

# **Application of Computational Modelling towards Digital Twinning for Human Physiological and Biochemical Processes**

*Thesis submitted for award of the degree*

*of*

**Doctor of Philosophy**

*by*

**Manoja Rajalakshmi Aravindakshan**

under the guidance of

**Dr Chittaranjan Mandal**



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING  
INDIAN INSTITUTE OF TECHNOLOGY KHARAGPUR  
DECEMBER 2024**

© 2024 Manoja Rajalakshmi Aravindakshan. All rights reserved.



*To*

*My Parents, Sister & Partner*





# APPROVAL OF THE VIVA-VOCE BOARD

Date: \_\_\_\_\_

Certified that the thesis entitled “**Application of Computational Modelling towards Digital Twinning for Human Physiological and Biochemical Processes**” submitted by **Manoja Rajalakshmi Aravindakshan** to the Indian Institute of Technology, Kharagpur, for the award of the degree Doctor of Philosophy has been accepted by the external examiners and that the student has successfully defended the thesis in the viva-voce examination held today.

\_\_\_\_\_  
(Member of the DSC)

\_\_\_\_\_  
(Member of the DSC)

\_\_\_\_\_  
(Member of the DSC)

\_\_\_\_\_  
(Supervisor)

\_\_\_\_\_  
(External Examiner)

\_\_\_\_\_  
(Chairman of the DSC)



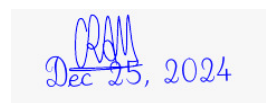
# Certificate

This is to certify that the thesis entitled “**Application of Computational Modelling towards Digital Twinning for Human Physiological and Biochemical Processes**” submitted by **Manoja Rajalakshmi Aravindakshan** to the Indian Institute of Technology Kharagpur, is a record of bonafide research work carried under my supervision and is worthy of consideration for the award of the Doctor of Philosophy of the Institute.

To the best of my knowledge, the results embodied in this thesis have not been submitted to any other University or Institute for the award of any other Degree or Diploma.

Date: Dec 25, 2024

Place: IIT Kharagpur



Dr Chittaranjan Mandal

Professor

Dept of Computer Science & Engg



# Declaration

I certify that

1. the work contained in the thesis is original and has been done by me under the guidance of my Supervisor;
2. the work has not been submitted to any other Institute for any degree or diploma;
3. I have followed the guidelines provided by the Institute in preparing the thesis;
4. I have conformed to ethical norms and guidelines while writing the thesis and;
5. whenever I have used material (data, models, figures and text) from other sources, I have given due credit to them by citing them in the text of the thesis, giving their details in the references and taking permission from the copyright owners of the sources, whenever necessary.

Date: Dec 24, 2024

Place: IIT Kharagpur



---

Manoja Rajalakshmi Aravindakshan



# Acknowledgements

The journey of my PhD has been a transformative experience, and I have been fortunate to be supported by many individuals who have played a crucial role in shaping this work. My academic path, from its beginnings at IIT Kharagpur to its near completion at Purdue University, has been made possible by their constant encouragement and guidance.

First, I would like to express my deepest gratitude to my advisor, *Chittaranjan Mandal*, for his invaluable guidance and mentorship throughout the course of my research. His expertise and constant encouragement have been a steady source of inspiration for me. I also extend my thanks to my doctoral thesis committee members - *Jayanta Mukhopadhyay*, *Santanu Chattopadhyay*, *Pabitra Mitra* and *Pralay Mitra* - for their time, engagement and insightful feedback. I am deeply thankful to all of my collaborators who have contributed to the success of this work. My passion for organ-on-chip technology was nurtured by the guidance of *Christian Maass*, a valued industry collaborator, while we worked on modelling the drug distribution in liver-on-chip systems. I had the privilege of working with *K V Venkatesh* on modelling the adult metabolism model and it greatly expanded my knowledge on physiological models. My work on the insulin-glucose regulatory system owes much to the collaboration with *Partha Chakrabarti* and *Jit Sarkar*, who provided data and clinical expertise for model interpretation.

I express my sincere appreciation to *Alex Pothén*, my adviser at Purdue University. His guidance over the past year and a half has been essential in shaping my progress as a researcher, particularly with his expertise in mathematics and computational modelling. I express my gratitude to *Pavlos Vlachos* and *Brett Meyers* for introducing me to the concept of lumped parameter modelling of the cardiovascular system. Working with them has greatly enriched my professional and scientific growth.

A special mention goes to my senior *Devleena*, who has been a constant source of support, from helping me navigate modelling techniques to offering guidance whenever I faced challenges. I also extend my thanks to my colleagues *Shankhadip*, *Amaljith* and *Nabin*, for their positivity and support. I am grateful to Prof. Pothén's lab members for fostering a collaborative environment and easing my transition into a new research environment.

I would like to thank all my friends and acquaintances at IIT Kharagpur and Purdue

University for their companionship and encouragement throughout this journey. I owe much to my wonderful roommates *Munmun*, *Srishti*, *Priya*, *Vaishnavi* and *Tanu*, who have been a constant source of support and understanding. I would also like to thank my dear friends *Renin*, *Krutarth*, *Soumya* and *Aswathi*, whose unwavering belief in me and their positive outlook have kept me going during tough times. I am deeply honoured to have received the SERB-OVDF 2022 fellowship, which has greatly enriched my research, introducing me to new opportunities and experiences at Purdue University. The sense of community and support at Purdue has made all the difference in my journey.

I am forever grateful to *Adharsh*, my partner, for standing by my side through every high and low and offering me the emotional support and care I needed. My heartfelt thanks go to my mother, whose unwavering love, blessings and support have been a constant source of strength, allowing me to grow and thrive. I am also profoundly grateful to my family — my sister, who has been a constant pillar of support; my niece, little Saanvi, for all the love; and to all who have supported me along the way. Finally, I acknowledge myself for persevering through the challenges, staying resilient and never giving up.



# Abstract

Mathematical models are valuable tools used to understand and predict the behaviour of complex physiological systems. In this thesis, ordinary differential equation (ODE) modelling is applied to the following two areas: (i) the development of digital twins for human physiological systems, specifically modelling drug metabolism in *in vitro* liver systems and simulating cardiovascular blood flow using lumped parameter models (LPMs); and (ii) modelling of biochemical processes, focusing on insulin-glucose (IG) regulatory systems in the context of type 2 diabetes mellitus (T2DM).

The drug testing process faces challenges such as high costs, lengthy timelines, and poor predictability of human responses, especially in estimating first-in-human doses. While advancements in microphysiological systems (MPS) and organ-on-chip (OoC) technologies offer more human-relevant testing, they still rely on conventional mathematical models that fail to capture biological processes adequately. This study aims to enhance the predictive power of MPS and OoC by integrating them into a digital twin framework (DigiLoCs) for better prediction of liver clearance. A compartmental physiological model of the liver using ODEs is developed to estimate pharmacokinetic parameters for *in vitro* liver-on-chip systems. Digital twinning is also applied to model the blood flow in the human circulatory system using zero-dimensional LPMs. LPMs approximate the cardiovascular network as a set of discrete compartments accounting for vascular pressure of blood flow, impedance to the flow of blood, the volume of blood in the vessels and the elasticity of the vessels. By applying electrical circuit principles, these models predict cardiovascular dynamics. The simulations facilitate the exploration of scenarios that are difficult to observe experimentally, such as blood flow in internal organs, making them useful for diagnosing and monitoring disease progression.

The thesis then explores biochemical processes starting with clustering-based methods to study T2DM. Using an unsupervised clustering algorithm three subgroups of uncontrolled T2DM patients were identified and characterised based on patho-clinical features. The significance of heterogeneity in T2DM is uncovered, challenging the assumption of its homogeneity and emphasising the need to reconsider uniform treatment protocols. Given obesity as a significant risk factor in T2DM, a distinction between estimated oral minimal model (OMM) parameters for obese and non-obese T2DM subjects is identified. Furthermore, the existing OMM is augmented to include the role of body mass index (BMI) and leptin (a hormone secreted by adipose tissue), leading to better simulation results.

**Keywords:** *Mathematical modelling, Parameter estimation, Organ-on-chip, Lumped parameter model, Cardiovascular modelling, Oral minimal model, Type-2 diabetes mellitus*



# Contents

<b>List of Abbreviations</b>	<b>xix</b>
<b>List of Figures</b>	<b>xxi</b>
<b>List of Tables</b>	<b>xxv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Introduction on topics explored in this thesis . . . . .	2
1.2 Contributions from this research . . . . .	4
1.2.1 Prediction of <i>in vivo</i> drug clearance in liver-on-a-chip device using PBPK modelling . . . . .	4
1.2.2 Lumped parameter modelling of cardiovascular system . . . . .	5
1.2.3 Clustering based methods to study T2DM sub groups . . . . .	6
1.2.4 Oral minimal model estimation of parameters for obese and non-obese T2DM . . . . .	6
1.2.5 Mathematical modelling of leptin in insulin-glucose regulatory system .	7
1.3 Organisation of the thesis . . . . .	7
<b>2 Background concepts and literature survey</b>	<b>9</b>
2.1 Organ-on-chip technology and its importance in drug development . . . . .	10
2.2 Physiologically based pharmacokinetic modelling and <i>in-vitro</i> systems . . . . .	11
2.3 Lumped parameter modelling of the cardiovascular system . . . . .	12
2.4 Insulin-glucose regulatory system and oral minimal model . . . . .	13
2.5 Optimisation-based parameter estimation . . . . .	16
<b>3 Mathematical modelling of drug distribution in liver-on-chip device</b>	<b>19</b>
3.1 Need for a digital twin to predict drug clearance in liver . . . . .	20
3.2 Methodology . . . . .	22
3.2.1 Dataset used for parameter estimation . . . . .	22
3.2.2 Mathematical model . . . . .	22

3.2.3	DigiLoCs: Digital Twins for cell-based liver assays . . . . .	24
3.2.4	Parameter estimation . . . . .	28
3.2.5	Translation to humans . . . . .	29
3.2.6	Prediction of human pharmacokinetics . . . . .	30
3.3	Results of DigiLoCs . . . . .	30
3.3.1	Simulating on-chip compound depletion . . . . .	31
3.3.2	Sensitivity analysis of three-compartment model . . . . .	31
3.3.3	Predicting human clearance . . . . .	34
3.3.4	Translation to human PK . . . . .	35
3.4	Discussion on DigiLoCs . . . . .	36
3.5	Conclusion . . . . .	39
<b>4</b>	<b>Lumped parameter modelling of cardiovascular system</b>	<b>41</b>
4.1	Need for a <i>comprehensive</i> cardiovascular model . . . . .	42
4.2	Methodology . . . . .	43
4.2.1	Overview of circulatory anatomy . . . . .	43
4.2.2	Model development . . . . .	44
4.2.3	Governing equations . . . . .	44
4.2.4	Parameter values and manual tuning . . . . .	47
4.2.5	Validation methods . . . . .	48
4.3	Results and validation . . . . .	49
4.3.1	Simulation . . . . .	52
4.3.2	Sensitivity analysis . . . . .	53
4.3.3	Validation . . . . .	55
4.4	Discussion . . . . .	58
<b>5</b>	<b>Computational methods to study T2DM</b>	<b>61</b>
5.1	Discovering subtypes of uncontrolled T2DM . . . . .	63
5.2	Study design and methodology . . . . .	63
5.3	Results of clustering . . . . .	66
5.3.1	Cluster characterisation . . . . .	68
5.3.2	Assessment of IR and $\beta$ -cell dysfunction within the clusters . . . . .	70
5.4	Discussion on uncontrolled T2DM clusters . . . . .	70
5.5	Clinical significance and summary of T2DM subtypes . . . . .	71
<b>6</b>	<b>Oral minimal model estimation</b>	<b>73</b>
6.1	Overview of insulin-glucose regulatory system . . . . .	73
6.2	Augmented oral minimal model . . . . .	75

6.3	Datasets used for parameter estimation . . . . .	76
6.4	Sensitivity indicators and parameter estimation . . . . .	76
6.5	Results from the oral minimal model using OGTT data . . . . .	78
6.6	Discussion on the findings . . . . .	82
6.7	Conclusion . . . . .	83
<b>7</b>	<b>Augmenting the insulin-glucose regulatory model with obesity indicators</b>	<b>85</b>
7.1	Association between obesity indicators and insulin-glucose dynamics in T2DM	86
7.2	Modelling mechanism . . . . .	87
7.3	Augmented model including leptin and BMI . . . . .	88
7.4	Parameter estimation . . . . .	91
7.5	Revision of indices derived from OGTT . . . . .	93
7.6	Results and validation . . . . .	95
7.7	Discussion on augmented model . . . . .	98
7.8	Conclusion . . . . .	100
<b>8</b>	<b>Conclusions</b>	<b>103</b>
8.1	Conclusions of present work . . . . .	103
8.2	Future work plan . . . . .	107
	<b>Bibliography</b>	<b>113</b>
<b>A</b>	<b>Supplementary information for Chapter 3</b>	<b>131</b>
A.1	Simulation of drug depletion data . . . . .	131
A.2	Human PBPK modelling of propranolol . . . . .	132
<b>B</b>	<b>Supplementary information for Chapter 4</b>	<b>147</b>
B.1	Compartments and ODEs . . . . .	147
B.2	Supplementary figures and tables . . . . .	149



# List of Abbreviations

ADME	Absorption, Distribution, Metabolism, Excretion
BMI	Body Mass Index
CL <sub>c</sub>	Clearance
CVS	Cardiovascular System
DBSCAN	Density-based Spatial Clustering of Applications with Noise
DigiLoCs	Digital Liver-on-Chip simulator
HOMA-B	Homeostatic Model Assessment of $\beta$ -cell function
HOMA-IR	Homeostatic Model Assessment of Insulin Resistance
IG	Insulin-Glucose
IR	Insulin Resistance
IVIVE	<i>In Vitro</i> to <i>In Vivo</i> Extrapolation
LPM	Lumped Parameter Model
MPS	Microphysiological System
ODE	Ordinary Differential Equation
OGTT	Oral Glucose Tolerance Test
OMM	Oral Minimal Model
OoC	Organ-on-chip
PBPK	Physiologically Based Pharmacokinetic
PK	Pharmacokinetics
T2DM	Type 2 Diabetes Mellitus
UMAP	Uniform Manifold Approximation and Projection





# List of Figures

2.1	Organ-on-chip model configurations . . . . .	11
3.1	Digital twin (DT) approach . . . . .	25
3.2	Common chip architectures: Chip 1 for CnBio and 3D spheroids and chip 2 is for Javelin architectures . . . . .	28
3.3	Translational workflow plan that integrates results from organ-on-chip with computer modelling to predict the kinetics of drugs in humans . . . . .	31
3.4	Digital twin-based model simulation of on-chip kinetics after fitting parameters for selected compounds . . . . .	32
3.5	Local and global sensitivity of the parameters with respect to output intracellular concentration . . . . .	33
3.6	Impact of DigiLoCs on predicting clinical clearance values compared to the conventional approach . . . . .	34
3.7	Correlation between observed and predicted <i>in vivo</i> intrinsic clearance ( $CL_{int}$ ) using the three-compartment liver chip for 12 drugs. . . . .	36
3.8	Simulated kinetics of propranolol after a single oral dose (80 mg) . . . . .	37
4.1	Human body anatomy marked with compartments in the comprehensive model .	45
4.2	Base model of cardiovascular system (adapted from [1]) . . . . .	45
4.3	Entire closed-loop electrical analog model of human cardiovascular system . . . .	46
4.4	Comparison of simulated waveforms from the comprehensive model with reference ranges. Twelve subplots show pressure (blue), volume (orange), and flow rates (green) across cardiac chambers and valves over one cardiac cycle. Simulated results (solid lines) and base model results (dashed lines) are plotted for left ventricle (LV), right ventricle (RV), left atrium (LA), right atrium (RA), and valves: mitral (MV), tricuspid (TV), aortic (AV), and pulmonary (PV). Reference regions are shaded: pressure and volume plots include reference minimum (lighter shade) and maximum (darker shade) bounds; flow plots indicate peak reference values. .	50

- 4.5 Pressure waveforms in the ascending aorta ( $P_{AscA}$ ), pulmonary artery ( $P_{PulAr}$ ), cerebral artery ( $P_{Cer}$ ), and renal artery ( $P_{Ren}$ ) over one cardiac cycle. Each colored line represents the pressure in the corresponding vessel. For each subplot, the shaded horizontal bands indicate the established physiological reference range: the upper reference limit is shown by the darker shade (Max ref), and the lower reference limit is shown by the lighter shade (Min ref), both in the color corresponding to each compartment. . . . . 51
- 4.6 Total Sobol sensitivity indices for six clinical variables, showing the five most influential parameters for each. Bar height indicates the overall impact of each parameter on the variable, with error bars representing confidence scores. Variables shown are: left ventricular volume ( $V_{LV}$ ), right ventricular volume ( $V_{RV}$ ), aortic pressure ( $P_{AA}$ ), pulmonary arterial pressure ( $P_{AR}^{PUL}$ ), aortic flow ( $Q_{AA}$ ), and superior vena cava pressure ( $P_{SVC}$ ). . . . . 54
- 4.7 Global sensitivity analysis for comprehensive model. Total Sobol indices for all model parameters and output variables. Darker colors indicate higher sensitivity. For readability, indices below a threshold of 0.1 have been omitted in the heatmap. 54
- 4.8 Stability of compartmental volumes in the comprehensive model. Shown are the heart, lung, and systemic vascular volumes during the last five cycles in 10,000 simulation cycles. The volumes remain stable and exhibit periodic oscillation, demonstrating that the model maintains dynamic steady state . . . . . 55
- 4.9 Pressure-volume (PV) loops for the left (left panel) and right (right panel) ventricles. Simulated model output (blue), standard reference data (black), and base model results (red) are shown for each ventricle. The simulated PV loops resemble both the shape and range of PV loops from literature [2, 3], indicating that the comprehensive model reproduces physiological ventricular function, while the base model exhibits significant deviations in both pressure and volume. . . . . 56
- 4.10 Distribution of cardiac output (%) across major arteries, veins, and organs as predicted by the model. Bars represent model predictions, grouped by compartment type: arteries (green), systemic veins (orange), and organs (blue). Grey bars indicate the physiological range from the literature. . . . . 56
- 4.11 Convergence of flow metrics over increasing cardiac cycles for (left) organs, (center) arteries, and (right) systemic vessels. Bars show mean absolute flow difference (green) and root mean square error (RMSE, lavender). Both metrics decrease rapidly with more cycles, approaching zero, and indicating that model flow predictions stabilize as the number of cycles increases. . . . . 57

5.1	Timeline diagram of the prospective study design of subjects with T2DM recruited from the Primary Health Clinic . . . . .	64
5.2	Clustering workflow with clinical data from 339 uncontrolled T2DM patients . . . .	65
5.3	Cluster formation using UMAP followed by DBSCAN . . . . .	66
5.4	Anti-diabetic drug combinations and the percentage of patients receiving them in three clusters . . . . .	68
5.5	Anti-hypertensive drug combinations and the percentage of patients receiving them for each individual cluster . . . . .	69
6.1	Parameter order in accordance with parameter importance index and collinearity index . . . . .	79
6.2	Sample simulation results for control and diabetic group . . . . .	80
6.3	Sample simulation results for obese and non-obese diabetic group . . . . .	81
6.4	Box plots of parameters for obese and non-obese diabetic . . . . .	82
7.1	<i>Leptin</i> model and the augmentations visually highlighted . . . . .	89
7.2	Sensitivity analysis of parameters with $\pm 30\%$ variation wrt glucose, insulin and leptin . . . . .	94
7.3	Variation of scaling factor used in revised HOMA-B and HOMA2 wrt correlation with ISI . . . . .	95
7.4	Simulation sample showing glucose, insulin and leptin concentrations for two subjects . . . . .	96
7.5	Simulation sample showing glucose and insulin (for 75g glucose) for <i>base</i> model and <i>leptin</i> model using Dataset I . . . . .	98
7.6	Simulation sample showing glucose and insulin (for 25g glucose) for <i>base</i> model and <i>leptin</i> model using Dataset II . . . . .	99
A.1	Digital twin-based model simulation of on-chip kinetics after fitting parameters for three compounds from Tsamandourous et al. . . . .	132
A.2	Digital twin-based model simulation of on-chip kinetics after fitting parameters for nine compounds from Docci et al. . . . .	137
A.3	Digital twin-based model simulation of on-chip kinetics after fitting parameters for eight compounds from Rajan et al. . . . .	139
A.4	Digital twin-based model simulation of on-chip kinetics after fitting parameters for 3D spheroids (Kanebratt et al., Bonn et al.) . . . . .	141
A.5	Correlation between observed and predicted <i>in vivo</i> intrinsic clearance ( $CL_{int}$ ) using the three-compartment model for four drugs from Kanebratt et al. and Carvedilol from Bonn et al. . . . .	142

A.6	Correlation between observed and predicted <i>in vivo</i> intrinsic clearance ( $CL_{int}$ ) using the three-compartment model for seven drugs from Hultman et al. . . . .	143
A.7	Correlation between observed and predicted <i>in vivo</i> intrinsic clearance ( $CL_{int}$ ) using three-compartment ODE liver chip for eight drugs from Rajan et al. . . . .	144
A.8	Representation of the generic structure of a whole-body PBPK model . . . . .	145
B1	Simulated pressure and flow rate for pulmonary circulation (left and right lung, pulmonary arteries and veins). . . . .	150
B2	Simulated pressure and flow rate for major arteries. . . . .	150
B3	Simulated pressure and flow rate for upper body organs. . . . .	150
B4	Simulated pressure and flow rate for abdominal organs. . . . .	150
B5	Simulated pressure and flow rate for abdominal organs. . . . .	151
B6	Simulated pressure and flow rate in major veins. . . . .	151
B7	Pressure, volume and flow rates during a heartbeat in heart chambers . . . . .	152
B8	Total Sobol indices for all model parameters and output variables for base model. Darker shades indicate higher sensitivity. . . . .	152
B9	Range of variables with the reference (Table B.2,B.3) for both base and comprehensive model. The base model overshoots the reference range for most parameters, including $Q_{PV}$ , $Q_{AV}$ and $Q_{MV}$ . . . . .	155

# List of Tables

3.1	Overview of literature reports providing on-chip pharmacokinetic information on compound clearance . . . . .	22
3.2	Names and description of all relevant variables and parameters in the three-compartment digital twin . . . . .	24
4.1	List of Acronyms . . . . .	51
4.2	Compartment names and parameter values used in the comprehensive cardiovascular model. . . . .	52
4.3	L2 norm difference of blood pressure, volume and flow in base and comprehensive model . . . . .	57
5.1	The post-clustering subject characteristics of the uncontrolled T2DM subgroups .	67
5.2	Summary statistics of the post-clustering features of all the T2DM clusters . . . . .	68
6.1	Sample rows from Dataset I and Dataset II used for parameter estimation of OMM	77
6.2	Comparison of parameters in control and diabetic group (Dataset I) . . . . .	80
6.3	Comparison of obese diabetic and non-obese diabetic group (Dataset II) . . . . .	81
7.1	Sample rows from Dataset I and Dataset II used for parameter estimation of <i>leptin</i> model . . . . .	91
7.2	Estimated parameters using the <i>leptin</i> model . . . . .	92
7.3	Correlation of indices: (a) Correlation of ISI to HOMA-B and HOMA2 and their revisions (b) Correlation of $DI_o$ to HOMA-IR and revised HOMA-IR . . . . .	97
A.1	Drug-specific information of 32 compounds involved in the study . . . . .	133
A.2	Overview of observed, conventional, on-chip and predicted clearance values of the investigated 32 drugs across different <i>in vitro</i> systems . . . . .	134
B.1	Initial values of variables. P, Q and V represent blood pressure ( $mmHg/s$ ), flow ( $mL/s$ ) and volume ( $mL$ ). . . . .	151
B.2	Hemodynamics for whole-body cardiovascular model. Simulated and reference ranges for volumes (mL) and pressures (mmHg) in all model compartments. . . . .	153

B.3 Hemodynamics for whole-body cardiovascular model: Simulated and reference mean flows in all compartments. The range of flow in the carotid arteries, the upper limb, and the lower limb is multiplied by two. . . . .	154
---	-----

# Introduction

## Chapter outline

- Introduction to topics relevant to this thesis
- Contributions of this thesis
  - Mathematical modelling of drug distribution in liver-on-chip device
  - Lumped parameter modelling of the cardiovascular system
  - Clustering-based methods to identify type 2 diabetes mellitus (T2DM) sub-groups
  - Oral minimal model estimation of parameters for obese and non-obese T2DM
  - Mathematical modelling of leptin in insulin-glucose regulatory system
- Organisation of the thesis

Mathematical models are valuable tools used to understand and predict the behaviour of complex physiological systems. This modelling process begins with a well-formulated hypothesis derived from prior observations, which is then translated into mathematical equations that describe the system's behaviour. The model formulation is an iterative process of developing an initial model, comparing its predictions with experimental data and refining the model based on the results and available theory to achieve satisfactory conformance [4]. These models are driven by data and act as digital twins, which are the virtual representations or simulations that mirror the dynamics of the modelled physiological system [5]. One of the advantages of computational modelling is the ability to bridge the gap between theoretical understanding and experimental observations, providing a systematic and quantitative framework for exploring physiological processes. This approach aids in understanding the system as a whole and predicts how it behaves in altered states (e.g., disease), leading to more

effective therapeutic strategies and personalised treatments. In this thesis, mathematical modelling has been applied to the following two areas (i) the development of digital twins for human physiological systems, specifically modelling drug metabolism in *in vitro* liver systems and simulating cardiovascular blood flow using lumped parameter models; and (ii) modelling of biochemical processes, focusing on the insulin-glucose regulatory system in the context of type 2 diabetes mellitus (T2DM).

