with CYP3A4 liver expression. Considering that the expressions of liver CYP3A4 and CYP2C8 are correlated, the drug plasma AUC would be correlated with CYP2C8 expression. This because the presence of a correlation between the two proteins and not because CYP2C8 plays a role in the metabolism of the drug. So, in these situations, performing a GSA using the correlation coefficient could give misleading results.

In the literature, different methodologies exist to perform GSA in case of correlated parameters, see for example [214, 215, 216, 217], but as reported in a recent review: *'this issue remains misunderstood'* [78]. Therefore, we choose to express the correlation between two parameters using an explicit relationship (in our case linear regression) with a residual error term and

then perform a variance based GSA on uncorrelated parameters [21].

Implementing the correlation between parameters can influence the results of the GSA. In fact, in figures 6.8 and 6.9 it can be seen how the OATP1B1 main and total effect change, even if not dramatically, throughout different correlation levels between CYP3A4 and CYP2C8. Nevertheless, with the chosen method for performing the GSA, one should be careful when interpreting the sensitivity indices. In figure 6.8 it is possible to observe that the importance of CYP3A4 and CYP2C8 increases and decreases, respectively, as the correlation between the two enzymatic expression increases. This effect is probably mainly due to CYP2C8 being dependent on CYP3A4, because the latter was arbitrarily chosen to be the independent variable in the linear regression and not because the CYP3A4 catalysed reaction acquires more importance. In this analysis we considered the abundances of CYP3A4 and CYP2C8 independent from the one of OATP1B1. If these abundances were correlated, the results would probably be different. It is important to note that the approach used in this paper to treat correlation in GSA is difficult to use when more than two parameters are correlated with each other.

To our knowledge, this is the first systematic analysis that investigates the impact of correlation between enzymatic abundances on drug clearance and metabolic DDIs with GSA in mind. Implementing the correlation between the enzymatic expression in population physiological models has the potential to impact the results of both predictions and GSA. Ignoring these correlations could lead to the generation of implausible parameter combinations and to incorrect estimation of parameters related to the PK (e.g., clearance). Thus, it is appropriate to assess experimentally if these correlations exist, their extent and how they differ with genetics, disease or physiological conditions. The approach presented in this study can be applied to highlight what correlations are of potential interest and therefore could be useful for informing experimental design. Our work informs the debate that is needed to take place in considering recent data generated by the proteomic analysis and regulatory interest in the use of sensitivity analysis in PBPK. We fully encourage continuation of investigation on inter-correlation not only for protein expressions relevant to drug disposi-

tion, but also for the various attributes of the gastro-intestinal tract where many potential/likely correlations related to physiology and biology are not fully considered in PBPK models yet. Inter-correlations may not be relevant when considering the model for an average individual, however, they are pertinent to all aspects of population-based projections from mechanistic models within pharmacometrics and systems pharmacology modelling.

Chapter 7

# Overall conclusions

The parameters of mechanistic models are inherently uncertain. Many drug-related parameters are measured, thus they are affected by measurement errors. Other parameters can be calibrated, so they are subject to estimation errors. Others can be predicted with other models and thus, they are affected by prediction errors. Moreover, if we consider a population, the parameters are variable as well as uncertain. The immediate consequence of the parameters being uncertain or variable is that the model outputs would be uncertain or variable too. The *quantification* of the model output variation is referred as uncertainty analysis. Instead, the act of apportioning the output variation to the sources of uncertainty or variability in the model inputs is referred as sensitivity analysis. Dependently on the choice of the input parameters variation, uncertainty and sensitivity analysis answer at different questions and thus, have different aims.

In this thesis, I showed some of the possible uses of uncertainty analysis and GSA for mechanistic models used in the field of pharmacology and, in particular, for PBPK models.

GSA can be used to understand what parameters mainly drive the variability of certain metrics of interest in a given population. This was the case of the mechanistic model describing gemcitabine PK, in chapter 3.

In that case, GSA showed that individual genetic factors affecting gemcitabine metabolism are mainly responsible for different accumulation of the active metabolites in the target tissue and so, for the different treatment responses. By doing this analysis, it was possible to highlight that a better characterization of the population distribution of the target enzymes is needed to improve the predicted variability.

GSA can be performed considering the 'between-drugs' parameters variability. In this case, we called this analysis '*inter-compounds* GSA'. This type of GSA helps in understanding the behaviour of the model as a function of parameters variation and in identifying limiting steps and bottlenecks. Inter-compound GSA was found to be particularly useful during the process of model development, as shown for the PBPK absorption model for inhaled compounds in chapter 5. In fact, this analysis helped in understanding if the model behaved as expected and, in case of discrepancies between the actual and the expected model behaviours, it gave useful information for identifying the reasons. Moreover, inter-compounds GSA gives an *a priori* information on what drug-specific parameters should be know with a lower degree of uncertainty in order to allow more confident model predictions.
This analysis was found to be useful also for well constructed, understood and characterized models, such the ones describing intestinal absorption. As shown in chapter 4, we identified the radius of the particle size of the formulation as one of the most sensitive parameters, especially for compounds administered at low dose levels. However, this parameter has also a high level of missingness in the OrBiTo database of compounds. Thus, we concluded that a performance evaluation fixing the value of one of the most sensitive parameters to an assumed or mean value, in certain situations, could result in an incomplete interpretation.

GSA can be performed also considering the parameters uncertainty related to a specific compound. This analysis was called '*intra-compound* GSA'. As shown in chapter 5, with this analysis is possible to understand how much the model output variation is apportioned to the model input uncertainty. If the model output uncertainty is considered too wide, and thus the model predictions unreliable, GSA helps in selecting what parameters should be know with lower degree of uncertainty in order to give more

precise predictions.

Uncertainty and sensitivity analysis helped also in informing experimental design. In fact, as explained in chapter 6, these analyses allowed to asses the impact of plausible correlations on some metrics of interest prior the data generation. For example, we identified that the correlation between CYP3A4 in the gut-wall and in the liver has little or no impact except for high intrinsic clearance drugs. Given that the process of experimentally assess the extent of this correlation can be expensive and challenging, knowing the results of this analysis could be useful when choosing what experiments to prioritize.

In our experience, uncertainty analysis and GSA showed their utility both in model development and use. In particular, they helped in understanding the model behaviour, identifying sensitive assumptions, uncovering technical errors and establishing priorities for research.

However, GSA showed some criticism. We observed that in certain situations identifying all the uncertain factors and determining their distribution was not an easy task. This is a well known concern, in fact it was already highlighted in [218, 219]. The choice of a given factor distribution introduces a subjective decision in the analysis [219]. This has the potential to impact the shape of the output distribution and the sensitivity indices. One alternative could be just to fix an uncertain factor to a given value, as often is done in PBPK models. However, even the choice of that particular value is subjective. For example, it could be the result of a given experiment rather than another one, or the mean (or median) of the results of multiple experiments. Despite the fact of being subjective too, fixing an uncertain factor before knowing if it has an impact on the model outputs, as extensively discussed in this thesis, is not a good practice. In fact, obtaining precise output through the arbitrary restriction of the input space of uncertainty and variability is a way of 'GIGOing' (from *garbage in, garbage out*, GIGO) [60, 218].

As highlighted by regulatory agencies and practitioners from multiple disciplines, sensitivity analysis is crucial for the quality assessment of model based inference [50, 51, 218]. However, guidances regarding PBPK models should further stress the use of *appropriate* sensitivity methods. To avoid a

perfunctory sensitivity analysis [66], they should dissuade the use of those methods that badly explore the parameter space and that do not allow the detection of interaction effects.

In conclusion, the aim of this thesis has been to show the utility of uncertainty analysis and GSA in PBPK modelling and simulation, with the view of seeing these techniques routinely applied in model development and use.

Appendix A

# Model equations, parameters and simulation results for the mechanistic model presented in chapter 3

## A.1 PBPK model equations

Given the hydrophilic nature of dFdC it was decided to model each organ with a permeability limited model. $R_{hCNT1,t}$, $R_{hENT1,t}$ and $R_{CDA,t}$ are hCNT1, hENT1 transport reaction and CDA metabolism in the tissue $t$, respectively [125].

$$R_{hCNT1,t} = V_{ext,t}\, e_{hCNT1,t}\, k_{hCNT1}\, \frac{V_{int,t}}{V_{ext,t}}\, Cu_{ext,t}$$

$$R_{hENT1,t} = e_{hENT1,t} \left( V_{ext,t}\, k_{hENT1,in}\, \frac{V_{int,t}}{V_{ext,t}}\, Cu_{ext,t} - V_{int,t}\, k_{hENT1,out}\, Cu_{int,t} \right)$$

$$R_{CDA,t} = V_{int,t}\, e_{CDA,t}\, CL_{CDA}\, Cu_{int,t}$$

$e_{x,t}$ is the enzyme or transporter relative expression in the tissue $t$, $k_x$ is the linear rate constant, assumed to be equal in all the organs and $Cu_{ext,t}$ and $Cu_{int,t}$ are the unbound extracellular and intracellular dFdC concentrations, respectively. Here are reported the differential equations relative to adipose tissue, bone, brain, stomach, small intestine, large intestine, heart, kidneys, muscle, skin and spleen.

$$V_{ext,t}\, \frac{dCu_{ext,t}}{dt} = Q_t \left( C_{art} - \frac{Cu_{ext,t}}{P_{t:p}/B:P} \right) - R_{hCNT1,t} - R_{hENT1,t}$$
$$V_{int,t}\, \frac{dCu_{int,t}}{dt} = R_{hCNT1,t} + R_{hENT1,t} - R_{CDA,t}$$

Liver differential equations (subscripts $s\ int$ and $l\ int$ represent small and large intestine, respectively).

$$\begin{aligned} V_{ext,liver}\, \frac{dCu_{ext,liver}}{dt} &= Q_{liver} \left( C_{art} - \frac{Cu_{ext,liver}}{P_{liver:p}/B:P} \right) \\ &+ Q_{stomach}\, \frac{Cu_{ext,stomach}}{P_{stomach:p}/B:P} + Q_{s\ int}\, \frac{Cu_{ext,s\ int}}{P_{s\ int:p}/B:P} \\ &+ Q_{l\ int}\, \frac{Cu_{ext,l\ int}}{P_{l\ int:p}/B:P} + Q_{panc}\, \frac{Cu_{ext,panc}}{P_{panc:p}/B:P} \\ &+ Q_{spleen}\, \frac{Cu_{ext,spleen}}{P_{spleen:p}/B:P} - R_{hCNT1,liver} - R_{hENT1,liver} \end{aligned}$$
$$V_{int,liver}\, \frac{dCu_{int,liver}}{dt} = R_{hCNT1,liver} + R_{hENT1,liver} - R_{CDA,liver}$$

Lung differential equations.

$$\begin{aligned} V_{ext,lungs}\,\frac{dCu_{ext,lungs}}{dt} &= Q_{TOT}\,\left(C_{ven} - \frac{Cu_{ext,lungs}}{P_{lungs:p}/B:P}\right) \\ &\quad - R_{hCNT1,lungs} - R_{hENT1,lungs} \\ V_{int,lungs}\,\frac{dCu_{int,lungs}}{dt} &= R_{hCNT1,lungs} + R_{hENT1,lungs} - R_{CDA,lungs} \end{aligned}$$

Venous blood differential equation ($\mathcal{T}$ contains all the organs and tissues except stomach, small intestine, large intestine, pancreas, spleen and arterial and venous blood).

$$V_{ven}\frac{dC_{ven}}{dt} = \sum_{t\in\mathcal{T}} Q_t\left(\frac{Cu_{ext,t}}{P_{t:p}/B:P}\right) - Q_{TOT}\,C_{ven}$$

Arterial blood differential equations.

$$V_{art}\frac{dC_{art}}{dt} = Q_{TOT}\,\left(\frac{Cu_{ext,lungs}}{P_{lungs:p}/B:P} - C_{art}\right)$$

Pancreas differential equations:

$$\begin{aligned} V_{ext,panc}\,\frac{dCu_{ext,panc}}{dt} &= Q_{panc}\,\left(C_{art} - \frac{Cu_{ext,panc}}{P_{panc:p}/B:P}\right) - R_{hCNT1,panc} \\ &\quad - R_{hENT1,panc} - \frac{\widetilde{V}_{max,hENT1}\,Cu_{ext,panc}\,V_{ext,panc}}{\widetilde{K}_{M,hENT1} + Cu_{ext,panc}\,V_{ext,panc}} \\ &\quad - \frac{\widetilde{V}_{max,CDA,ext}\,Cu_{ext,panc}\,V_{ext,panc}}{\widetilde{K}_{M,CDA,ext} + Cu_{ext,panc}\,V_{ext,panc}} \\ V_{int,panc}\frac{dCu_{int,panc}}{dt} &= R_{hCNT1,panc} + R_{hENT1,panc} - R_{CDA,panc} \end{aligned}$$

Intracellular tumour equations.

$$\begin{aligned}
\frac{d\,dFdC_{int}}{dt} &= \frac{\widetilde{V}_{max,hENT1}\,Cu_{ext,panc}\,V_{ext,panc}}{\widetilde{K}_{M,hENT1} + Cu_{ext,panc}\;V_{ext,panc}} - \frac{\widetilde{V}_{max,CDA,int}\,dFdC_{int}}{\widetilde{K}_{M,CDA,int} + dFdC_{int}} \\
&\quad - \frac{\widetilde{V}_{max,dCK}\;dFdC_{int}}{\widetilde{K}_{M,dCK} + dFdC_{int}} + \widetilde{K}_{mpc}\;dFdCMP_{int} \\
\frac{d\;dFdCMP_{int}}{dt} &= \frac{\widetilde{V}_{max,dCK}\,dFdC_{int}}{\widetilde{K}_{M,dCK} + dFdC_{int}} - \left(\widetilde{K}_{MPC} + \widetilde{K}_{NMPK}\right)\;dFdCMP_{int} \\
&\quad - \frac{\widetilde{K}_{CMPD}\;dFdCMP_{int}}{1 + \widetilde{INH}\,dFdCTP_{int}} + \widetilde{K}_{DPMP}\;dFdCDP_{int} \\
\frac{d\;dFdCDP_{int}}{dt} &= \widetilde{K}_{NMPK}\,dFdCMP_{int} - \left(\widetilde{K}_{NDPK} + \widetilde{K}_{DPMP}\right)\;dFdCDP_{int} \\
&\quad + \widetilde{K}_{TPDP}\;dFdCTP_{int} \\
\frac{d\,dFdCTP_{int}}{dt} &= \widetilde{K}_{NDPK}\;dFdCDP_{int} - \left(\widetilde{K}_{TPDP} + \widetilde{K}_{DNA}\right)\;dFdCTP_{int}
\end{aligned}$$

All the parameters of the *in vivo* metabolic network are derived from the *in vitro* ones as explained in section 3.2.3.

## A.2 dFdC pharmacokinetic model

The *Zhang model*[136] used to simulate dFdC pharmacokinetics for a typical male is here reported. This model was used for generating data used to identify the PBPK model parameters.

$$\begin{aligned}
\frac{d\ dFdC_c}{dt} &= -\left(k_{12} + k_{10}\right)\ dFdC_c + k_{21}\, dFdC_p \\
\frac{d\ dFdC_p}{dt} &= k_{12}\, dFdC_c - k_{21}\, dFdC_p \\
\frac{d\ dFdCTP_{WBC}}{dt} &= \frac{V_{max}[dFdC_c]}{K_M + [dFdC_c]} - k_{30}\, dFdCTP_{WBC}
\end{aligned}$$

$dFdC_c$, $dFdC_p$ and $dFdCTP_{WBC}$ are the dFdC amount in central compartment, peripheral compartment and dFdCTP white blood cells (WBC) amount in $mmol/m^2$. $[dFdC_c]$ represents the drug concentration in central compartment. All the parameters values used for the simulations are reported in table A.1.

Table A.1: Zhang model parameters

| **Parameters** | value | units |
|---|---|---|
| $CL$ | 92.2 | $L/h/m^2$ |
| $Q$ | 125 | $L/h/m^2$ |
| $V_1$ | 17.5 | $L/m^2$ |
| $V_2$ | 47.4 | $L/m^2$ |
| $V_{max}$ | 2 | $mmol/h/m^2$ |
| $K_M$ | $7.5 \cdot 10^{-3}$ | $mmol/L$ |
| $K_{30}$ | 0.058 | $1/h$ |

## A.3 Multiple dosages

Here, 3.34 $mmol/m^2$ of dFdC were infused weekly for 20 weeks; results are shown in figure A.1.

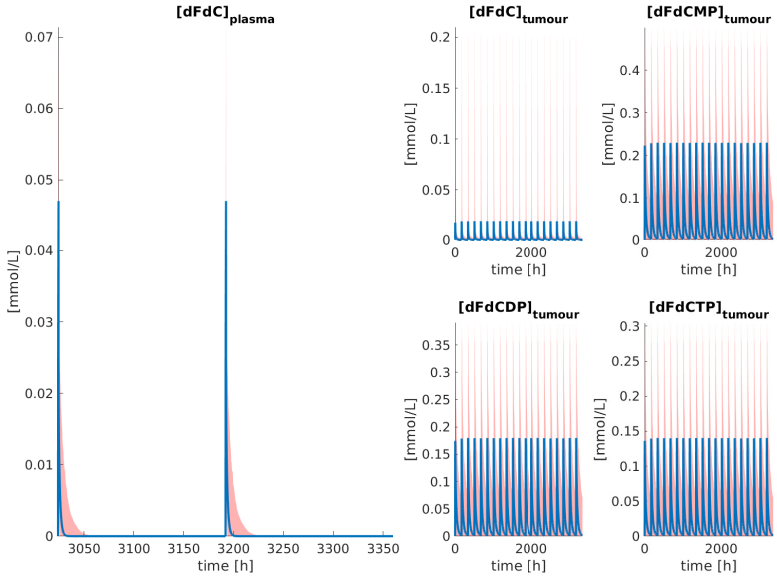



Figure A.1: dFdC and its metabolites PK after multiple administrations.

## A.4 System-specific parameters for PBPK building

In table A.2 the relative expressions of hCNT1, hENT1 and CDA for all the organs considered in the PBPK model are reported. These values were taken from the *Open Systems Pharmacology Suite*, version 7.1. In table A.3 the parameters relative to the organ compositions are reported, while in table A.4 the organs weight, blood flow and blood content are reported.

Table A.2: Relative expression of hCNT1, hENT1 and CDA, in %

| **Organs** | hCNT1[a] | hENT1[a] | CDA[b] |
|---|---|---|---|
| Adipose | 0 | 2.53[c] | 0 |
| Bone | 0.02 | 11.59 | 100 |
| Brain | 0.11 | 3.44 | 1.45 |
| Heart | 0.75 | 49.56 | 2.50 |
| Muscle | 0.62 | 100 | 1.02 |
| Skin | 0 | 2.53[c] | 6.4 |
| Spleen | 0.02 | 2.53 | 16.47 |
| Kidney | 100 | 4.69 | 4.52 |
| Gonads | 0.45 | 9.54 | 1.29 |
| Lung | 0.03 | 5.9 | 1.1 |
| Stomach | 0.52 | 12.36 | 0.9 |
| Small intestine | 14.01 | 6.65 | 3.66 |
| Large intestine | 0.03 | 8.44 | 16.27 |
| Liver | 44.73 | 13.52 | 6.4 |
| Pancreas | 0.02 | 5.88 | 1.3 |

[a] RT-PCR values.
[b] Array values.
[c] In the *Open Systems Pharmacology* suite this value was not provided, so it was fixed to its lowest value.