## 1.1 Introduction on topics explored in this thesis

Traditional drug development relies heavily on animal models and clinical trials, which are costly and time-consuming and often fail to predict human responses accurately. *In vitro* models have emerged as a promising alternative, offering a controlled, ethically acceptable and cost-effective means to study biological processes. The liver plays an important role in the clearance of drugs from the body. However, studying this in a human population using traditional *in vitro* methods is challenging. Recent developments in organ-on-chip technology have opened up new vistas of exploration. Liver-on-chip devices provide a physiologically reliable platform to investigate liver function predicting *in vivo* drug clearance. By leveraging kinetic data from liver-on-chip experiments, a computational model is developed that effectively simulates drug distribution and clearance in the liver. This is an example of digital twinning, where a computational model is used to mimic certain aspects of a biological system. In this case, the mathematical model for drug clearance mimics the drug clearance behaviour of actual liver cells.

Another example of digital twinning is the mathematical model that mimics the behaviour of the human circulatory system, accounting for vascular pressure of blood flow, impedance to the flow of blood, the volume of blood in the vessels and the elasticity of the vessels. This can be modelled at various levels of detail, from simple zero-dimensional (0D) lumped parameter models (LPM) to more complex one-dimensional (1D) and three-dimensional (3D) models that simulate blood flow dynamics with increasing levels of spatial and structural detail [6, 7]. The LPMs approximate the cardiovascular network as a set of discrete compartments, which makes the model computationally efficient and practical for real-time simulations or quick assessments. These models can be enhanced to simulate various cardiovascular diseases by tuning the model parameters to match observed physiological behaviour, allowing for the replication of blood flow patterns in different compartments. These simulations facilitate the exploration of scenarios that are difficult to observe experimentally, such as blood flow in internal organs, making them useful in personalised medicine for diagnosing and monitoring disease progression.

Application of computational modelling to the insulin-glucose regulatory system to inves-



Investigate the dynamics of T2DM has also been carried out as part of this work. T2DM affects a significant number of people worldwide. Achieving adequate glycemic control is important to prevent diabetic complications such as retinopathy, nephropathy and neuropathy, which remain a significant challenge globally, particularly in low-income and middle-income countries. T2DM has traditionally been viewed as a homogeneous disease entity characterised by insulin resistance (IR) and  $\beta$ -cell dysfunction. However, recent research suggests that T2DM is a heterogeneous mix of pathological conditions, with IR and  $\beta$ -cell dysfunction varying across subtypes. To further understand obesity as a risk factor in T2DM, parameter estimation on obese and non-obese T2DM data is performed with established insulin-glucose regulatory models. Understanding the differences between obese and non-obese T2DM by estimating key parameters allows for identifying distinct characteristics of each group, ultimately informing personalised treatment strategies. The existing insulin-glucose model is augmented to include the role of obesity indicators, namely BMI and leptin (a hormone secreted by adipose tissue), to enhance our understanding of insulin-glucose dynamics in the body. The model is built on several theoretical studies on the association of leptin with insulin and glucose concentrations.

Common mathematical modelling approaches for physiological systems include differential equations, difference equations, delay differential equations and agent-based models [8, 9]. The work in this thesis focuses on modelling using ordinary differential equations (ODEs). A key aspect of model development is validating the model's behaviour by comparing its predictions with observed data from the system being studied. The models are usually parameterised, and the parameters need to be determined based on available observations and may require re-estimation using subject-specific data. The parameters could represent physical constants, biological rates, coefficients of a system, or any other factors that govern the model's behaviour. Parameter estimation infers the most likely values of these parameters by comparing the model predictions with observed data. Exploring the parameter space through model simulations facilitates the identification of optimal conditions and critical factors that influence system behaviour. Optimisation methods are used to systematically explore the best combination of parameters that maximise or minimise an objective function, typically representing the goodness of fit between model simulations and observed data. Alongside parameter estimation, sensitivity analysis is performed to assess how changes in these values affect model outcomes, providing insights into the model's robustness and reliability.

Background concepts, a study of literature on the topics discussed above and research gaps on which the work in this thesis is based are presented in Chapter 2. The contributions from this thesis are given in the next Section.

## 1.2 Contributions from this research

The first part of this thesis explores the realm of digital twins by developing a realistic liver chip using data and mathematical modelling. A compartmental physiological model of the liver using ODEs is developed to estimate pharmacokinetic (PK) parameters for liver-on-chip. Published data on metabolism and toxicology studies are used to determine parameters. The model is used to predict *in vivo* liver clearance, assessing *in vitro* to *in vivo* extrapolation (IVIVE) of estimated on-chip clearance through time series kinetic data.

The following segment of the thesis investigates lumped parameter modelling of the cardiovascular system. It is a mathematical approach that simplifies the dynamics of blood flow and pressure by representing the system as a network of interconnected compartments. This method aggregates spatially distributed parameters into discrete elements, enabling the analysis of cardiovascular dynamics through simplified mathematical equations.

Moving forward, this thesis discusses clustering-based methods to study T2DM subgroups, aiming to identify distinct patterns and characteristics within the patient population. This study addresses a gap in T2DM management by recognising the heterogeneity of the disease, leading to effective standardised treatments. Focusing on uncontrolled T2DM patients in rural India, this work considers clustering on longitudinal data to identify subtypes and lay the groundwork for a more personalised T2DM management protocol.

The following section of the thesis focuses on parameter estimation of the oral minimal model (OMM), explicitly considering the differences between obese and non-obese individuals with T2DM. This study uses Oral Glucose Tolerance Test (OGTT) data to classify model parameters between control groups and subjects with T2DM. Parameter variations in obese and non-obese individuals with T2DM are also examined. Through sensitivity analysis and statistical tests, the study offers a better understanding of the role of obesity in T2DM, correlating specific parameters with physiological findings. The final section of this thesis is on the mathematical modelling of leptin in the insulin-glucose regulatory system. This ODE-based model, incorporating both leptin and BMI, extends insulin-glucose modelling to accommodate the role of obesity.

### 1.2.1 Prediction of *in vivo* drug clearance in liver-on-a-chip device using PBPK modelling

This study explores the concept of digital twins, driven by data and mechanistic modelling, to develop a liver chip that serves as a realistic digital twin mimicking the functionalities of the human liver. A compartmental physiological model of the liver is developed using ODEs and physiologically based pharmacokinetic (PBPK) modelling. This digital liver-on-chip simulator (DigiLoCs) aims to estimate PK parameters associated with on-chip liver clearance using

time-series data. DigiLoCs uses published data [10, 11] on metabolism-related PK parameters and toxicology studies (IC50) to estimate parameters.

This model has been applied to predict the *in vitro* liver clearance for thirty-two drugs using a three-compartment model of the liver-on-chip. The study also explores IVIVE, where the on-chip clearance values are scaled to human values. Three ODEs in the model define drug concentrations in media, interstitium and intracellular compartments based on biological, hardware and physicochemical information. A key issue in determining liver clearance appears to be insufficient drug concentration within the intracellular compartment. The digital twin establishes a connection between the hardware chip structure and an advanced mapping of the underlying biology, specifically focusing on the intracellular compartment. This modelling approach offers several advantages, including a superior prediction of intrinsic liver clearance compared to single-compartment or conventional models, an explainable behaviour based on physiological parameters, and the ability to predict intracellular drug concentration, even for difficult-to-measure cases. This work may represent the most extensive cross-organ-on-chip platform study, analysing and predicting human clearance values through data from diverse *in vitro* liver-on-chip systems.

### 1.2.2 Lumped parameter modelling of cardiovascular system

The cardiovascular system (CVS) is a network of organs and vessels responsible for pumping and transporting blood throughout the body. One common method for modelling its hemodynamics is the zero-dimensional (0D) lumped parameter model (LPM), which simplifies the system by dividing it into compartments that represent different aspects of blood flow, pressure and volume dynamics [6]. These models use hydraulic-electrical analogies, where pressure gradients drive blood flow similarly to how voltage drives current in electrical circuits. Blood flow is governed by the principles of mass, momentum and energy conservation. This analogy allows for the modelling of vascular resistance, compliance and inertance, enabling a detailed analysis of pressure and flow rate over time. By applying electrical circuit principles, these models offer a mathematical framework for predicting cardiovascular dynamics.

Modelling the entire CVS is a complex task due to its intricate nature. The complexity of the model can range from zero dimensions to higher dimensions, depending on the study goals and the required accuracy. LPMs abstractly model sets of body organs as electrical analogues, a process known as compartmental modelling. These models assume uniform variations in key variables (pressure, flow and volume) within each compartment (such as an organ, vessel or part of a vessel) at any given time. In contrast, higher-dimensional models provide a distributed variation of these variables across space and time [12]. LPMs are suitable for real-time simulations and can be adapted to model individual organs to entire circulatory networks, making them versatile across various research and clinical applications.

### 1.2.3 Clustering based methods to study T2DM sub groups

The study addresses a gap in T2DM management by recognising the heterogeneity within the disease that challenges the effectiveness of a uniform treatment protocol. Despite the diverse nature of T2DM subtypes, current guidelines for T2DM management remain standardised, leading to inadequate glycemic control in a significant number of patients. This research aims to bridge this gap by employing an unsupervised clustering workflow on longitudinal cohort data from T2DM patients in rural India. The objective is to identify and characterise subtypes among uncontrolled T2DM patients, thereby opening the possibility for a more personalised and precise T2DM management protocol. By exploring the heterogeneity of the uncontrolled T2DM population, the study seeks to extract translational insights that can reform the approach to diabetes management and contribute to the emerging era of precision medicine in diabetes care.

In this study, the traditional view of T2DM as a homogenous disease entity with insulin resistance (IR) as the primary pathology is investigated. Recent research suggests that T2DM is a heterogeneous mix of different pathological conditions, with the contributions of IR and  $\beta$ -cell dysfunction varying among subtypes. The precise determination of the underlying pathology is important as the field transitions to precision medicine for diabetes management. Inadequate glycemic control poses a significant challenge, particularly in low and middle income countries, leading to an increase in diabetic complications. The primary aim of this work is to investigate the heterogeneity among the uncontrolled T2DM population and identify patho-clinical features that contribute to this diversity. The ultimate goal is to pave the way toward precision therapy by developing a more tailored T2DM management protocol based on the identified subtypes among uncontrolled T2DM patients.

### 1.2.4 Oral minimal model estimation of parameters for obese and non-obese T2DM

The Oral Glucose Tolerance Test (OGTT) is a standard diagnostic tool for identifying T2DM in clinical settings. Clinical OGTT data from obese and non-obese individuals are used to study the influence of obesity in the insulin-glucose regulatory system using the simplest model, the Oral Minimal Model (OMM). The OMM is based on the Bergman minimal model [13], where glucose analysis is done using an intravenous glucose tolerance test (IVGTT). Identification of subject-specific oral minimal models is carried out for the insulin-glucose regulatory system from the observed glucose and insulin readings in the OGTT data and inferring the rate of appearance of ingested glucose from the OMM. The model parameters of the participating subjects are studied to suggest differences in obese and non-obese diabetic physiology. By studying the differences in model parameters between control groups and those with T2DM,

the study aims to understand the behaviour and physiology of T2DM. Sensitivity analysis is used to examine the relationship between parameter changes and model outputs, and then statistical tests are applied to validate and confirm the robustness of these findings.

By correlating sensitive parameters with established physiological findings, the study enhances our understanding of the complexities associated with T2DM. Furthermore, sensitivity analysis and statistical tests add value to the investigation, ensuring that the identified parameter differences are clinically significant. Overall, this study marks a significant step in the application of the OMM for providing a comprehensive framework to investigate T2DM and its relationships with both obesity and physiological parameters.

### 1.2.5 Mathematical modelling of leptin in insulin-glucose regulatory system

This study focuses on understanding the insulin-glucose regulatory system, specifically addressing the limitations of existing models that overlook the influence of body weight indicators, such as body mass index (BMI) and leptin. Leptin, a hormone derived from fat cells, is recognised for its role in satiety, body weight regulation, and its regulatory impact on glucose metabolism and energy homeostasis in the insulin-glucose system. To bridge this gap, the ODEs of the OMM are augmented by including the role of leptin and BMI. Through the estimation of model parameters using data from the OGTT, the augmented model conforms better with observed glucose, insulin and leptin levels in individuals with T2DM. The study also introduces revised indices formulated from OGTT data, including BMI and fasting leptin values, revealing a stronger correlation with existing indices. Furthermore, parameter sensitivity analysis results show the impact of model parameters on observed variables, emphasising the associations between leptin, glucose and insulin concentrations.

The validation of the augmented model with clinical data, even in the absence of leptin measurements, demonstrates a superior fit to glucose and insulin data compared to the model without augmentation. This significant improvement underscores the potential of the proposed model to enhance the understanding of the interplay between leptin, glucose and insulin concentrations. This work introduces the first ODE-based model incorporating both leptin and BMI in the insulin-glucose system, paving the way for developing targeted interventions and therapies for T2DM.

## 1.3 Organisation of the thesis

The thesis is organised as follows.

- **Chapter 1 [Introduction]** presents the motivation and objectives of this work along with a summary of the contributions from this research.

- **Chapter 2 [Background concepts and literature survey]** gives an overview of the basic concepts used in the research along with a summary of the background study and literature review related to each contribution.
- **Chapter 3 [Mathematical modelling of drug distribution in liver-on-chip device]** provides data-driven mechanistic approaches to predict human *in vivo* drug clearance using a digital liver-on-chip simulator (DigiLoCs).
- **Chapter 4 [Lumped parameter cardiovascular modelling]** develops and applies zero-dimensional models to simulate and analyse cardiovascular hemodynamics.
- **Chapter 5 [Clustering based methods to study T2DM sub groups]** identifies and characterises subtypes among uncontrolled T2DM patients in rural India, leading to personalised and precise T2DM treatment protocols.
- **Chapter 6 [Oral minimal model (OMM) estimation of parameters for obese and non-obese T2DM]** explores relationships between obesity and insulin-glucose dynamics through parameter estimation of OMM.
- **Chapter 7 [Mathematical modelling of leptin in insulin-glucose pathway]** augments the OMM by incorporating obesity indicators, leptin and BMI, to better understand the insulin-glucose regulatory system.
- **Chapter 8 [Conclusion]** summarises the work done and concludes the thesis while identifying possible directions for future work.

# Background concepts and literature survey

## Chapter outline

- Organ-on-chip technology and its importance in drug development
- Physiologically based pharmacokinetic modelling of *in vitro* systems
- Lumped parameter modelling of the cardiovascular system
- Insulin-glucose regulatory system and oral minimal model
- Parameter estimation and optimisation methods

This chapter covers some background topics for the work presented in this thesis. Organ-on-chip has found several biomedical applications. The application of organ-on-chip towards expediting drug development by mimicking human physiology and providing more accurate predictions of pharmacokinetics and pharmacodynamics is discussed in Sections 2.1 and 2.2. The approach followed here may be considered a form of digital twinning to study the clearance of drugs in the liver. Another form of digital twinning, focusing on the circulatory system through a network of electrical components, has also been explored. The lumped parameter modelling approach for the cardiovascular system to study the pressure and flow in different compartments is discussed in Section 2.3. An overview of the insulin-glucose regulatory system and the oral minimal model are given in Section 2.4. Parameter estimation and optimisation methods are discussed in Section 2.5.

## 2.1 Organ-on-chip technology and its importance in drug development

Organ-on-chip (OoC) is a microfluidic cell culture device, which reproduces the physiological and pathological characteristics of organs or tissues by reconstructing the structure and function *in vitro*. OoC has reformed the field of drug development by providing a novel, robust and accurate platform for testing drug efficacy and toxicity. OoCs are microfluidic devices that capture the functionality of human organs along structural lines, allowing for the creation of a personalised, miniaturised model of human physiology. By mimicking the complex interactions between cells, tissues and organs, OoCs enable the simulation of human disease mechanisms and the testing of drug candidates in a realistic and controlled environment. OoC technology is of significant importance in drug development. Traditional drug development methods rely on animal models and two-dimensional cell cultures, which often fail to accurately predict human responses to drugs. This has led to high attrition rates in clinical trials, resulting in significant financial and resource demands. OoCs offer a promising solution to this problem by providing a more accurate and predictive model of human drug responses, enabling researchers to identify potential drug failures earlier in the development process and accelerate the discovery of new, more effective treatments.

One of the key focuses of this thesis is on the liver-on-chip system. The liver is the main site of drug metabolism and is composed of a series of complex hepatic lobules. Maintaining the long-term physiological function of liver cells is a challenging problem. The liver-on-chip technology has exhibited significant applicability in the realm of clinical trials. Kanebratt et al. [14], Hultman et al. [15] and Bonn et al. [16] have used regression curve fitting to analyse kinetic data in 3D spheroid models and calculate intrinsic clearance by determining the slope of the fitted curve. In the works of Docci et al. [10] and Tsamandourous et al. [11] the kinetic data from a liver-on-chip device is used to predict the *in vivo* intrinsic clearance.

A key challenge for the development of a new pharmaceutical compound is the accurate prediction of *in vivo* clearance from *in vitro* data. *In vivo* clearance, or the rate at which a drug is eliminated from the body, is a critical parameter that influences the dosing regimen, efficacy and safety of a drug. It is a fundamental aspect of drug development, as an inadequate understanding of the clearance of a compound can lead to unexpected and undesirable outcomes in clinical trials, ranging from underdosing to toxicity. However, the prediction of *in vivo* clearance is challenging and frequently described as giving an underprediction even when scaling factors or regression lines are applied. This is primarily due to the inherent complexity of biological systems, the influence of various physiological factors and the dynamic nature of drug metabolism and disposition within the human body. To address these



issues, researchers are turning to digital twinning [17, 18, 19], computer models employing mathematical equations as tools for simulating real-world systems.

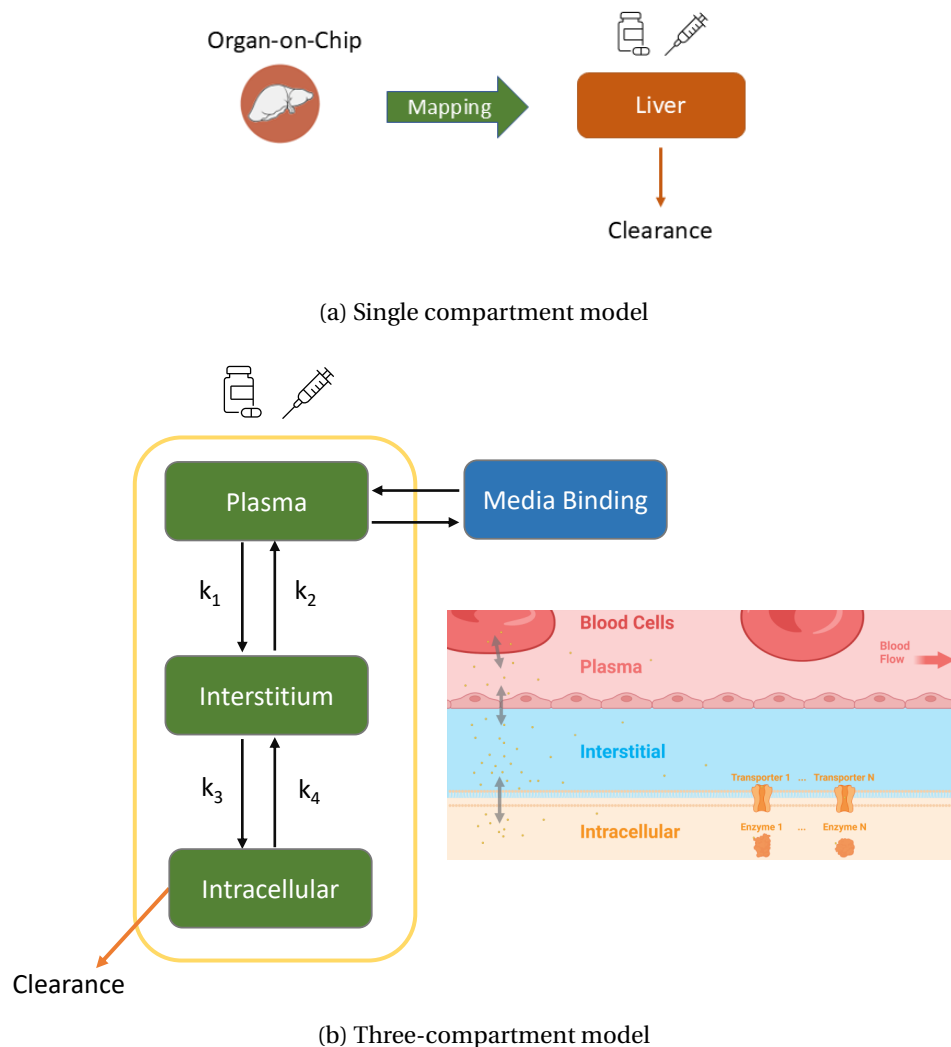


Figure 2.1: Organ-on-chip model configurations

## 2.2 Physiologically based pharmacokinetic modelling and *in-vitro* systems

Physiologically based pharmacokinetic modelling (PBPK) studies drug pharmacokinetics by viewing human organs as separate compartments and integrating them into a system according to physiological and anatomical knowledge. Compared to the traditional pharmacokinetic model, the compartments and model parameters of the PBPK model have physiological meaning. Thus, PBPK modelling is believed to reflect the absorption, distribution, metabolism and excretion (ADME) of drugs or other substances in the body *in vivo* more accurately. These models use differential equations to describe the physiological processes that govern the

concentration of substances in the body, such as drug clearance, blood flow and enzyme kinetics. PBPK models have several key features. These are mechanistic and quantitative in nature, meaning they are based on the underlying physiology of the body and they provide numerical predictions of drug concentrations in different tissues and organs. Additionally, PBPK models are dynamic, simulating the time course of drug absorption, distribution and elimination.

PBPK models have a wide range of applications in drug development, drug safety, personalised medicine, regulatory submissions and research. In drug development, PBPK models can be used to predict drug pharmacokinetics and pharmacodynamics in different populations. In drug safety, PBPK models can be used to predict drug interactions and toxicities. In personalised medicine, these models can be used to predict drug response in individual patients based on their physiological characteristics.

The development of *in vitro* liver-on-chip models, coupled with PBPK modelling, has emerged as a cutting-edge approach to replicate and study liver function in a more physiologically relevant manner. Our approach is to build a PBPK model for determining *in vitro* liver clearance that uses ordinary differential equations (ODEs) and parameter estimation. Firstly, we use an empirical single ODE model by considering the whole liver as a single compartment (Figure 2.1(a)) that describes the dynamics of the drug within the whole liver compartment. Subsequently, we develop a more complex three-compartment liver model, which includes distinct compartments for plasma, interstitium and intracellular spaces (Figure 2.1(b)). This model is explained in detail in Chapter 3. These models will utilise published data on pharmacokinetics (metabolism) or toxicology studies (IC<sub>50</sub>) to more effectively determine PK parameters than current state-of-the-art methods. By achieving this, we anticipate better predictions of clinical outcomes, leading to reduced time, cost and patient burden associated with drug development [20, 21].

## 2.3 Lumped parameter modelling of the cardiovascular system

The cardiovascular system (CVS), also known as the circulatory system, is a complex network of organs and vessels responsible for transporting blood throughout the body. Comprising the heart, arteries, veins and blood vessels, this vital system plays a crucial role in maintaining the body's homeostasis and overall health. The lumped parameter or zero-dimensional (0D) cardiovascular system is a simplified model that describes the hemodynamics of the CVS. This approach simplifies the complex distributed system into a finite number of compartments, enabling the analysis of blood flow, pressure and volume dynamics. A typical cardiovascular model consists of the heart chamber compartments that simulate blood pumping and vascular networks that distribute blood throughout the body [22]. This system can be represented

electrically using a combination of resistors, capacitors and inductors to mimic its physical properties. Specifically, resistors represent blood's viscous resistance in vessels, capacitors account for vessel compliance and inductors represent blood inertia. By modifying these elements, models of varying complexity can be created. Additionally, this approach enables the simulation of various cardiovascular diseases, including peripheral and vascular disease, coronary artery disease, valvular disease and others.

Ohm's and Kirchhoff's laws govern the electrical equivalents of fluid behaviour, enabling mathematical representation of each compartment through ODEs that capture pressure-volume relationships based on compliance. Numerical methods solve these equations, allowing examination of blood flow and pressure throughout the network. The lumped parameter modelling approach offers key advantages, including minimal computation time for solving ODEs and algebraic equations and simple model structure facilitating automated personalised simulations. 0D cardiovascular modelling has its roots in the Windkessel model of arterial flow and has since been expanded to include the heart, heart valves and veins [23]. Various 0D models have been developed to capture the unique characteristics of each circulatory subsystem, which are explained in detail by Yubing Shi et al. [6] and Louis Garber et al. [24].

Lumped parameter models (LPMs) have been applied to various cardiovascular research areas, including hemodynamic simulation, cardiovascular disease diagnosis and cardiac assist device development. While lumped parameter modelling simulates global hemodynamics and cardiac function, other advanced modelling approaches, such as one-dimensional (1D) and three-dimensional (3D) computational models of fluid dynamics [25, 26, 27], are better suited for studying local flow dynamics. These higher-order models use numerical methods like finite difference methods to solve Navier-Stokes equations and offer enhanced spatial detail. Unlike LPMs, which vary only with time (zero-dimensional), higher-order models account for variations in time and three-dimensional space within specific structures. Mathematically, this distinction translates to higher-order models represented by non-linear partial differential equations (PDEs), whereas LPMs rely on ODEs.

## 2.4 Insulin-glucose regulatory system and oral minimal model

The human body maintains a balance of glucose levels in the bloodstream through complex processes involving various hormones and physiological processes. Central to this process is the insulin-glucose regulatory system, a network of interactions between insulin, glucose and other metabolic factors. Dysfunction in this system causes metabolic disorders, namely type 2 diabetes mellitus (T2DM), where the body struggles to regulate blood glucose levels effectively. To understand the dynamics of the insulin-glucose system, studies with tests such as intravenous glucose tolerance test (IVGTT), oral glucose tolerance test (OGTT) and

short-term modelling using ODEs are performed.

OGTT involves measuring plasma glucose and insulin levels at specific intervals (0, 30, 60, 90 and 120 minutes) after consuming a standardised 75g oral glucose load. This test is commonly used in clinical settings to diagnose glucose intolerance and T2DM. The OGTT minimal model, an extension of the model initially proposed for IVGTT, adapts the basic framework by replacing the bolus dose of glucose with an input function representing the rate of appearance of oral glucose in plasma. This modification allows for a more accurate assessment of glucose-insulin dynamics during the oral test, for evaluating the ability to metabolise exogenous glucose and return to baseline glucose levels (glucose tolerance [28]) and the level of insulin that is necessary to keep a constant blood glucose level (insulin sensitivity [29]).

ODE modelling started with the *minimal model* for the IVGTT. Bergman's minimal model is a well-accepted coupled model from the original model [13]. Bolie [30] introduced another simpler model with two coupled ODEs. A detailed feedback model with fourteen equations, introduced by Sturis et al. [31, 32], is regulated by three variables, namely the amount of glucose in plasma and intercellular space, the amount of insulin in the plasma and the amount of insulin in the intercellular space. The Hovorka model consists of a glucose subsystem to model absorption, distribution and disposal of glucose, an insulin subsystem (to model insulin absorption, distribution, disposal) and an insulin action subsystem (insulin action on glucose transport, disposal and endogenous production) [33].

Even though these models are widely used, a quantitative physiological model for the insulin glucose regulatory system is more suitable for epidemiological studies. This is achieved in the oral minimal model (OMM) developed by Dallaman et al. [34] based on OGTT. The simplest form of OMM [34] represented using ODE are as follows

$$\frac{dG}{dt} = -k_1 \cdot (G - G_b) - X \cdot G + \frac{R_a(t)}{V}, \quad (2.1)$$

$$\frac{dX}{dt} = -k_2 \cdot X + k_3 \cdot (I - I_b), \quad (2.2)$$

$$\frac{dI}{dt} = \gamma \cdot (G - G_T)^+ - k_4 \cdot I, \quad (2.3)$$

where  $G$ ,  $I$  denotes blood glucose concentration and blood insulin concentration (suffix  $b$  denotes basal values) at time  $t$ , respectively. Here,  $X$  is an auxiliary variable to model the time delay in insulin-dependent glucose uptake activity. The initial values of  $G$  and  $I$  are taken at the time when the glucose is orally ingested. The initial value of  $X$  is taken as 0 and  $(G - G_T)^+ = (G - G_T)$ , if  $G > G_T$  and 0 otherwise.  $R_a$  is the rate of absorption of ingested glucose;  $V$  is the volume of distribution;  $k_1$  is the fractional glucose effectiveness, i.e. the ability to promote glucose disposal and inhibit glucose production,  $k_2$  is the insulin-dependent increase in glucose uptake ability in tissue per unit of insulin concentration above

$I_b$ ,  $k_3$  is a scale factor for the amplitude of insulin action and  $k_4$  is the first-order decay rate for insulin in plasma. There are several complex IG models that have evolved over time ([33, 32]). The OMM is based on the Bergman minimal model [13] in which the glucose analysis is done using IVGTT. The variable  $X$  in the minimal model could be more accurately understood as representing the incremental insulin response within the interstitial space in response to a bolus stimulus rather than being directly proportional to the interstitial insulin level, as it was conventionally assumed [35]. A quantitative physiological model for insulin glucose regulatory system will be more suitable for epidemiological studies and this is achieved in the OMM developed by Dallaman et al. [34] based on OGTT.

The parameter  $R_a$ , for the rate of appearance of oral glucose in plasma, is coupled with a minimal model of glucose kinetics. Various  $R_a$  models have been developed based on gastric emptying data. Lehmann and Deutsch [36] proposed a trapezoidal gastric emptying function with a single compartment for the intestine, whereas Elashoff et al. [37] considered exponential gastric emptying. Another linear model (Model 1 in [38]) with two compartments for the stomach and a single compartment for the intestine was proposed where the gastric emptying rate was constant. Yet, another parametric description of  $R_a$  as a piece-wise linear function was proposed [34] with a known number of break times where there is a shorter interval towards the beginning and longer intervals towards the end.

$$\frac{q_{sto}}{dt} = -k_{21} \cdot q_{sto1}(t) + D \cdot \delta(t), \quad (2.4)$$

$$\frac{q_{sto2}}{dt} = -k_{empt} \cdot q_{sto2}(t) + k_{21} \cdot q_{sto1}(t), \quad (2.5)$$

$$\frac{q_{gut}}{dt} = -k_{abs} \cdot q_{gut}(t) + k_{empt} \cdot q_{sto2}(t), \quad (2.6)$$

$$R_a(t) = f \cdot k_{abs} \cdot q_{gut}, \quad (2.7)$$

where  $q_{sto1}$  and  $q_{sto2}$  are the amounts of glucose in the stomach (solid and liquid phase, respectively),  $\delta(t)$  is the impulse function,  $D$  is the amount of ingested glucose,  $q_{gut}$  is the glucose mass in the intestine,  $k_{21}$  the rate of grinding,  $k_{empt}$  is the rate of gastric emptying,  $k_{abs}$  is the rate constant of intestinal absorption and  $f$  is the fraction of the intestinal absorption which actually appears in plasma.

$$k_{empt}(q_{sto}) = k_{min} + \left( \frac{k_{max} - k_{min}}{2} \right) \cdot \{ \tanh(\alpha \cdot q_{sto} - b \cdot D) - \tanh(\beta \cdot (q_{sto} - c \cdot D)) + 2 \}, \quad (2.8)$$

$$\alpha = \frac{5}{2 \cdot D \cdot (1 - b)},$$

$$\beta = \frac{5}{2 \cdot D \cdot c},$$

$$q_{sto} = q_{sto1} + q_{sto2}.$$

The gastric emptying rate,  $k_{\text{empt}}$  is a function of the amount of glucose in the stomach ( $q_{\text{sto}}$ , Eqn 2.8). It equals  $k_{\text{max}}$  when the stomach contains the amount of the ingested glucose,  $D$ , then it decreases to a minimum value of  $k_{\text{min}}$ . The constant  $b$  is the percentage of the dose for which  $k_{\text{empt}}$  decreases at  $(k_{\text{max}} - k_{\text{min}})/2$  and  $c$  is the percentage of the dose for which  $k_{\text{empt}}$  is back to  $(k_{\text{max}} - k_{\text{min}})/2$ .

## 2.5 Optimisation-based parameter estimation

The unknown model parameters are estimated using experimental and observational data to adequately identify the modelled biochemical system or processes. A common estimation mechanism is to simulate the model with a prospective set of parameters and determine the gap (distance) between values predicted by the model and actual experimental observations. Parameters are iteratively re-estimated to reduce the gap between predicted and actual observations. Optimisation involves finding the optimal values of model parameters that maximise or minimise a specific objective function. Optimisation techniques, such as gradient descent or genetic algorithms are used to search the parameter space and identify the optimal values. In many cases, parameter estimation and optimisation are performed iteratively, with the results of parameter estimation serving as inputs to the optimisation process. The distance is usually taken as the L2 norm. Consider a dynamic system

$$\frac{dy}{dt} = f(y, t, p), \quad (2.9)$$

where  $f$  describes how  $y$  changes over time  $t$  and  $p$  is the parameters in the system. The objective function for optimisation is as follows

$$\underset{p}{\text{minimise}} \sum_{i=1}^n (y_{i_{\text{obs}}} - y_{i_{\text{comp}}})^2,$$

where  $n$  is the number of observed data points,  $y_{i_{\text{obs}}}$  is the observed data point, and  $y_{i_{\text{comp}}}$  is the computed value at time  $i$ , which can be estimated using analytical methods (using eigenvalues and eigenvectors for simple models) or numerical methods.

Some parameters can be measured experimentally, while others need to be estimated from available observations as it may not be feasible to determine them experimentally. The parameters which are not directly measurable can be estimated using least squares or any other fitting methods to analyse the model quantitatively. This way, the behaviour of the model is captured effectively.

A good parameter estimation algorithm is to be chosen when fitting a model to experimental data. The challenge is that no single optimal estimation technique exists for all models. Many different estimation methods have been developed to determine the best strategy for a given problem. The commonly used parameter estimation methods include maximum

likelihood estimation [39], least squares estimation and Nelder-Mead optimisation [40]. The Nelder-Mead algorithm [41] is a simplex-based optimisation method that uses a set of vertices to search for the minimum of a function, iteratively replacing the worst vertex with a better one through reflection, expansion, and contraction operations. This process continues until convergence, allowing the algorithm to find the optimal solution without requiring derivative information. Least squares estimation minimises the sum of the squared differences between the observed data and the model's predictions. Other approaches are Bayesian estimation, which uses Bayes' theorem to update the probability of the parameter values based on the observed data and maximum a posteriori estimation, which finds the parameter values that maximize the posterior probability given the data. Many other evolutionary methods of parameter estimation, namely genetic algorithms, also exist.

**Sensitivity Analysis** One of the crucial aspects of modelling is to understand how parameter values change with model output. The parameter variations affect the output of a model when dealing with critical data specific to biochemical systems. The interpretation of parameter values in these systems is important to understand the relationship to their physiological properties. To achieve this, sensitivity analysis is performed. Two broad categories of sensitivity analysis are the local and global approaches. To obtain the local response of each parameter to a model, local approaches are used which is specific to a region in parameter space. For global response of parameters a global approach which spans the entire parameter space is used.

Local sensitivity analysis is a simple technique which involves calculating the partial derivative of output with respect to the parameter considered. Since single point derivatives are taken, this method is suitable for less complex cost functions. The parameter interactions are not considered. The local sensitivity is estimated by changing one input parameter by 10% at a time while the other parameters are held constant, and the changes in output variable  $V$  are compared. Eqn 2.10 shows the local sensitivity index for  $V$  with respect to the varying model parameter ( $P_i$ ), which is approximated by a small perturbation  $\Delta P_i$ ,

$$\frac{\delta V}{\delta P_i} = \lim_{\Delta P_i \rightarrow +0} \frac{V(P_i + \Delta P_i, P_{n \neq i}) - V(P_i)}{\Delta P_i}, \quad (2.10)$$

where  $V(P)$  is the model prediction of the intracellular concentration for parameter set  $P$ . The local sensitivity index is normalised to eliminate the effect of units,

$$S_i = \frac{\delta V}{\delta P_i} \frac{P_i}{V(P)}. \quad (2.11)$$

Global sensitivity analysis captures sensitivity over the entire parameter range. In biological systems, most of the models are non-linear, and variability and uncertainty of the inputs are to be considered over a wide range to estimate actual model sensitivities. As the global approach



relies on sampling the parameter space, this method becomes computationally expensive, but the method captures the relative ranking of parameter influence. Global sensitivity analysis evaluates the effect of potential interactions of the input parameters in an output variable.

Morris's screening method [42] is a sampling-based global sensitivity analysis technique. In this method, elementary effects are averaged to rank the parameter importance by computing sensitivity measures. The individual effects of parameters are evident from the mean and variance of elementary effects. Fisher Information Matrix (FIM) is another measure to capture changes in parameter values which is based on calculated output sensitivities. FIM also gives an idea of how much information can be extracted by the parameters from the experimental data [43].

The general formulation of a non-linear dynamic system is shown as

$$\dot{x} = f(x, t, p, u); y = g(x, t, p, u); x(0) = u,$$

where  $\dot{x} \in \mathbb{R}^{n_s}$  is the vector of  $n_s$  number of state variables,  $u \in \mathbb{R}^{n_u}$  is the vector of  $n_u$  number of initial values of state variables,  $p \in \mathbb{R}^{n_p}$  is the parameter vector of length  $n_p$  and  $y \in \mathbb{R}^{n_y}$  is the output vector of length  $n_y$ . The first-order sensitivity function is a sampling time dependent matrix  $S_{it}(t_k)$  such that the sensitivity value of  $i^{th}$  output variable concerning  $j^{th}$  parameter at time  $t_k$  is defined as

$$S_{it}(t_k) = \frac{\delta y_i(t_k)}{\delta p_j}.$$

To determine which of the parameters has a higher influence on the output, the Morris screening method [42] is applied. It is a derivative-based approach to analyse the sensitivity of a model function  $h : \mathbb{R}^m \rightarrow \mathbb{R}$  to variations of  $m$  parameters  $p_j \in \mathbb{R}$  where  $j = 1, \dots, m$ . Under the assumption that  $h$  is at least once differentiable, partial derivatives  $\frac{\delta h}{\delta p_j}(p)$  can be computed to define elementary effect directly. The elementary effect of the  $j^{th}$  parameter  $p_j$  on  $h$  is defined as

$$d_j(p) = \frac{h(p + \Delta e_j) - h(p)}{\Delta},$$

with step size  $\Delta \in \left\{ \frac{1}{l-1}, \frac{1}{l-2}, \dots, \frac{(l-1)-1}{l-1} \right\}$ ,  $e_j \in \mathbb{R}^m$  is a small perturbation to parameter  $p_j$  and  $l$  can be any even number.

Sensitivity analysis helps in understanding model behaviour by identifying which parameters have the most impact on model output. Parameters can be prioritised, and model robustness can be evaluated in different scenarios. More sensitive parameters are those for which small changes significantly impact the model's output, making them critical to the model's behaviour and accuracy. Less sensitive parameters are those for which changes have a minimal impact on the model's output, indicating the model's relative insensitivity to these parameters and their limited influence on its behavior and accuracy.



# Mathematical modelling of drug distribution in liver-on-chip device

## Chapter outline

- PBPK modelling of drug distribution in *in vitro* systems
- Three-compartment model for liver-on-chip device
- Prediction and translation of *in vitro* clearance to human pharmacokinetics

The drug testing dilemma presents a significant challenge in pharmaceutical development, marked by high costs and a distressing attrition rate in accurately predicting human responses [44, 45]. A pivotal element in preclinical drug development is the accurate estimation of the first-in-human dose and different dosing regimens to keep drug levels within a therapeutic range. This demands precise assessments of hepatic clearance of the drug and human pharmacokinetics [46, 47]. Typically, the gold standard in drug development is the use of simpler *in vitro* systems to study drug metabolism, including liver microsomes [48] and suspension or plated hepatocytes [49]. The drug depletion data (time-concentration profile) are then analysed to determine the *in vitro* drug clearance rate. A simple mathematical model has been employed in earlier work, and it considers the *in vitro* system as a single compartment, the one-compartment PK model [50]. Well-mixing and instantaneous drug distribution is assumed and all biological processes, e.g., permeability and partitioning from cell culture media into intracellular milieu are lumped into drug clearance. This approach also cannot differentiate between compounds actually being metabolised and compounds bound to media proteins or hardware. The so determined *in vitro* clearance value is then extrapolated to humans (*in vitro* – *in vivo* extrapolation) and integrated into human physiologically-based pharmacokinetic (PBPK) models [47, 51] to predict human pharmacokinetics (absorption,

distribution, metabolism and excretion (ADME)), before actually testing a new compound in humans. Although this approach is well-established in drug development and easy to use, it also systematically underpredicts human PK [52] by 5-10 fold across studies and compounds.

Microphysiological systems (MPS) and organ-on-chips, as well as 3D organoids, hold great promise to address more complex *in vitro* ADME (Absorption, Distribution, Metabolism and Excretion), toxicology and pharmacology questions offering miniature, biomimetic systems that replicate key aspects of human organ physiology [53, 54, 55]. These technologies create an environment where human cells can grow and interact in an organ-specific context, providing insights into human biology and disease that were previously unattainable in conventional *in vitro* models or animal studies. Organ-on-chip (OoC) and MPS-based systems are not only used in today's drug development for pharmacokinetics (PK) but also for assessing drug efficacy and toxicity [10, 53, 56, 57]. While the emulated *in vitro* biology of MPS and OoCs is getting ever more complex and produces more human-relevant data, these systems still fall short in considerably improving the prediction power of in-human situations, like PK [10, 58]. However, MPS and OoC data are also still analysed using the conventional mathematical model (one-compartment) that does not account for advanced biology. It remains unclear whether the OoC and MPS biology is still not human-relevant enough (and thus not producing human-relevant data) or the conventional mathematical analysis is the cause for the under-prediction. Potentially, a digital twin framework that enables the mapping of complex on-chip biology to advanced mathematical models could provide a useful approach to enable OoC and MPS translation to humans and increase the prediction power, but it is currently lacking.

### 3.1 Need for a digital twin to predict drug clearance in liver

The drug testing dilemma presents a significant challenge in pharmaceutical development, marked by high costs and a distressing attrition rate in accurately predicting human responses [44, 45]. A pivotal element in preclinical drug development is the accurate estimation of the first-in-human dose and different dosing regimens to keep drug levels within a therapeutic range. This demands precise assessments of hepatic clearance of the drug and human pharmacokinetics [46, 47]. Typically, the gold standard in drug development is the use of simpler *in vitro* systems to study drug metabolism, including liver microsomes [48] and suspension or plated hepatocytes [49]. The drug depletion data (time-concentration profile) are then analysed to determine the *in vitro* drug clearance rate. A simple mathematical model has been employed in earlier work, and it considers the *in vitro* system as a single compartment, the one-compartment PK model [50]. Well-mixing and instantaneous drug distribution is assumed, and all biological processes, e.g., permeability and partitioning from cell culture media into intracellular milieu, are lumped into drug clearance. This approach also cannot

differentiate between compounds actually being metabolised and compounds bound to media proteins or hardware. The so determined *in vitro* clearance value is then extrapolated to humans (*in vitro* – *in vivo* extrapolation) and integrated into human physiologically-based pharmacokinetic (PBPK) models [47, 51] to predict human PK (ADME), before actually testing a new compound in humans. Although this approach is well-established in drug development and easy to use, it also systematically underpredicts human PK [52] by 5-10 fold across studies and compounds.

The current study aims to develop a digital twin approach integrating MPS and OoC data within advanced computational models of biology to improve the prediction of clinical clearances. DigiLoCs, our developed digital liver-on-chip simulator, facilitates the accurate description of on-chip complex biology to disentangle biological processes, namely clearance, permeability, and partitioning. The tool comprises and utilises information on complex biological processes (clearance, permeability, partitioning), hardware-specific information from the studied *in vitro* system, and compound-specific information. By accounting for more multi-dimensional information, the tool enables differentiation between active biological processes, such as metabolism, and passive ones, such as permeability and partitioning of a compound from cell culture media into the cellular environment. This tool offers a significant advancement over conventional approaches, which fail to explicitly consider passive biological processes and conflate them into a singular clearance mechanism. By providing a more detailed understanding of biological processes, our tool has the potential to reveal better insights into liver-chip biology. Drug depletion kinetics of 32 compounds were taken from literature covering commercially available liver-on-chips (CnBio [10, 11], Javelin [59]), and 3D spheroids [14, 16], including fast and slow-cleared compounds. According to these studies, DigiLoCs outperform the conventional prediction approach considerably. The impact of a more accurate description of clinical clearance values on predicting human PK was investigated in a proof-of-concept study using propranolol. The kinetics of propranolol was predicted in humans using the conventional approach, DigiLoCs and literature approach. The results obtained from DigiLoCs for propranolol in the proof-of-concept study were much closer to the actual observed human values than those of other approaches.

From the survey of the literature, this appears to be the first and most extensive study so far, comparing head-to-head the performance of different hepatic *in vitro* systems to predict human clearance and demonstrating the impact OoC and MPS systems can have in the drug development process, enhanced through the modelling and prediction features of DigiLoCs.

## 3.2 Methodology

In this section, the following are described: *i)* data used in the study for predicting human clearance, *ii)* DigiLoCs, digital twin for liver-on-chip, *iii)* mathematical model, parameter estimation and sensitivity analysis for DigiLoCs and *iv)* translation to humans and prediction of human pharmacokinetics.

### 3.2.1 Dataset used for parameter estimation

In this work, published data on pharmacokinetics (metabolism) or toxicology studies of 32 drugs are used (See Table 3.1) to predict human PK.

Study	Drugs	Used in this study	<i>In vitro</i> system	cell number [a.u.]	Media volume [ml]	Flow	On-chip compartments
Docci et al. [10]	9	9	CnBio Liver-Chip	3E5	1.60	Recirculation	1
Tsamandouras et al. [11]	6	3	CnBio Liver-Chip	3E5	1.60	Recirculation	1
Rajan et al. [59]	12	8	Javelin Liver-Chip	2.15E5	1.30	Recirculation	2
Kanebratt et al. [14]	4	4	3D Spheroid (Hurel)	6E3	0.05	No	1
Bonn et al. [16]	8	8	3D Spheroid (Hurel)	3E4	0.10	No	1
Total	38	32					

Table 3.1: Overview of literature reports providing on-chip pharmacokinetic information on compound clearance

### 3.2.2 Mathematical model

A typical single-compartment model is described as follows. Let  $C(t)$  be the drug concentration in the chip at time  $t$ ,  $V$  the volume of the chip, and  $CL_c$  the clearance parameter. Then we have

$$\frac{dC}{dt} = -\frac{CL_c}{V} \cdot C,$$

with initial value  $C(t = 0) = C_0$ . The solution to this ODE is

$$C(t) = C_0 \cdot e^{-\frac{CL_c}{V} \cdot t}. \quad (3.1)$$

Taking the logarithm on both sides,

$$\log C(t) = -\frac{CL_c}{V} \cdot t + \log C_0, \quad (3.2)$$

which is equivalent to regression on log-transformed kinetic data.

A digital twin of liver-on-chip with three compartments that incorporate much more information on parameters related to both on-chip characteristics and drug-specific properties was developed. The three-compartment model considering media, interstitium and intracellular compartment is described as follows: Let  $C_m(t)$ ,  $C_i(t)$ ,  $C_c(t)$  and  $V_m$ ,  $V_i$ ,  $V_c$  be the

concentration of the drug at time  $t$ , and volume of the media, interstitium, and intracellular compartment respectively.  $CL_c$  is the clearance parameter. Then we have

$$\frac{dC_m}{dt} = -\frac{k_1}{V_m} \cdot C_m + \frac{k_2}{V_m} \cdot C_i; \quad (3.3)$$

$$\frac{dC_i}{dt} = \frac{k_1}{V_i} \cdot C_m - \left( \frac{k_2 + k_3}{V_i} \right) \cdot C_i + \frac{k_4}{V_i} \cdot C_c; \quad (3.4)$$

$$\frac{dC_c}{dt} = \frac{k_3}{V_c} \cdot C_i - \left( \frac{k_4 + CL_c}{V_c} \right) \cdot C_c. \quad (3.5)$$

The parameters here are defined as follows,

$$k_1 = F_{\text{unbound}} \cdot P_{\text{endothelial}} \cdot SA_{\text{med\_int\_liver}}, \quad (3.6)$$

$$k_2 = \frac{k_1}{K_{\text{int\_med}}},$$

$$k_3 = K_{\text{water\_int}} \cdot PA_{\text{int\_cell}} \text{ and}$$

$$k_4 = K_{\text{water\_cell}} \cdot PA_{\text{cell\_int}}.$$

where  $K_{\text{int\_med}}$  is the partition coefficient for the transfer of drug between the media and interstitium,  $P_{\text{endothelial}}$  is the permeability coefficient of the drug between the endothelial layer,  $SA_{\text{med\_int\_liver}}$  is the surface area of the interstitium,  $K_{\text{water\_int}}$  is the partition coefficient for water exchange or movement within the interstitium,  $K_{\text{water\_cell}}$  is the partition coefficient for water exchange or movement within the intracellular compartment,  $PA_{\text{int\_cell}}$  is the product of the surface area and permeability coefficient of the cellular membrane in the interstitium compartment,  $PA_{\text{cell\_int}}$  is the product of the surface area and permeability coefficient of the cellular membrane in the intracellular compartment and  $F_{\text{unbound}}$  is the fraction unbound (media, reference value). All the relevant variables and parameters with descriptions are given in Table 3.2.

We can write this system of linear ordinary differential equations (ODEs) in matrix form,

$$C' = A \cdot C, \text{ where} \quad (3.7)$$

$$A = \begin{bmatrix} -\frac{k_1}{V_m} & \frac{k_2}{V_m} & 0 \\ \frac{k_1}{V_i} & -\left(\frac{k_2+k_3}{V_i}\right) & \frac{k_4}{V_i} \\ 0 & \frac{k_3}{V_c} & -\left(\frac{k_4+CL_c}{V_c}\right) \end{bmatrix}.$$

The general solution for the system of ODEs at time  $t$  is:

$$C(t) = e^{A \cdot t} = e^{X \cdot \Lambda t \cdot X^{-1}} \cdot C_0 = X \cdot e^{\Lambda t} \cdot X^{-1} \cdot C_0, \quad (3.8)$$

where  $X$  is the matrix of eigenvectors of  $A$ ,  $\Lambda$  is the diagonal matrix with  $\lambda_1, \lambda_2, \lambda_3$  (eigenvalues of  $A$ ) as diagonal entries, and  $C_0$  is the initial value of variables at time 0.

Parameter	Description	Unit	Status
$K_{\text{int\_med}}$	Partition coefficient for the transfer of drug between the medium and interstitium	dimensionless	calculated
$K_{\text{water\_int}}$	Partition coefficient for water exchange or movement within the interstitium	dimensionless	calculated
$K_{\text{water\_cell}}$	partition coefficient for water exchange or movement within the intracellular compartment	dimensionless	calculated
$P_{\text{endothelial}}$	The rate of passive diffusion across the endothelial layer, which is the product of surface area and the permeability coefficient of the drug	cm/min	estimated
$PA_{\text{int\_cell}}$	Permeability coefficient of the cellular membrane in the interstitium compartment	mL/min	calculated
$PA_{\text{cell\_int}}$	Permeability coefficient of the cellular membrane in the intracellular compartment	mL/min	calculated
$F_{\text{unbound}}$	Fraction unbound (medium, reference value)	dimensionless	calculated
$SA_{\text{med\_int\_liver}}$	Surface area of the interstitium	cm <sup>2</sup>	calculated
$CL_c$	Intrinsic on-chip clearance	mL/min	estimated
<b>Variables</b>			
$C_m$	Concentration of the drug in the media compartment	$\mu\text{mol/mL}$	simulated
$C_i$	Concentration of the drug in the interstitium compartment	$\mu\text{mol/mL}$	simulated
$C_c$	Concentration of the drug in the intracellular compartment	$\mu\text{mol/mL}$	simulated

Table 3.2: Names and description of all relevant variables and parameters in the three-compartment digital twin

The objective function for optimisation is as follows:

$$\underset{CL_c}{\text{minimise}} \sum_{i=1}^n (y_{i_{\text{obs}}} - y_{i_a})^2 \left\{ \begin{array}{l} y_{i_a} = e^{A \cdot t_i} \cdot C_0 \\ \text{computed using eigenvalues} \\ \text{and eigenvectors,} \end{array} \right.$$

where  $y_{i_{\text{obs}}}$  is the observed data point and  $y_{i_a}$  is the computed value at time  $i$  respectively and  $n$  is the number of observed data points.