Table A.3: Organs composition

| **Organs** | neutral lipids fraction [41] | phospholipids fraction [41] | extracellular water fraction [40] | intracellular water fraction [40] | lipoprotein tissue to plasma ratio [41] | albumine tissue to plasma ratio [41] | acidic phospholipids amount [40] | organ density[d] |
|---|---|---|---|---|---|---|---|---|
| Adipose | 0.0016 | 0.853 | 0.135 | 0.017 | 0.068 | 0.049 | 0.4 | 0.916 |
| Bone | 0.0174 | 0.0016 | 0.1 | 0.346 | 0.05 | 0.1 | 0.67 | 1.4303 |
| Brain | 0.0391 | 0.0015 | 0.162 | 0.62 | 0.041 | 0.048 | 0.4 | 1.0365 |
| Heart | 0.0135 | 0.0106 | 0.32 | 0.456 | 0.16 | 0.157 | 2.25 | 1.03 |
| Muscle | 0.01 | 0.0072 | 0.118 | 0.63 | 0.059 | 0.064 | 1.53 | 1.041 |
| Skin | 0.0603 | 0.0044 | 0.382 | 0.291 | 0.096 | 0.277 | 1.32 | 1.1754 |
| Spleen | 0.0071 | 0.0107 | 0.207 | 0.579 | 0.207 | 0.097 | 3.18 | 1.054 |
| Kidney | 0.0121 | 0.024 | 0.273 | 0.483 | 0.137 | 0.13 | 5.03 | 1.05 |
| Gonads[a] | 0.0048 | 0.0100 | 0.06 | 0.88 | 0.04 | 0.05 | 2.45 | 1[e] |
| Lung | 0.0215 | 0.0123 | 0.336 | 0.446 | 0.168 | 0.212 | 3.91 | 1.0515 |
| Stomach[b] | 0.0375 | 0.0124 | 0.282 | 0.475 | 0.141 | 0.158 | 2.41 | 1.046 |
| Small intestine[b] | 0.0375 | 0.0124 | 0.282 | 0.475 | 0.141 | 0.158 | 2.41 | 1.046 |
| Large intestine[b] | 0.0375 | 0.0124 | 0.282 | 0.475 | 0.141 | 0.158 | 2.41 | 1.046 |
| Liver | 0.0135 | 0.0238 | 0.161 | 0.573 | 0.161 | 0.086 | 4.56 | 1.08[f] |
| Pancreas | 0.0403 | 0.009 | 0.12 | 0.664 | 0.06 | 0.06 | 1.67 | 1.045 |
| Plasma | 0.0023 | 0.0013 | - | - | - | - | 0.5[c] | 1[e] |

[a] Values taken from *Open Systems Pharmacology suite* version 7.1.
[b] Values for stomach, small and large intestine were supposed equal.
[c] Refers to blood cells.
[d] Calculated using specific gravity values from [171], considering that water density is 1 $kg/L$.
[e] Gonads and blood density were fixed to 1.
[f] Value taken from [220].

Table A.4: Organs weight, blood flows and blood content

| **Organs** | weight fraction[b] | | blood flow fraction[a] | | blood fraction[c] | |
|---|---|---|---|---|---|---|
| | male | female | male | female | male | female |
| Adipose | 0.2040 | 0.3220 | 0.0530 | 0.0900 | 0.05 | 0.0850 |
| Bone | 0.1620 | 0.1520 | 0.0530 | 0.0500 | 0.07 | 0.07 |
| Brain | 0.0210 | 0.0230 | 0.1280 | 0.130 | 0.012 | 0.012 |
| Heart | 0.0057 | 0.0055 | 0.0430 | 0.05 | 0.01 | 0.01 |
| Muscle | 0.4430 | 0.3380 | 0.1810 | 0.12 | 0.14 | 0.105 |
| Skin | 0.0520 | 0.0450 | 0.0530 | 0.05 | 0.03 | 0.03 |
| Spleen | 0.0033 | 0.0037 | 0.0320 | 0.03 | 0.014 | 0.0104 |
| Kidney | 0.0060 | 0.0067 | 0.2170 | 0.2 | 0.02 | 0.02 |
| Gonads | 0.0006 | 0.0002 | 0.0005 | 0.0002 | 0.0004 | 0.0002 |
| Lung | 0.0180 | 0.0170 | 1 | 1 | 0.1050 | 0.1050 |
| Stomach | 0.0023 | 0.0027 | 0.0110 | 0.01 | 0.01 | 0.01 |
| Small intestine | 0.0100 | 0.0120 | 0.1060 | 0.12 | 0.038 | 0.038 |
| Large intestine | 0.0056 | 0.0069 | 0.0430 | 0.05 | 0.022 | 0.022 |
| Liver | 0.0320 | 0.0320 | 0.0690 | 0.07 | 0.1 | 0.1 |
| Pancreas | 0.0026 | 0.0028 | 0.0110 | 0.01 | 0.006 | 0.006 |
| Blood | 0.0767[d] | 0.0683[d] | - | - | (0,06,0.18)[e] | (0.06,0.18)[e] |

[a] Organ weight fraction (including blood content) on total body weight [137].
[b] Fraction of cardiac output directed to each organ [137].
[c] Fraction of blood weight (relative to total blood weight) [170].
[d] Blood fraction on total body weight [170].
[e] (arterial fraction, venous fraction) [170].

Appendix B

# GSA results for the physiological intestinal absorption models presented in chapter 4

## B.1 GSA sensitivity indices

Below are presented the results of the analysis on the CAT based model for neutral, acidic and basic compounds, for basic compounds in presence of precipitation and on the mixing tank (MT) based model for neutral compounds, in that order. The reported sensitivity indices are the mean of those calculated on bootstrap samples. Each figure is composed of four panels containing one heatmap each: A, B, C and D. These heatmaps represent the results for the Biopharmaceutical Classification System (BCS) class I, II, III and IV, respectively. Each panel contains a heatmap that has the input parameters on the vertical axis and the different dose levels on the horizontal axis. Each heatmap cell contains the value of the main

and total effects relative to a parameter and dose level.
The parameters are:

- *GET*: gastric emptying time.
- *SITT*: small intestine transit time.
- *CYP3A4_liv*: concentration of CYP3A4 enzyme per mg of liver microsomal protein.
- *CYP3A4_ent*: amount of CYP3A4 in gut wall.
- *CYP3A4_CL*: drug intrinsic clearance.
- *MPPGL*: microsomal protein per g of liver.
- *rho*: density of the formulation.
- *r*: formulation radius of the particle.
- *mw*: drug molecular weight.
- *Peff*: drug effective permeability through the gut wall.
- *D0*: dose number.
- *pKa*: acid dissociation constant.
- *Rss*: supersaturation ratio.
- *kp*: precipitation time constant.

### B.1.1 CAT based model - neutral compounds

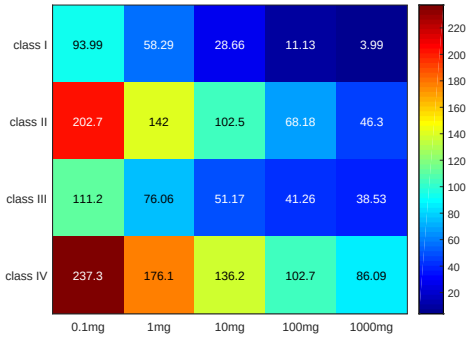



Figure B.1: $f_a$ CV for CAT based model - neutral compounds

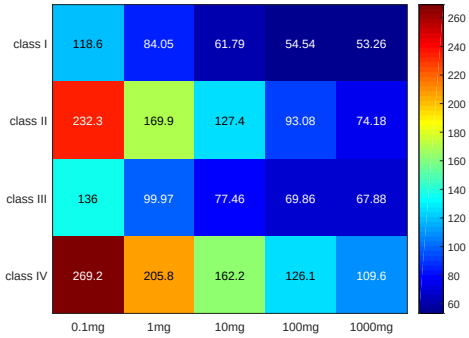



Figure B.2: $F_{oral}$ CV for CAT based model - neutral compounds

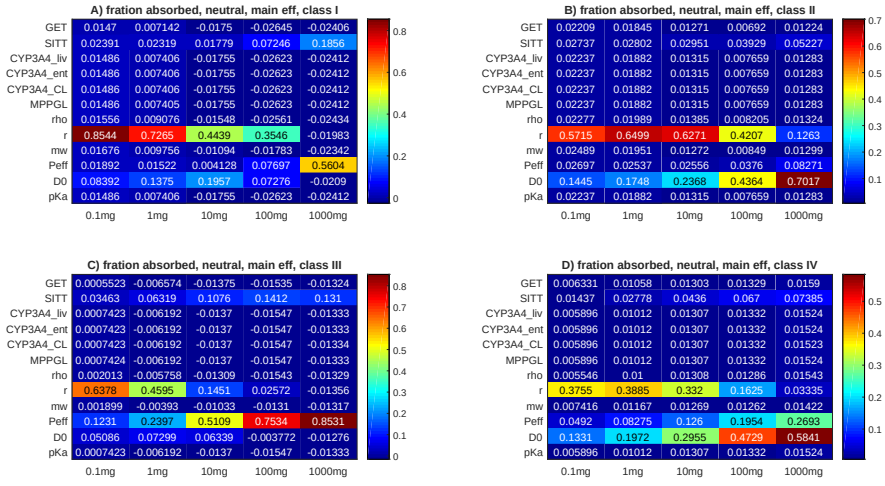



Figure B.3: $f_a$ main effect for CAT based model - neutral compounds

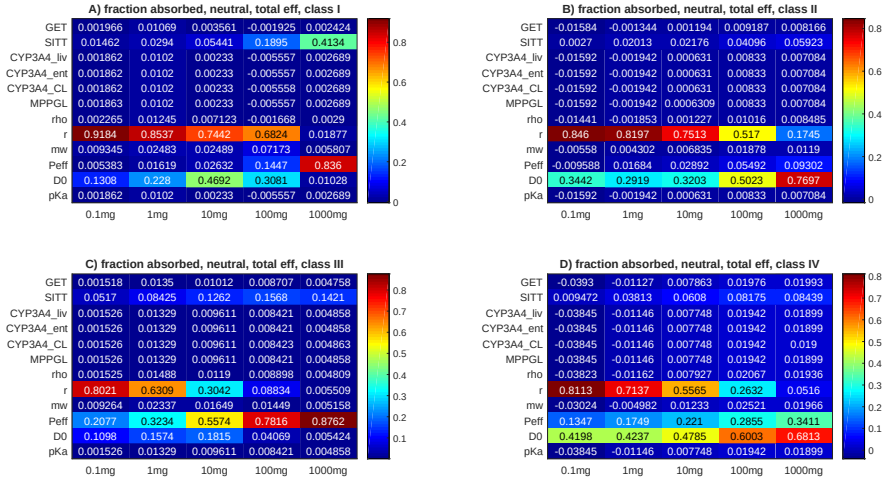



Figure B.4: $f_a$ total effect for CAT based model - neutral compounds

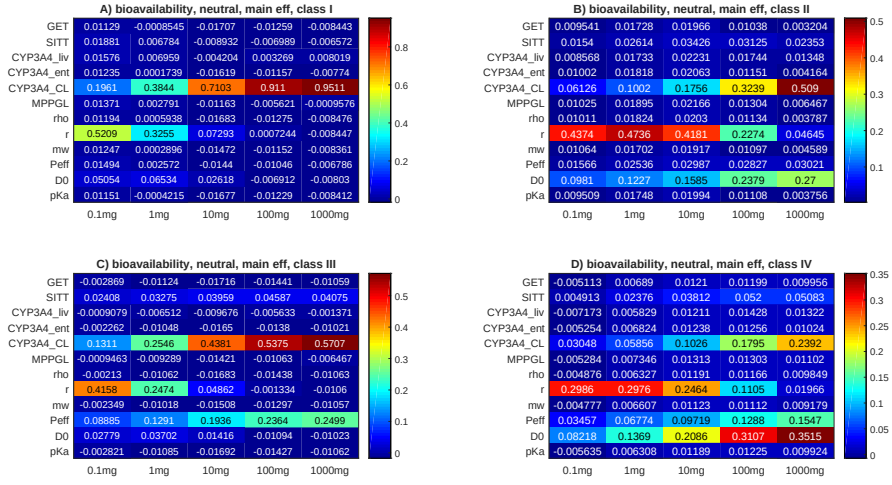



Figure B.5: $F_{oral}$ main effect for CAT based model - neutral compounds

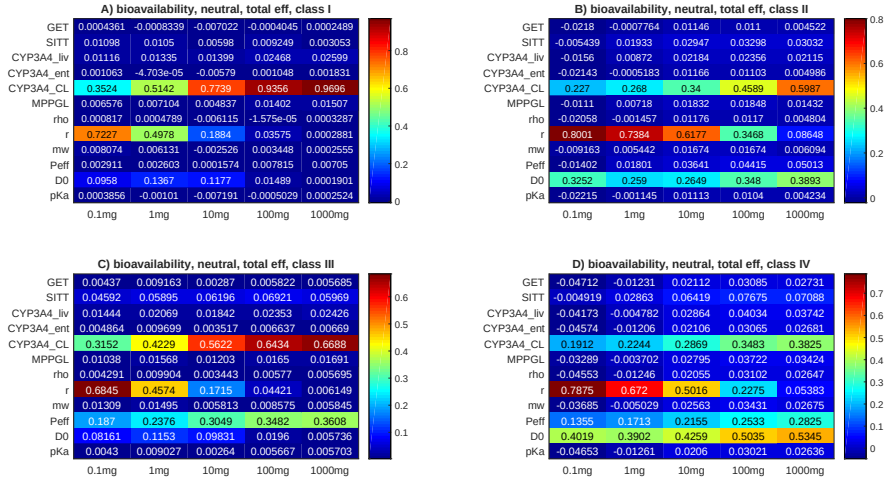



Figure B.6: $F_{oral}$ total effect for CAT based model - neutral compounds

### B.1.2 CAT based model - acidic compounds

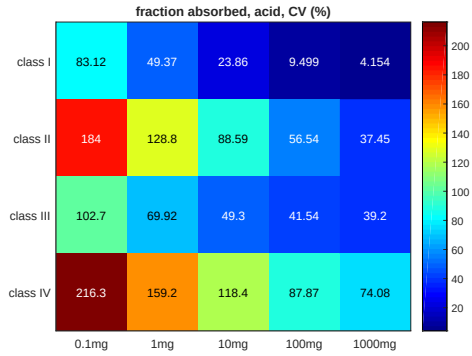



Figure B.7: $f_a$ CV for CAT based model - acidic compounds

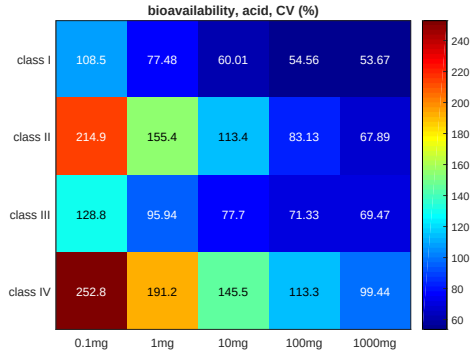



Figure B.8: $F_{oral}$ CV for CAT based model - acidic compounds

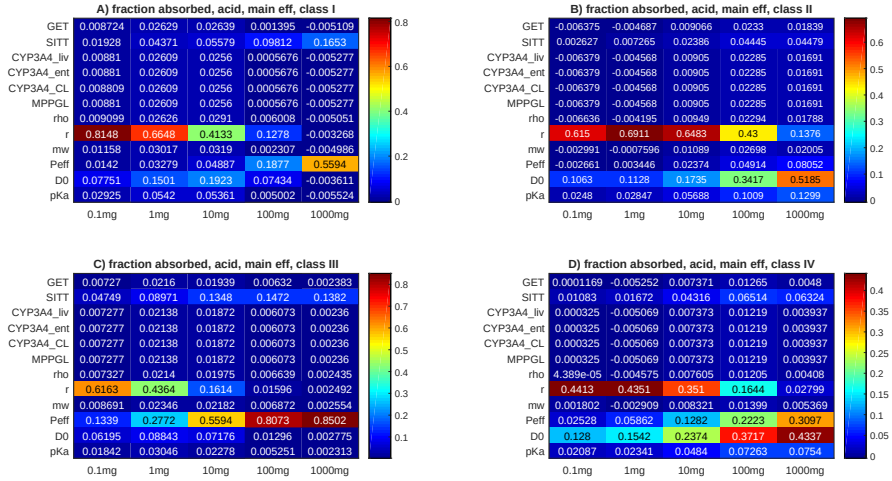



Figure B.9: $f_a$ main effect for CAT based model - acidic compounds

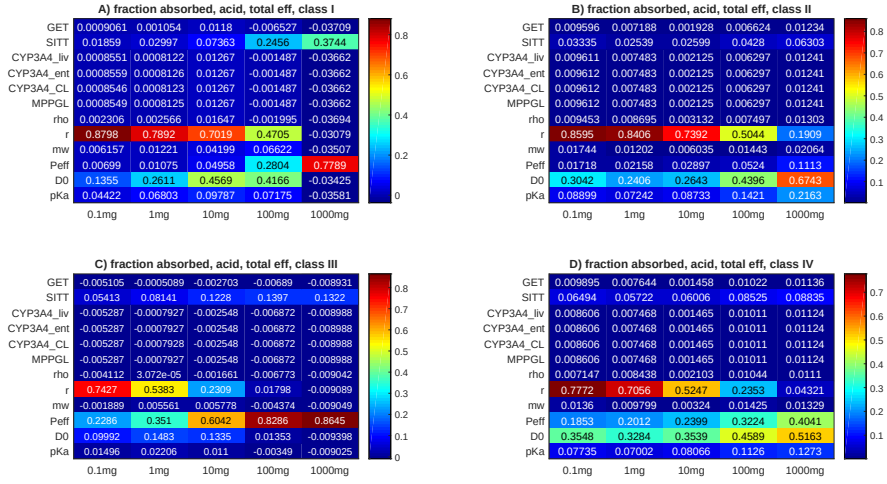



Figure B.10: $f_a$ total effect for CAT based model - acidic compounds

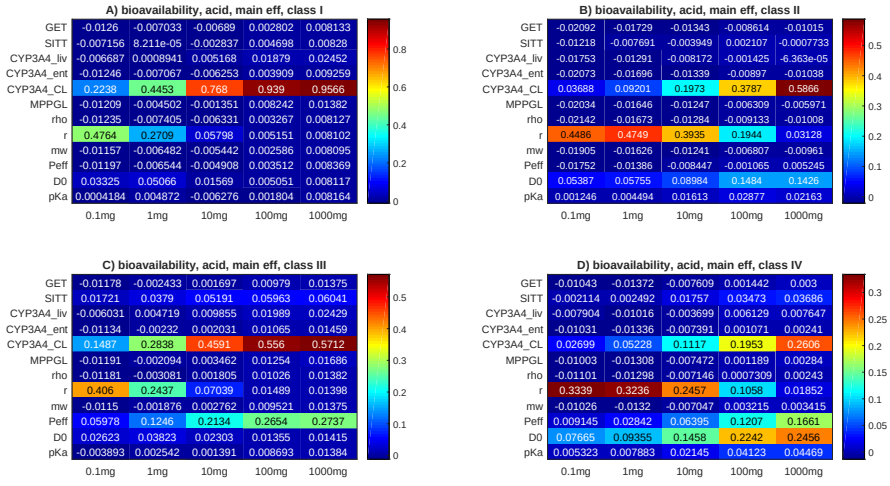



Figure B.11: $F_{oral}$ main effect for CAT based model - acidic compounds

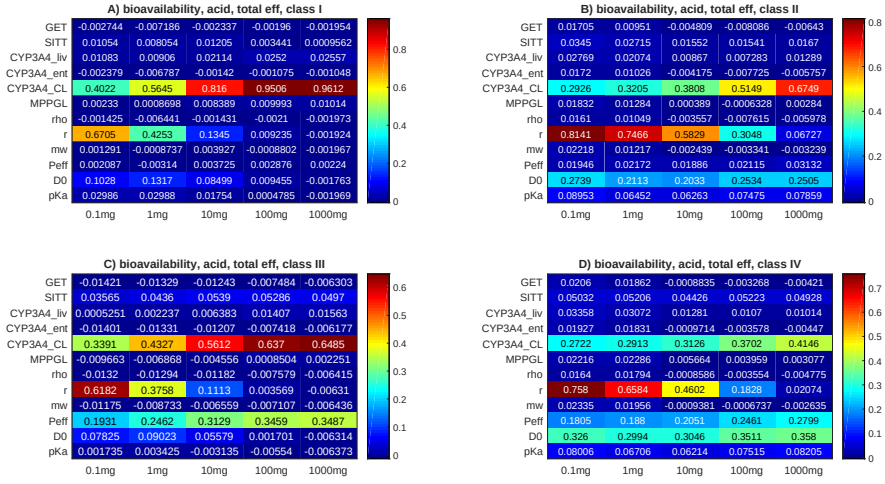



Figure B.12: $F_{oral}$ total effect for CAT based model - acidic compounds

### B.1.3 CAT based model - basic compounds

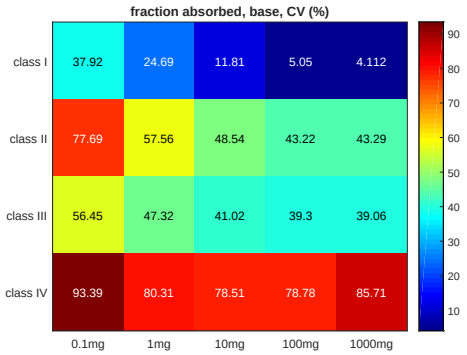



Figure B.13: $f_a$ CV for CAT based model - basic compounds

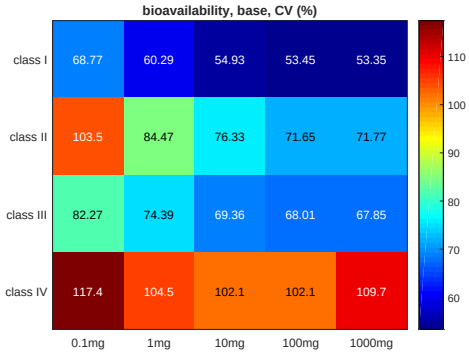



Figure B.14: $F_{oral}$ CV for CAT based model - basic compounds

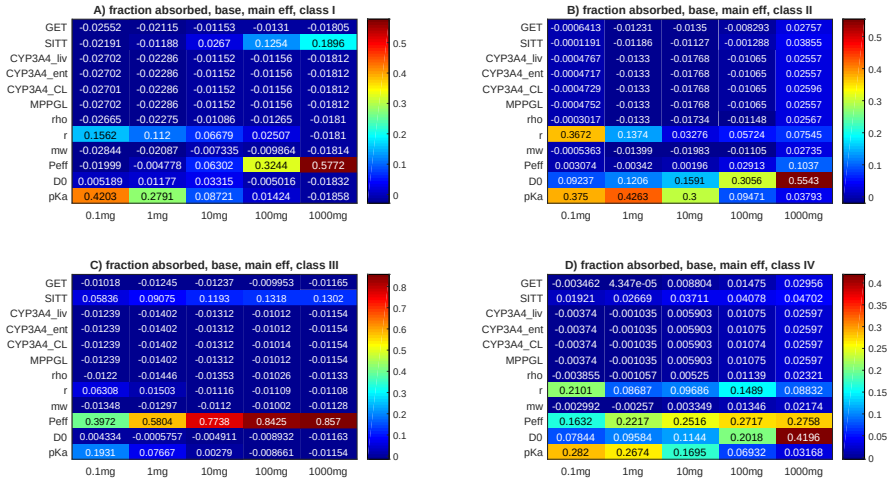



Figure B.15: $f_a$ main effect for CAT based model - basic compounds

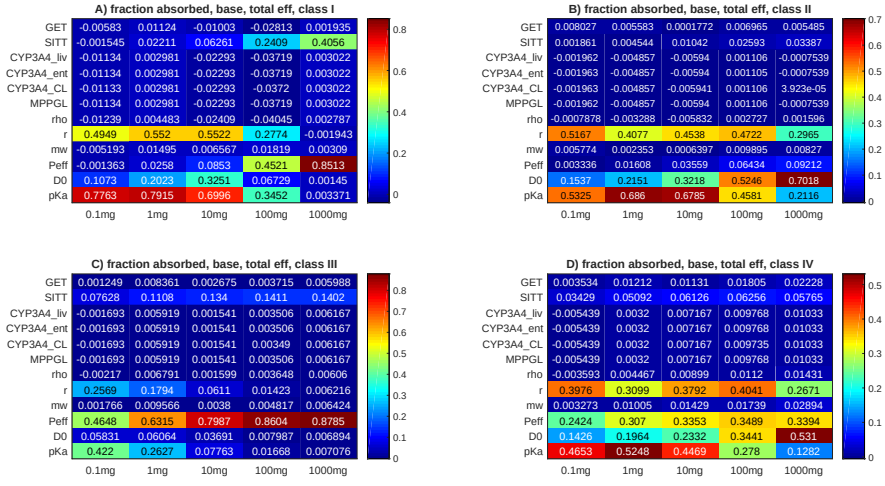



Figure B.16: $f_a$ total effect for CAT based model - basic compounds

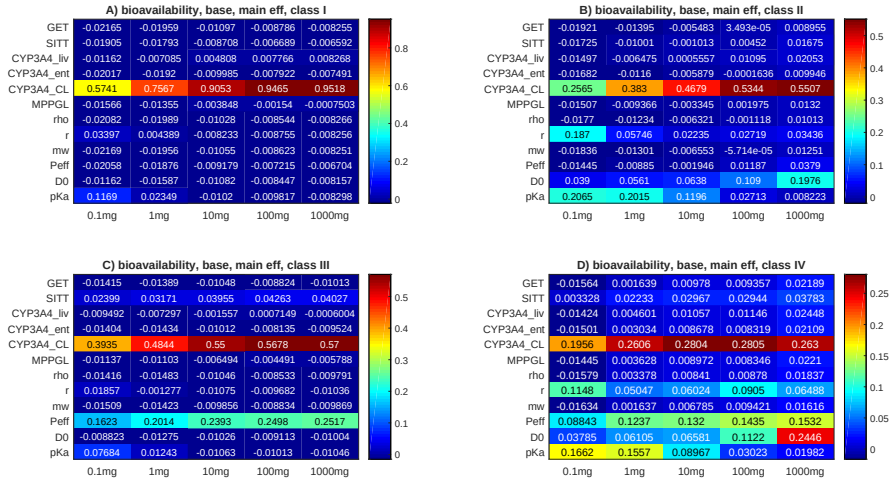



Figure B.17: $F_{oral}$ main effect for CAT based model - basic compounds

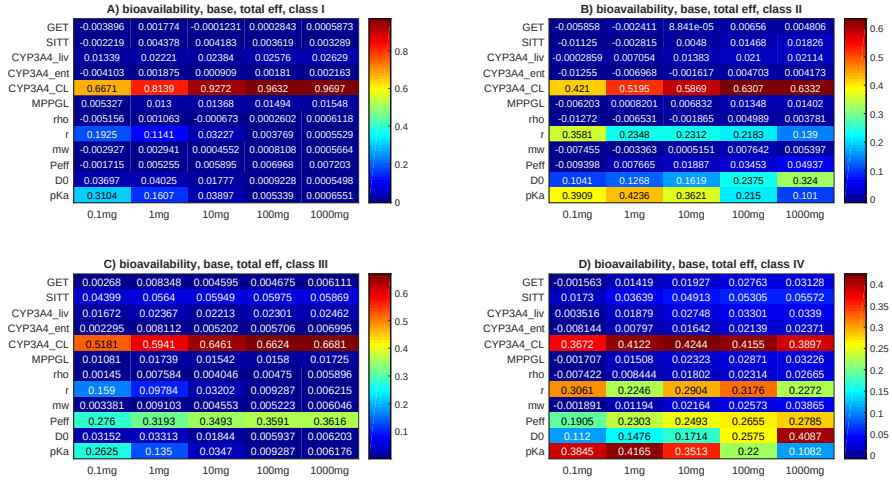



Figure B.18: $F_{oral}$ total effect for CAT based model - basic compounds

### B.1.4 CAT based model - basic compounds with precipitation

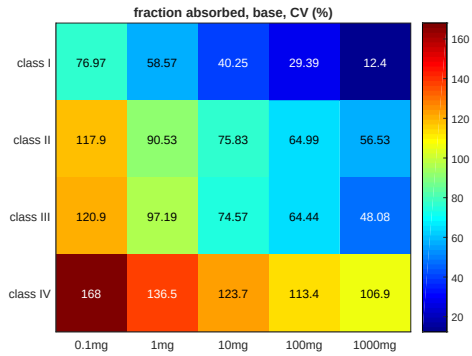



Figure B.19: $f_a$ CV for CAT based model - basic compounds with precipitation

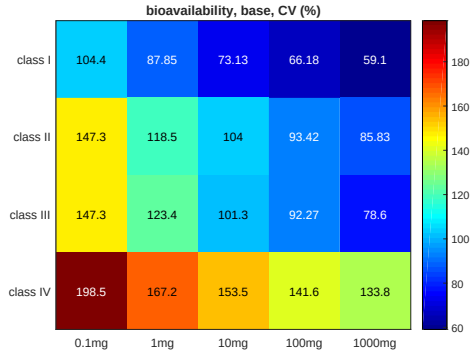



Figure B.20: $F_{oral}$ CV for CAT based model - basic compounds with precipitation

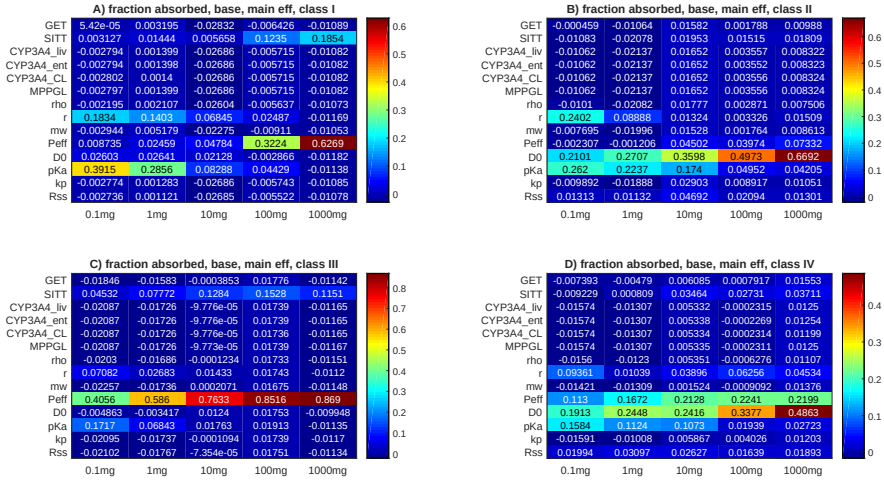



Figure B.21: $f_a$ main effect for CAT based model - basic compounds with precipitation

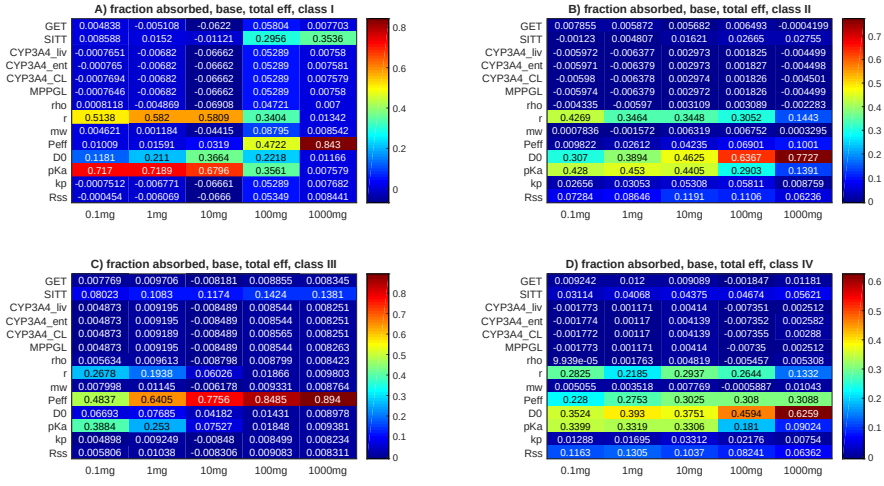



Figure B.22: $f_a$ total effect for CAT based model - basic compounds with precipitation

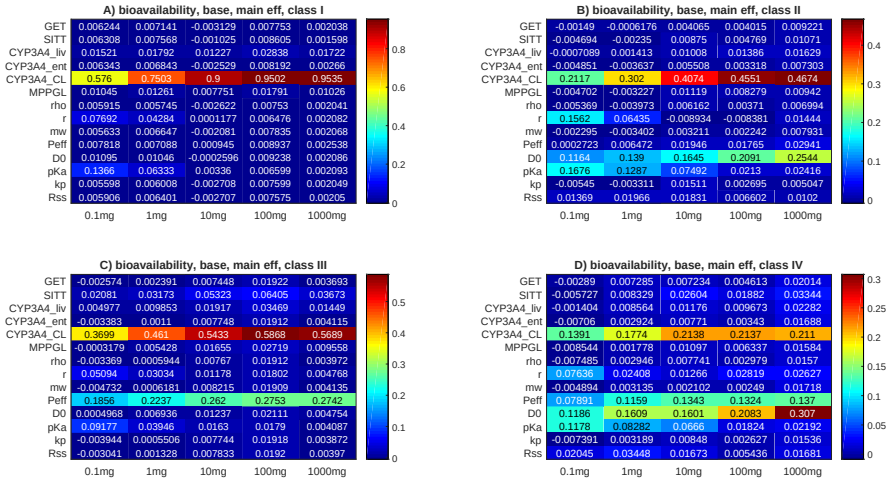



Figure B.23: $F_{oral}$ main effect for CAT based model - basic compounds with precipitation

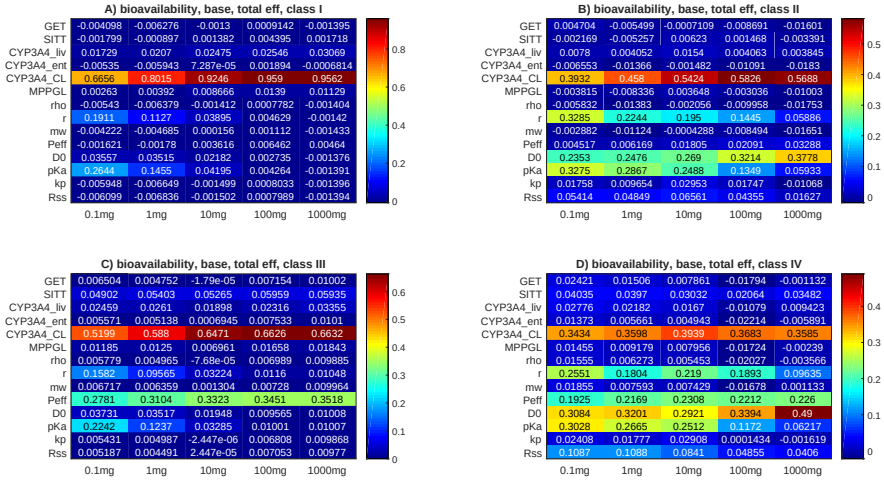



Figure B.24: $F_{oral}$ total effect for CAT based model - basic compounds with precipitation