### 3.2.3 DigiLoCs: Digital Twins for cell-based liver assays

DigiLoCs is a software tool, developed within this work that describes the on-chip complex biology more accurately in the context of use to predict clinical clearance values. The software comprises (Figure 3.1):

- modelling of complex biological processes (clearance, permeability, partitioning),
- hardware-specific information from the studied *in vitro* system and
- compound-specific information.

The tool differentiates between active biological processes, such as metabolism, and passive ones, like permeability and partitioning of a compound from cell culture media into the cellular milieu. This contrasts with conventional approaches, where passive biological processes are not considered especially and lumped together into a single process, i.e., clearance.

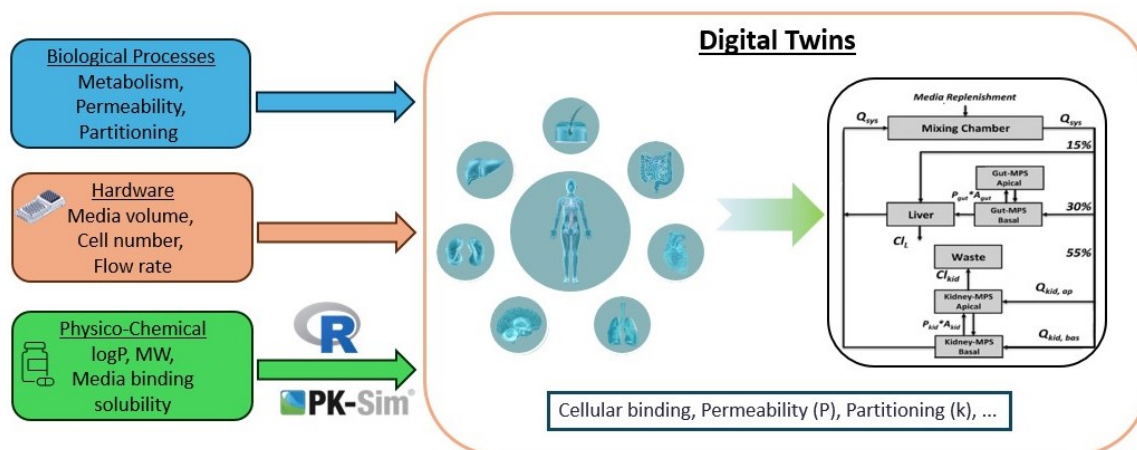


Figure 3.1: Digital twin (DT) approach. Contrasting conventional approach, the DT approach uses biological, hardware, and physicochemical information to map the biological processes on liver-on-chip more accurately to *in silico*, thereby maximising the information leveraged. This results in the disentanglement of active (metabolism) and passive (permeability, partitioning) processes. Image adapted from [60].

Liver-on-chip technology provides a more physiologically relevant environment compared to traditional cell cultures or animal models, enhancing the simulation of drug responses using mathematical models. Hence, a more accurate mathematical description is needed. The three primary compartments of the liver chip considered in the model are media, interstitium, and intracellular space, which serve as dynamic environments where drugs are distributed, metabolised, and interact with hepatic cells. This compartmentalisation is based on concepts applied in human whole-body PBPK modelling.

The software tool is developed in the open-source programming language R [61] and seamlessly communicates with PK-Sim (<https://www.open-systems-pharmacology.org/>) via in-house developed functions. For more information, see esqlabsR package (<https://github.com/esqLABS/esqlabsR>). All analysis and plotting were also done in R. The proposed workflow does not interfere with existing wet lab standard operating procedures (SOPs) for performing biological experiments and does not add an extra considerable burden to the user. DigiLoCs uses existing biological data, and its performance may be improved by measuring cell-associated compound concentrations in addition to the compound media depletion time course, which would add a minor extra step in the lab SOP. This, however, is negligible given the improvement in performance power and the confidence in the prediction.

**Implementation of hardware specifications** DigiLoCs map the chip architecture to a compartmental model to describe the time-dependent distribution of a compound on-chip. The compartment models use time-dependent ODEs and assume well-mixing within compartments. These are generally accepted to describe the distribution of exogenous and endoge-



nous compounds and molecules. A physical chamber separated by a membrane or connected by flow to another chamber is represented by a compartment in the software. Serial compartments are connected via media flow rates (typically in  $ml/min$ ) between the compartments describing mass transport via gradients and diffusion and normalised by the volume of the originating compartment.

**Implementation of biological specifications** The biology (more precisely, the cell type exerting the biological function under investigation; here: metabolism) is mapped by two additional compartments representing the interstitial and intracellular space of the investigated biology (Figure 3.1). Transport rates from the cell culture media into the interstitial and intracellular milieu are described by two core processes:

- passive diffusion (driven by concentration gradients between compartments)
- thermodynamic equilibrium (distribution of compound between two phases)

These processes are then described by two main parameters in the computational model: permeability (how fast is a compound taken up?) and partitioning (how much of the compound is taken up by cells?). Lastly, the metabolism rate is reallocated to the intracellular compartment and corrected by the unbound fraction of the compound in the intracellular compartment. Similarly, in pharmacokinetics, the distribution of compounds is often described using a compartmental framework, which involves dividing a system into distinct compartments and modelling the processes that govern the movement of compounds between them.

**Implementation of compound-specific information** The following physicochemical properties of the investigated compounds are used in the software: *i*) lipophilicity ( $\log P$ ), *ii*) molecular weight (MW) and *iii*) fraction unbound ( $f_u$ ); to calculate up to six dependent downstream parameters (listed below). These parameters describe the partitioning from the main media compartment into the interstitial space and between the water fraction and both interstitial and intracellular space. Additionally, permeability across the endothelial barrier and between interstitial and intracellular spaces are calculated.

Partitioning: *i*)  $K_{int\_med}$  *ii*)  $K_{water\_cell}$  *iii*)  $K_{water\_int}$

Permeability: *i*)  $P_{endothelial}$  *ii*)  $PA_{cell\_int}$  *iii*)  $PA_{int\_cell}$

Here *int* refers to interstitial space, *med* refers to media, *water* refers to water exchange fraction, and *cell* refers to intracellular space. These parameters are calculated based on well-established and documented equations implemented in PK-Sim software [46], which is a comprehensive software tool for whole-body PBPK modelling. It enables rapid access to all relevant anatomical and physiological parameters for humans and common laboratory animals contained in the integrated database for model building and parameterisation. The



same partition coefficient calculation methods as implemented in PK-Sim are also readily available and can be investigated:

- PK-Sim standard
- Poulin and Theil
- Rodgers and Rowland
- Schmitt
- Berezhkovskiy

Further, only the unbound fraction of a compound can be taken up by cells and be metabolised by cells. The unbound fraction in the cell culture media is typically informed by biological experiments. However, the intracellular unbound fraction is not often available or measured. Thus, two established QSAR models (quantitative structure-activity relationship) are implemented in the software to predict the unbound intracellular fraction of the investigated compound as a function of its physicochemical properties [62, 63].

**Implementation of software** Methodologically, DigiLoCs is implemented in the open-source programming environment R with its own package. A library of two common chip architectures and two cell types with six different chip-specific settings has already been implemented.

1. One chamber, no media flow
2. Two chamber, recirculating flow
3. Organ-on-chip (hepatocytes)
  - a. CnBio
  - b. Hurel 1
  - c. Hurel 2
  - d. Dynamic42
  - e. Javelin

The most straightforward system is a single, perfused microfluidic chamber containing one kind of cultured cell (e.g., hepatocytes) that exhibits functions of one tissue type linked to channels for fluid transport. In more complex designs, two or more microchannels are connected by porous membranes, lined on opposite sides by different cell types, to recreate interfaces between different tissues. The CnBio and Javelin chip settings are shown in Figure 3.2. For detailed information on the processes and modelling, readers are directed to Bhatia et al. [65]. These building blocks can be interchangeably used and connected, similar to the

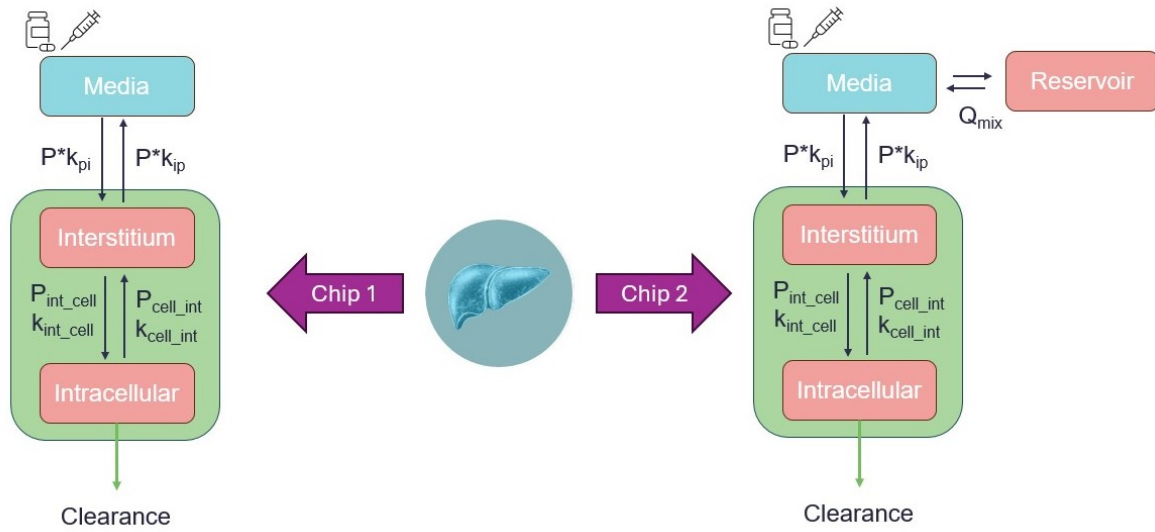


Figure 3.2: Common chip architectures: Chip 1 is for CnBio [10] and 3D spheroids [64] and chip 2 is for Javelin [59] architectures.  $Q_{mix}$  is the mixing flow rate in mL/min.

building blocks in PK-Sim. While the R code and its package provide a step-by-step guide to generate and run a simulation, the code communicates seamlessly with a generic PK-Sim model to determine partitioning and permeability values as described above, which are used in the simulation.

### 3.2.4 Parameter estimation

Parameter estimation aims to find unknown parameters in a computational model and is estimated using experimental data collected from well-defined and standard conditions. By minimising the distance of theoretical function values and experimentally known data, the set of parameters in the model can be estimated. The parameters which are not directly measurable can be estimated using least squares or any other fitting methods to analyse the model quantitatively. Nominal parameter values are obtained from PK-Sim software.

Parameter estimation in DigiLoCs is a two-step process. Firstly, a customised cost function is implemented. This cost function calculates the weighted difference (ssq) between the model simulation (pred) from a specific compartment and the corresponding observed data (obs) for each time point according to the equation

$$ssq = \frac{obs - pred}{pred}. \quad (3.9)$$

Common parameter estimation methods include maximum likelihood estimation and Nelder-Mead optimisation. Nelder-Mead, a non-linear optimisation method, is used to find the minima of the objective function in this work. Additionally, the partition coefficient between the intracellular (IC) and the main media compartment is estimated using the area under the

curve (AUC) of simulated time-concentration profile of the IC and interstitial (Ist) compartment. This value is corrected for by the QSAR-predicted cellular unbound fraction ( $fu_{cell}$ ) and the unbound fraction in the media (measured,  $fu_{media}$ ) as follows

$$K_{P_{uu,pred}} = \frac{AUC(Ist + IC)}{AUC(media)} \cdot \frac{fu_{cell}}{fu_{media}}. \quad (3.10)$$

$K_{P_{uu,obs}}$  is calculated from literature [66], where media and intracellular concentrations in hepatocytes were determined. Initially investigated for suspension hepatocytes in a 2D setting, the authors provide a scaling factor ( $\sim 4.9$ ) to apply to human hepatocytes. Further, the ionisation state of the investigated compound (-1, 0, 1) results in a different partitioning. Otherwise, a range of possible partition coefficients are investigated. This is an additional anchor point for estimating the cost function value and links the simulated intracellular and main media compartment concentrations. Eventually, both differences are squared and summed up, resulting in the final sum of residuals. Based on this, a compound-specific scaling factor (SF) is calculated and used to scale the predicted human clearance:

$$SF(drug) = \frac{K_{P_{uu,obs}}}{K_{P_{uu,pred}}}. \quad (3.11)$$

Specifically for the liver use cases, on-chip liver clearance and surface area between the main media compartment and the cell layer are estimated. It is possible to estimate other parameters, such as pre-calculated permeability or partition coefficient values.

### 3.2.5 Translation to humans

Drug-related parameters extracted from OoC or any other *in vitro* studies can be scaled to predict clinical parameters using *in vitro-in vivo* extrapolation (IVIVE) [10, 11]. The typical value of unbound intrinsic clearance  $CL_{int(u)}$  determined for each drug from the pharmacokinetic analysis of the *in vitro* depletion data is scaled up to a human liver equivalent unbound intrinsic clearance  $CL_{int(u),H}$  using

$$CL_{int(u),H} = \frac{CL_{int(u)} \cdot HC \cdot LW}{fu_{inc}}, \quad (3.12)$$

where HC is the human hepatocellularity of 120 million cells / g of liver, LW is the average human liver weight of 25.7g / kg of body weight [11] and  $fu_{inc}$  is the unbound fraction of drug in the incubation medium. The hepatic clearance (referring to whole blood concentrations) is then predicted ( $CL_{H,pred}$ ) using the Well-Stirred (WS) model:

$$CL_H(pred, WS) = \frac{Q_H \cdot fu_b \cdot CL_{int(u),H}}{Q_H + fu_b \cdot CL_{int(u),H}}, \quad (3.13)$$

where  $Q_H$  is the average hepatic blood flow of 20.7 mL/min/kg of body weight and  $fu_b$  is the fraction of the drug unbound in blood. The fraction unbound in the blood was calculated for

each compound from the known fraction unbound in the media ( $f_{u_p}$ ) and blood-to-plasma ratio ( $R_{bp}$ ) according to the equation  $f_{u_b} = f_{u_p}/R_{bp}$  or directly used, if available from the literature. The predicted hepatic clearance ( $CL_{H,pred}$ ) values were then compared to observed hepatic clearance ( $CL_{H,obs}$ ) values (referring to whole blood concentrations). Following, the ratio of predicted (either via conventional or digital twin approach) and observed clinical clearance values for all investigated compounds was calculated. Using these ratios, a density distribution function was computed (function `geom_density` from `ggplot` [67]) for visual purposes only.

### 3.2.6 Prediction of human pharmacokinetics

Initially, a PBPK model is developed using qualified installations of the software PK-Sim that ensures the software has been installed correctly and thoroughly validated according to established procedures. A whole-body PBPK model includes an explicit representation of the organs most relevant to the uptake, distribution, excretion, and metabolism of the drug. These typically include the heart, lungs, brain, stomach, spleen, pancreas, intestine, liver, kidney, gonads, thymus, adipose tissue, muscles, bones, and skin. More information can be found in Appendix A.

The tissues are interconnected by arterial and venous blood compartments, and each is characterised by an associated blood flow rate, volume, tissue partition coefficient, and permeability. If applicable, R (Distribution 4.0) and RStudio (Version 1.2.5) are used in the analysis for preprocessing and post-processing of data and model outputs [68]. The analytical approach is based on the principles set out in the EMA, FDA, and/or OECD guidelines for reporting on PBK M&S [54]. The developed PBPK model is used to describe the human kinetics of propranolol. Key kinetic parameters are informed by either clinical data, literature values or on-chip predictions. The translational workflow that integrates organ-on-chip results to predict human pharmacokinetics is shown in Figure 3.3.

## 3.3 Results of DigiLoCs

We report the following results in this section. First, results from a simulation model based on a digital twin for selected compounds; second, sensitivity analyses; third, prediction of human clearance values; and finally, translation to human PK using propranolol as a proof-of-concept study. The liver clearance and surface area of the chip are estimated after fitting the drug kinetic data. The Poulin and Theil method of partition coefficient calculation was used due to its superior fit to observed drug kinetics, which outperformed alternative methods.

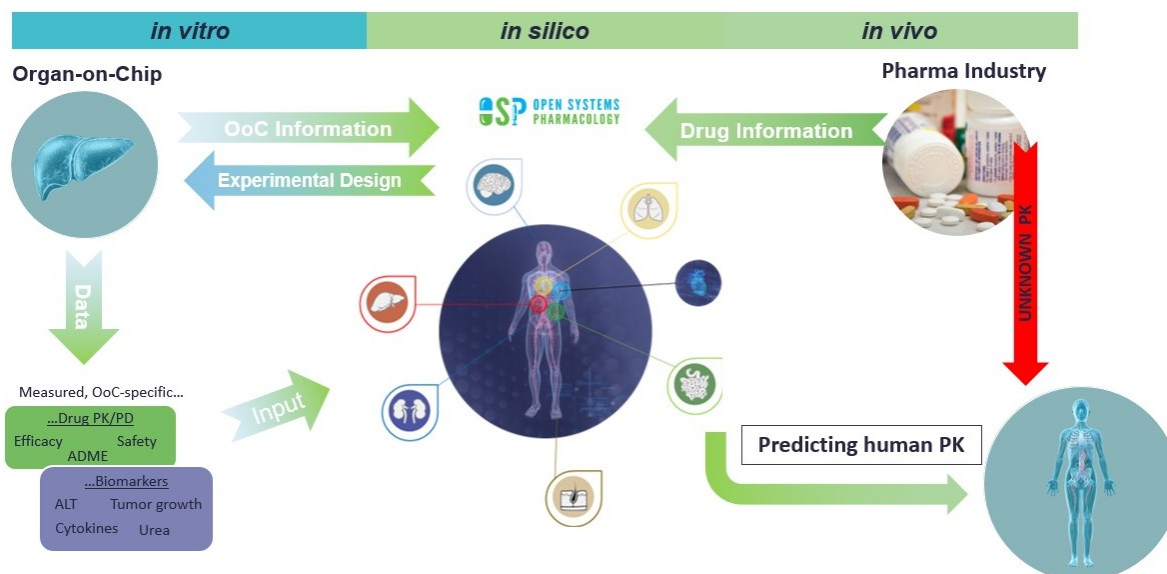


Figure 3.3: Translational workflow plan that integrates results from organ-on-chip with computer modelling to predict the kinetics of drugs in humans. The digital twins of the humanised organ-on-chip systems, together with chip-specific information and physicochemical information, are developed in R.

### 3.3.1 Simulating on-chip compound depletion

The digital twins for the investigated *in vitro* liver systems were successfully implemented in R and used to simulate the on-chip kinetics. After parameter estimation, the resulting model simulations describing the observed compound depletion data were visually inspected. The final parameter values can be found in Table A.2. Additionally, the squared sum of residuals was evaluated and deemed acceptable as it was less than  $< 0.01$ , which was the case for all simulations (data not shown). An example of on-chip kinetics is presented in Figure 3.4. As can be seen, the digital twin approach (violet line) captures the on-chip kinetics (blue dots) very well. Simultaneously, the intracellular (IC) kinetics are plotted (red lines), clearly highlighting the difference in compound uptake and, thus, clearance rates. The remaining figures are presented in Appendix (See Figures A.1, A.2, A.3 and A.4).

### 3.3.2 Sensitivity analysis of three-compartment model

The sensitivity analysis, both local and global, was conducted to quantify the sensitivity of model output intracellular concentration with input parameters. The input parameters  $K_{\text{int\_med}}$ ,  $P_{\text{endothelial}}$ ,  $SA_{\text{med\_int\_liver}}$ ,  $K_{\text{water\_int}}$ ,  $K_{\text{water\_cell}}$ ,  $PA_{\text{cell\_int}}$ ,  $PA_{\text{int\_cell}}$ ,  $f_u$  and  $CL_c$  are varied to evaluate the sensitivity of the output variable intracellular concentration ( $C_c$ ). The Sobol sensitivity analysis of the SALib package in Python is used to perform the global sensitivity analysis. Input parameters are sampled using the Saltelli sampler. The lower and upper bound

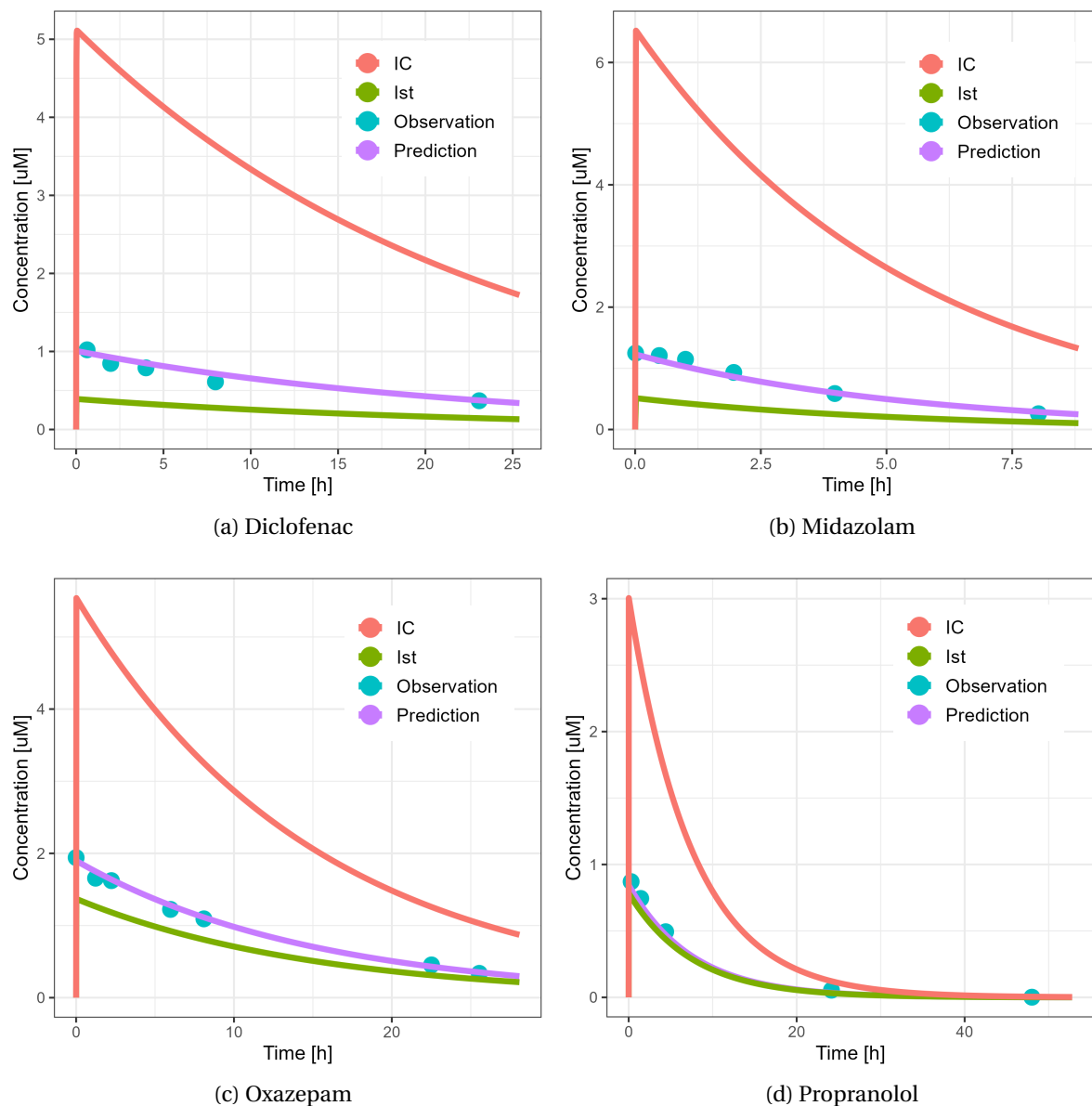
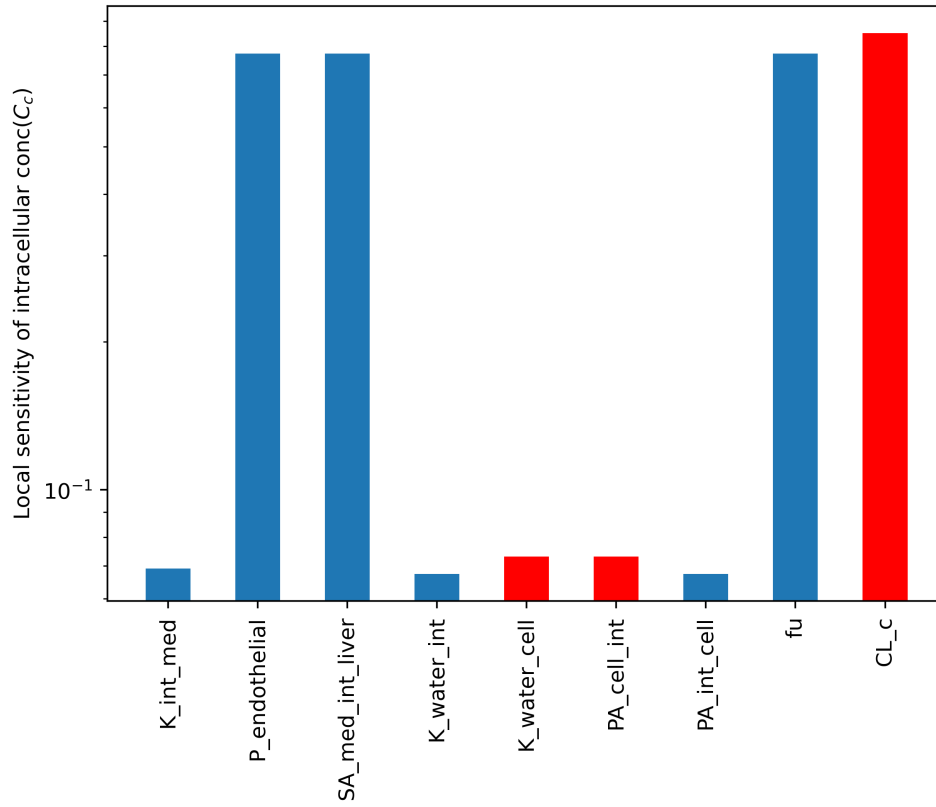
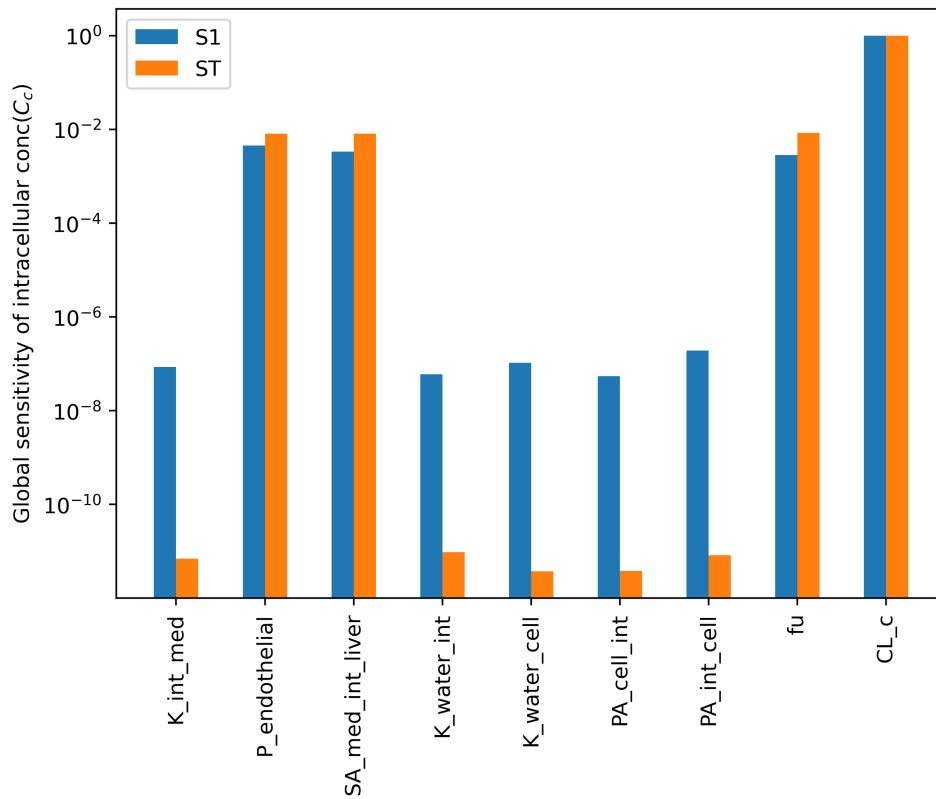