### B.1.5 MT based model - neutral compounds

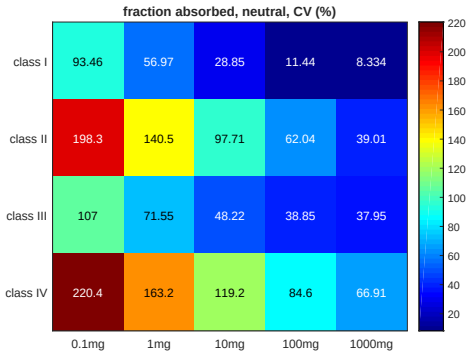



Figure B.25: $f_a$ CV for MT based model - neutral compounds

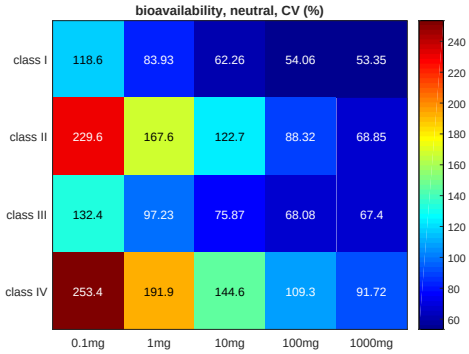



Figure B.26: $F_{oral}$ CV for MT based model - neutral compounds

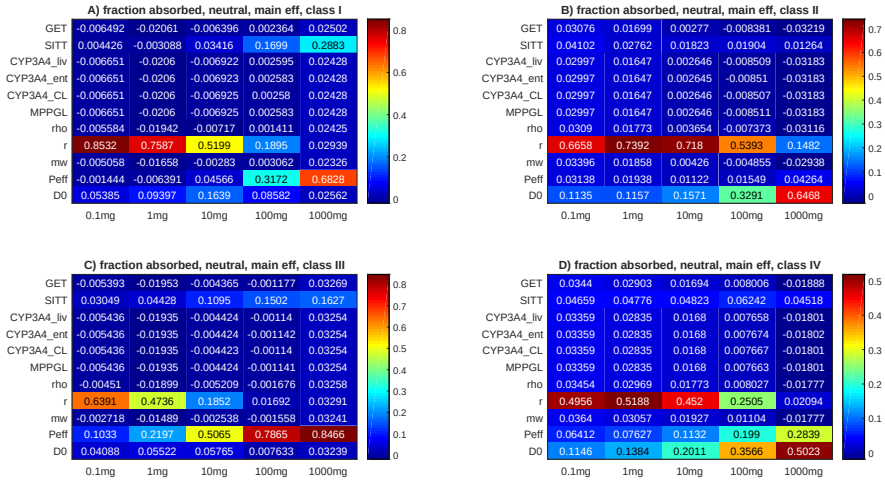



Figure B.27: $f_a$ main effect for MT based model - neutral compounds

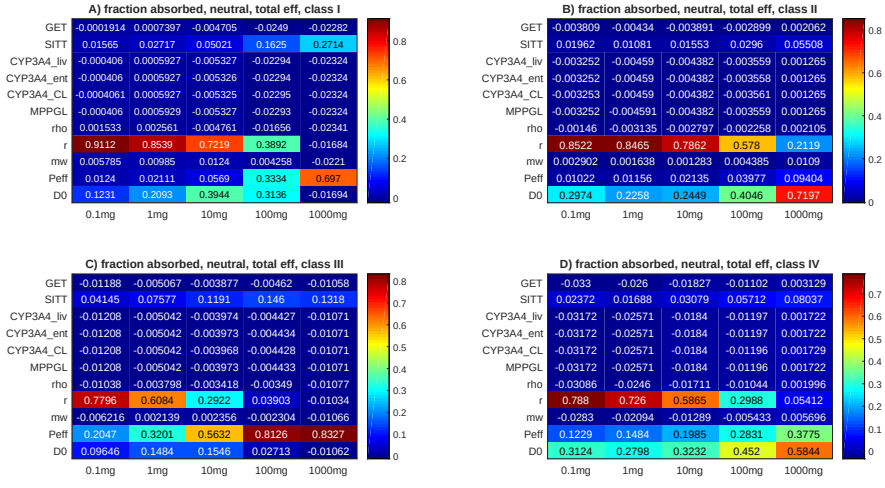



Figure B.28: $f_a$ total effect for MT based model - neutral compounds

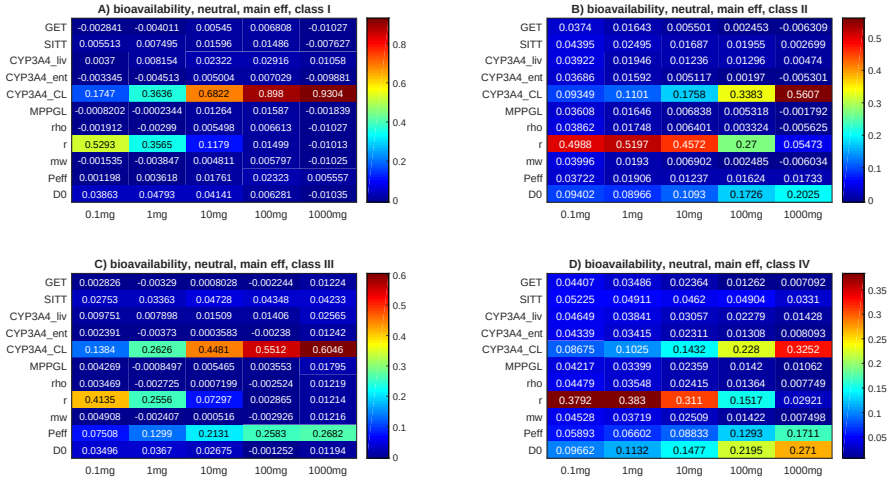



Figure B.29: $F_{oral}$ main effect for MT based model - neutral compounds

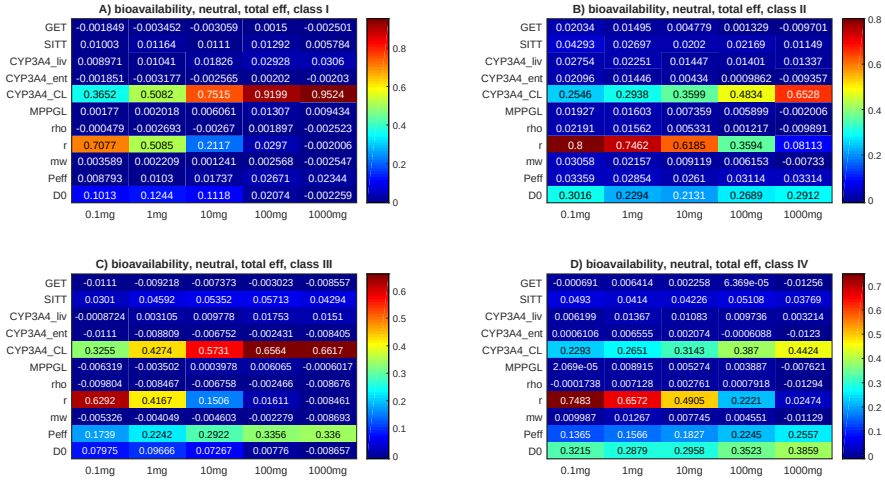



Figure B.30: $F_{oral}$ total effect for MT based model - neutral compounds

## B.2 GSA indices uncertainty

CV, expressed in percentage, of the most sensitive parameter (the one with higher total effect) for a given dose and a given BCS class, obtained using 1000 bootstrap samples.

Table B.1: CAT based model, neutral compounds, $f_a$, main effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 0.88 | 1.33 | 4.29 | 6.82 | 4.07 |
| **BCS II** | 3.72 | 2.12 | 1.75 | 2.97 | 1.52 |
| **BCS III** | 1.81 | 2.65 | 2.1 | 1.69 | 1.5 |
| **BCS IV** | 5.97 | 4.28 | 4.5 | 2.98 | 2.28 |

Table B.2: CAT based model, neutral compounds, $f_a$, total effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 1.55 | 1.75 | 2.36 | 3.31 | 2.08 |
| **BCS II** | 2.30 | 1.96 | 1.96 | 2.33 | 1.75 |
| **BCS III** | 1.93 | 2.25 | 2.59 | 1.78 | 1.61 |
| **BCS IV** | 2.96 | 3.10 | 3.24 | 2.55 | 2.18 |

Table B.3: CAT based model, neutral compounds, $F_{oral}$, main effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 2.56 | 3.38 | 1.62 | 1.04 | 0.99 |
| **BCS II** | 5.45 | 3.55 | 3.45 | 4.34 | 2.49 |
| **BCS III** | 3.7 | 5.72 | 3.01 | 2.29 | 2.19 |
| **BCS IV** | 9.40 | 6.55 | 7.17 | 5.3 | 4.39 |

Table B.4: CAT based model, neutral compounds, $F_{oral}$, total effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 2.17 | 2.53 | 1.76 | 1.53 | 1.43 |
| **BCS II** | 2.85 | 2.55 | 2.7 | 3.08 | 2.28 |
| **BCS III** | 2.69 | 3.37 | 2.60 | 2.30 | 2.26 |
| **BCS IV** | 3.68 | 4.02 | 4.70 | 3.68 | 3.1 |

Table B.5: CAT based model, acidic compounds, $f_a$, main effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 0.99 | 1.68 | 4.8 | 18.91 | 4.48 |
| **BCS II** | 3.01 | 1.75 | 1.67 | 3 | 2.7 |
| **BCS III** | 2.03 | 2.87 | 2.43 | 1.55 | 1.45 |
| **BCS IV** | 5.00 | 3.83 | 4.15 | 3.77 | 3.06 |

Table B.6: CAT based model, acidic compounds, $f_a$, total effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 1.62 | 1.93 | 2.8 | 7.68 | 3.04 |
| **BCS II** | 2.09 | 1.71 | 1.83 | 2.46 | 2.29 |
| **BCS III** | 2.09 | 2.55 | 2.32 | 1.64 | 1.64 |
| **BCS IV** | 3.01 | 2.65 | 3.17 | 3.01 | 2.69 |

Table B.7: CAT based model, acidic compounds, $F_{oral}$, main effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 2.94 | 3.01 | 1.51 | 1.02 | 1 |
| **BCS II** | 4.9 | 3.34 | 3.45 | 3.69 | 2.05 |
| **BCS III** | 3.51 | 4.96 | 2.88 | 2.23 | 2.21 |
| **BCS IV** | 7.07 | 6.09 | 6.50 | 7.54 | 5.75 |

Table B.8: CAT based model, acidic compounds, $F_{oral}$, total effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 2.21 | 2.36 | 1.84 | 1.48 | 1.47 |
| **BCS II** | 2.55 | 2.37 | 2.67 | 2.62 | 2.05 |
| **BCS III** | 2.82 | 3.31 | 2.58 | 2.22 | 2.25 |
| **BCS IV** | 3.67 | 3.52 | 4.49 | 4.53 | 3.82 |

Table B.9: CAT based model, basic compounds, $f_a$, main effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 4.43 | 8.10 | 24.81 | 6.94 | 4.08 |
| **BCS II** | 3.42 | 3.42 | 5.03 | 4.78 | 2.36 |
| **BCS III** | 3.55 | 2.42 | 1.69 | 1.47 | 1.45 |
| **BCS IV** | 4.68 | 5.00 | 8.16 | 9.57 | 3.19 |

Table B.10: CAT based model, basic compounds, $f_a$, total effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 2.08 | 2.71 | 4.67 | 12.34 | 2.02 |
| **BCS II** | 2.29 | 1.99 | 2.27 | 2.79 | 2.07 |
| **BCS III** | 2.99 | 2.40 | 1.87 | 1.56 | 1.57 |
| **BCS IV** | 2.97 | 2.66 | 3.09 | 3.39 | 2.68 |

Table B.11: CAT based model, basic compounds, $F_{oral}$, main effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 2.31 | 1.49 | 1.13 | 0.96 | 0.99 |
| **BCS II** | 5.6 | 3.70 | 2.72 | 2.43 | 2.31 |
| **BCS III** | 3.46 | 2.75 | 2.26 | 2.23 | 2.15 |
| **BCS IV** | 8.5 | 8.93 | 5.47 | 5.56 | 6.09 |

Table B.12: CAT based model, basic compounds, $F_{oral}$, total effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 2.07 | 1.73 | 1.47 | 1.42 | 1.46 |
| **BCS II** | 3.19 | 2.44 | 2.33 | 2.17 | 2.25 |
| **BCS III** | 2.72 | 2.54 | 2.29 | 2.17 | 2.22 |
| **BCS IV** | 4.46 | 4.09 | 3.44 | 3.7 | 4.02 |

Table B.13: CAT based model, basic compounds with precipitation, $f_a$, main effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 4.88 | 7.78 | 25.02 | 7.47 | 4.39 |
| **BCS II** | 5.37 | 6.47 | 3.89 | 2.56 | 1.68 |
| **BCS III** | 3.67 | 2.42 | 1.76 | 1.48 | 1.44 |
| **BCS IV** | 7.68 | 5.94 | 6 | 3.96 | 2.75 |

Table B.14: CAT based model, basic compounds with precipitation, $f_a$, total effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 2.32 | 3.26 | 5.24 | 11.90 | 2.27 |
| **BCS II** | 3.01 | 2.94 | 2.84 | 2.18 | 1.92 |
| **BCS III** | 2.75 | 2.42 | 1.85 | 1.68 | 1.58 |
| **BCS IV** | 4.2 | 3.6 | 3.44 | 3.01 | 2.43 |

Table B.15: CAT based model, basic compounds with precipitation, $F_{oral}$, main effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 2.28 | 1.57 | 1.15 | 1.03 | 1.02 |
| **BCS II** | 7.24 | 4.8 | 3.32 | 2.85 | 2.68 |
| **BCS III** | 3.71 | 3.04 | 2.21 | 2.08 | 2.21 |
| **BCS IV** | 11.25 | 8.68 | 7.09 | 7.19 | 5.08 |

Table B.16: CAT based model, basic compounds with precipitation, $F_{oral}$, total effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 2.18 | 1.78 | 1.57 | 1.47 | 1.48 |
| **BCS II** | 3.68 | 2.96 | 2.54 | 2.30 | 2.49 |
| **BCS III** | 2.73 | 2.47 | 2.19 | 2.28 | 2.21 |
| **BCS IV** | 5.48 | 4.72 | 4.03 | 4.66 | 3.56 |

Table B.17: MT based model, neutral compounds, $f_a$, main effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 0.95 | 1.24 | 3.17 | 10.34 | 2.41 |
| **BCS II** | 3.17 | 1.71 | 1.41 | 2.17 | 1.84 |
| **BCS III** | 1.72 | 2.55 | 2.7 | 1.61 | 1.46 |
| **BCS IV** | 4.37 | 3.08 | 3.05 | 3.81 | 2.52 |

Table B.18: MT based model, neutral compounds, $f_a$, total effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 1.56 | 1.72 | 2.33 | 5.39 | 2.21 |
| **BCS II** | 2.29 | 1.95 | 1.94 | 2.23 | 1.94 |
| **BCS III** | 2.03 | 2.36 | 2.58 | 1.77 | 1.69 |
| **BCS IV** | 3.16 | 2.77 | 3.06 | 3.04 | 2.38 |

Table B.19: MT based model, neutral compounds, $F_{oral}$, main effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 2.61 | 3.59 | 1.68 | 1.07 | 0.99 |
| **BCS II** | 4.92 | 3.40 | 3.19 | 4.35 | 2.1 |
| **BCS III** | 3.48 | 5.46 | 2.86 | 2.24 | 1.99 |
| **BCS IV** | 6.41 | 5.01 | 5.39 | 6.69 | 4.48 |

Table B.20: MT based model, neutral compounds, $F_{oral}$, total effect, CV (%)

| | **0.1 mg** | **1 mg** | **10 mg** | **100 mg** | **1000 mg** |
|---|---|---|---|---|---|
| **BCS I** | 2.23 | 2.68 | 1.90 | 1.55 | 1.5 |
| **BCS II** | 2.93 | 2.58 | 2.69 | 2.97 | 2.19 |
| **BCS III** | 2.89 | 3.60 | 2.56 | 2.23 | 2.21 |
| **BCS IV** | 3.99 | 3.66 | 4.39 | 4.57 | 3.92 |

Appendix C

# Derivation of central and peripheral fraction and lung permeability for the PBPK model presented in chapter 5

## C.1 Derivation of central and peripheral fraction

The central fraction ($F_C$) of the total inhaled dose was calculated from the values of $MMAD$ and $GSD$. The software MPPD2.11 [221, 222] was used for this purpose, with the hypotheses of drug true density ($\rho$) equal to one. In particular, $F_C$ was derived as in equation C.1.

$$F_C = \frac{F_{TB}}{F_{TB} + F_{PUL}} \tag{C.1}$$

$F_{TB}$ and $F_{PUL}$ are the tracheobronchial fraction and pulmonary fraction predicted by using MPPD. This was done because the *precise inhale technology* [223] was used for compound administration in rats, and so, com-

plete inhalation was supposed. During sensitivity analysis the model is repetitively evaluated for different parameter sets. Calling MPPD at each model evaluation for computing $F_C$ would not be an optimal solution from a computational point of view. To reduce the computational cost, we pre-calculated the values of $F_C$ with MPPD for a 15x15 grid of $MMAD$ and $GSD$ values, thus, generating a surface, reported in figure C.1. $F_C$ was then calculated for given values of $MMAD$ and $GSD$ through linear interpolation of this surface.

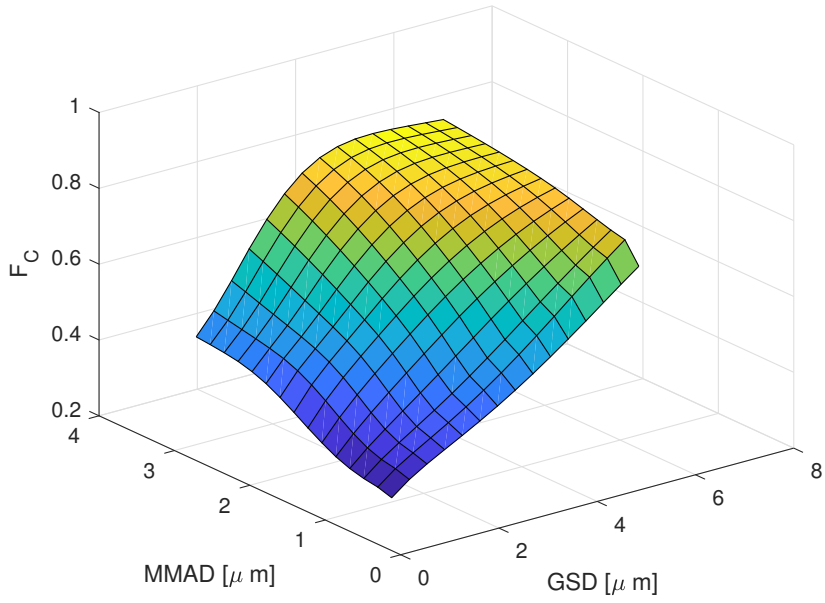



Figure C.1: $F_C$ values calculated from $MMAD$ and $GSD$ by using MPPD.

## C.2 *In vitro* permeability model

Due to the capability of reaching high trans epithelial electric resistance values and for their transporter expression similar to those of the respiratory tissues, the Calu3 cell lines are commonly used to assess the permeability of inhaled drugs [224]. This type of cell is robust and easy to culture, so, it can be used for routine screening purposes. Furthermore, good correlation between the permeability values and the in vivo drug absorption in drug lung was shown [225, 226]. During the experiment, the cells are exposed to a liquid medium in both the apical and the basolateral surface. A 10 $\mu M$ compound solution is added in the liquid on a certain side (e.g., apical) and the concentration is monitored in the opposite side (e.g., basolateral) two hours later. The monitored concentration allows to obtain the apparent permeability $P_{app}$ (for the apical-basolateral and basolateral-apical directions) that is typically calculated with the following equation:

$$P_{app} = \frac{\Delta Q/\Delta t}{C_0\, A}, \tag{C.2}$$

where $\Delta Q/\Delta t$ is the change in the amount of drug during the considered interval in the receiving fluid, $C_0$ is the initial concentration in the medium in which the compound is administered, $A$ is the cell layer surface area. The obtained value is the net result of passive and active transports.

To allow a discrimination of the passive and active transport, in [178] a mechanistic model of the in vitro Calu3 permeability system was built. The model is composed by three compartments: the apical media, the cells and the basolateral media. It is assumed that the main fluxes in the system are due to the passive transcellular bidirectional transport between the cells and the fluids and to the monodirectional efflux from the tissue to the epithelial lining fluid. The model is reported in equation system C.3.

$$\begin{aligned}
\frac{dc_{AF}}{dt} &= \frac{1}{V_{AF}}(-P_p\,A\,c_{AF} + P_p\,A\,c_c\,f_{u,lung} + P_a\,A\,c_c\,f_{u,lung}) \\
\frac{dc_c}{dt} &= \frac{1}{V_C}(-2P_p\,A\,c_c\,f_{u,lung} - P_a\,A\,c_c\,f_{u,lung} + P_p\,A\,c_{AF} + P_p\,A\,c_{BF}) \\
\frac{dc_{BF}}{dt} &= \frac{1}{V_{BF}}(-P_p\,A\,c_{BF} + P_p\,A\,c_c\,f_{u,lung})
\end{aligned} \tag{C.3}$$

$c_{AF}$, $c_{BF}$ and $c_c$ are the drug concentrations in the apical and basolateral fluid and in the cell layer, respectively; $V_{AF}$, $V_{BF}$ and $V_C$ are the apical and basolateral fluid and cells volumes, respectively; $A$ is the area of the cell culture, $P_p$ and $P_a$ are the passive and active permeabilities. $V_{AF}$ and $V_{BF}$ are equal to 100 and 600 $\mu L$, respectively and $V_C$ was calculated by multiplying $A$, equal to 0.33 $cm^2$, for the cell layer width, equal to 17.5 $\mu m$ (internal data). The values of $P_p$ and $P_a$ are estimated by simultaneously fitting the basolateral media concentration data obtained with the apical-basolateral experiment and the apical medium concentration data after obtained with the basolateral-apical permeability experiment.

Appendix D

# Derivation of $F_{oral}$ for the semi-mechanistic oral absorption model presented in chapter 6

The differential equations of the semi-mechanistic model describing the absorption of orally administered compounds metabolized by CYP3A4 in gut wall and liver, are reported in equation system 6.5. The model is represented in figure 6.1 (b). The model describes the bioavailability of a generic drug following oral administration. Our purpose was to derive analytically the expression of the $F_{oral}$ following a bolus in the compartment representing the small intestine.

The solution of $x_{lum}$ is the one of a single compartment with two elimination rates, so:

$$x_{lum} = M_0 e^{-(k_t + k_a)t},$$

with $M_0$ the dose. It is possible to substitute the expression of $x_{lum}$ in the equation representing the dynamics of $x_{ent}$ in the equation system 6.5. By

multiplying $\exp((Q_{ent} + CL_{ent})/V_{ent} \cdot t)$ in both the sides of that equation, is possible to derive the expressions below and thus the analytical solution of $x_{ent}$.

$$\frac{d}{dt}\left(x_{ent} e^{\left(\frac{Q_{ent}}{V_{ent}} + \frac{CL_{ent}}{V_{ent}}\right)t}\right) = M_0 \; k_a \; e^{\left(-(k_t+k_a)t + \left(\frac{Q_{ent}}{V_{ent}} + \frac{CL_{ent}}{V_{ent}}\right)t\right)}$$

$$x_{ent} \; e^{\left(\frac{Q_{ent}}{V_{ent}} + \frac{CL_{ent}}{V_{ent}}\right)t} = M_0 \; k_a \int_0^t e^{\left(-(k_t+k_a)t + \left(\frac{Q_{ent}}{V_{ent}} + \frac{CL_{ent}}{V_{ent}}\right)t\right)} dt$$

$$x_{ent} = \frac{M_0 k_a}{\left(\frac{Q_{ent}}{V_{ent}} + \frac{CL_{ent}}{V_{ent}}\right) - (k_t + k_a)} \left[e^{-(k_t+k_a)t} - e^{-\left(\frac{Q_{ent}}{V_{ent}} + \frac{CL_{ent}}{V_{ent}}\right)t}\right]$$

For readability purpose, let us define the following variables.

$$\begin{aligned} k_{lum} &= (k_t + k_a) \\ k_{ent} &= \left(\frac{Q_{ent}}{V_{ent}} + \frac{CL_{ent}}{V_{ent}}\right) \\ k_{liv} &= \left(\frac{Q_{liv,ven}}{V_{liv}} + \frac{CL_{liv}}{V_{liv}}\right) \\ \psi_1 &= \frac{M_0 k_a}{k_{ent} - k_{lum}} \end{aligned}$$

Now, by substituting the $x_{ent}$ expression in $x_{liv}$ differential equation and multiplying both equation sides for $\exp(k_{liv} \cdot \; t)$, is possible to derive the analytical solution of $x_{liv}$.

$$\frac{d}{dt}\left(x_{liv} e^{(k_{liv}t)}\right) = \frac{Q_{ent}}{V_{ent}} \psi_1 \left(e^{(-k_{lum}t + k_{liv}t)} - e^{(-k_{ent}t + k_{liv}t)}\right)$$

$$x_{liv} e^{k_{liv}t} = \frac{Q_{ent}}{V_{ent}} \psi_1 \int_0^t \left(e^{(-k_{lum}t + k_{liv}t)} - e^{(-k_{ent}t + k_{liv}t)}\right) dt$$

$$x_{liv} = \frac{Q_{ent}}{V_{ent}}\psi_1\Bigg[\frac{1}{k_{liv}-k_{lum}}\ \left(e^{-k_{lum}t}-e^{-k_{liv}t}\right) - \frac{1}{k_{liv}-k_{ent}}\left(e^{-k_{ent}t}-e^{-k_{liv}t}\right)\Bigg]$$

Now, by substituting $x_{liv}$ in the last differential equation of system 6.5, is possible to derive directly the analytical solution of $x_{sys}$.

$$x_{sys} = \frac{Q_{liv,ven}}{V_{liv}}\frac{Q_{ent}}{V_{ent}}\psi_1\int_0^t\Bigg[\frac{1}{k_{liv}-k_{lum}}\ \left(e^{-k_{lum}t}-e^{-k_{liv}t}\right) - \frac{1}{k_{liv}-k_{ent}}\left(e^{-k_{ent}t}-e^{-k_{liv}t}\right)\Bigg]dt$$

$$x_{sys} = \frac{Q_{liv,ven}}{V_{liv}}\frac{Q_{ent}}{V_{ent}}\psi_1 \cdot\Bigg[\frac{1}{k_{liv}-k_{lum}}\left(\frac{1}{k_{lum}}\left(1-e^{-k_{lum}t}\right)-\frac{1}{k_{liv}}\left(1-e^{-k_{liv}t}\right)\right) - \frac{1}{k_{liv}-k_{ent}}\left(\frac{1}{k_{ent}}\left(1-e^{-k_{ent}t}\right)-\frac{1}{k_{liv}}\left(1-e^{-k_{liv}t}\right)\right)\Bigg]$$

Finally, it is possible to derive $F_{oral}$.

$$F_{oral} = \lim_{t\to+\infty}\frac{x_{sys}}{M_0} = \frac{k_a(Q_{ent}/V_{ent})(Q_{liv,ven}/V_{liv})}{k_{ent}-k_{lum}} \cdot\Bigg[\frac{1}{k_{liv}-k_{lum}}\left(\frac{1}{k_{lum}}-\frac{1}{k_{liv}}\right)-\frac{1}{k_{liv}-k_{ent}}\left(\frac{1}{k_{ent}}-\frac{1}{k_{liv}}\right)\Bigg]$$

$$F_{oral} = \frac{k_a(Q_{ent}/V_{ent})(Q_{liv,ven}/V_{liv})}{k_{ent}-k_{lum}}\ \cdot\ \frac{k_{ent}-k_{lum}}{k_{lum}k_{ent}k_{liv}}$$

$$F_{oral} = \frac{k_a\cdot(Q_{ent}/V_{ent})\cdot(Q_{liv,ven}/V_{liv})}{(k_a+k_t)\left(\frac{Q_{ent}}{V_{ent}}+\frac{CL_{ent}}{V_{ent}}\right)\left(\frac{Q_{liv,ven}}{V_{liv}}+\frac{CL_{liv}}{V_{liv}}\right)}$$

# Bibliography

[1] Steven M. Paul, Daniel S. Mytelka, Christopher T. Dunwiddie, Charles C. Persinger, Bernard H. Munos, Stacy R. Lindborg, and Aaron L. Schacht. How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nature Reviews Drug Discovery*, 9(3):203–214, March 2010.

[2] Joseph A. DiMasi, Henry G. Grabowski, and Ronald W. Hansen. Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics*, 47:20–33, May 2016.

[3] Lawrence J. Lesko, Malcolm Rowland, Carl C. Peck, and Terrence F. Blaschke. Optimizing the Science of Drug Development: Opportunities for Better Candidate Selection and Accelerated Evaluation in Humans. *The Journal of Clinical Pharmacology*, 40(8):803–814, 2000.

[4] The Drug Development Process. `http://www.fda.gov/patients/learn-about-drug-and-device-approvals/drug-development-process`. Accessed 3 September 2019.

[5] Malcolm Rowland and Thomas N. Tozer. *Clinical Pharmacokinetics: Concepts and Applications*. Lippincott Williams & Wilkins, third edition, 1995.

[6] Paul J. Williams and Ene I. Ette. Pharmacometrics: Impacting Drug Development and Pharmacotherapy. In *Pharmacometrics: The Science of Quantitative Pharmacology*, pages 1–21. John Wiley & Sons, Ltd, 2006.

[7] John Arrowsmith. Trial watch: Phase II failures: 2008-2010. *Nature Reviews Drug Discovery*, 10:328–329, April 2011.

[8] John Arrowsmith. Trial watch: Phase III and submission failures: 2007-2010. *Nature Reviews Drug Discovery*, 10:87–87, February 2011.

[9] John Arrowsmith and Philip Miller. Trial Watch: Phase II and Phase III attrition rates 2011-2012. *Nature Reviews Drug Discovery*, 12:569–569, August 2013.

[10] Satyaprakash Nayak, Oliver Sander, Nidal Al-Huniti, Dinesh de Alwis, Anne Chain, Marylore Chenel, Soujanya Sunkaraneni, Shruti Agrawal, Neeraj Gupta, and Sandra A. G. Visser. Getting Innovative Therapies Faster to Patients at the Right Dose: Impact of Quantitative Pharmacology Towards First Registration and Expanding Therapeutic Use. *Clinical Pharmacology & Therapeutics*, 103(3):378–383, 2018.

[11] S. F. Marshall, R. Burghaus, V. Cosson, S. Y. A. Cheung, M. Chenel, O. DellaPasqua, N. Frey, B. Hamrén, L. Harnisch, F. Ivanow, T. Kerbusch, J. Lippert, P. A. Milligan, S. Rohou, A. Staab, J. L. Steimer, C. Tornøe, and S. a. G. Visser. Good Practices in Model-Informed Drug Discovery and Development: Practice, Application, and Documentation. *CPT: Pharmacometrics & Systems Pharmacology*, 5(3):93–122, 2016.

[12] Scott Marshall, Rajanikanth Madabushi, Efthymios Manolis, Kevin Krudys, Alexander Staab, Kevin Dykstra, and Sandra A. G. Visser. Model-Informed Drug Discovery and Development: Current Industry Good Practice and Regulatory Expectations and Future Perspectives. *CPT: Pharmacometrics & Systems Pharmacology*, 8(2):87–96, 2019.

[13] P. A. Milligan, M. J. Brown, B. Marchant, S. W. Martin, P. H. van der Graaf, N. Benson, G. Nucci, D. J. Nichols, R. A. Boyd, J. W. Mandema, S. Krishnaswami, S. Zwillich, D. Gruben, R. J. Anziano, T. C. Stock, and R. L. Lalonde. Model-Based Drug Development: A Rational Approach to Efficiently Accelerate Drug Development. *Clinical Pharmacology & Therapeutics*, 93(6):502–514, 2013.

[14] S. R. B. Allerheiligen. Impact of Modeling and Simulation: Myth or Fact? *Clinical Pharmacology & Therapeutics*, 96(4):413–415, 2014.

[15] S. a. G. Visser, D. P. de Alwis, T. Kerbusch, J. A. Stone, and S. R. B. Allerheiligen. Implementation of Quantitative and Systems Pharmacology in Large Pharma. *CPT: Pharmacometrics & Systems Pharmacology*, 3(10):142, 2014.

[16] US Food and Drug Administration (FDA). Pdufa reauthorization performance goals and procedures fiscal years 2018 through 2022. `https://www.fda.gov/media/99140/download`. Accessed 3 September 2019.

[17] Yaning Wang, Hao Zhu, Rajanikanth Madabushi, Qi Liu, Shiew-Mei Huang, and Issam Zineh. Model-Informed Drug Development: Current US Regulatory Practice and Future Considerations. *Clinical Pharmacology & Therapeutics*, 105(4):899–911, 2019.

[18] Malcolm Rowland and Leslie Benet. Pharmacometrics: a new journal section. *Journal of Pharmacokinetics and Biopharmaceutics*, (10):349–350, 1982.

[19] Peter K. Sorger, Sandra R. B. Allerheiligen, Darrell R. Abernethy, Russ B. Altman, Kim L. R. Brouwer, Andrea Califano, David Z. D'Argenio, Ravi Iyengar, William J. Jusko, Richard Lalonde, Douglas A. Lauffenburger, Brian Shoichet, James L. Stevens, Shankar Subramaniam, Piet Van der Graaf, and Paolo Vicini. Quantitative and systems pharmacology in the post-genomic era: new approaches to discovering drugs and understanding therapeutic mechanisms. NIH

White Paper by the QSP Workshop Group, National Institutes of Health, October 2011.

[20] Piet Van der Graaf. CPT: Pharmacometrics and Systems Pharmacology. *CPT: Pharmacometrics & Systems Pharmacology*, 1(9):8, 2012.

[21] Andrea Saltelli, Marco Ratto, Terry Andres, Francesca Campolongo, Jessica Cariboni, Debora Gatelli, Michaela Saisana, and Stefano Tarantola. *Global Sensitivity Analysis. The Primer*. John Wiley & Sons, Ltd, January 2008.

[22] P. Vicini and P. H. van der Graaf. Systems Pharmacology for Drug Discovery and Development: Paradigm Shift or Flash in the Pan? *Clinical Pharmacology & Therapeutics*, 93(5):379–381, 2013.

[23] Ene I. Ette and Paul J. Williams. *Pharmacometrics: The Science of Quantitative Pharmacology*. John Wiley & Sons, Inc., May 2006.

[24] Naomi Oreskes, Kristin Shrader-Frechette, and Kenneth Belitz. Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science*, 263(5147):641–646, February 1994.

[25] Arthur A. Noyes and Willis R. Whitney. The rate of solution of solid substances in their own solutions. *Journal of the American Chemical Society*, 19(12):930–934, December 1897.

[26] Michaelis L. and Menten M.L. Die Kinetik der Invertinwirkung. *Biochemistry Zeitung*, 49:333 – 369, 1913.

[27] Peter Kohl and Denis Noble. Systems biology and the virtual physiological human. *Molecular Systems Biology*, 5(1):292, January 2009.

[28] Issam Zineh. Quantitative Systems Pharmacology: A Regulatory Perspective on Translation. *CPT: Pharmacometrics & Systems Pharmacology*, 8(6):336–339, 2019.

[29] Amin Rostami-Hodjegan. Physiologically Based Pharmacokinetics Joined With In Vitro-In Vivo Extrapolation of ADME: A Marriage Under the Arch of Systems Pharmacology. *Clinical Pharmacology & Therapeutics*, 92(1):50–61, 2012.

[30] Amin Rostami-Hodjegan. Reverse Translation in PBPK and QSP: Going Backwards in Order to Go Forward With Confidence. *Clinical Pharmacology & Therapeutics*, 103(2):224–232, 2018.

[31] Malcolm Rowland, Carl Peck, and Geoffrey Tucker. Physiologically-Based Pharmacokinetics in Drug Development and Regulatory Science. *Annual Review of Pharmacology and Toxicology*, 51(1):45–73, 2011.

[32] H. M. Jones and K. Rowland-Yeo. Basic Concepts in Physiologically Based Pharmacokinetic Modeling in Drug Discovery and Development. *CPT: Pharmacometrics & Systems Pharmacology*, page 63, 2013.

[33] Torsten Teorell. Kinetics of distribution of substances administered to the body. I. The extravascular modes of administration. *Archives internationales de pharmacodynamie et de therapie*, 57:202–5, 1937.

[34] Italo Poggesi, Jan Snoeys, and Achiel Van Peer. The successes and failures of physiologically based pharmacokinetic modeling: there is room for improvement. *Expert Opinion on Drug Metabolism & Toxicology*, 10(5):631–635, May 2014.

[35] Masoud Jamei. Recent Advances in Development and Application of Physiologically-Based Pharmacokinetic (PBPK) Models: a Transition from Academic Curiosity to Regulatory Acceptance. *Current Pharmacology Reports*, 2(3):161–169, June 2016.

[36] Hannah M. Jones, Maurice Dickins, Kuresh Youdim, James R. Gosset, Neil J. Attkins, Tanya L. Hay, Ian K. Gurrell, Y. Raj Logan, Peter J. Bungay, Barry C. Jones, and Iain B. Gardner. Application

of PBPK modelling in drug discovery and development at Pfizer. *Xenobiotica*, 42(1):94–106, January 2012.

[37] H. M. Jones, Y. Chen, C. Gibson, T. Heimbach, N. Parrott, S. A. Peters, J. Snoeys, V. V. Upreti, M. Zheng, and S. D. Hall. Physiologically based pharmacokinetic modeling in drug discovery and development: A pharmaceutical industry perspective. *Clinical Pharmacology & Therapeutics*, 97(3):247–262, 2015.

[38] Jennifer E. Sager, Jingjing Yu, Isabelle Ragueneau-Majlessi, and Nina Isoherranen. Physiologically Based Pharmacokinetic (PBPK) Modeling and Simulation Approaches: A Systematic Review of Published Models, Applications, and Model Verification. *Drug Metabolism and Disposition*, 43(11):1823–1837, November 2015.

[39] Patrick Poulin and Frank-Peter Theil. Prediction of pharmacokinetics prior to in vivo studies. 1. Mechanism-based prediction of volume of distribution. *Journal of Pharmaceutical Sciences*, 91(1):129–156, January 2002.

[40] Trudy Rodgers, David Leahy, and Malcolm Rowland. Physiologically Based Pharmacokinetic Modeling 1: Predicting the Tissue Distribution of Moderate-to-Strong Bases. *Journal of Pharmaceutical Sciences*, 94(6):1259–1276, June 2005.

[41] Trudy Rodgers and Malcolm Rowland. Physiologically based pharmacokinetic modelling 2: Predicting the tissue distribution of acids, very weak bases, neutrals and zwitterions. *Journal of Pharmaceutical Sciences*, 95(6):1238–1257, June 2006.

[42] Leonid M. Berezhkovskiy. Volume of Distribution at Steady State for a Linear Pharmacokinetic System with Peripheral Elimination. *Journal of Pharmaceutical Sciences*, 93(6):1628–1640, June 2004.

[43] Neil Parrott and Thierry Lave. Applications of Physiologically Based Absorption Models in Drug Discovery and Development. *Molecular Pharmaceutics*, 5(5):760–775, October 2008.

[44] Pankaj R. Daga, Michael B. Bolger, Ian S. Haworth, Robert D. Clark, and Eric J. Martin. Physiologically Based Pharmacokinetic Modeling in Lead Optimization. 1. Evaluation and Adaptation of GastroPlus To Predict Bioavailability of Medchem Series. *Molecular Pharmaceutics*, 15(3):821–830, March 2018.

[45] Ivan A. Nestorov, Leon J. Aarons, Philip A. Arundel, and Malcolm Rowland. Lumping of Whole-Body Physiologically Based Pharmacokinetic Models. *Journal of Pharmacokinetics and Biopharmaceutics*, 26(1):21–46, February 1998.

[46] Ivelina Gueorguieva, Ivan A. Nestorov, and Malcolm Rowland. Reducing Whole Body Physiologically Based Pharmacokinetic Models Using Global Sensitivity Analysis: Diazepam Case Study. *Journal of Pharmacokinetics and Pharmacodynamics*, 33(1):1–27, February 2006.

[47] S.-M. Huang and M. Rowland. The Role of Physiologically Based Pharmacokinetic Modeling in Regulatory Review. *Clinical Pharmacology & Therapeutics*, 91(3):542–549, 2012.

[48] P. Zhao, M. Rowland, and S.-M. Huang. Best Practice in the Use of Physiologically Based Pharmacokinetic Modeling and Simulation to Address Clinical Pharmacology Regulatory Questions. *Clinical Pharmacology & Therapeutics*, 92(1):17–20, 2012.

[49] Manuela Grimstein, Yuching Yang, Xinyuan Zhang, Joseph Grillo, Shiew-Mei Huang, Issam Zineh, and Yaning Wang. Physiologically Based Pharmacokinetic Modeling in Regulatory Science: An Update From the U.S. Food and Drug Administration's Office of Clinical Pharmacology. *Journal of Pharmaceutical Sciences*, 108(1):21–25, January 2019.

[50] CHMP (EMA). Guideline on the reporting of physiologically based pharmacokinetic (PBPK) modelling and simulation. Technical Report EMA/CHMP/458101/2016, Committee for Medicinal Products

for Human Use (CHMP), European Medicines Agency (EMA), London, UK, December 2018.

[51] CDER (FDA). Physiologically Based Pharmacokinetic analyses - format and content: guidance for industry. Technical report, Center for Drug Evaluation and Research (CDER), U.S. Food and Drug Administration (FDA), Rockville, MD, August 2018.

[52] Patrick Poulin, Rhys D. O. Jones, Hannah M. Jones, Christopher R. Gibson, Malcolm Rowland, Jenny Y. Chien, Barbara J. Ring, Kimberly K. Adkison, M. Sherry Ku, Handan He, Ragini Vuppugalla, Punit Marathe, Volker Fischer, Sandeep Dutta, Vikash K. Sinha, Thorir Björnsson, Thierry Lavé, and James W. T. Yates. 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. *Journal of Pharmaceutical Sciences*, 100(10):4127–4157, October 2011.