Figure 3.4: Digital twin-based model simulation of on-chip kinetics after fitting parameters for selected compounds. Observed data are shown in blue dots, where the data for diclofenac, midazolam, and oxazepam are from Docci et al. [10], while propranolol is from Tsamandouras et al. [11]. The red, violet and green curve plots the drug concentration in the intracellular, medium and interstitium compartments, respectively: IC = intracellular, Ist = interstitium.

of the parameters are set as 0.1-fold and 10-fold of the baseline parameter values, respectively. The first-order and total-order indices are estimated using the Sobol sensitivity analysis. The analyses were performed for various parameters, and the results indicated that the output is more sensitive to parameters such as the permeability coefficient of the endothelial layer, surface area of the liver sinusoids and clearance. These parameters were estimated or calculated from clinical data/ experimental results. Clearance ( $CL_c$ ) is identified as the most sensitive parameter with respect to intracellular concentration. The results imply that accurate values of these sensitive parameters are crucial for the model's accuracy.



(a) Local sensitivity analysis



(b) Global sensitivity analysis

Figure 3.5: Local and global sensitivity of the parameters with respect to output intracellular concentration. (a) Blue bars indicate that the output and the input change in the same direction, and the red bar indicates that the output decreases when the input increases. (b) The blue and orange bars represent first-order and total-order indices, respectively.

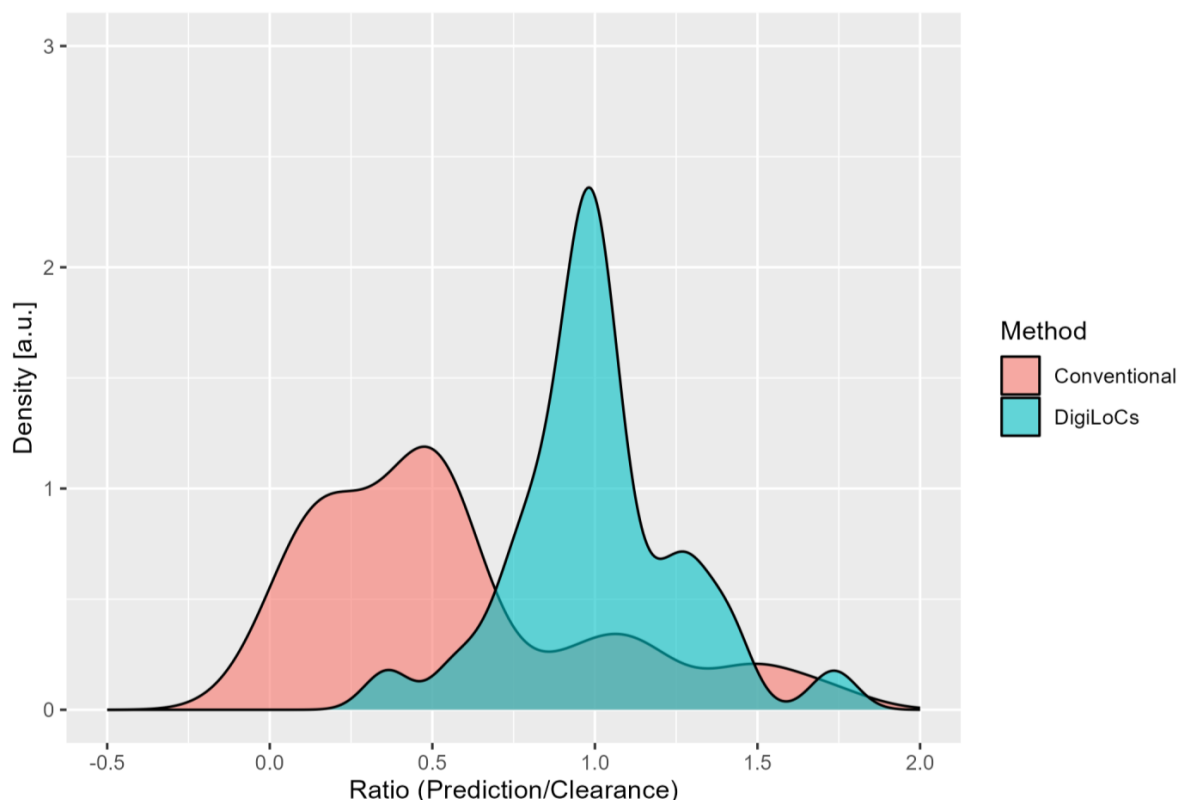


Figure 3.6: Impact of DigiLoCs on predicting clinical clearance values compared to the conventional approach. In total, a set of 32 compounds across three different *in vitro* liver systems have been investigated. The x-axis presents the ratio of predicted/observed clinical clearance values using either the DigiLoCs or the conventional approach, and the y-axis shows the frequency of the ratio observed.

The normalised local sensitivity indices (Figure 3.5(a)) and the first-order and total-order global sensitivity indices (Figure 3.5(b)) for intracellular concentration across the input parameter set are shown. The results from both local and global sensitivity analyses show that the output is more sensitive to parameters  $P_{\text{endothelial}}$ ,  $SA_{\text{med\_int\_liver}}$ ,  $f_u$  and  $CL_c$ .  $SA_{\text{med\_int\_liver}}$  and  $CL_c$  were estimated and nominal values were used for all other parameters. SA results imply that we need correct values of the constants  $P_{\text{endothelial}}$ ,  $f_u$  as they are more sensitive.

### 3.3.3 Predicting human clearance

The on-chip estimated clearance values were translated to total human clearance according to Eqn 3.13. More detailed information is available in the Appendix (see Tables A.1 and A.2). Likewise, from the investigated studies (Table 3.1), *in vitro* unbound clearance values were available and scaled to human equivalents.

The conventional approach is a single-compartment model explained in Section 3.2.2 using a single ODE. The ratio of clinical observed human clearance values and either predicted human clearances using conventional mathematical modelling or the digital twin approach



were estimated and converted into a density function for easier graphical visualisation as described in Section 3.2.5. As can be seen in Figure 3.6, the digital twin approach (DigiLoCs) outperforms the conventional approach considerably. The centre of the distribution is around 1, indicating a non-biased prediction of clinical clearance values, while the width of the distribution is very small. Quantitatively, the ratio for the digital twin approach overall compounds is  $1.04 \pm 0.31$ , with a coefficient of variation of 30%. In contrast, the conventional approach (red curve) majorly under-predicts the clearance values while maintaining a broad distribution and thereby adding to uncertainty in the prediction ( $0.56 \pm 0.44$ , CV = 79.3%). The correlation plot between the observed and herein predicted clinical clearance values highlights the improved prediction performance of the DigiLoCs approach on an individual drug level. As can be seen in Figure 3.7 (example graph for CnBio Liver-on-Chip data), most of the compounds fall within the 1.5-fold line (Average fold error, AFE = 0.965). Similar correlation plots are presented in the Appendix (See Figures A.5, A.6 and A.7) for the other *in vitro* systems.

### 3.3.4 Translation to human PK

The impact of accurately predicting human clearance values based on *in vitro* cell-based assays on predicting human PK was assessed using propranolol as a proof-of-concept case study. First, a human PBPK model describing the human kinetics of propranolol was implemented in PK-Sim and qualified with clinical observations. Next, the predicted human clearance value using either the conventional modelling approach or based on the same on-chip kinetic data was implemented in the human PBPK model, simulating the kinetics after a single oral dose. Further, a population of  $n = 1000$  patients was simulated to account for inter-patient variability. As shown in Figure 5, the implemented human PBPK model describes observed clinical data well (using clinical clearance values). When substituting only the clearance value with the conventional or the digital twin-based values, the impact on simulating human PK becomes apparent, while the conventional approach would overpredict (i.e., the on-chip clearance is underpredicted) the human PK (3-fold  $C_{\max}$ , up to 6-fold overprediction of AUC). Moreover, this approach would actually simulate non-negligible concentrations of propranolol left over after 24 h. Repeated daily dosing would result in the accumulation of propranolol in this hypothetical setting, which would have immediate implications for potential toxicity or efficacy considerations. On the other hand, the digital-twin-based approach still slightly overpredicts the AUC and  $C_{\max}$ ; however, only by 1.5-fold and captures the terminal phase correctly.

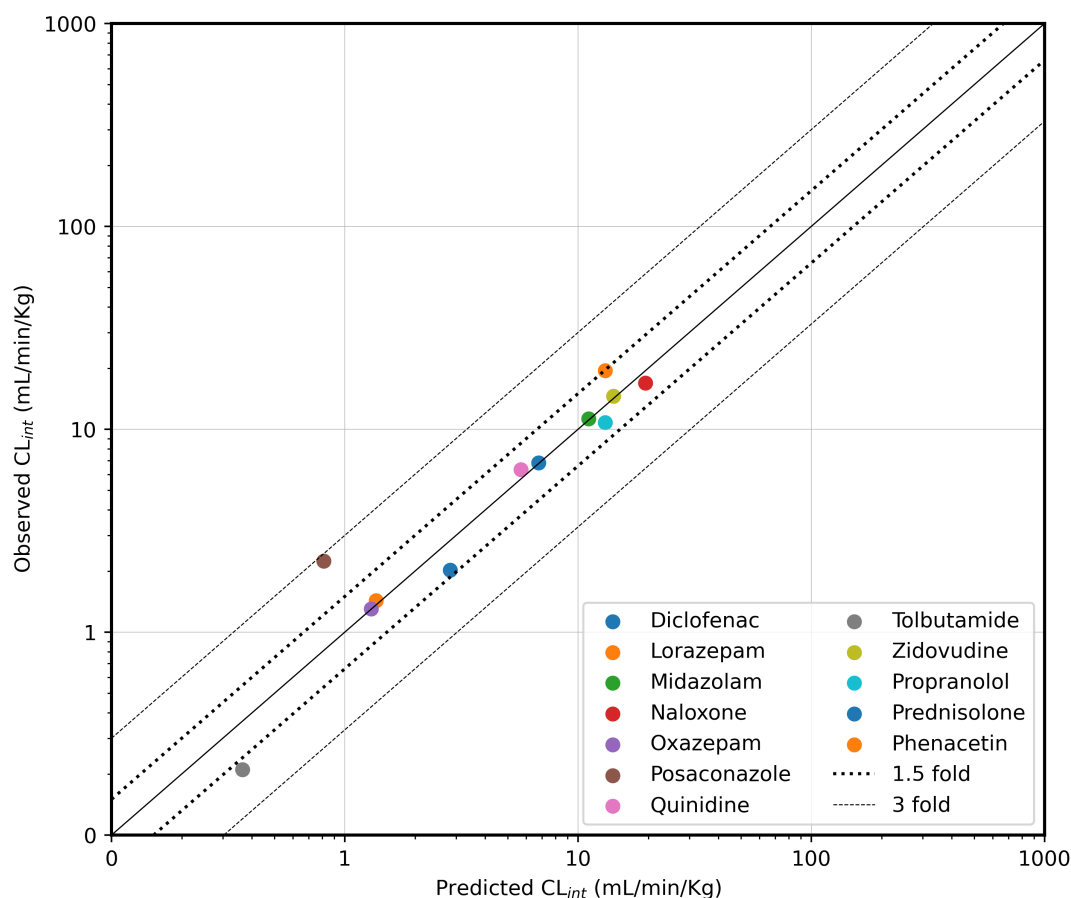


Figure 3.7: Correlation between observed and predicted *in vivo* intrinsic clearance ( $CL_{int}$ ) using the three-compartment liver chip for 12 drugs (Docci et al. 2022; Tsamandouras et al. 2017). The solid line shows the line of unity, while the dotted line is 1.5-fold, and the dashed line has a 3-fold deviation.

### 3.4 Discussion on DigiLoCs

The aim of this work is to improve the current prediction of human clearance values and to present a framework for translating *in vitro* findings to relevant clinical situations. The presented integrated translational approach combined quantitative OoC and cell-based assay compound depletion kinetics with an OoC-digital twin to simulate drug kinetics in humans.

Initial investigations revealed the potential to describe clinical clearance values more appropriately than is currently possible with the conventional approach. This simpler approach lumps biological processes together into a single process – clearance – and uses only minimal information available, e.g., only the cell number and media volume. While biological systems have evolved rapidly in the last decade, especially in the field of organ-on-chip and microphysiological systems, the applied mathematical models to analyse the quantitative

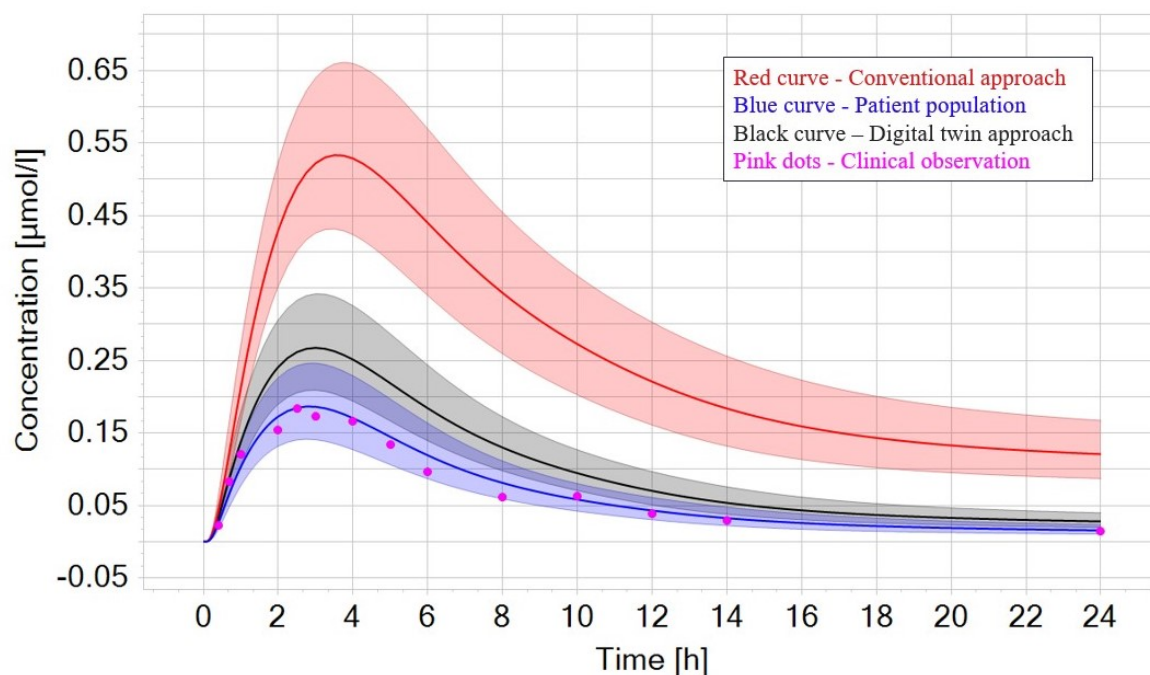


Figure 3.8: Simulated kinetics of propranolol after a single oral dose (80 mg). Pink dots are clinical observations (digitised from Borgström et al. [69]), while the blue solid line represents the mean of the patient population using the clinical observed clearance value. When using the clearance value (red line) based on conventional approaches, the area under the curve is 6-fold overpredicted. In contrast, using the digital twin-based clearance, the AUC is only 1.5-fold overpredicted, also simulating the right kinetics at 24 h (black curve). Shaded areas represent  $\pm 1$  SD.

complex biological data have been the same for decades (the early concept of clearance was introduced by Möllers in 1928, while Well-stirred model was introduced in 1971).

In contrast, the digital twin approach for the organ-on-chip and 3D spheroids comprises three building blocks described here: biological, hardware, and physicochemical information. The distinction between active and passive processes is achieved by an explicit description of uptake, distribution, and metabolism involved in the biological processes. Further, the digital twin links the architecture of the hardware chip with an advanced mapping of the underlying biology (intracellular compartment). The on-chip kinetics for 32 compounds (six compounds were removed from the initial set due to missing information) was well described, highlighting the drug-specific effects on cellular uptake and, hence, metabolism. Note that this analysis used the same biological information as used in the conventional approach and that no additional biological experiments were needed or performed to improve the outcome of the digital twin approach. The predictive power of organ-on-chip and 3D spheroids over conventional approaches was revealed when the depletion data was analysed with the digital twins (Figure 3.6). Not only was the systematic underprediction issue resolved, but the uncertainty in prediction was also reduced by a factor of 3 (comparing CVs).

Lastly, we aimed to demonstrate the clinical impact of this approach by translating the results to humans using propranolol as a proof-of-concept example. Here, the head-to-head comparison clearly demonstrated the superior power of both quantitative biological data from OoCs and digital twins over conventional approaches in predicting human PK more appropriately (Figure 3.8). Although only one compound was used to demonstrate clinical impact, the workflow and process are easily applicable to other compounds. To the best of our knowledge, this study is the biggest comprehensive report to systematically assess the predictive power of organ-on-chip in the context of the use of liver clearance.

The mathematical algorithm to determine liver clearance depends on the time-concentration profiles and, if available, on intracellular or cell-associated compound concentrations. The algorithm minimises a cost function by identifying a clearance value such that model prediction and data observation match. The cost function takes both the PK profile from cell culture media and the cell-associated levels into account, which is not the case for the conventional approach. Further, binding of the compound to the plastic/hardware of the chips, to proteins contained in the cell culture media, or any other intracellular lipids can be accounted for to accurately determine liver clearance. DigiLocs, further, also does not depend on a scaling factor, which overcomes the systematic underprediction of conventional approaches (5-10 fold on average across multiple studies).

So far, limited information is available from the literature or in-house measurements on the observed partitioning of compounds into the intracellular or cell-associated milieu of hepatocytes. If that data becomes available, it may be incorporated compared to the adjustments made in the software to match the clinical clearance values. If the predicted and observed  $K_{p_{uu}}$  values match, the digital twin approach truly improves the prediction. If there is a discrepancy between these values, the fitting process can be re-run including the observed  $K_{p_{uu}}$  value. This would inform the maximum capacity of the system to metabolise the compound. If this final rate is still lower than the observed clinical clearance value, two options are possible to understand and improve the prediction:

1. Calculate a correction factor, which is compound-specific and chip-specific and not generic like in the conventional approaches.
2. Investigate other model-specific parameters to optimise, e.g., permeability or partitioning.

Although initially developed for hepatic clearance, the mathematical model can be employed for toxicity or efficacy-related questions depending on the context of use. In such a setting, time-concentration profiles will be simulated and linked to other measured biomarkers (e.g., ATP (adenosine triphosphate), TEER (barrier integrity)) to determine IC<sub>50</sub> (drug effectiveness) or EC<sub>50</sub> (toxic potency of drugs) values, and parameters to assess toxicity

and efficacy, respectively. Likewise, the same integration of complex biological processes, hardware-, and drug-specific information can be used to model other cell and chip types, e.g., a blood-brain-barrier-chip, which is used to determine the permeability of compounds across the barrier. Eventually, we envision DigiLoCs support the pharmaceutical decision-making process by reducing animal testing and ultimately streamlining the drug development process.

### 3.5 Conclusion

The development of digital twins for organ-on-chips, reported here, incorporating systems of differential equations-based models and leveraging published data, holds great potential to enhance our understanding of drug behaviour and clinical outcomes. The *in vitro* liver clearance for 32 drugs was predicted using DigiLoCs and a proof-of-concept (translation to human pharmacokinetics) study on propranolol was done. DigiLoCs are envisioned to serve as a decision-support tool for pharmaceutical research, aiding in estimating first-in-human doses, evaluating human pharmacokinetics, and importantly, diminishing reliance on animal experimentation, thereby fostering more efficient, expedited and sustainable drug development processes. Our approach is generalisable across various physiological contexts and not limited to liver metabolism but may be extended to other organs as well, such as gut metabolism and barrier models such as the brain or placenta.



# Lumped parameter modelling of cardiovascular system

## Chapter outline

- Need for a comprehensive lumped parameter model of the circulatory system
- Governing ODEs and simulation for whole-body cardiovascular system
- Quantitative analysis and validation of the developed model

The cardiovascular system (CVS) is a complex network of organs and vessels responsible for pumping and transporting blood throughout the body. At its core, the CVS is a remarkable transportation system that delivers oxygen and nutrients to cells and removes waste products. Comprising the heart, arteries, veins and blood vessels, this vital system maintains body homeostasis and general health. One common method for modelling its hemodynamics is the zero-dimensional (0D) lumped parameter model (LPM), which simplifies the system by dividing it into compartments that represent different aspects of blood flow, pressure and volume dynamics [6]. These models use hydraulic-electrical analogies, where pressure gradients drive blood flow similarly to how voltage drives current in electrical circuits. This analogy allows for the modelling of vascular resistance, compliance and inertance, enabling a detailed analysis of pressure and flow rate over time [6]. By applying electrical circuit principles, these models offer a mathematical framework to predict cardiovascular dynamics. Blood flow is described by equations similar to those governing electric flow, allowing for the application of established electric circuit analysis methods to investigate cardiovascular dynamics. 0D models simulate global hemodynamics in the circulatory system using hydraulic-electrical analogues. While the electrical analogue simplifies nonlinear aspects of cardiovascular mechanics, such as nonlinear pressure-volume relationships, these limitations do not undermine

the general analytical approach. This work aims to expand on existing compartment models by developing a *comprehensive* whole-body cardiovascular model.

## 4.1 Need for a *comprehensive* cardiovascular model

Lumped parameter cardiovascular models draw parallels with hydraulic systems, representing the cardiovascular system as a network of resistors, capacitors and inductors. The development of LPMs has been motivated by the need for a more extensive understanding of cardiovascular physiology and the desire to simulate various disease states and therapeutic interventions. Existing LPMs are mostly disease-specific and lack a comprehensive approach. The pressure and flow are studied only in particular regions of interest, and the effects on other organs/whole-body are often not considered. There is a need for a whole-body model that incorporates most vital organs to study diseases and their effects in detail, addressing the limitations of current models. This study presents a *comprehensive* 0D model of the CVS, incrementally developed and validated through a thorough review of existing literature.

Modelling the human CVS is crucial for understanding cardiovascular disorders and can aid in developing simulations that explore pressure and flow dynamics [7, 70, 71]. LPMs have evolved from the Windkessel model of arterial flow [23] now to include the heart, heart valves, and veins. Various 0D models have been developed to capture the unique characteristics of each circulatory subsystem, which are reviewed in detail in Yubing Shi et al. [6]. These models are widely applied in various areas of cardiovascular research to simulate hemodynamics, characterise diseases, and develop cardiac medical devices [70, 24]. Advances such as the combination of LPMs and one-dimensional (1D) models have improved vascular hemodynamic simulations [72], and more complex models have been created to simulate systemic and pulmonary circulation, as well as arrhythmias and diseases like COVID-19 [73, 74, 75]. Recent developments also include non-invasive monitoring models for continuous cardiac output, enhancing the accuracy of lumped parameter models through validation [6, 76]. Subject-specific CVS model parameters are estimated non-invasively from signals like photoplethysmograph using LPM [12].