[53] Alison Margolskee, Adam S. Darwich, Xavier Pepin, Leon Aarons, Aleksandra Galetin, Amin Rostami-Hodjegan, Sara Carlert, Maria Hammarberg, Constanze Hilgendorf, Pernilla Johansson, Eva Karlsson, Dónal Murphy, Christer Tannergren, Helena Thörn, Mohammed Yasin, Florent Mazuir, Olivier Nicolas, Sergej Ramusovic, Christine Xu, Shriram M. Pathak, Timo Korjamo, Johanna Laru, Jussi Malkki, Sari Pappinen, Johanna Tuunainen, Jennifer Dressman, Simone Hansmann, Edmund Kostewicz, Handan He, Tycho Heimbach, Fan Wu, Carolin Hoft, Loic Laplanche, Yan Pang, Michael B. Bolger, Eva Huehn, Viera Lukacova, James M. Mullin, Ke X. Szeto, Chester Costales, Jian Lin, Mark McAllister, Sweta Modi, Charles Rotter, Manthena Varma, Mei Wong, Amitava Mitra, Jan Bevernage, Jeike Biewenga, Achiel Van Peer, Richard Lloyd, Carole Shardlow, Peter Langguth, Irina Mishenzon, Mai Anh Nguyen, Jonathan Brown, Hans Lennernäs, and Bertil Abrahamsson. Imi - oral biopharmaceutics tools project - evaluation of bottom-up pbpk prediction success part

2: An introduction to the simulation exercise and overview of results. *European Journal of Pharmaceutical Sciences*, 96:610–625, January 2017.

[54] Alison Margolskee, Adam S. Darwich, Xavier Pepin, Shriram M. Pathak, Michael B. Bolger, Leon Aarons, Amin Rostami-Hodjegan, Jonas Angstenberger, Franziska Graf, Loic Laplanche, Thomas Müller, Sara Carlert, Pankaj Daga, Dònal Murphy, Christer Tannergren, Mohammed Yasin, Susanne Greschat-Schade, Wolfgang Mück, Uwe Muenster, Dorina van der Mey, Kerstin Julia Frank, Richard Lloyd, Lieve Adriaenssen, Jan Bevernage, Loeckie De Zwart, Dominique Swerts, Christophe Tistaert, An Van Den Bergh, Achiel Van Peer, Stefania Beato, Anh-Thu Nguyen-Trung, Joanne Bennett, Mark McAllister, Mei Wong, Patricia Zane, Cèline Ollier, Pascale Vicat, Markus Kolhmann, Alexander Marker, Priscilla Brun, Florent Mazuir, Stèphane Beilles, Marta Venczel, Xavier Boulenc, Petra Loos, Hans Lennernäs, and Bertil Abrahamsson. Imi - oral biopharmaceutics tools project - evaluation of bottom-up pbpk prediction success part 1: Characterisation of the orbito database of compounds. *European Journal of Pharmaceutical Sciences*, 96:598–609, January 2017.

[55] Adam S. Darwich, Alison Margolskee, Xavier Pepin, Leon Aarons, Aleksandra Galetin, Amin Rostami-Hodjegan, Sara Carlert, Maria Hammarberg, Constanze Hilgendorf, Pernilla Johansson, Eva Karlsson, Dónal Murphy, Christer Tannergren, Helena Thörn, Mohammed Yasin, Florent Mazuir, Olivier Nicolas, Sergej Ramusovic, Christine Xu, Shriram M. Pathak, Timo Korjamo, Johanna Laru, Jussi Malkki, Sari Pappinen, Johanna Tuunainen, Jennifer Dressman, Simone Hansmann, Edmund Kostewicz, Handan He, Tycho Heimbach, Fan Wu, Carolin Hoft, Yan Pang, Michael B. Bolger, Eva Huehn, Viera Lukacova, James M. Mullin, Ke X. Szeto, Chester Costales, Jian Lin, Mark McAllister, Sweta Modi, Charles Rotter, Manthena Varma, Mei Wong, Amitava Mitra, Jan Bevernage, Jeike Biewenga, Achiel Van Peer, Richard Lloyd, Carole Shardlow, Peter Langguth, Irina

Mishenzon, Mai Anh Nguyen, Jonathan Brown, Hans Lennernäs, and Bertil Abrahamsson. IMI - Oral biopharmaceutics tools project - Evaluation of bottom-up PBPK prediction success part 3: Identifying gaps in system parameters by analysing In Silico performance across different compound classes. *European Journal of Pharmaceutical Sciences*, 96:626–642, January 2017.

[56] Nicola Melillo, Adam S. Darwich, Paolo Magni, and Amin Rostami-Hodjegan. Accounting for inter-correlation between enzyme abundance: a simulation study to assess implications on global sensitivity analysis within physiologically-based pharmacokinetics. *Journal of Pharmacokinetics and Pharmacodynamics*, 46(2):137–154, April 2019.

[57] Sheila Annie Peters. Variability, Uncertainty, and Sensitivity Analysis. In *Physiologically-Based Pharmacokinetic (PBPK) Modeling and Simulations*, pages 161–181. John Wiley & Sons, Ltd, 2012.

[58] Duxin Sun, Hans Lennernäs, Lynda S. Welage, Jeffery L. Barnett, Christopher P. Landowski, David Foster, David Fleisher, Kyung-Dall Lee, and Gordon L. Amidon. Comparison of Human Duodenum and Caco-2 Gene Expression Profiles for 12,000 Gene Sequences Tags and Correlation with Permeability of 26 Drugs. *Pharmaceutical Research*, 19(10):1400–1416, October 2002.

[59] Jong Bong Lee, Atheer Zgair, Dhiaa A. Taha, Xiaowei Zang, Leonid Kagan, Tae Hwan Kim, Min Gi Kim, Hwi-yeol Yun, Peter M. Fischer, and Pavel Gershkovich. Quantitative analysis of lab-to-lab variability in Caco-2 permeability assays. *European Journal of Pharmaceutics and Biopharmaceutics*, 114:38–42, May 2017.

[60] Andrea Saltelli, Stefano Tarantola, Francesca Campolongo, and Marco Ratto. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. John Wiley & Sons, Ltd, June 2004.

[61] Emanuele Borgonovo and Elmar Plischke. Sensitivity analysis: A review of recent advances. *European Journal of Operational Research*, 248(3):869–887, February 2016.

[62] Alisa Bokulich and Naomi Oreskes. Models in Geosciences. In Lorenzo Magnani and Tommaso Bertolotti, editors, *Springer Handbook of Model-Based Science*, Springer Handbooks, pages 891–911. Springer International Publishing, Cham, 2017.

[63] Andrea Saltelli, Ksenia Aleksankina, William Becker, Pamela Fennell, Federico Ferretti, Niels Holst, Sushan Li, and Qiongli Wu. Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling & Software*, 114:29–39, April 2019.

[64] Federico Ferretti, Andrea Saltelli, and Stefano Tarantola. Trends in sensitivity analysis practice in the last decade. *Science of The Total Environment*, 568:666–670, October 2016.

[65] D. M. Hamby. A review of techniques for parameter sensitivity analysis of environmental models. *Environmental Monitoring and Assessment*, 32(2):135–154, September 1994.

[66] Andrea Saltelli and Paola Annoni. How to avoid a perfunctory sensitivity analysis. *Environmental Modelling & Software*, 25(12):1508–1517, December 2010.

[67] G. E. B. Archer, A. Saltelli, and I. M. Sobol. Sensitivity measures, anova-like Techniques and the use of bootstrap. *Journal of Statistical Computation and Simulation*, 58(2):99–120, May 1997.

[68] Andrea Saltelli. Sensitivity Analysis for Importance Assessment. *Risk Analysis*, 22(3):579–590, 2002.

[69] Ilya M. Sobol. Sensitivity Estimates for Nonlinear Mathematical Models. *Mathematical modelling and computational experiments*, 1(4):407–414, 1993.

[70] George M. Hornberger and Robert C. Spear. An Approach to the Preliminary Analysis of Environmental Systems. *Journal of Environmental Management*, 12:7–18, 1981.

[71] Ivelina Gueorguieva, Ivan A. Nestorov, Leon Aarons, and Malcolm Rowland. Uncertainty Analysis in Pharmacokinetics and Pharmacodynamics: Application to Naratriptan. *Pharmaceutical Research*, 22(10):1614–1626, October 2005.

[72] Kevin McNally, Richard Cotton, and George D. Loizou. A Workflow for Global Sensitivity Analysis of PBPK Models. *Frontiers in Pharmacology*, 2:31, 2011.

[73] X-Y Zhang, Mn Trame, Lj Lesko, and S Schmidt. Sobol Sensitivity Analysis: A Tool to Guide the Development and Evaluation of Systems Pharmacology Models. *CPT: Pharmacometrics & Systems Pharmacology*, 4(2):69–79, February 2015.

[74] Frederique Fenneteau, Jun Li, and Fahima Nekka. Assessing drug distribution in tissues expressing P-glycoprotein using physiologically based pharmacokinetic modeling: identification of important model parameters through global sensitivity analysis. *Journal of Pharmacokinetics and Pharmacodynamics*, 36(6):495, December 2009.

[75] Pankaj R. Daga, Michael B. Bolger, Ian S. Haworth, Robert D. Clark, and Eric J. Martin. Physiologically Based Pharmacokinetic Modeling in Lead Optimization. 2. Rational Bioavailability Design by Global Sensitivity Analysis To Identify Properties Affecting Bioavailability. *Molecular Pharmaceutics*, 15(3):831–839, March 2018.

[76] Nan-Hung Hsieh, Brad Reisfeld, Frederic Y. Bois, and Weihsueh A. Chiu. Applying a Global Sensitivity Analysis Workflow to Improve the Computational Efficiencies in Physiologically-Based Pharmacokinetic Modeling. *Frontiers in Pharmacology*, 9, 2018.

[77] Kevin McNally, Alex Hogg, and George Loizou. A Computational Workflow for Probabilistic Quantitative in Vitro to in Vivo Extrapolation. *Frontiers in Pharmacology*, 9, 2018.

[78] Bertrand Iooss and Paul Lemaître. A review on global sensitivity analysis methods. In C. Meloni and G. Dellino, editors, *Uncertainty management in Simulation-Optimization of Complex Systems: Algorithms and Applications*. Springer, 2015.

[79] Francesca Pianosi, Keith Beven, Jim Freer, Jim W. Hall, Jonathan Rougier, David B. Stephenson, and Thorsten Wagener. Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, 79:214–232, May 2016.

[80] Markus Joerger, Stefanie Kraff, Alwin D. R. Huitema, Gary Feiss, Berta Moritz, Jan H. M. Schellens, Jos H. Beijnen, and Ulrich Jaehde. Evaluation of a Pharmacology-Driven Dosing Algorithm of 3-Weekly Paclitaxel Using Therapeutic Drug Monitoring. *Clinical Pharmacokinetics*, 51(9):607–617, September 2012.

[81] Andrea Henrich, Markus Joerger, Stefanie Kraff, Ulrich Jaehde, Wilhelm Huisinga, Charlotte Kloft, and Zinnia Patricia Parra-Guillen. Semimechanistic Bone Marrow Exhaustion Pharmacokinetic/Pharmacodynamic Model for Chemotherapy-Induced Cumulative Neutropenia. *Journal of Pharmacology and Experimental Therapeutics*, 362(2):347–358, August 2017.

[82] Lena E. Friberg, Anja Henningsson, Hugo Maas, Laurent Nguyen, and Mats O. Karlsson. Model of Chemotherapy-Induced Myelosuppression With Parameter Consistency Across Drugs. *Journal of Clinical Oncology*, 20(24):4713–4721, December 2002.

[83] Fanny Sarrazin, Francesca Pianosi, and Thorsten Wagener. Global Sensitivity Analysis of environmental models: Convergence and validation. *Environmental Modelling & Software*, 79:135–152, May 2016.

[84] Robert Tibshirani. Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996.

[85] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*. Springer Series in Statistics. Springer-Verlag, New York, 2 edition, 2009.

[86] B. Efron and C. Stein. The Jackknife Estimate of Variance. *The Annals of Statistics*, 9(3):586–596, 1981.

[87] Andrea Saltelli. Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications*, 145(2):280–297, May 2002.

[88] S. C. Hora and R. L. Iman. Comparison of Maximus/Bounding and Bayes/Monte Carlo for fault tree uncertainty analysis. Technical Report SAND-85-2839, Hawaii Univ., Hilo (USA); Sandia National Labs., Albuquerque, NM (USA), March 1986.

[89] Karl Pearson. On the general theory of skew correlation and non-linear regression. In *Mathematical contributions to the theory of evolution*, volume XIV. Drapers' Company Research Memoirs: Dulau & Co, London, UK, 1905.

[90] Toshimitsu Homma and Andrea Saltelli. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, 52(1):1–17, April 1996.

[91] Harvey M. Wagner. Global Sensitivity Analysis. *Operations Research*, 43(6):948–969, December 1995.

[92] Ronald L. Iman and Stephen C. Hora. A Robust Measure of Uncertainty Importance for Use in Fault Tree System Analysis. *Risk Analysis*, 10(3):401–406, 1990.

[93] R. I Cukier, H. B Levine, and K. E Shuler. Nonlinear sensitivity analysis of multiparameter model systems. *Journal of Computational Physics*, 26(1):1–42, January 1978.

[94] A. Saltelli, S. Tarantola, and K. P.-S. Chan. A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. *Technometrics*, 41(1):39–56, February 1999.

[95] Jean-Yves Tissot and Clémentine Prieur. Bias correction for the estimation of sensitivity indices based on random balance designs. *Reliability Engineering & System Safety*, 107:205–213, November 2012.

[96] E. Borgonovo. A new uncertainty importance measure. *Reliability Engineering & System Safety*, 92(6):771–784, June 2007.

[97] Elmar Plischke, Emanuele Borgonovo, and Curtis L. Smith. Global sensitivity measures from given data. *European Journal of Operational Research*, 226(3):536–550, May 2013.

[98] Francesca Pianosi and Thorsten Wagener. A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. *Environmental Modelling & Software*, 67:1–11, May 2015.

[99] Dennis Fisher and Steven Shafer. Fisher/Shafer NONMEM Workshop: Pharmacokinetic and Pharmacodynamic Analysis with NONMEM, March 2007.

[100] R. L. Burkes and F. A. Shepherd. Gemcitabine in the treatment of non-small-cell lung cancer. *Annals of Oncology*, 6(suppl_3):S57–S60, January 1995.

[101] S. W. Hansen. Gemcitabine in the treatment of ovarian cancer. *International Journal of Gynecological Cancer*, 11(s1):39–41, 2001.

[102] Volker Heinemann. Gemcitabine in metastatic breast cancer. *Expert Review of Anticancer Therapy*, 5(3):429–443, June 2005.

[103] Helmut Oettle, Dirk Arnold, Christine Hempel, and Hanno Riess. The role of gemcitabine alone and in combination in the treatment of pancreatic cancer. *Anti-cancer Drugs*, 11(10):771–786, November 2000.

[104] Nha Le, Malin Sund, Alessio Vinci, George Beyer, M. Ashan Javed, Sebastian Krug, Albrecht Neessee, and Marvin Schober. Prognostic and predictive markers in pancreatic adenocarcinoma. *Digestive and Liver Disease*, 48(3):223–230, March 2016.

[105] Dietrich A. Ruess, Kivanc Görgülü, Sonja M. Wörmann, and Hana Algül. Pharmacotherapeutic Management of Pancreatic Ductal Adenocarcinoma: Current and Emerging Concepts. *Drugs & Aging*, 34(5):331–357, May 2017.

[106] Andries M Bergman, Herbert M Pinedo, and Godefridus J Peters. Determinants of resistance to 2',2'-difluorodeoxycytidine (gemcitabine). *Drug Resistance Updates*, 5(1):19–33, February 2002.

[107] John R. Mackey, Rajam S. Mani, Milada Selner, Delores Mowles, James D. Young, Judith A. Belt, Charles R. Crawford, and Carol E. Cass. Functional Nucleoside Transporters Are Required for Gemcitabine Influx and Manifestation of Toxicity in Cancer Cell Lines. *Cancer Research*, 58(19):4349–4357, October 1998.

[108] A. M. Storniolo, S. R. Allerheiligen, and H. L. Pearce. Preclinical, pharmacologic, and phase I studies of gemcitabine. *Seminars in oncology*, 24(2 Suppl 7):S7–2–S7–7, April 1997.

[109] W. Plunkett, P. Huang, Y. Z. Xu, V. Heinemann, R. Grunewald, and V. Gandhi. Gemcitabine: metabolism, mechanisms of action, and self-potentiation. *Seminars in oncology*, 22(4 Suppl 11):3–10, August 1995.

[110] Jessica A. Roseberry Baker, Enaksha R. Wickremsinhe, Claire H. Li, Olukayode A. Oluyedun, Anne H. Dantzig, Stephen D. Hall, Yuewei Qian, Barbara J. Ring, Steven A. Wrighton, and Yingying Guo.

Pharmacogenomics of Gemcitabine Metabolism: Functional Analysis of Genetic Variants in Cytidine Deaminase and Deoxycytidine Kinase. *Drug Metabolism and Disposition*, 41(3):541–545, March 2013.

[111] Lucas de Sousa Cavalcante and Gisele Monteiro. Gemcitabine: Metabolism and molecular mechanisms of action, sensitivity and chemoresistance in pancreatic cancer. *European Journal of Pharmacology*, 741:8–16, October 2014.

[112] T. Neff and C. A. Blau. Forced expression of cytidine deaminase confers resistance to cytosine arabinoside and gemcitabine. *Experimental hematology*, 24(11):1340–1346, September 1996.

[113] Elisa Giovannetti, Mario Del Tacca, Valentina Mey, Niccola Funel, Sara Nannizzi, Sergio Ricci, Cinzia Orlandini, Ugo Boggi, Daniela Campani, Marco Del Chiaro, Mauro Iannopollo, Generoso Bevilacqua, Franco Mosca, and Romano Danesi. Transcription Analysis of Human Equilibrative Nucleoside Transporter-1 Predicts Survival in Pancreas Cancer Patients Treated with Gemcitabine. *Cancer Research*, 66(7):3928–3935, April 2006.

[114] Ken Ohmine, Kei Kawaguchi, Sumio Ohtsuki, Fuyuhiko Motoi, Shinichi Egawa, Michiaki Unno, and Tetsuya Terasaki. Attenuation of Phosphorylation by Deoxycytidine Kinase is Key to Acquired Gemcitabine Resistance in a Pancreatic Cancer Cell Line: Targeted Proteomic and Metabolomic Analyses in PK9 Cells. *Pharmaceutical Research*, 29(7):2006–2016, July 2012.

[115] C. Bengala, V. Guarneri, E. Giovannetti, M. Lencioni, E. Fontana, V. Mey, A. Fontana, U. Boggi, M. Del Chiaro, R. Danesi, S. Ricci, F. Mosca, M. Del Tacca, and P. F. Conte. Prolonged fixed dose rate infusion of gemcitabine with autologous haemopoietic support in advanced pancreatic adenocarcinoma. *British Journal of Cancer*, 93(1):35–40, July 2005.

[116] Valeria Sebastiani, Francesca Ricci, Belen Rubio-Viquiera, Piotr Kulesza, Charles J. Yeo, Manuel Hidalgo, Alison Klein, Daniel Laheru, and Christine A. Iacobuzio-Donahue. Immunohistochemical and Genetic Evaluation of Deoxycytidine Kinase in Pancreatic Cancer: Relationship to Molecular Mechanisms of Gemcitabine Resistance and Survival. *Clinical Cancer Research*, 12(8):2492–2497, April 2006.

[117] Salaheldin S. Hamed, Robert M. Straubinger, and William J. Jusko. Pharmacodynamic modeling of cell cycle and apoptotic effects of gemcitabine on pancreatic adenocarcinoma cells. *Cancer Chemotherapy and Pharmacology*, 72(3):553–563, September 2013.