The advancements of LPMs underscore the ongoing evolution in cardiovascular research and their potential for simulating disease states and personalising treatments. However, there is a need for a unified whole-body model that includes vital organs to study the systemic effects of diseases and therapies. For instance, patients with complex congenital heart defects (CHDs) [77] often experience multi-organ failure due to altered circulatory systems. A whole-body CVS model could simulate the long-term effects of CHDs on liver, kidney and brain function. A unified model could also predict systemic consequences of heart failure [78] on other organs, mainly lungs and liver, guiding more effective treatment strategies. Our work



addresses this gap by developing and validating a *comprehensive* 0D cardiovascular model that will serve as a foundation for studying drug transport and cardiovascular dynamics in the context of personalised medicine.

The *base* model [75], comprising heart chambers, pulmonary circulation, and systemic circulation, is expanded to include relevant organs as compartments to form a complete whole-body closed circuit. The following sections provide a description of the whole-body model, including its mathematical formulation using ODEs, parameter setting, sensitivity analysis and validation using L2 norm and a review of existing literature.

## 4.2 Methodology

The whole-body LPM for the cardiovascular system is developed as a closed-loop comprising compartments representing body vasculature. The subsections briefly describe the circulatory anatomy, followed by governing equations, parameter sensitivity, and validation.

### 4.2.1 Overview of circulatory anatomy

The cardiovascular system were abstracted to form compartments that correspond to specific organ systems or vascular regions: the four *heart chambers*, *pulmonary circulation*, *aortic circulation*, *cerebral circulation*, *facial circulation*, *upper body circulation*, *renal circulation*, *hepatic circulation*, *mesenteric circulation* and *lower body circulation*. We developed a model with major vasculature, including the ascending aorta, aortic arch, subclavian, common carotid artery, internal and external carotid arteries, descending and abdominal aorta, celiac, renal, mesenteric, splenic, hepatic, and iliac arteries, as well as the inferior and superior vena cava and portal vein. An illustration of the human body with the compartments (organs) we have modeled is given in Figure 4.1 and the acronyms used in this work are described in Table 4.1.

The heart acts as a pump, with its four chambers - the left and right atria and ventricles - working together in a regular rhythm to circulate blood throughout the body. The pulmonary circulation in our model is divided into pulmonary arteries, left lung, right lung and pulmonary veins. Oxygen is absorbed from the blood by the tissue in systemic circulation and the deoxygenated blood is sent back to right heart for re-oxygenation through lungs. The aorta, arising from the left ventricle, is anatomically divided into the ascending aorta, aortic arch, descending aorta and abdominal aorta. From the aortic arch, it ascends to supply the common carotid artery, which directs blood to the cerebral (via internal carotid artery) and facial circulation (via external carotid artery), while another branch (via subclavian) supplies the upper limbs. It then continues downward as the descending aorta and supplies blood to the lower body organs.

The upper body circulation involves cerebral, facial and upper limb circulation which are connected to the superior vena cava on the other end. The lower body circulation is through the abdominal aorta which extensively branches to renal circulation, hepatic circulation, splanchnic circulation and lower limbs, connected on the other end to the inferior vena cava. Hepatic tissue receives blood from two sources: the hepatic artery and the portal vein. The portal vein is formed by the joining of the splenic and superior mesenteric veins, which together collect blood from the spleen, intestines, and other abdominal organs, and then deliver this nutrient-rich blood to the liver. The portal vein carries about 75% of the blood to the liver, while the hepatic artery supplies the rest with oxygen-rich blood. Circulation within a compartment is modeled by representing smaller arteries as an RLC circuit system, while the tissue is represented by a resistor. The circuit diagram for the entire network is shown in Figure 4.3. The color scheme in the diagram is as follows. The heart chambers are displayed within green boxes, and the lung compartments are shown in salmon-colored boxes. Blue represents deoxygenated blood, red indicates oxygenated blood, and purple represents areas where the two blood types mix.

#### 4.2.2 Model development

Time-varying elastances represent heart chambers and the valves are represented using diodes which ensure unidirectional flow. The model is built by adding compartments incrementally (say one by one) to the base model with eight compartments described in Regazzoni et al. [75]. After addition of a compartment the model parameters are manually tuned to ensure physiological matches. The blood pressure, flow and volume in each compartment are represented using  $P$ ,  $Q$  and  $V$ , respectively. Each compartment is represented using an RLC (resistor, capacitor, inductor) circuit. The combination of the three elements captures the dynamics of a compartment in terms of blood storage, flow, and dissipation. Table 4.2 lists the compartments, parameter acronyms, resistance, capacitance and inductance. Parameter values are primarily obtained from the literature when available; otherwise, they are manually adjusted to align with physiological behavior.

#### 4.2.3 Governing equations

The dynamics of the four heart chambers are described by time-dependent elastance [79]. These chambers are denoted by the right and left atria and ventricles, RA, RV, LA and LV, respectively. We will denote a generic chamber by XX. To run the model we need to capture the dynamic state of the heart, one where the heart allows passive filling of blood which is the diastolic phase and other one where the heart pushes blood to the lungs which is the systolic phase. The elastance  $E$  is the sum of these terms involving the active and passive phases. At

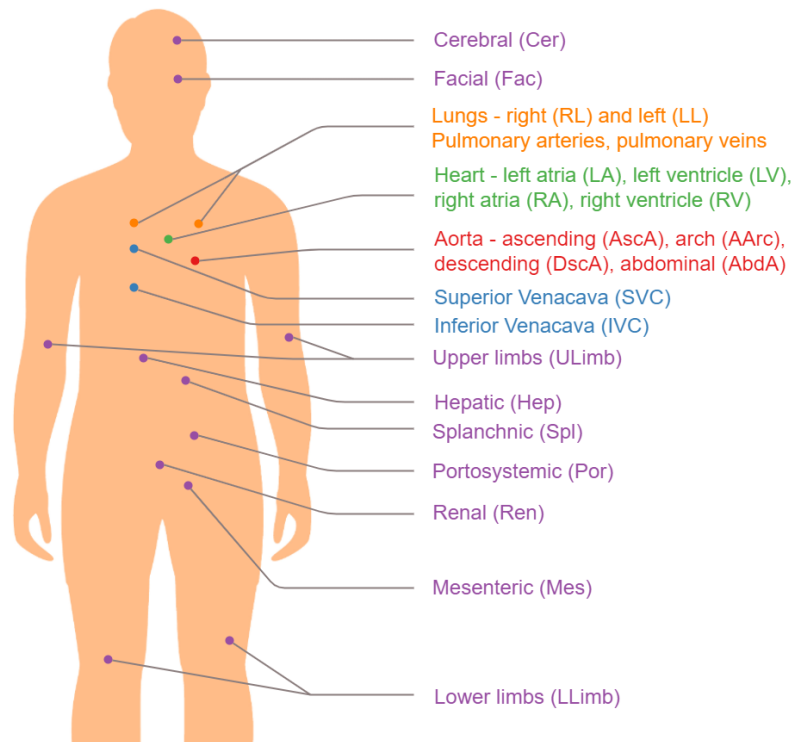


Figure 4.1: Human body anatomy marked with compartments in the comprehensive model

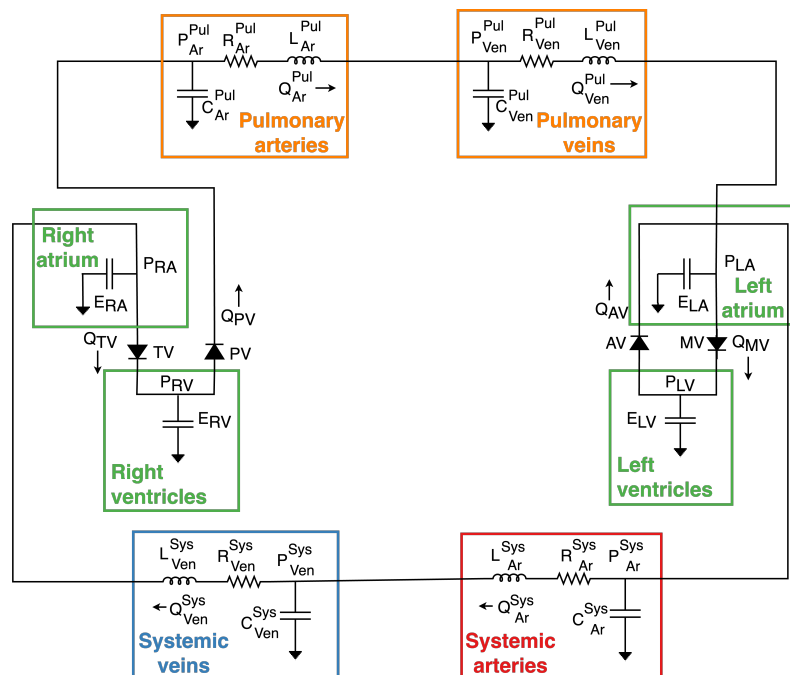


Figure 4.2: Base model of cardiovascular system (adapted from [1])

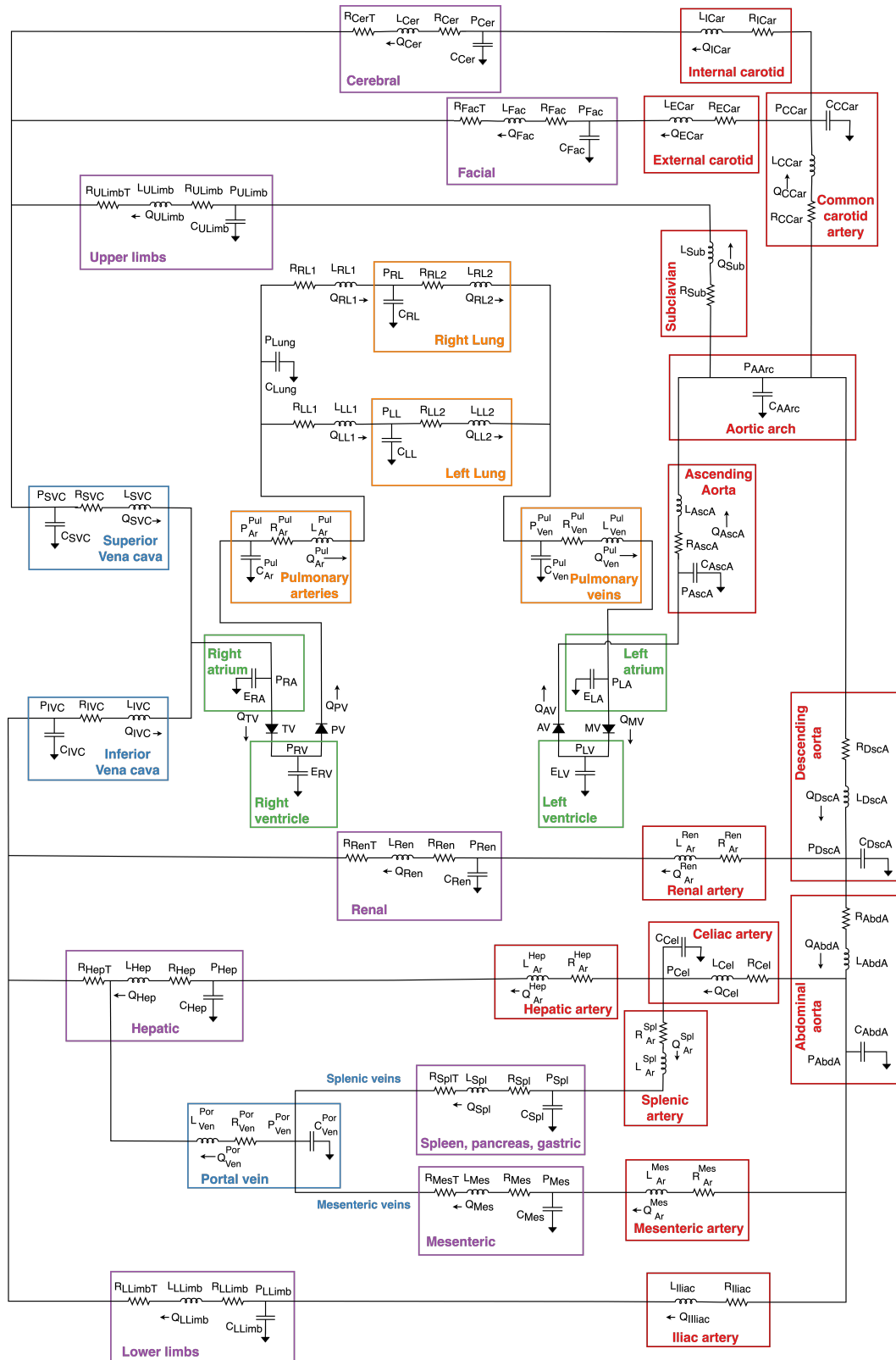


Figure 4.3: Entire closed-loop electrical analog model of human cardiovascular system

time  $t$ , the elastance  $E_{XX}(t)$  is

$$E_{XX}(t) = EB_{XX} + EA_{XX} \cdot f_{XX}(t), \quad (4.1)$$

where  $f_{XX}(t)$  is a periodic (with a period of a heartbeat) time-dependent function ranging from 0 to 1 that accounts for the two phases of the heart. The change in volume with time in each chamber is the difference in inward and outward blood fluxes. For instance, volume of blood contained in the left atrium ( $V_{LA}$ ) is represented by

$$\frac{dV_{LA}(t)}{dt} = Q_{Ven}^{Pul}(t) - Q_{MV}(t), \quad (4.2)$$

where  $Q_{Ven}^{Pul}$ ,  $Q_{MV}$  is the blood flow in the pulmonary vein and mitral valve, respectively.

The pressure inside a specific heart chamber ( $P_{XX}(t)$ ) is computed using  $E_{XX}$  and difference in volume of blood in the chamber ( $V_{XX}$ ) and the unloaded volume ( $V0_{XX}$ ), represented by

$$P_{XX}(t) = E_{XX}(t)(V_{XX}(t) - V0_{XX}). \quad (4.3)$$

**Windkessel model for compartments** Let us consider the formulation of the ascending aorta (AscA) compartment. The ODEs for the pressure and flow in the ascending aorta are represented using a simple RLC circuit as follows

$$\begin{aligned} C_{AscA} \frac{dP_{AscA}}{dt} &= Q_{AV}(t) - Q_{AscA}(t), \\ L_{AscA} \frac{dQ_{AscA}}{dt} &= -R_{AscA} \cdot Q_{AscA}(t) + P_{AscA}(t) - P_{AArc}(t). \end{aligned}$$

The upstream blood flow is from the aortic valve ( $Q_{AV}$ ), and the downstream blood flow is through the ascending aorta ( $Q_{AscA}$ ). A total of 56 ODEs represent the whole-body cardiovascular model. The equations were solved using the `diffurx` library [80] with a Just in Time (JIT) compiled solver using automatic differentiation. We used an initial time step of 1 ms and set both absolute and relative error tolerances to  $1 \times 10^{-8}$ . Detailed descriptions of compartments and ODEs are provided in Appendix B.1.

#### 4.2.4 Parameter values and manual tuning

The parameters of the whole-body model, summarized in Table 4.2, consist of values from existing literature (marked with \*) and others that are manually tuned. The compliance for the superior (SVC) and inferior vena cava (IVC) are set high ( $C_{SVC} = 50.0$ ,  $C_{IVC} = 160.0$ ) in our model as it includes all the compliances of the veins in organs. Veins are highly compliant vessels and can accommodate large changes in blood volume with minimal changes in pressure, stabilizing venous return to the heart. The aorta ( $C_{AscA} = 0.3$ ) is assigned a higher compliance compared to smaller arteries ( $C_{DscA} = 0.1$ ,  $C_{AbdA} = 0.08$ ,  $C_{CCar} = 0.07$ ) to reflect the elastic

properties. As we move from the aorta to smaller arteries, the compliance of individual vessels decreases. However, the sum of the compliance across all smaller arteries increases. This distributed compliance is important for dampening pressure fluctuations for smooth blood flow to tissues.

The resistance of arteries increases as we move from the aorta to smaller arteries and arterioles. The aorta has the lowest resistance due to its large diameter, while smaller arteries and arterioles have progressively higher resistance. The resistance values for organs (for instance,  $R_{ULimbT}$ ,  $R_{FacT}$ ,  $R_{CerT}$ ,  $R_{RenT}$ , etc.) are set relatively high. This is because the main resistance to blood flow in the systemic circulation comes from the arterioles and small arteries within organ tissues.

Large arteries (e.g., aorta, carotids) have higher inductance values, which allow pressure waves generated by the heart to propagate more rapidly along arterial walls. Smaller arteries and veins have lower inductance because the inertia of blood is less significant in these vessels due to lower flow velocities and smaller diameters. The parameters for pulmonary arteries, veins, and lung compartments are characterised by lower resistance and higher compliance compared to systemic arteries, reflecting the low-pressure, high-compliance nature of the pulmonary circuit. Initial values of hemodynamic variables (flow, pressure and volume) for all the compartments in the model are presented in Appendix B.1 Table B.1.

#### 4.2.5 Validation methods

To ensure stability and physiological relevance of our comprehensive model, we analysed quantitative metrics and physiological comparisons as explained below.

1. Volume conservation - the conservation of blood volume across all compartments of the model is established. The simulation is run for a large number of heart cycles ( > 1000) and checked for periodic volume stability.
2. Flow conservation metrics - two quantitative metrics are presented: the mean absolute flow difference and the root mean square error (RMSE) expressed as percentages and are calculated based on the difference between compartmental inflow and outflow.

- Mean absolute flow difference quantifies the average net imbalance between inflow and outflow, normalized as a percentage of the mean flow:

$$\left( \frac{\sum |\text{inflow} - \text{outflow}|}{\sum \text{mean flow}} \right) \times 100$$

- Root mean square error (RMSE) captures standard deviation of the flow imbalance, also normalized to the mean flow:

$$\left( \frac{\sqrt{\text{mean}(\text{inflow} - \text{outflow})^2}}{\text{mean flow}} \right) \times 100$$

Both metrics are expressed as percentages and are calculated for each compartment.

3. L2 norm comparison with base model - We quantified the difference in flow rates and pressures using L2 norm to evaluate the consistency between the comprehensive and base model. The L2 norm difference is normalized with respect to the base model variables for comparison.

$$\text{L2 norm difference} = \frac{|\theta_{comprehensive} - \theta_{base}|_2}{|\theta_{base}|_2}$$

where  $\theta_{base}$  and  $\theta_{comprehensive}$  are vectors of physiological variables (e.g., flow rates and pressures) from the base and comprehensive models, respectively.

4. Pressure-volume (PV) loop - simulated pressure-volume (PV) loops for the left and right ventricles were compared with base model and reference PV loops from the literature. PV loops are a gold standard for assessing cardiac cycle dynamics, and close agreement indicates that the model reproduces ventricular hemodynamics.
5. Cardiac output validation - Cardiac output of each compartment is calculated percent-wise to establish physiological consistency.

$$\text{Cardiac output percent per compartment} = \frac{\text{mean flow in compartment}}{\text{stroke volume}} * 100$$

$$\text{Stroke volume(mL/s)} = \frac{(EDV - ESV) * \text{heart rate (beats per min)}}{60}$$

where EDV and ESV are end diastolic and end-systolic volume, respectively.

6. Physiological range validation - The range of blood pressure, flow, and volume is validated with the available literature, ensuring consistency with reported physiological values. However, an exhaustive list of values is unavailable, especially considering blood flow in internal organs, for instance, in hepatic and splanchnic circulation.

## 4.3 Results and validation

The comprehensive model is run for 10000 cardiac cycles, with a duration of 0.8 sec per cycle, and the results from the final cycle are presented. This section presents the outcomes of the comprehensive model, comparing its performance to the base model through quantitative and qualitative analyses. The systemic circulation in the comprehensive model is divided into several compartments, unlike the base model, which has one compartment each for the systemic artery and veins. Section 3.2 examines the sensitivity of model parameters and Section 3.3 validates model using quantitative metrics and physiological ranges from the literature.

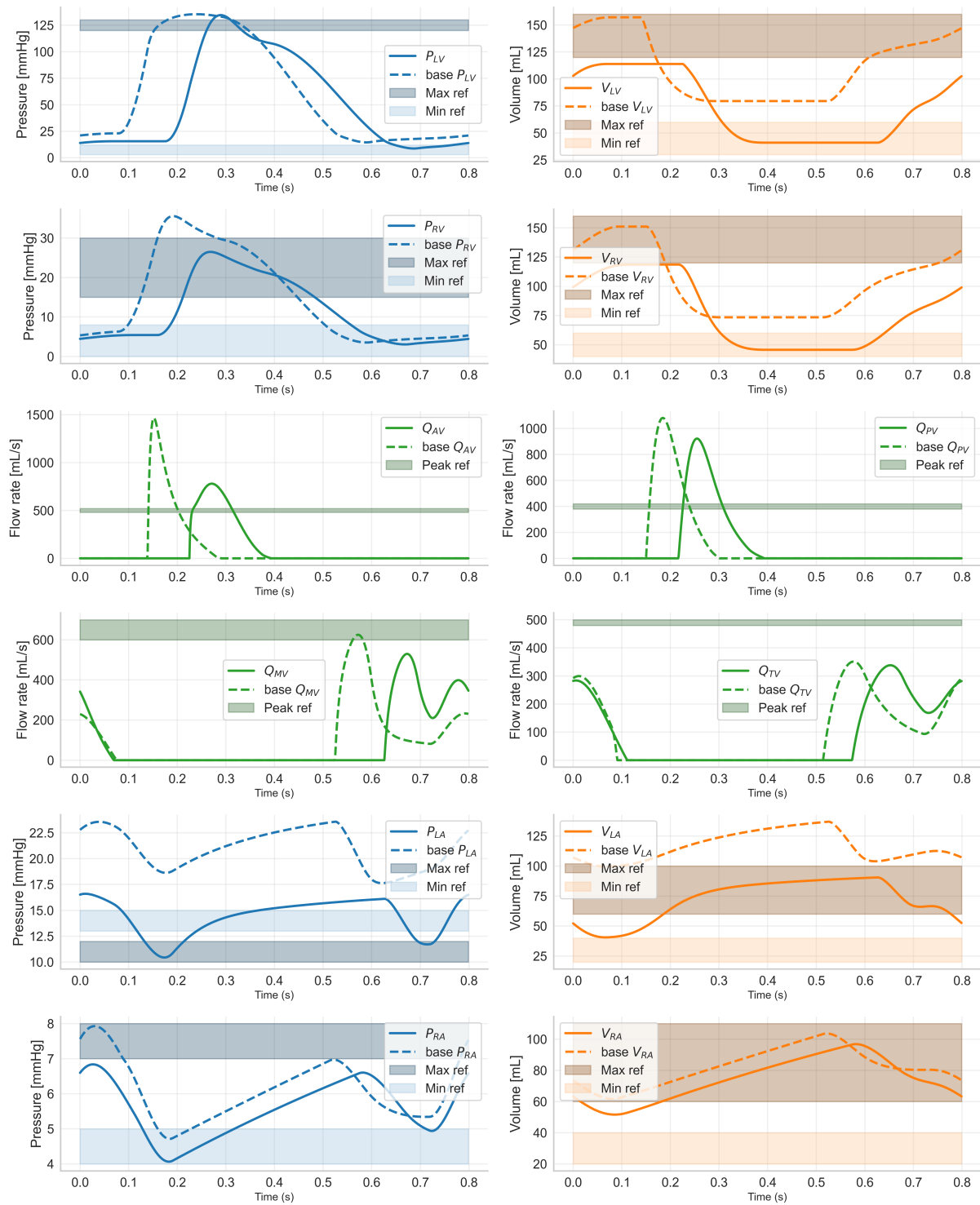


Figure 4.4: Comparison of simulated waveforms from the comprehensive model with reference ranges. Twelve subplots show pressure (blue), volume (orange), and flow rates (green) across cardiac chambers and valves over one cardiac cycle. Simulated results (solid lines) and base model results (dashed lines) are plotted for left ventricle (LV), right ventricle (RV), left atrium (LA), right atrium (RA), and valves: mitral (MV), tricuspid (TV), aortic (AV), and pulmonary (PV). Reference regions are shaded: pressure and volume plots include reference minimum (lighter shade) and maximum (darker shade) bounds; flow plots indicate peak reference values.