[118] Xin Miao, Gilbert Koch, Robert M. Straubinger, and William J. Jusko. Pharmacodynamic modeling of combined chemotherapeutic effects predicts synergistic activity of gemcitabine and trabectedin in pancreatic cancer cells. *Cancer Chemotherapy and Pharmacology*, 77(1):181–193, January 2016.

[119] Xu Zhu, Robert M. Straubinger, and William J. Jusko. Mechanism-based mathematical modeling of combined gemcitabine and birinapant in pancreatic cancer cells. *Journal of Pharmacokinetics and Pharmacodynamics*, 42(5):477–496, October 2015.

[120] Xu Zhu, Xiaomeng Shen, Jun Qu, Robert M. Straubinger, and William J. Jusko. Multi-Scale Network Model Supported by Proteomics for Analysis of Combined Gemcitabine and Birinapant Effects in Pancreatic Cancer Cells. *CPT: Pharmacometrics & Systems Pharmacology*, 7(9):549–561, 2018.

[121] Maria Garcia-Cremades, Celine Pitou, Philip W. Iversen, and Iñaki F. Troconiz. Characterizing Gemcitabine Effects Administered as Single Agent or Combined with Carboplatin in Mice Pancreatic and Ovarian Cancer Xenografts: A Semimechanistic Pharmacokinetic/Pharmacodynamics Tumor Growth-Response Model. *Jour-*

*nal of Pharmacology and Experimental Therapeutics*, 360(3):445–456, March 2017.

[122] Xu Zhu, Sheryl Trueman, Robert M. Straubinger, and William J. Jusko. Physiologically-based pharmacokinetic and pharmacodynamic models for gemcitabine and birinapant in pancreatic cancer xenografts. *Journal of Pharmacokinetics and Pharmacodynamics*, 45(5):733–746, October 2018.

[123] M. A. Battaglia Jr and R. S. Parker. Pharmacokinetic/pharmacodynamic modelling of intracellular gemcitabine triphosphate accumulation: translating in vitro to in vivo. *IET Systems Biology*, 5(1):34–43, January 2011.

[124] Maria Garcia-Cremades, Celine Pitou, Philip W. Iversen, and Iñaki F. Troconiz. Predicting tumour growth and its impact on survival in gemcitabine-treated patients with advanced pancreatic cancer. *European Journal of Pharmaceutical Sciences*, 115:296–303, March 2018.

[125] L. Kuepfer, C. Niederalt, T. Wendl, J.-F. Schlender, S. Willmann, J. Lippert, M. Block, T. Eissing, and D. Teutonico. Applied Concepts in PBPK Modeling: How to Build a PBPK/PD Model. *CPT: Pharmacometrics & Systems Pharmacology*, 5(10):516–531, 2016.

[126] Michaela Meyer, Sebastian Schneckener, Bernd Ludewig, Lars Kuepfer, and Joerg Lippert. Using expression data for quantification of active processes in physiologically-based pharmacokinetic modeling. *Drug Metabolism and Disposition*, page dmd.111.043174, January 2012.

[127] MATLAB R2019a, 2019.

[128] Christian Igel, Nikolaus Hansen, and Stefan Roth. Covariance Matrix Adaptation for Multi-objective Optimization. *Evolutionary Computation*, 15(1):1–28, March 2007.

[129] Masao Kobari, Hirotake Hisano, Seiki Matsuno, Toshio Sato, Mikio Kan, and Takehiko Tachibana. Establishment of Six Human Pancreatic Cancer Cell Lines and Their Sensitivities to Anti-Tumor Drugs. *The Tohoku Journal of Experimental Medicine*, 150(3):231–248, 1986.

[130] Takahiro Yoshida, Yoshio Endo, Tohru Obata, Yuri Kosugi, Kazuki Sakamoto, and Takuma Sasaki. Influence of Cytidine Deaminase on Antitumor Activity of 2'-Deoxycytidine Analogues in Vitro and in Vivo. *Drug Metabolism and Disposition*, page dmd.110.034397, June 2010.

[131] Anna Pisania, Gordon C. Weir, John J. O'Neil, Abdulkadir Omer, Vaja Tchipashvili, Ji Lei, Clark K. Colton, and Susan Bonner-Weir. Quantitative analysis of cell composition and purity of human pancreatic islet preparations. *Laboratory Investigation*, 90(11):1661–1675, November 2010.

[132] M. Jamei, F. Bajot, S. Neuhoff, Z. Barter, J. Yang, A. Rostami-Hodjegan, and K. Rowland-Yeo. A Mechanistic Framework for In Vitro-In Vivo Extrapolation of Liver Membrane Transporters: Prediction of Drug-Drug Interaction Between Rosuvastatin and Cyclosporine. *Clinical Pharmacokinetics*, 53(1):73–87, January 2014.

[133] Stephan A. Veltkamp, Jos H. Beijnen, and Jan H. M. Schellens. Prolonged Versus Standard Gemcitabine Infusion: Translation of Molecular Pharmacology to New Treatment Strategy. *The Oncologist*, 13(3):261–276, January 2008.

[134] E. Mini, S. Nobili, B. Caciagli, I. Landini, and T. Mazzei. Cellular pharmacology of gemcitabine. *Annals of Oncology*, 17(suppl_5):v7–v12, May 2006.

[135] Patrick Poulin, Yung-Hsiang Chen, Xiao Ding, Stephen E. Gould, Cornelis Eca Hop, Kirsten Messick, Jason Oeh, and Bianca M. Liederer. Prediction of Drug Distribution in Subcutaneous Xenografts

of Human Tumor Cell Lines and Healthy Tissues in Mouse: Application of the Tissue Composition-Based Model to Antineoplastic Drugs. *Journal of Pharmaceutical Sciences*, 104(4):1508–1521, April 2015.

[136] Liping Zhang, Vikram Sinha, S. Thomas Forgue, Sophie Callies, Lan Ni, Richard Peck, and Sandra R. B. Allerheiligen. Model-Based Drug Development: The Road to Quantitative Pharmacology. *Journal of Pharmacokinetics and Pharmacodynamics*, 33(3):369–393, June 2006.

[137] Stefan Willmann, Karsten Höhn, Andrea Edginton, Michael Sevestre, Juri Solodenko, Wolfgang Weiss, Jörg Lippert, and Walter Schmitt. Development of a Physiology-Based Whole-Body Population Model for Assessing the Influence of Individual Variability on the Pharmacokinetics of Drugs. *Journal of Pharmacokinetics and Pharmacodynamics*, 34(3):401–431, June 2007.

[138] Howard J. Burt, Arian Emami Riedmaier, Matthew D. Harwood, H. Kim Crewe, Katherine L. Gill, and Sibylle Neuhoff. Abundance of Hepatic Transporters in Caucasians: A Meta-Analysis. *Drug Metabolism and Disposition*, 44(10):1550–1561, October 2016.

[139] Ken Ohmine, Kei Kawaguchi, Sumio Ohtsuki, Fuyuhiko Motoi, Hideo Ohtsuka, Junichi Kamiie, Takaaki Abe, Michiaki Unno, and Tetsuya Terasaki. Quantitative Targeted Proteomics of Pancreatic Cancer: Deoxycytidine Kinase Protein Level Correlates to Progression-Free Survival of Patients Receiving Gemcitabine Treatment. *Molecular Pharmaceutics*, 12(9):3282–3291, September 2015.

[140] Cem Parlak, Erkan Topkan, Cem Onal, Mehmet Reyhan, and Ugur Selek. Prognostic value of gross tumor volume delineated by FDG-PET-CT based radiotherapy treatment planning in patients with locally advanced pancreatic cancer treated with chemoradiotherapy. *Radiation Oncology*, 7(1):37, March 2012.

[141] Herlinde Dumez, Gunther Guetens, Gert De Boeck, Martin S. Highley, Ernst A. de Bruijn, Allan T. van Oosterom, and Robert A. A.

Maes. *in vitro* partition of docetaxel and gemcitabine in human volunteer blood: the influence of concentration and gender. *Anti-cancer Drugs*, 16(8):885–891, September 2005.

[142] David Y. Bouffard, Josée Laliberté, and Richard L. Momparler. Kinetic studies on 2',2'-difluorodeoxycytidine (gemcitabine) with purified human deoxycytidine kinase and cytidine deaminase. *Biochemical Pharmacology*, 45(9):1857–1861, May 1993.

[143] S. M. de Lange, K. van der Born, J. R. Kroep, H. A. Jensen, P. Pfeiffer, A. Cleverly, C. J. van Groeningen, and G. J. Peters. No evidence of gemcitabine accumulation during weekly administration. *European Journal of Clinical Pharmacology*, 61(11):843–849, December 2005.

[144] H. P. Ting-Beall, D. Needham, and R. M. Hochmuth. Volume and osmotic properties of human neutrophils. *Blood*, 81(10):2774–2780, May 1993.

[145] Rosa F. Hwang, Todd Moore, Thiruvengadam Arumugam, Vijaya Ramachandran, Keith D. Amos, Armando Rivera, Baoan Ji, Douglas B. Evans, and Craig D. Logsdon. Cancer-Associated Stromal Fibroblasts Promote Pancreatic Tumor Progression. *Cancer Research*, 68(3):918–926, February 2008.

[146] Nicola Melillo, Leon Aarons, Paolo Magni, and Adam S. Darwich. Variance based global sensitivity analysis of physiologically based pharmacokinetic absorption models for bcs i-iv drugs. *Journal of Pharmacokinetics and Pharmacodynamics*, 46(1):27–42, February 2019.

[147] Gordon L. Amidon, Hans Lennernäs, Vinod P. Shah, and John R. Crison. A Theoretical Basis for a Biopharmaceutic Drug Classification: The Correlation of in Vitro Drug Product Dissolution and in Vivo Bioavailability. *Pharmaceutical Research*, 12(3):413–420, March 1995.

[148] Arik Dahan, Jonathan M. Miller, and Gordon L. Amidon. Prediction of Solubility and Permeability Class Membership: Provisional BCS Classification of the World's Top Oral Drugs. *The AAPS Journal*, 11(4):740–746, December 2009.

[149] CDER (FDA). Guidance for Industry: Waiver of In Vivo Bioavailability and Bioequivalence Studies for Immediate-release Solid Oral Dosage Forms Based on a Biopharmaceutics Classification System. Technical report, Center for Drug Evaluation and Research (CDER), U.S. Food and Drug Administration (FDA), Rockville, MD, 2000.

[150] H. Lennernäs, L. Aarons, P. Augustijns, S. Beato, M. Bolger, K. Box, M. Brewster, J. Butler, J. Dressman, R. Holm, K. Julia Frank, R. Kendall, P. Langguth, J. Sydor, A. Lindahl, M. McAllister, U. Muenster, A. Müllertz, K. Ojala, X. Pepin, C. Reppas, A. Rostami-Hodjegan, M. Verwei, W. Weitschies, C. Wilson, C. Karlsson, and B. Abrahamsson. Oral biopharmaceutics tools - Time for a new initiative - An introduction to the IMI project OrBiTo. *European Journal of Pharmaceutical Sciences*, 57:292–299, June 2014.

[151] Balaji Agoram, Walter S. Woltosz, and Michael B. Bolger. Predicting the impact of physiological and biochemical processes on oral drug bioavailability. *Advanced Drug Delivery Reviews*, 50:S41–S67, October 2001.

[152] D. Cong, M. Doherty, and K. S. Pang. A new physiologically based, segregated-flow model to explain route-dependent intestinal metabolism. *Drug Metabolism and Disposition: The Biological Fate of Chemicals*, 28(2):224–235, February 2000.

[153] A. S. Darwich, S. Neuhoff, Masoud Jamei, and Amin Rostami-Hodjegan. Interplay of Metabolism and Transport in Determining Oral Drug Absorption and Gut Wall Metabolism: A Simulation Assessment Using the 'Advanced Dissolution, Absorption, Metabolism (ADAM)' Model. *Current Drug Metabolism*, 11(9):716–729, October 2010.

[154] Michael Gertz, J. Brian Houston, and Aleksandra Galetin. Physiologically Based Pharmacokinetic Modeling of Intestinal First-Pass Metabolism of CYP3a Substrates with High Intestinal Extraction. *Drug Metabolism and Disposition*, 39(9):1633–1642, September 2011.

[155] Masoud Jamei, David Turner, Jiansong Yang, Sibylle Neuhoff, Sebastian Polak, Amin Rostami-Hodjegan, and Geoffrey Tucker. Population-Based Mechanistic Prediction of Oral Drug Absorption. *The AAPS Journal*, 11(2):225–237, June 2009.

[156] Stefan Willmann, Walter Schmitt, Jörg Keldenich, Jörg Lippert, and Jennifer B. Dressman. A Physiological Model for the Estimation of the Fraction Dose Absorbed in Humans. *Journal of Medicinal Chemistry*, 47(16):4022–4031, July 2004.

[157] Hans Lennernäs. Human in Vivo Regional Intestinal Permeability: Importance for Pharmaceutical Drug Development. *Molecular Pharmaceutics*, 11(1):12–23, January 2014.

[158] J. B. Dressman and D. Fleisher. Mixing-Tank Model for Predicting Dissolution Rate Control of Oral Absorption. *Journal of Pharmaceutical Sciences*, 75(2):109–116, February 1986.

[159] Lawrence X. Yu and Gordon L. Amidon. A compartmental absorption and transit model for estimating oral drug absorption. *International Journal of Pharmaceutics*, 186(2):119–125, September 1999.

[160] CERTARA L.P. Simcyp Simulator - Version 17, April 2017.

[161] Xavier J. H. Pepin, Talia R. Flanagan, David J. Holt, Anna Eidelman, Don Treacy, and Colin E. Rowlings. Justification of Drug Product Dissolution Rate and Drug Substance Particle Size Specifications Based on Absorption PBPK Modeling for Lesinurad Immediate Release Tablets. *Molecular Pharmaceutics*, 13(9):3256–3269, July 2016.

[162] Rebecca J. Hintz and Kevin C. Johnson. The effect of particle size distribution on dissolution rate and oral absorption. *International Journal of Pharmaceutics*, 51(1):9–17, April 1989.

[163] Kiyohiko Sugano. Theoretical comparison of hydrodynamic diffusion layer models used for dissolution simulation in drug discovery and development. *International Journal of Pharmaceutics*, 363(1):73–77, November 2008.

[164] H. Z. Bu. A Literature Review of Enzyme Kinetic Parameters for CYP3a4-Mediated Metabolic Reactions of 113 Drugs in Human Liver Microsomes: Structure- Kinetics Relationship Assessment. *Current Drug Metabolism*, 7(3):231–249, April 2006.

[165] David T. Manallack. The pKa Distribution of Drugs: Application to Drug Discovery. *Perspectives in Medicinal Chemistry*, 1, January 2007.

[166] J. Valetin. Human Alimentary Tract Model for Radiological Protection. Technical Report 100, International Commission on Radiological Protection (ICRP), 2006.

[167] Simulations Plus. Gastroplus - Version 9.5, 2017.

[168] W. S. Snyder, M. J. Cook, E. S. Nasset, L. R. Karhausen, G. Parry Howells, and I. H. Tipton. Report of the Task Group on Reference Man. Technical Report 23, International Commission on Radiological Protection (ICRP), 1975.

[169] Andrés Olivares-Morales, Avijit Ghosh, Leon Aarons, and Amin Rostami-Hodjegan. Development of a Novel Simplified PBPK Absorption Model to Explain the Higher Relative Bioavailability of the OROS® Formulation of Oxybutynin. *The AAPS Journal*, 18(6):1532–1549, November 2016.

[170] J. Valetin. Basic Anatomical and Physiological Data for Use in Radiological Protection: Reference Values. Technical Report 89, International Commission on Radiological Protection (ICRP), 2002.

[171] Ronald P. Brown, Michael D. Delp, Stan L. Lindstedt, Lorenz R. Rhomberg, and Robert P. Beliles. Physiological Parameter Values for Physiologically Based Pharmacokinetic Models. *Toxicology and Industrial Health*, 13(4):407–484, July 1997.

[172] H. Lennernäs, K. Palm, U. Fagerholm, and P. Artursson. Comparison between active and passive drug transport in human intestinal epithelial (caco-2) cells in vitro and human jejunum in vivo. *International Journal of Pharmaceutics*, 127(1):103–107, January 1996.

[173] Megerle L. Scherholz, James Forder, and Ioannis P. Androulakis. A framework for 2-stage global sensitivity analysis of GastroPlus™ compartmental models. *Journal of Pharmacokinetics and Pharmacodynamics*, 45(2):309–327, April 2018.

[174] CDER (FDA). Physiologically Based Pharmacokinetic analyses - format and content: guidance for industry, draft. Technical report, Center for Drug Evaluation and Research (CDER), U.S. Food and Drug Administration (FDA), Rockville, MD, December 2016.

[175] Ben Forbes, Bahman Asgharian, Lea Ann Dailey, Douglas Ferguson, Per Gerde, Mark Gumbleton, Lena Gustavsson, Colin Hardy, David Hassall, Rhys Jones, Ruth Lock, Janet Maas, Tim McGovern, Gary R. Pitcairn, Graham Somers, and Ron K. Wolff. Challenges in inhaled product development and opportunities for open innovation. *Advanced Drug Delivery Reviews*, 63(1):69–87, January 2011.

[176] N. R. Labiris and M. B. Dolovich. Pulmonary drug delivery. Part I: Physiological factors affecting therapeutic effectiveness of aerosolized medications. *British Journal of Clinical Pharmacology*, 56(6):588–599, 2003.

[177] Peter Strong, Kazuhiro Ito, John Murray, and Garth Rapeport. Current approaches to the discovery of novel inhaled medicines. *Drug Discovery Today*, 23(10):1705–1717, October 2018.

[178] Silvia Grandoni, Nicola Cesari, Nicola Melillo, Giandomenico Brogin, Paola Puccini, and Paolo Magni. Development and evaluation of a PBPK model to study the pharmacokinetics of inhaled drugs in rats. In *PAGE 28, Submitted Abstract*, Stockholm, Sweden., 2019.

[179] E. Boger, N. Evans, M. Chappell, A. Lundqvist, P. Ewing, A. Wigenborg, and M. Fridén. Systems Pharmacology Approach for Prediction of Pulmonary and Systemic Pharmacokinetics and Receptor Occupancy of Inhaled Drugs. *CPT: Pharmacometrics & Systems Pharmacology*, 5(4):201–210, 2016.

[180] Jens Markus Borghardt, Benjamin Weber, Alexander Staab, and Charlotte Kloft. Pharmacometric Models for Characterizing the Pharmacokinetics of Orally Inhaled Drugs. *The AAPS Journal*, 17(4):853–870, July 2015.

[181] Silvia Grandoni, Nicola Cesari, Giandomenico Brogin, Paola Puccini, and Paolo Magni. Building in-house PBPK modelling tools for oral drug administration from literature information. *ADMET and DMPK*, 7(1):4–21, February 2019.

[182] Jayne E. Hastedt, Per Bäckman, Andrew R. Clark, William Doub, Anthony Hickey, Guenther Hochhaus, Phil J. Kuehl, Claus-Michael Lehr, Peter Mauser, Jason McConville, Ralph Niven, Masahiro Sakagimi, and Jeffry G. Weers. Scope and relevance of a pulmonary biopharmaceutical classification system AAPS/FDA/USP Workshop March 16-17th, 2015 in Baltimore, MD. *AAPS Open*, 2(1):1, January 2016.

[183] Mireille Hassoun, Paul G. Royall, Mark Parry, Richard D. Harvey, and Ben Forbes. Design and development of a biorelevant simulated

human lung fluid. *Journal of Drug Delivery Science and Technology*, 47:485–491, October 2018.