Acronym	Description	Acronym	Description
RA	Right atrium	Fac	Facial
RV	Right ventricle	Cer	Cerebral
LA	Left atrium	ULimb	Upper limb
LV	Left ventricle	Cel	Celiac
MV	Mitral valve	Ren	Renal
PV	Pulmonary valve	Spl	Splanchnic
AV	Aortic valve	Mes	Mesenteric
TV	Tricuspid valve	Por	Portal vein
RL	Right lung	Hep	Hepatic
LL	Left lung	Iliac	Iliac artery
AscA	Ascending aorta	LLimb	Lower limb
AArc	Aortic arch	SVC	Superior vena cava
DscA	Descending aorta	IVC	Inferior vena cava
AbdA	Abdominal aorta	Pul	Pulmonary
Sub	Subclavian	Ven	Vein
CCar	Common carotid	Ar	Artery
ICar	Internal carotid	$R_{\min}$	Open valve resistance
ECar	External carotid	$R_{\max}$	Closed valve resistance

Table 4.1: List of Acronyms

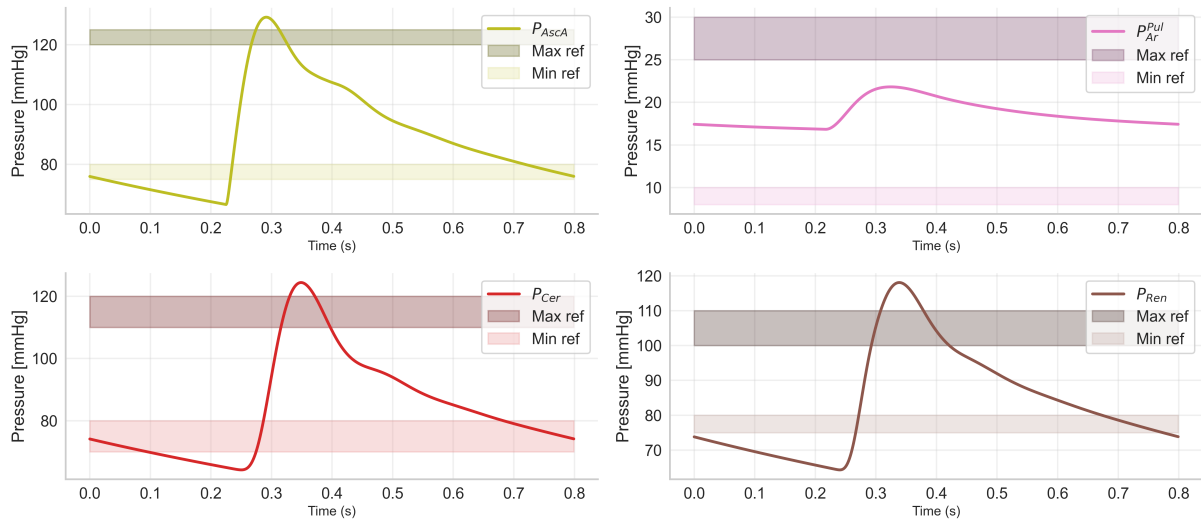


Figure 4.5: Pressure waveforms in the ascending aorta ( $P_{AscA}$ ), pulmonary artery ( $P_{PulAr}$ ), cerebral artery ( $P_{Cer}$ ), and renal artery ( $P_{Ren}$ ) over one cardiac cycle. Each colored line represents the pressure in the corresponding vessel. For each subplot, the shaded horizontal bands indicate the established physiological reference range: the upper reference limit is shown by the darker shade (Max ref), and the lower reference limit is shown by the lighter shade (Min ref), both in the color corresponding to each compartment.

Compartment	Active elastance* [mmHg mL <sup>-1</sup> ]	Passive elastance* [mmHg mL <sup>-1</sup> ]	Time of contraction* [s]	Time of relaxation* [s]	Duration of contraction* [s]	Unloaded volume* [mL]
Left atrium	0.22	0.18	0.17	0.17	0.90	2.00
Left ventricle	3.35	0.20	0.25	0.40	0.22	20.00
Right atrium	0.06	0.07	0.17	0.17	0.90	5.00
Right ventricle	0.55	0.05	0.05	0.05	0.20	20.00

(a)

Compartment	Abbreviation(s)	Resistance (R) [mmHg s mL <sup>-1</sup> ]	Compliance (C) [mL mmHg <sup>-1</sup> ]	Inductance (L × 10 <sup>-3</sup> ) [mmHg s <sup>2</sup> mL <sup>-1</sup> ]
Pulmonary Artery	$X_{Ar}^{Pul}$	0.008	10.0	0.87
Pulmonary Vein	$X_{Ven}^{Pul}$	0.0123	30.0	0.0087
Right Pulmonary Artery	$X_{Ar}^{RPul}$	0.012	—	0.0296
Left Pulmonary Artery	$X_{Ar}^{LPul}$	0.016	—	0.0331
Right Lung (RL)	$X_{RL}$	0.012	10.0	0.0296
Left Lung (LL)	$X_{LL}$	0.012	8.0	0.0296
Ascending Aorta	$X_{AscA}$	0.02	0.3	1.25
Aortic Arch	$X_{AArc}$	—	0.2	—
Common Carotid	$X_{CCar}$	0.06	0.07	1.79
Internal carotid	$X_{ICar}$	0.08	—	1.79
External carotid	$X_{ECar}$	0.60	—	5.36
Cerebral Artery	$X_{Ar}^{Cer}$	0.1	0.1	1.79
Cerebral circulation	$X_{Cer}, X_{CerT}$	0.1, 4.0	0.1	1.79
Facial Artery	$X_{Ar}^{Fac}$	0.30	0.10	5.36
Facial circulation	$X_{Fac}, X_{FacT}$	0.3, 6.0	0.10	5.36
Subclavian	$X_{Sub}$	1.0	—	2.68
Upper Limb	$X_{ULimb}, X_{ULimbT}$	2.10, 6.0	0.20	2.68
Descending Aorta	$X_{DscA}$	0.03	0.1	1.25
Abdominal Aorta	$X_{AbdA}$	0.04	0.08	1.25
Celiac artery	$X_{Cel}$	0.1	0.05	1.07
Renal Artery	$X_{Ar}^{Ren}$	0.1	0.005	1.07
Renal circulation	$X_{Ren}, X_{RenT}$	0.9, 3.0	0.005	1.07
Splenic Artery	$X_{Ar}^{Spl}$	4.0	0.07	4.29
Splanchnic circulation	$X_{Spl}, X_{SplT}$	4.0, 4.0	0.07	4.29
Mesenteric Artery	$X_{Ar}^{Mes}$	1.2	0.35	1.07
Mesenteric circulation	$X_{Mes}, X_{MesT}$	1.2, 2.0	0.35	1.07
Superior Vena Cava	$X_{SVC}$	0.18	50.0	0.08
Inferior Vena Cava	$X_{IVC}$	0.173	160.0	0.04
Portal Vein	$X_{Ven}^{Por}$	0.6	8.0	1.07
Hepatic Artery	$X_{Ar}^{Hep}$	1.2	0.035	1.07
Hepatic circulation	$X_{Hep}, X_{HepT}$	1.2, 6.0	0.035	1.07
Iliac Artery	$X_{Iliac}$	0.4	—	0.86
Lower Limb	$X_{LLimb}, X_{LLimbT}$	0.8, 6.0	0.38	0.86

(b)

Table 4.2: Compartment names and parameter values used in the comprehensive cardiovascular model. Each compartment is associated with parameters denoted by  $X$ , representing resistance (R), inductance (L), and compliance (C) within the circulatory network. Values marked with an asterisk (\*) are from literature [75]; all other parameters have been manually tuned. A '—' indicates that the corresponding component is not present in that compartment.

### 4.3.1 Simulation

Figure 4.4 presents the simulated pressure, flow, and volume waveforms across the heart chambers, overlaid with physiological minimum and maximum bounds for reference. Figure

4.5 illustrates the simulated pressures in major vessels, including the pulmonary vein and artery, the ascending aorta, and the cerebral artery. Results for the remaining compartments are provided in the Appendix (Figures B1–B6).

The simulated hemodynamic variables closely match the physiological reference values. The left ventricular end-systolic pressure ( $P_{LV}$ ) peaks at  $\sim 125$  mmHg and the right ventricular pressure ( $P_{RV}$ ) reaches a peak of 28 mmHg both within the normal physiological range. The peak flow through the mitral valve ( $Q_{MV}$ ) is about 600 mL/s, and through the aortic valve ( $Q_{AV}$ ) is  $\sim 950$  mL/s both aligning within the physiological value. The ascending aortic pressure ( $P_{AA}$ ) peaks at  $\sim 125$  mmHg and 80 mmHg, matching the standard aortic pressure of  $\sim 120/80$  mmHg. The mean pulmonary artery pressure ( $P_{Ar}^{Pul}$ ) is around 25 mmHg, which is within the normal range (12–25 mmHg), and the pulmonary venous pressure ( $P_{Ven}^{Pul}$ ) peaks at about 15 mmHg (reference: 8–15 mmHg).

### 4.3.2 Sensitivity analysis

A global sensitivity analysis using Sobol indices is performed to quantify the influence of whole body model parameters on hemodynamic variables - pressure, volume and flow. Each parameter is sampled uniformly between 0.9 and 1.1 times its nominal value, with 230,000 samples used in the analysis. The influence of each parameter is quantified by measuring the area under the curve for each variable over a cardiac cycle. Results are presented as bar plots for selected clinically measurable outputs and as a heatmap summarizing parameter influence on all variables. Figure 4.6 presents the total Sobol indices for clinically measurable variables, including left and right ventricular volumes ( $V_{LV}$ ,  $V_{RV}$ ), aortic and pulmonary artery pressures ( $P_{AA}$ ,  $P_{Ar}^{Pul}$ ), flow in ascending aorta ( $Q_{AA}$ ), and superior vena cava pressure ( $P_{SVC}$ ).

For both  $V_{LV}$  and  $V_{RV}$ , heart rate (BPM) is the dominant parameter, with total Sobol indices exceeding 0.6, indicating a strong influence on ventricular filling and ejection. The elastance of the left and right ventricles ( $EA_{LV}$ ,  $EA_{RV}$ ) and right ventricular resistance ( $R_{SVC}$ ) also contribute, but to a much lesser extent (indices  $\sim 0.1$ ). Pressure in ascending aorta ( $P_{AA}$ ) is influenced by BPM (index  $\sim 0.35$ ) and  $R_{SVC}$  ( $\sim 0.15$ ), whereas pulmonary artery pressure ( $P_{Ar}^{Pul}$ ) is more evenly affected by time period of left atrial contraction ( $tC_{LA}$ ),  $R_{SVC}$ , and left ventricular passive elastance ( $EB_{LV}$ ), each with indices around 0.2. Flow in ascending aorta ( $Q_{AA}$ ) is mostly determined by BPM (index  $> 0.85$ ). In contrast, superior vena cava pressure ( $P_{SVC}$ ) is most sensitive to  $R_{SVC}$  and  $R_{IVC}$  (indices  $\sim 0.25$ ).

An overview of parameter influence across all model outputs is provided in Figure 4.7. Heart rate (BPM) demonstrates consistently high sensitivity across a wide range of outputs, particularly for ventricular volumes, ascending aorta flow, and other pressure variables, confirming its global impact on the whole model. In contrast, certain resistance parameters (e.g.,  $R_{SVC}$ ,  $R_{Ren}$ ) exhibit high sensitivity for specific outputs, such as venous and renal pressures,

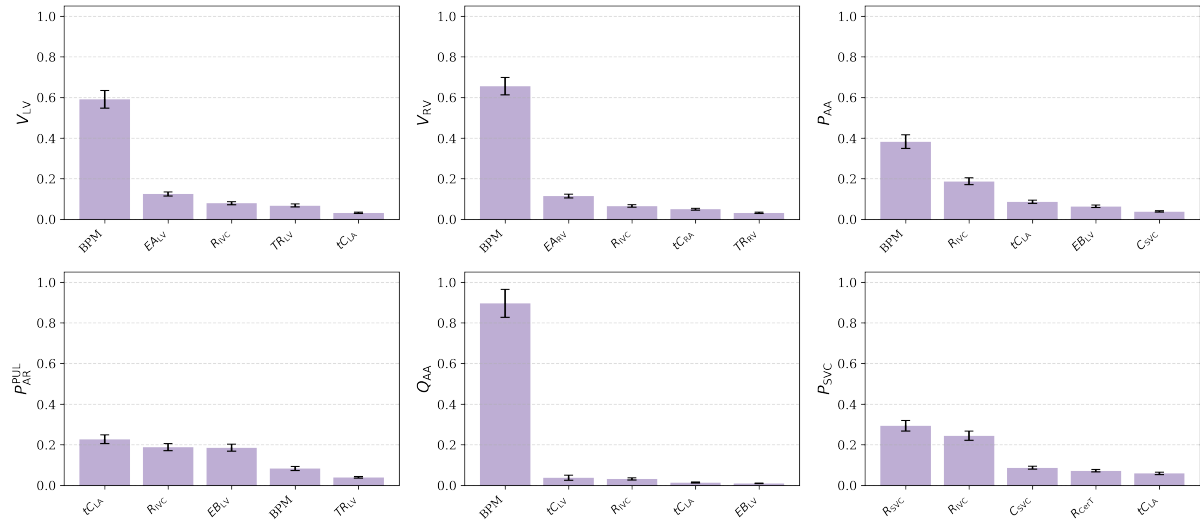


Figure 4.6: Total Sobol sensitivity indices for six clinical variables, showing the five most influential parameters for each. Bar height indicates the overall impact of each parameter on the variable, with error bars representing confidence scores. Variables shown are: left ventricular volume ( $V_{LV}$ ), right ventricular volume ( $V_{RV}$ ), aortic pressure ( $P_{AA}$ ), pulmonary arterial pressure ( $P_{AR}^{PUL}$ ), aortic flow ( $Q_{AA}$ ), and superior vena cava pressure ( $P_{SVC}$ ).

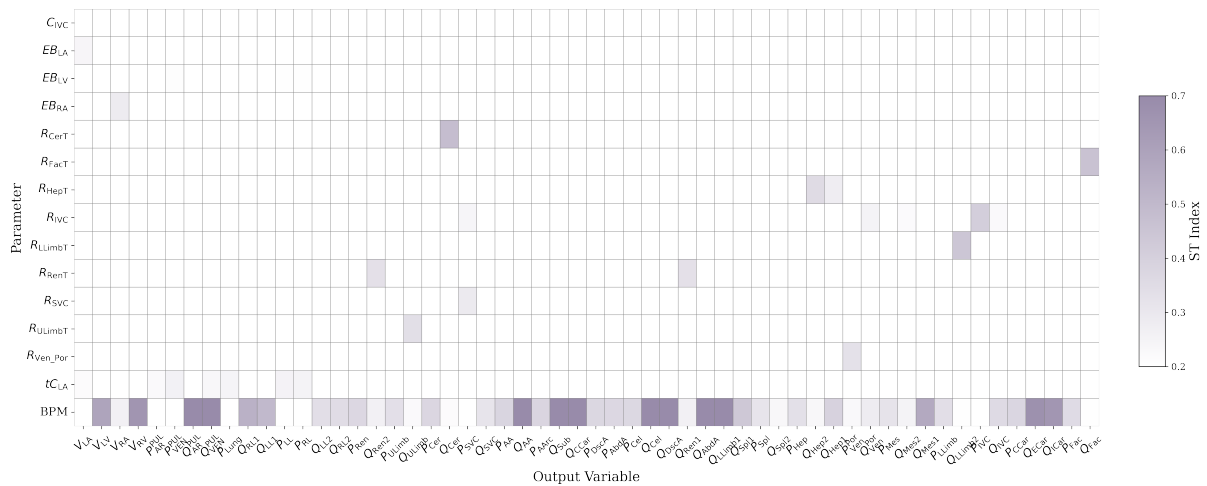


Figure 4.7: Global sensitivity analysis for comprehensive model. Total Sobol indices for all model parameters and output variables. Darker colors indicate higher sensitivity. For readability, indices below a threshold of 0.1 have been omitted in the heatmap.

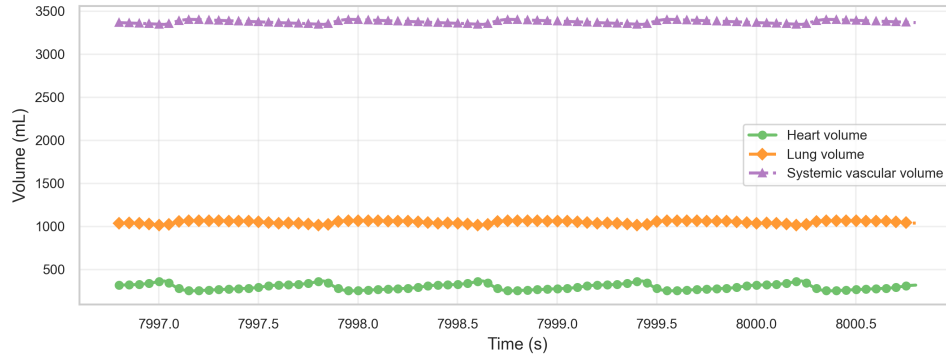


Figure 4.8: Stability of compartmental volumes in the comprehensive model. Shown are the heart, lung, and systemic vascular volumes during the last five cycles in 10,000 simulation cycles. The volumes remain stable and exhibit periodic oscillation, demonstrating that the model maintains dynamic steady state

but have minimal influence elsewhere. The timing parameters ( $t_{CLA}$ ,  $t_{CRA}$ ) and ventricular elastances ( $EA_{LV}$ ,  $EA_{RV}$ ) show moderate output-specific effects particularly on pulmonary variables. Several parameters display low sensitivity, indicating limited influence within the tested parameter ranges. A similar sensitivity analysis is performed for the base model, with the corresponding results included in the Appendix.

### 4.3.3 Validation

To assess the stability and periodicity of our comprehensive model, we monitored the volumes in the heart, lungs, and systemic vasculature over 10,000 simulation cycles. As shown in Figure 4.8 (last five cycles), the volumes for all compartments remain stable and shows periodic oscillation. This demonstrates that the model achieves a dynamic steady state, with no instability in compartment volumes. The periodic constancy of these volumes indicates conservation of mass and confirms that the model maintains physiological homeostasis over extended simulations.

Pressure-volume (PV) loops for the left and right ventricles, comparing the simulated model (blue), standard reference (black), and base model (red) are shown in Figure 4.9. The simulated PV loops resemble both shape and range of PV loop from literature [2], indicating that the comprehensive model reproduces physiological ventricular function. In contrast, the base model shows clear differences in the size and shape of the loops, highlighting the improvements achieved with the current model.

The distribution of cardiac output (as a percentage) across vascular compartments as predicted by the model is presented in Figure 4.10. The y-axis shows the percent of total cardiac output delivered to each compartment, while the x-axis lists the main compartments. The model outputs are grouped and color-coded: arteries (green), systemic veins (orange), and organs (blue). For each compartment, the model's predicted value is shown as a bar, with

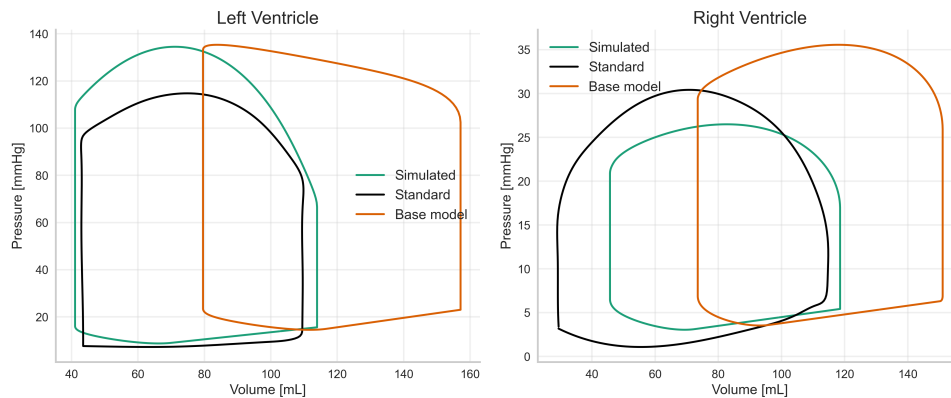


Figure 4.9: Pressure-volume (PV) loops for the left (left panel) and right (right panel) ventricles. Simulated model output (blue), standard reference data (black), and base model results (red) are shown for each ventricle. The simulated PV loops resemble both the shape and range of PV loops from literature [2, 3], indicating that the comprehensive model reproduces physiological ventricular function, while the base model exhibits significant deviations in both pressure and volume.

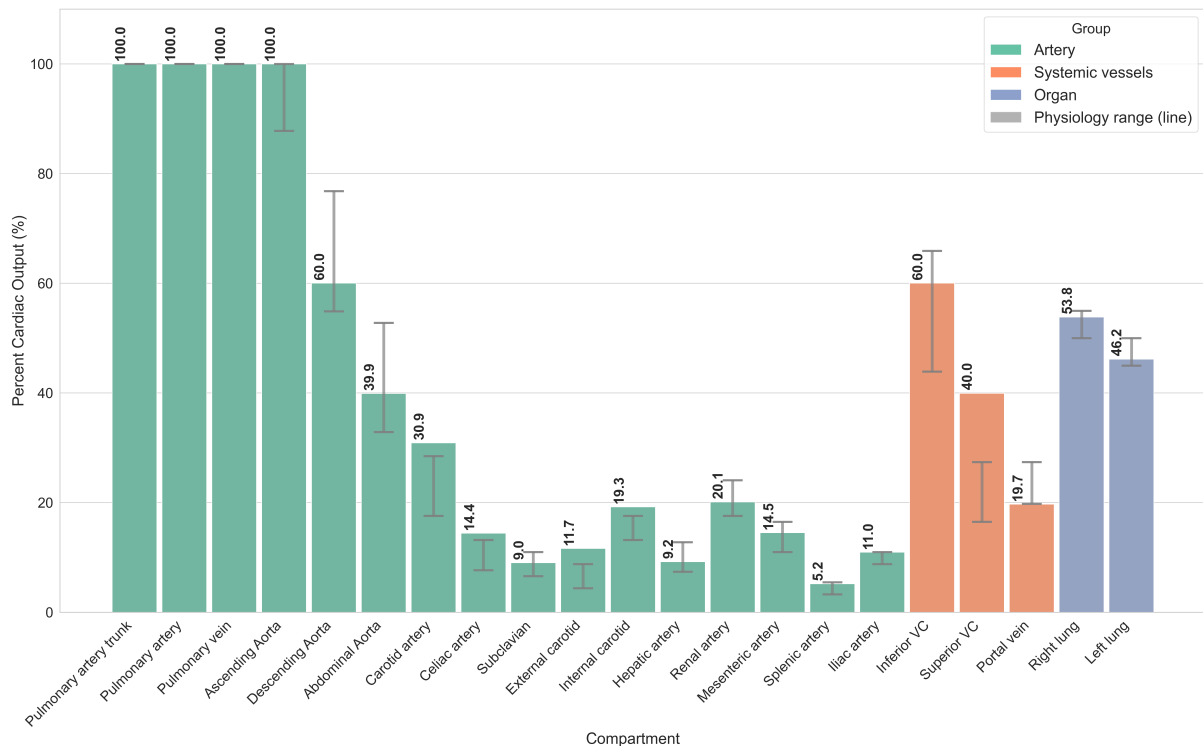


Figure 4.10: Distribution of cardiac output (%) across major arteries, veins, and organs as predicted by the model. Bars represent model predictions, grouped by compartment type: arteries (green), systemic veins (orange), and organs (blue). Grey bars indicate the physiological range from the literature.

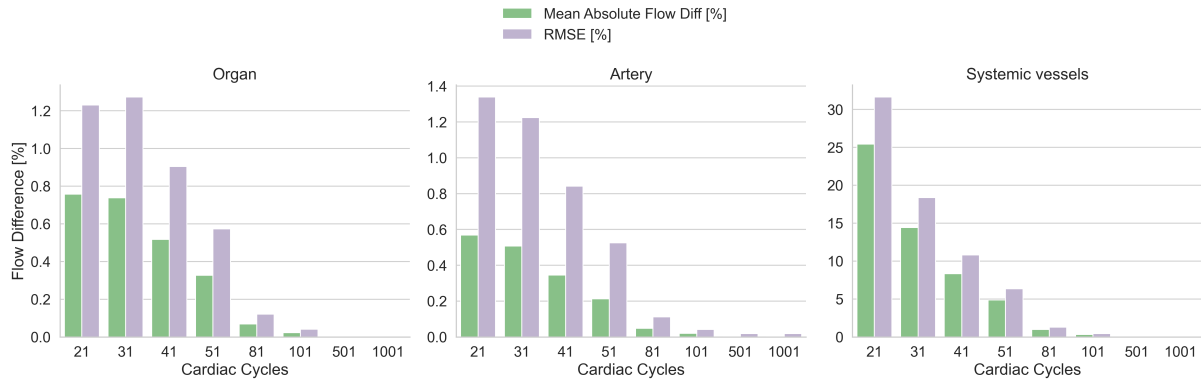


Figure 4.11: Convergence of flow metrics over increasing cardiac cycles for (left) organs, (center) arteries, and (right) systemic vessels. Bars show mean absolute flow difference (green) and root mean square error (RMSE, lavender). Both metrics decrease rapidly with more cycles, approaching zero, and indicating that model flow predictions stabilize as the number of cycles increases.

Pressure	Value	Volume	Value	Flow	Value
$P_{LV}$	0.511	$V_{LV}$	0.399	$Q_{MV}$	1.101
$P_{LA}$	0.323	$V_{LA}$	0.394	$Q_{AV}$	1.221
$P_{RV}$	0.458	$V_{RV}$	0.292	$Q_{TV}$	0.639
$P_{RA}$	0.113	$V_{RA}$	0.115	$Q_{PV}$	1.162

Table 4.3: L2 norm difference of blood pressure, volume and flow in base and comprehensive model

bars representing the physiological range reported in the literature. The results demonstrate that the model reproduces the expected distribution of cardiac output, with major vessels such as the pulmonary artery, pulmonary vein, and aorta receiving 100% of the output, as expected. Downstream compartments and organs receive fractions of the total output in line with physiological norms. For instance, the celiac artery, renal artery, and mesenteric artery receive 30.9%, 9.2%, and 14.6% of the cardiac output, respectively. Minor deviations are discussed in detail in the Discussion section.

The convergence of model-predicted flow rates across organs, arteries, and systemic vessels over increasing numbers of cardiac cycles is shown in Figure 4.11. As shown in all three panels, both metrics decrease sharply with increasing cycle number. In the early cycles (21–51), both the mean absolute flow difference and RMSE are relatively high, reflecting initial model transients and adaptation. By 101 cycles and beyond, both metrics approach zero for all compartments, indicating that the model has reached a stable periodic solution with negligible cycle-to-cycle variability. This convergence is most pronounced in systemic vessels, which initially exhibit higher variability but stabilize with continued simulation. These results confirm that the model achieves numerical stability, with flow metrics becoming consistent after sufficient cardiac cycles.

The L2 norm differences for flow rates and pressures for comprehensive and base models

are reported (after adjusting for phase-shift in curves) in Table 4.3. The larger L2 norm values for certain variables are justified when considering their reference ranges, which align more closely with the comprehensive model than the base model. The comparison in base and comprehensive model simulations with reference ranges are shown in Appendix Figure B9. Table B.2 in Appendix presents the variables (pressure, volume, and flow) with descriptions, simulation ranges, reference ranges, and the status of their physiological acceptance. The status is marked as acceptable/satisfactory if the simulated range is within or close to the reference range.