[184] Mariarita Stefani, Nicola Cesari, Roberto Corsaletti, Alessandro Fioni, Francesca Saccani, Roberta Volta, Giandomenico Brogin, and Paola Puccini. Development of an in vitro solubility test as a tool for predicting lung retention of poorly water soluble compounds. In *The Aerosol Society*, Edinburgh, 2018.

[185] Alessandro Fioni, Ewa Selg, Valentina Cenacchi, Fernando Acevedo, Giandomenico Brogin, Per Gerde, and Paola Puccini. Investigation of Lung Pharmacokinetic of the Novel PDE4 Inhibitor CHF6001 in Preclinical Models: Evaluation of the PreciseInhale Technology. *Journal of Aerosol Medicine and Pulmonary Drug Delivery*, 31(1):61–70, August 2017.

[186] Ryosei Kawai, Diane Mathew, Chiaki Tanaka, and Malcolm Rowland. Physiologically Based Pharmacokinetics of Cyclosporine A: Extension to Tissue Distribution Kinetics in Rats and Scale-up to Human. *Journal of Pharmacology and Experimental Therapeutics*, 287(2):457–468, November 1998.

[187] Werner Hofmann and Bahman Asgharian. The Effect of Lung Structure on Mucociliary Clearance and Particle Retention in Human and Rat Lungs. *Toxicological Sciences*, 73(2):448–456, June 2003.

[188] Alonso-Galicia Magdalena, Brands Michael W., Zappe Dion H., and Hall John E. Hypertension in Obese Zucker Rats. *Hypertension*, 28(6):1047–1054, December 1996.

[189] David A. Edwards, Justin Hanes, Giovanni Caponetti, Jeffrey Hrkach, Abdelaziz Ben-Jebria, Mary Lou Eskew, Jeffrey Mintzes, Daniel Deaver, Noah Lotan, and Robert Langer. Large Porous Particles for Pulmonary Drug Delivery. *Science*, 276(5320):1868–1872, June 1997.

[190] P. Gehr. Annexe A. Anatomy and morphology of the respiratory tract. *Annals of the ICRP*, 24(1-3):121–166, September 1994.

[191] R R Mercer, M L Russell, V L Roggli, and J D Crapo. Cell number and distribution in human and rat airways. *American Journal of Respiratory Cell and Molecular Biology*, 10(6):613–624, June 1994.

[192] Yi-Chu Nie, Hao Wu, Pei-Bo Li, Yu-Long Luo, Kang Long, Li-Ming Xie, Jian-Gang Shen, and Wei-Wei Su. Anti-Inflammatory Effects of Naringin in Chronic Pulmonary Neutrophilic Inflammation in Cigarette Smoke-Exposed Rats. *Journal of Medicinal Food*, 15(10):894–900, September 2012.

[193] Richard A. Parent. *Treatise on Pulmonary Toxicology: Comparative Biology of the Normal Lung*. CRC Press, February 1992.

[194] Kent E. Pinkerton, Peter Gehr, Alejandro Castañeda, and James D. Crapo. Architecture and Cellular Composition of the Air-Blood Tissue Barrier. In *Comparative Biology of the Normal Lung: Second Edition*, pages 105–117. Elsevier Inc., March 2015.

[195] Charles G. Plopper, Lila H. Hill, and Andrew T. Mariassy. Ultrastructure of the Nonciliated Bronchiolar Epithelial (Clara) Cell of Mammalian Lung. III. A Study of Man with Comparison of 15 Mammalian Species. *Experimental Lung Research*, 1(2):171–180, January 1980.

[196] Susan D. Reynolds, Kent E. Pinkerton, and Andrew T. Mariassy. Epithelial Cells of Trachea and Bronchi. In *Comparative Biology of the Normal Lung: Second Edition*, pages 61–81. Elsevier Inc., March 2015.

[197] Kevin McNally, Richard Cotton, Alex Hogg, and George Loizou. Reprint of PopGen: A virtual human population generator. *Toxicology*, 332(Supplement C):77–93, June 2015.

[198] Nikolaos Tsamandouras, Thierry Wendling, Amin Rostami-Hodjegan, Aleksandra Galetin, and Leon Aarons. Incorporation of stochastic variability in mechanistic population pharmacokinetic models: handling the physiological constraints using normal transformations. *Journal of Pharmacokinetics and Pharmacodynamics*, 42(4):349–373, August 2015.

[199] Masoud Jamei, Gemma L. Dickinson, and Amin Rostami-Hodjegan. 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 Metabolism and Pharmacokinetics*, 24(1):53–75, January 2009.

[200] Hajar Al Feteisi, Brahim Achour, Amin Rostami-Hodjegan, and Jill Barber. Translational value of liquid chromatography coupled with tandem mass spectrometry-based quantitative proteomics for in vitro-in vivo extrapolation of drug metabolism and transport and considerations in selecting appropriate techniques. *Expert Opinion on Drug Metabolism & Toxicology*, 11(9):1357–1369, September 2015.

[201] Doki Kosuke, Darwich Adam S., Achour Brahim, Tornio Aleksi, Backman Janne T., and Rostami-Hodjegan Amin. Implications of intercorrelation between hepatic CYP3a4-CYP2c8 enzymes for the evaluation of drug-drug interactions: a case study with repaglinide. *British Journal of Clinical Pharmacology*, 84(5):972–986, March 2018.

[202] Brahim Achour, Matthew R. Russell, Jill Barber, and Amin Rostami-Hodjegan. Simultaneous Quantification of the Abundance of Several Cytochrome P450 and Uridine 5'-Diphospho-Glucuronosyltransferase Enzymes in Human Liver Microsomes Using Multiplexed Targeted Proteomics. *Drug Metabolism and Disposition*, 42(4):500–510, April 2014.

[203] Oliver von Richter, Oliver Burk, Martin F. Fromm, Klaus P. Thon, Michel Eichelbaum, and Kari T. Kivistö. Cytochrome P450 3a4 and P-glycoprotein Expression in Human Small Intestinal Enterocytes and Hepatocytes: A Comparative Analysis in Paired Tissue Specimens. *Clinical Pharmacology & Therapeutics*, 75(3):172–183, March 2004.

[204] Marek Drozdzik, Diana Busch, Joanna Lapczuk, Janett Müller, Marek Ostrowski, Mateusz Kurzawski, and Stefan Oswald. Protein Abundance of Clinically Relevant Drug-Metabolizing Enzymes in the Human Liver and Intestine: A Comparative Analysis in Paired Tissue Specimens. *Clinical Pharmacology & Therapeutics*, 0(0), December 2017.

[205] CHMP (EMA). Guideline on the qualification and reporting of physiologically based pharmacokinetic (PBPK) modelling and simulation - draft. Technical Report EMA/CHMP/458101/2016, Committee for Medicinal Products for Human Use (CHMP), European Medicines Agency (EMA), London, UK, July 2016.

[206] Rhys Do Jones, Hannah M. Jones, Malcolm Rowland, Christopher R. Gibson, James W. T. Yates, Jenny Y. Chien, Barbara J. Ring, Kimberly K. Adkison, M. Sherry Ku, Handan He, Ragini Vuppugalla, Punit Marathe, Volker Fischer, Sandeep Dutta, Vikash K. Sinha, Thorir Björnsson, Thierry Lavó, and Patrick Poulin. PhRMA CPCDC initiative on predictive models of human pharmacokinetics, part 2: Comparative assessment of prediction methods of human volume of distribution. *Journal of Pharmaceutical Sciences*, 100(10):4074–4089, October 2011.

[207] Lauri I. Kajosaari, Jouko Laitila, Pertti J. Neuvonen, and Janne T. Backman. Metabolism of Repaglinide by CYP2c8 and CYP3a4 in vitro: Effect of Fibrates and Rifampicin. *Basic & Clinical Pharmacology & Toxicology*, 97(4):249–256, 2005.

[208] Michael Gertz, Nikolaos Tsamandouras, Carolina Säll, J. Brian Houston, and Aleksandra Galetin. Reduced Physiologically-Based Pharmacokinetic Model of Repaglinide: Impact of OATP1b1 and CYP2c8 Genotype and Source of In Vitro Data on the Prediction of Drug-Drug Interaction Risk. *Pharmaceutical Research*, 31(9):2367–2382, September 2014.

[209] Karelle Ménochet, Kathryn E. Kenworthy, J. Brian Houston, and Aleksandra Galetin. Simultaneous Assessment of Uptake and Metabolism in Rat Hepatocytes: A Comprehensive Mechanistic Model. *Journal of Pharmacology and Experimental Therapeutics*, 341(1):2–15, April 2012.

[210] Masoud Jamei, Steve Marciniak, Duncan Edwards, Kris Wragg, Kairui Feng, Adrian Barnett, and Amin Rostami-Hodjegan. The Simcyp Population Based Simulator: Architecture, Implementation, and Quality Assurance. *In Silico Pharmacology*, 1(1):9, December 2013.

[211] Yoshihisa Shitara, Kazuya Maeda, Kazuaki Ikejiri, Kenta Yoshida, Toshiharu Horie, and Yuichi Sugiyama. Clinical significance of organic anion transporting polypeptides (OATPs) in drug disposition: their roles in hepatic clearance and intestinal absorption. *Biopharmaceutics & Drug Disposition*, 34(1):45–78, January 2013.

[212] Narciso Couto, Zubida M. Al-Majdoub, Brahim Achour, Phillip C. Wright, Amin Rostami-Hodjegan, and Jill Barber. Quantification of Proteins Involved in Drug Metabolism and Disposition in the Human Liver Using Label-Free Global Proteomics. *Molecular Pharmaceutics*, 16(2):632–647, February 2019.

[213] David J. Greenblatt, Lisa L. von Moltke, Jerold S. Harmatz, Gengsheng Chen, James L. Weemhoff, Cheng Jen, Charles J. Kelley, Barbara W. LeDuc, and Miguel A. Zinny. Time course of recovery of cytochrome p450 3a function after single doses of grapefruit juice. *Clinical Pharmacology & Therapeutics*, 74(2):121–129, 2003.

[214] Sebastien Da Veiga, Francois Wahl, and Fabrice Gamboa. Local Polynomial Estimation for Sensitivity Analysis on Models With Correlated Inputs. *Technometrics*, 51(4):452–463, November 2009.

[215] Genyuan Li, Herschel Rabitz, Paul E. Yelvington, Oluwayemisi O. Oluwole, Fred Bacon, Charles E. Kolb, and Jacqueline Schoendorf. Global Sensitivity Analysis for Systems with Independent and/or Correlated Inputs. *The Journal of Physical Chemistry A*, 114(19):6022–6032, May 2010.

[216] Andrea Saltelli and Stefano Tarantola. On the Relative Importance of Input Factors in Mathematical Models: Safety Assessment for Nuclear Waste Disposal. *Journal of the American Statistical Association*, 97(459):702–709, 2002.

[217] C. Xu and G. Gertner. Extending a global sensitivity analysis technique to models with correlated parameters. *Computational Statistics & Data Analysis*, 51(12):5579–5590, August 2007.

[218] Andrea Saltelli, Ângela Guimaraes Pereira, Jeroen P. Van der Sluijs, and Silvio Funtowicz. What do I make of your latinorum? Sensitivity auditing of mathematical modelling. *International Journal of Foresight and Innovation Policy*, 9(2/3/4):213, 2013.

[219] F. Pappenberger and K. J. Beven. Ignorance is bliss: Or seven reasons not to use uncertainty analysis. *Water Resources Research*, 42(5), 2006.

[220] Axel Heinemann, Friedel Wischhusen, Klaus Püschel, and Xavier Rogiers. Standard liver volume in the caucasian population. *Liver Transplantation and Surgery*, 5(5):366–368, 1999.

[221] Satish Anjilvel and Bahman Asgharian. A Multiple-Path Model of Particle Deposition in the Rat Lung. *Fundamental and Applied Toxicology*, 28(1):41–50, November 1995.

[222] National Institute for Public Health and the Environment (RIVM). Multiple Path Particle Dosimetry Model (MPPD v 1.0): A Model for Human and Rat Airway Particle Dosimetry. Bilthoven, The Netherlands. Technical report, RIVA Report 650010030., Bilthoven, The Netherlands., 2002.

[223] Per Gerde, Pär Ewing, Lena Låstbom, Åke Ryrfeldt, Juri Waher, and Göran Lidén. A Novel Method to Aerosolize Powder for Short Inhalation Exposures at High Concentrations: Isolated Rat Lungs Exposed to Respirable Diesel Soot. *Inhalation Toxicology*, 16(1):45–52, January 2004.

[224] Claudia Meindl, Sandra Stranzinger, Neira Dzidic, Sharareh Salar-Behzadi, Stefan Mohr, Andreas Zimmer, and Eleonore Fröhlich. Permeation of Therapeutic Drugs in Different Formulations across the Airway Epithelium In Vitro. *PLOS ONE*, 10(8):e0135690, August 2015.

[225] Cynthia Bosquillon, Michaela Madlova, Nilesh Patel, Nicola Clear, and Ben Forbes. A Comparison of Drug Transport in Pulmonary Absorption Models: Isolated Perfused rat Lungs, Respiratory Epithelial Cell Lines and Primary Cell Culture. *Pharmaceutical Research*, 34(12):2532–2540, December 2017.

[226] Neil R. Mathias, Julita Timoszyk, Paul I. Stetsko, John R. Megill, Ronald L. Smith, and Doris A. Wall. Permeability Characteristics of Calu-3 Human Bronchial Epithelial Cells: In Vitro - In Vivo Correlation to Predict Lung Absorption in Rats. *Journal of Drug Targeting*, 10(1):31–40, January 2002.

# List of publications

## Articles in peer reviewed journals

- **N. Melillo**, L. Aarons, P. Magni, A.S. Darwich. Variance based global sensitivity analysis of physiologically based pharmacokinetic absorption models for BCS I-IV drugs. *Journal of Pharmacokinetics and Pharmacodynamics*, 46(1):27-42, February 2019. `https://doi.org/10.1007/s10928-018-9615-8`.

- **N. Melillo**, A.S. Darwich, P. Magni, A. Rostami-Hodjegan. Accounting for inter-correlation between enzyme abundance: a simulation study to assess implications on global sensitivity analysis within physiologically-based pharmacokinetics. *Journal of Pharmacokinetics and Pharmacodynamics*, 46(2):137-154, April 2019. `https://doi.org/10.1007/s10928-019-09627-6`.

- M. García-Cremades, **N. Melillo**, I.F. Tróconiz, P. Magni. Mechanistic multi-scale pharmacokinetic model for the anticancer drug gemcitabine in pancreatic cancer. *Clinical and Translational Science*, 2020. `https://doi.org/10.1111/cts.12747`.

## Manuscript in preparation

- **N. Melillo**, S. Grandoni, N. Cesari, G. Brogin, P. Puccini, P. Magni. Global sensitivity analysis of a physiological model for pulmonary absorption of inhaled compounds. *Manuscript in preparation.*

## Contributed talks

- **N. Melillo**, L. Aarons, P. Magni, A.S. Darwich. Variance based Global Sensitivity Analysis of a Mechanistic Physiological Absorption model for BCS I-IV compounds. *Abstract for contributed talk, 2nd Exchange workshop, UK Quantitative System Pharmacology network*, Reading (UK), July 2-4, 2018

- **N. Melillo**, L. Aarons, P. Magni, A.S. Darwich. Global sensitivity analysis of a physiological absorption model for BCS I-IV compounds. *Abstract for a contributed talk, PKUK 2018, Manchester (UK)*

## Poster presentations

### 2019

- **N. Melillo**, S. Grandoni, N. Cesari, G. Brogin, P. Puccini, P. Magni. Variance based global sensitivity analysis of a pharmacokinetic mechanistic model for inhaled compounds in rats. *Ninth International Conference on Sensitivity Analysis of Model Output, SAMO (2019)*

- **N. Melillo**, S. Grandoni, N. Cesari, G. Brogin, P. Puccini, P. Magni. Global sensitivity analysis of a physiologically based pulmonary absorption model. *PAGE 28 (2019), Abstr 9072* [`www.page-meeting.org/?abstract=9072`]

- S. Grandoni, N. Cesari, **N. Melillo**, G. Brogin, P. Puccini, P. Magni. Development and evaluation of a PBPK model to study the pahrma-

cokinetics of inhaled drugs in rats. *PAGE 28 (2019), Abstr 9047* [www.page-meeting.org/?abstract=9047]

- R. Bartolucci, S. Grandoni, **N. Melillo**, G. Nicora, E. Sauta, E. M. Tosca, P. Magni. Artificial intelligence and machine learning: just a hype or a new opportunity for pharmacometrics? *PAGE 28 (2019), Abstr 9148* [www.page-meeting.org/?abstract=9148]

**2018**

- **N. Melillo**, L. Aarons, P. Magni and A.S. Darwich. Variance based Global Sensitivity Analysis of a Physiological model for compounds in different BCS classes. *PAGE 27 (2018) Abstr 8559* [www.page-meeting.org/?abstract=8559]

- **N. Melillo**, L. Aarons, P. Magni and A.S. Darwich. Variance based Global Sensitivity Analysis on Models describing Drug Absorption following Oral Administration. *GNB 2018, June 25th-27th 2018, Milan, Italy*

- **N. Melillo**, A.S. Darwich, P. Magni and A. Rostami-Hodjegan. Impact of enzyme inter-correlation on global sensitivity analysis of PBPK models: a case study with repaglinide. *Abstract for poster presentation, 9th International Marbach Castle DDI Workshop, May 27-29, 2018*

- R. Bartolucci, **N. Melillo**, P. Magni. Whole-body PBPK model for a translational study of Azathioprine pharmacokinetics. *Proceedings of ACOP9 meeting, 07-10/10/2018, San Diego, California, USA*

- R. Bartolucci, S.M. Lavezzi, E.M. Tosca, **N. Melillo**, S. Grandoni, E. Borella, L. Pasotti, G. De Nicolao, P. Magni. Evaluation of software tools for Bayesian estimation on population models: an update based on current software versions. *PAGE 27 (2018) Abstr 8690* [www.page-meeting.org/?abstract=8690]

**2017**

- **N. Melillo**, L. Pasotti and P. Magni. Multiscale mechanistic models in Systems Pharmacology: development of a model describing Atorvastatin pharmacokinetics through integration of metabolic network in Physiologically Based Pharmacokinetic models. *PAGE 26 (2017) Abstr 7147* [www.page-meeting.org/?abstract=7147]

- M. García-Cremades, **N. Melillo**, P. Magni, I.F. Tróconiz. Mechanistic multi-scale systems pharmacokinetics model applied for the anticancer drug gemcitabine in pancreatic cancer. *PAGE 26 (2017) Abstr 7368* [www.page-meeting.org/?abstract=7368]

# Acknowledgements

I sincerely believe that my PhD studentship was a beautiful period. So, I would like to acknowledge:

My supervisor Professor Paolo Magni, from "blackboard questions" during the statistics course to nowadays, for his guidance, inspiration, support and humour.

Dr Adam Darwich, Professor Leon Aarons and Professor Amin Rostami-Hodjegan of the University of Manchester, for making my abroad period an unique experience of personal and professional growth.

Dr María García-Cremades and Professor Iñaki Tróconiz of the Universidad de Navarra, for our fruitful collaboration and for all the nice moments in Pavia and Pamplona.

Nicola Cesari and Giandomenico Brogin of Chiesi Farmaceutici for the work together and the useful and professional discussions.

All my friends of the BMSlab at the Università di Pavia, for the nice and fun lunches, lab and group meetings, discussions, trips to conferences and for making our inhospitable and dark lab a "warm" place.

All my friends of the CAPKR at the University of Manchester, for all the nice moments, like the karaoke nights, Friday pub and trips through

the UK.

My parents, family and friends for everything else.