## 4.4 Discussion

The simulations demonstrate that our comprehensive model simulations agree with the reference range for most of the model variables. The comprehensive model fits the physiological conditions better than the simplified base model. The extensive and adaptable nature of our model enables easy replication and modification to simulate various disease states by selectively removing or adjusting specific compartments or variables. This flexibility is important in capturing the disease state interactions between multiple organs, such as those observed in sepsis, a multi-organ dysfunction syndrome.

The volume of the left ventricle (LV) is undervalued compared to the normal range, but this is not a critical issue. Instead, it highlights the challenge of tuning the large system within the given constraints. Similarly, the flow rate in mitral and pulmonary valves is high in the range of 1500 mL/s, whereas the comprehensive model captures this at 700-800 mL/s. The volume in the left ventricle and right ventricle for the base model peaks around 150 mL, whereas the physiological range is 100- 120 mL, where the comprehensive model is close.

The larger L2 norm values (4.3) for certain variables are understandable, given that the base model simplifies physiological conditions, resulting in notable deviations from typical physiological states. For instance, the base model's  $P_{LV}$  peak pressure reaches approximately 150 mmHg, which is 36% higher than the reference range peak of 110 mmHg in Table B.2. In contrast,  $P_{LV}$  peak pressure in the comprehensive model is closer to the reference range, with a peak value of approximately 120 mmHg. The base model's peak aortic valve flow rate ( $Q_{AV}$ ) exceeds the physiological reference value (700 mL/s) by 114% (1500 mL/s), whereas the comprehensive model's  $Q_{AV}$  peak (600 mL/s) shows a modest 14% difference. This trend persists across other quantities, highlighting that the comprehensive model is more physiologically consistent.

The present sensitivity analysis reveals that heart rate (BPM) exerts the most influence across a range of clinically observable variables, including ventricular volumes and ascending aortic flow, as evidenced by high Sobol indices. Vascular resistances, particularly those affect-



ing venous return, and chamber elastances are also significant contributors to pressure and flow dynamics, though their impact was more localized compared to the global effect of heart rate. These findings highlight that future parameter estimation for the comprehensive model should focus on heart rate, resistances, and elastances, as these parameters are most likely to drive changes in model outputs. This way, model calibration can be made more efficient, while less sensitive parameters may be fixed at nominal values.

Several difficulties make parameter setting a challenging task. These include the invasive nature of many of the measurements, restricted access to the required measurement sites due to anatomical configuration, and practical difficulties in the orientation of flow probes (particularly invasive ones). In this work, parameter estimation is not considered due to the computational complexity arising from the large number of parameters, which exceeds 118. Finding a suitable parameter combination for the model is quite difficult. Our solution in this situation is to first determine reasonable ranges for each parameter. We then assign various combinations of values within these ranges to the model parameters and assess whether the resulting model outputs align with the measured pressure and flow data. This trial-and-error procedure is repeated until an acceptable value combination of parameters is obtained. Parameter fitting optimization and aggressive tuning can refine the comprehensive model to better match the physiology.

The comprehensive model effectively captures blood circulation within organs such as the lungs and liver, and can be extended to simulate disease states involving multiple organ failure. Future work will focus on fitting the model to real-world data from a maximum number of observational quantities collected from patients and applying optimization techniques to refine the model performance.

The present results demonstrate that the comprehensive cardiovascular model successfully reproduces physiological pressure-volume (PV) relationships for both the left and right ventricles, closely matching the shape and range of PV loops reported in the literature. The base model exhibits notable deviations in both the size and shape of the PV loops, which highlights the improvements achieved by the comprehensive model. The simulated right ventricular volume is lower than physiological values, but is better representative of the actual physiology compared to the base model. The model effectively captures the systemic distribution of cardiac output, supporting its use for investigating regional hemodynamics and organ perfusion. The cardiac output to the external carotid artery is higher in our model due to the inclusion of the muscles of the neck in the facial circulation compartment. Similarly, the increased cardiac output to the celiac artery can be attributed to the inclusion of both muscular and adipose tissues in the abdominal compartments. The decline in both mean absolute flow difference and RMSE with increasing numbers of simulated cardiac cycles indicates that the model quickly overcomes initial transients and achieves a stable periodic

solution (Figure 4.11). This is particularly evident in systemic vessels, which initially show greater variability but stabilize as the simulation progresses. The larger L2 norms observed for certain variables in the base model can be attributed to its less detailed representation of cardiovascular mechanics and flow distribution. The validation against reference ranges, as visualized in Figure B9(Appendix), further supports the comprehensive model simulation.

To effectively treat disease states, incorporating pharmacokinetic attributes, such as drug diffusion, into modeling systems enables analysis of how drug distribution affects the cardiovascular system response. Pharmacokinetics can be introduced to the current whole-body network to enhance understanding of the effects of drugs by applying simulation-based testing. Complex models have been built for single-ventricle patients, which mimic vasculature but fall short in capturing the dynamic responses to various therapies. This complex model needs to be fabricated in the first place for potential use cases, for instance, in the case of sepsis, which greatly impacts several compartments. A potential future implementation would involve calibrating the model using real-world data, using optimization.

Although ours is an expanded model, it still has limitations. A significant limitation is the lack of extensive literature providing information on how these compartments should be initially configured (in terms of resistance, inductance and capacitance values). However, as the model becomes more complex with additional variables and parameters, it opens opportunities for optimizing the system to achieve proper tuning. The coronary circulation is not included in the current model. As a result, a portion of the cardiac output is effectively unaccounted for due to its distribution to the coronary vessels. We still modeled a simplified version focusing solely on critical compartments, omitting other relevant components in the cardiovascular system. This deliberate simplification results in a lower-order representation of the system, potentially overlooking subtle interactions.

In conclusion, our comprehensive model provides an adaptable framework for simulating blood flow and pressure dynamics in a healthy individual. The developed model acts as a digital twin for the cardiovascular system, which can be tuned extensively for patient-specific hemodynamic simulations. The challenges in optimizing this network can be effectively addressed by fitting the model to clinical data, ensuring better alignment with real-world physiological conditions. This foundational model has significant potential for future research applications, including the simulation of disease states, evaluation of pharmaceutical interventions, and personalised medicine approaches.

# Computational methods to study T2DM

## Chapter outline

- Identifying subtypes of uncontrolled T2DM
- Clustering workflow using clinical data
- Cluster characterisation and assessment
- Clinical significance of T2DM subtypes

Type 2 Diabetes Mellitus (T2DM) is the most common type of diabetes, contributing to 90% of the global diabetic population [81]. Despite the rapidly increasing diabetic population across the entire globe [81], inadequate glycemic control still pose an enormous challenge to the healthcare delivery system [81, 82, 83, 84], especially in the low-income and middle-income countries. In 2019, a TIGHT (The Investigation of Glycosylated Hemoglobin on Therapy in Indian diabetics) study conducted in India reported over 75% of the T2DM population have inadequate glycemic control with a glycated haemoglobin level (HbA1c) over 7% [84]. However, whether this uncontrolled T2DM population is a heterogeneous mix of disease subtypes still remains unknown. Heterogeneity among the uncontrolled T2DM population challenges the efficacy of a uniform treatment protocol [85, 86] recommended for all T2DM patients.

Drug therapy in T2DM majorly involves modulation of insulin resistance (IR) and  $\beta$ -cell dysfunction with different groups of anti-diabetic drugs [85] for which treatment algorithms exist [87]. As improvement in glycemic control reduces the complications of diabetes (such as retinopathy, nephropathy and neuropathy) thereby translating into an increase of life expectancy [88], there is an urgent and immense need to re-evaluate the T2DM management protocol in the context of heterogeneity among T2DM patients.

Until now, studies have identified T2DM clusters based on a few selective variables from the first clinical visit and the association with diabetic complications on longitudinal follow-up [89, 90]. However, studies to identify clusters among uncontrolled T2DM patients receiving different combinations of anti-diabetic drugs remain undone. This sort of study will not only reveal the pattern of heterogeneity among uncontrolled T2DM patients but will also help in the selection of features associated with the identified clusters. Identification of clusters among uncontrolled T2DM patients may result in a more customised T2DM management protocol, thereby paving the way towards precision management in T2DM.

In order to explore the patterns in the clinical data, the key step is the dimensionality reduction technique uniform manifold approximation and projection (UMAP), followed by a density-based clustering technique. The data collected from a sample of individuals comprises of characteristics that vary among themselves. A variable is a characteristic that takes different values in different individuals, places or things [91]. The information on each diabetic patient is collected using 129 features. Features or variables may be quantitative or qualitative based on measurements. Quantitative variables may be discrete or continuous, whereas qualitative variables are categorical variables classified as nominal or ordinal. The features or variables are classified mainly into three categories in this analysis:

- **Continuous features:** A continuous variable assumes any value within a specified range of values assumed by the variables. They are quantitative variables as they can be measured numerically. Continuous variables can take infinitely many values within a range. Say, for a measurement of 10-20 kg weight, any value, say 10.1, 12.005, 18.99 kg...etc, can be taken depending on the accuracy required. Some examples of continuous variables that can be measured in individuals are height, weight, glucose level and waist circumference.
- **Ordinal features:** Ordinal variables form rank-ordered categories. These categories can be assigned numerical values as it is easier in coding and data analysis. Some examples of ordinal variables are the educational status of subjects (below class 10, above class 10 or illiterate, matriculation, graduate, professional) and the severity of disease (mild, moderate, severe). The exact difference between the categories cannot be measured, say, the distinction between mild disease and moderate disease.
- **Nominal features:** These are qualitative measurements as they can't be measured numerically and can't be rank ordered. Typical examples of nominal features are sex, blood group, religion, etc. For nominal variables, there is no intrinsic ordering. For identification, nominal variables are assigned values (Female=0, Male=1). When there are two categories (alcoholic or non-alcoholic, male or female) for a qualitative variable, it is called a binary or dichotomous variable [92].

## 5.1 Discovering subtypes of uncontrolled T2DM

T2DM is a heterogeneous mix of disease subtypes. However, the guidelines for T2DM management remain uniform for all patients. Inadequate glycemic control among a great majority of T2DM patients calls for subtype identification in uncontrolled T2DM. Identifying these subtypes would result in a customized T2DM management protocol, paving the way towards precision therapy. Employing an unsupervised clustering workflow on longitudinal cohort data of T2DM patients from rural India, we discovered subtypes among uncontrolled T2DM patients and further characterized them to extract translational insights.

For a long time, T2DM has been considered a homogeneous disease entity where IR is the primary pathology followed by  $\beta$ -cell dysfunction [93]. However, recent studies report T2DM [89, 90] to be a heterogeneous mix of different pathological conditions with the contribution of IR and  $\beta$ -cell dysfunction varying among T2DM subtypes [89, 90]. Precise determination of this underlying pathology is of prime importance as we look forward to entering the era of precision medicine for managing diabetes [94, 95, 96, 97]. Hence, the aim of this study is to investigate whether the uncontrolled T2DM population is a heterogeneous mix of disease subtypes and to further identify the patho-clinical features underlying the heterogeneity. Identification of patient subtypes among uncontrolled T2DM would result in a more customized T2DM management protocol, thereby paving the way towards precision therapy and potentially improving long-term health outcomes for patients.

## 5.2 Study design and methodology

Anonymised electronic health records from March 2016 to December 2020 were collected from a community-based primary health clinic located in the district of North 24 Parganas of West Bengal, India. The clinic is run as a vertical of the Community-based metabolic health screening programme 'From Food to Nutrition Security' [98] with a focus on rendering clinical services in non-communicable diseases (NCDs), including T2DM, hypertension, ischaemic heart disease (IHD), dyslipidemia etc. A detailed clinical history of each patient, including present illness, history of past illness, family history, operative history, personal history, menstrual and pregnancy history, allergy history and past investigations, are recorded during the first visit. Enrolled patients are then regularly followed up, and examination data (height, weight, pulse and blood pressure), investigation data (blood and imaging investigations) and management data (preventive advice and drugs with dosage) are entered into the digital database during each visit.

460 T2DM patients either presenting with a known history of T2DM or diagnosed as T2DM during screening are initially selected among whom 370 patients are followed up regularly and

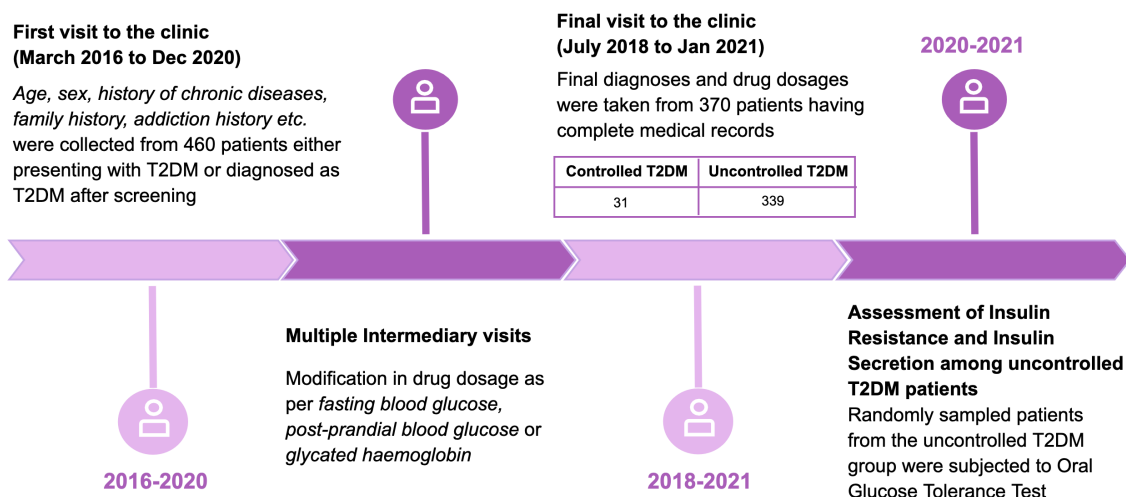


Figure 5.1: Study design. Timeline diagram of the prospective study design of subjects with type 2 diabetes mellitus recruited from the Primary Health Clinic.

has complete medical records. During the last visit, these patients are screened for glycemic control, out of which 339 patients are found to have uncontrolled T2DM (diagnosed by fasting plasma glucose over 126 mg/dl, postprandial plasma glucose over 180 mg/dl or glycated haemoglobin (HbA1c) over 7%). Data of these 339 patients are taken for clustering. The study design with the timeline is represented in Figure 5.1. For all T2DM patients, anti-diabetic drugs are modified by a clinician following the American Diabetic Association guideline [86]. The anti-hypertensive drug dosage for those diagnosed with hypertension is modified according to the measured blood pressure value during each visit. Co-morbidities like dyslipidemia and IHD are also managed. For the management of all these co-morbid conditions, the American Diabetic Association Standard of Medical Care in Diabetes 2016 is followed. All the features from the first clinic visit combined with the diagnoses (including all the co-morbid conditions) and drug combinations from the final visit are used in the final analysis. A total of 339 uncontrolled T2DM patients followed up for a median duration of 14 months (25th percentile: 7 months, 75th percentile: 24 months) are used in this study.

Both continuous and categorical variables are present in the dataset. The continuous variables in the dataset are age, duration of known hypertension (in months) and known T2DM (in months), systolic and diastolic blood pressure, Body Mass Index (BMI) and body fat percentage. The categorical variables are sex, education, occupation (shopkeepers, farmers, homemaker), diseases (IHD, dyslipidemia, hypertension, hypothyroid) and prescribed drugs (metformin, glimepiride, amlodipine etc), family history of diabetes and hypertension. The dosage of a particular drug is added up and combined into a single column to calculate the total intake of the drug.

Dimensionality reduction is applied first to the data, followed by clustering. This strategy is

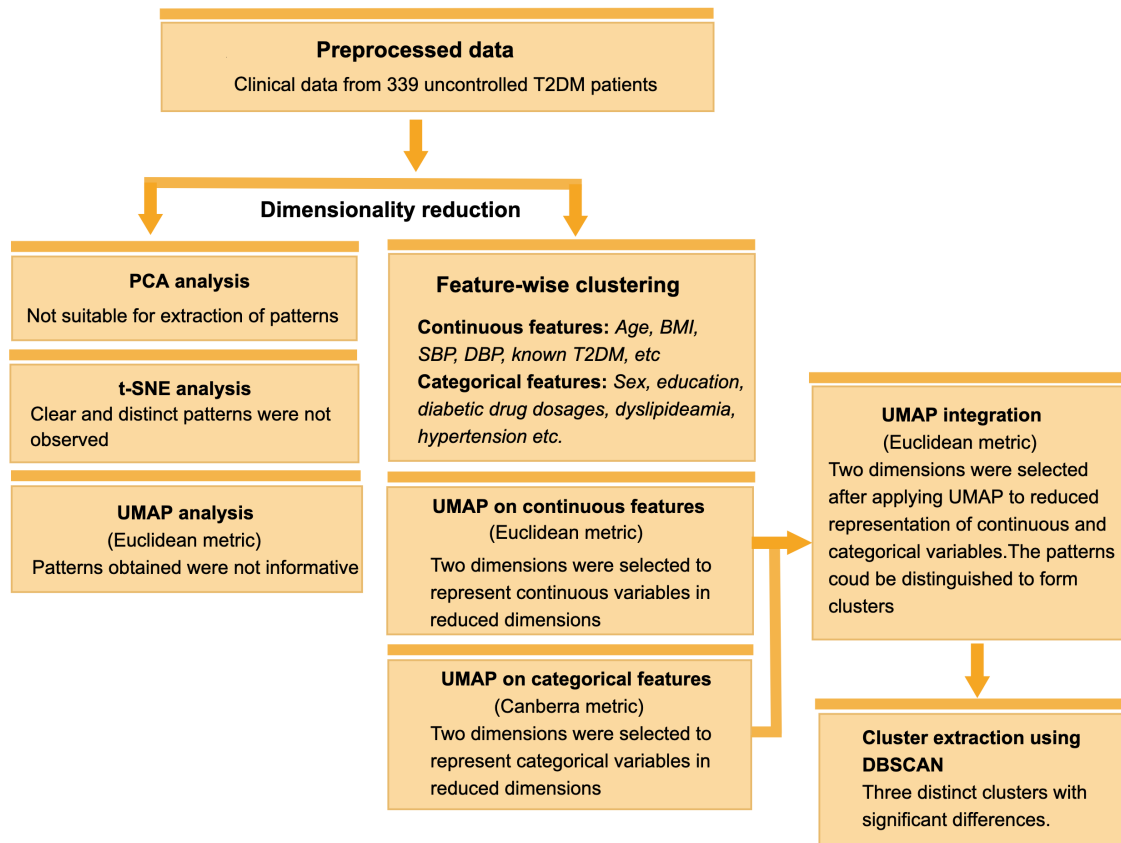


Figure 5.2: Clustering workflow with clinical data from 339 uncontrolled T2DM patients. PCA: Principal Component Analysis, t-SNE: t-Distributed Stochastic Neighbour Embedding, UMAP: Uniform Manifold Approximation and Projection, DBSCAN: Density-Based Spatial Clustering of Applications with Noise.

used because the performance of clustering algorithms can degrade in the presence of a large number of input features. Several dimensionality reduction algorithms; principal component analysis (PCA), t-distributed stochastic neighbour embedding (t-SNE) and uniform manifold approximation and projection (UMAP) are used for this purpose [99]. The dimensionality reduction technique with the best silhouette score is finally accepted. After the data is visualised in lower dimensions and reasonable patterns are observed, the non-parametric clustering algorithm, density-based spatial clustering of applications with noise (DBSCAN), is applied to this dimensionally reduced data. The number of clusters need not be specified explicitly in DBSCAN as it finds nonlinear structures based on the density of data points, which can vary in both size and shape.

An oral glucose tolerance test (OGTT) is done in 100 randomly selected patients from the obtained clusters. Blood samples are collected at fasting and 30 min after intake of 75 gm of anhydrous glucose dissolved in 200 ml of water. All the patients gave proper informed consent. The procedure is approved by the Institutional Human Ethics Committee of CSIR-IICB, Kolkata. HOMA-IR [100], Insulinogenic Index and Matsuda Index [101] are reported



as indices of IR, HOMA-B [100] and Disposition Index [102] are reported as indices of  $\beta$ -cell dysfunction.

### 5.3 Results of clustering

The cluster formations obtained using UMAP followed by DBSCAN is shown in Figure 5.3. The continuous features are summarised as mean with standard error and the categorical features as count with percentage contribution for all the clusters separately (Table 5.1). The post-clustering features for all the clusters obtained from OGTT are summarised as mean with standard error in Table 5.2. Comparison among the clusters is done using Analysis of Variance. Variables with non-normal distribution are log-transformed before comparison.

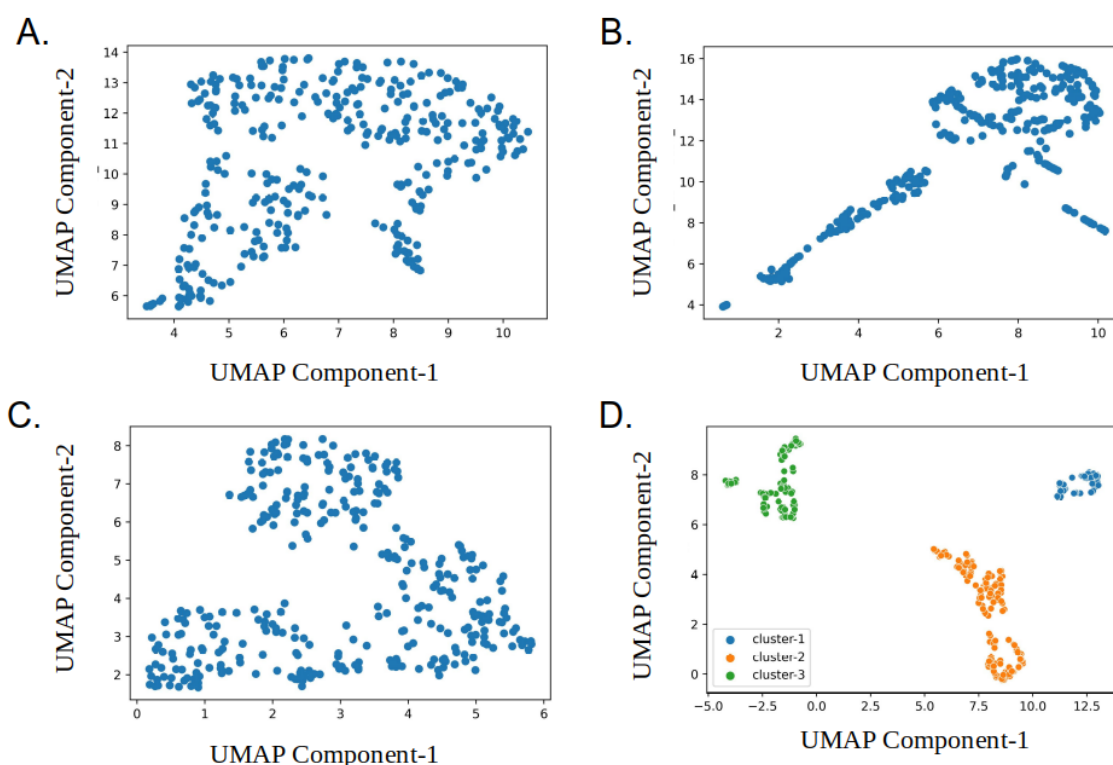


Figure 5.3: Cluster formation using UMAP followed by DBSCAN (A) UMAP visualisation on using the Euclidean metric on the entire data (B) UMAP clustering of the continuous variables only, using the Euclidean metric (C) UMAP clustering of the categorical variables only, using the Canberra metric (D) UMAP applied on the four dimensional reduced representation followed by DBSCAN to obtain clusters.

UMAP outperformed all the other dimensionality reduction techniques in the given dataset shown by comparing the silhouette score obtained after performing DBSCAN. The entire clustering workflow is detailed in Figure 5.2. DBSCAN algorithm ( $\text{eps} = 1.3$ ) is performed on this UMAP reduced data and three main clusters are obtained with a silhouette score of



Identified clusters		Cluster 1	Cluster 2	Cluster 3	p-value
Cluster size (N)		62	167	107	
Variables included from first visit					
Continuous variables					
Age (years)		43.56 ± 0.47	51.15 ± 0.27	49.68 ± 0.34	0.000
BMI (kg/m <sup>2</sup> )		24.58 ± 0.14	25.04 ± 0.12	24.26 ± 0.12	0.261
Systolic Blood Pressure (mm of Hg)		119.38 ± 0.58	143.47 ± 0.65	131.15 ± 0.73	0.000
Diastolic Blood Pressure (mm of Hg)		78.72 ± 0.42	83.57 ± 0.32	81.4 ± 0.33	0.009
T2DM duration (in months)		2.24 ± 0.19	20.93 ± 1.43	73.68 ± 1.75	0.000
Hypertension duration (in months)		0.00 ± 0.00	27.85 ± 1.53	0.00 ± 0.00	
Body fat percentage		18.48 ± 0.22	20.2 ± 0.18	18.87 ± 0.19	0.072
Category	Value	Count (%)			
Sex	Female	62(100.0)	93(55.69)	73(68.22)	0.000
	Male	0(0.0)	74(44.31)	34(31.78)	
Education	Below class 10	9(14.52)	43(25.75)	24(22.43)	0.196
Homemaker	Yes	60(96.77)	74(44.31)	61(57.01)	0.000
Shopkeepers	Yes	0(0.0)	8(4.79)	1(0.93)	0.055
Farmers	Yes	4(6.45)	19(11.38)	12(11.21)	0.527
Alcoholic	Yes	0	2(2.7)	3(8.82)	0.317
Smoker	Yes	0	40(54.05)	16(47.06)	0.000
Bowel regularity	Yes	19(30.65)	46(27.54)	31(28.97)	0.893
Ligation	Yes	5(8.06)	6(6.45)	10(13.7)	0.128
Appendectomy	Yes	4(6.45)	7(4.19)	11(10.28)	0.139
T2D parent	Yes	23(37.1)	42(25.15)	31(28.97)	0.204
Hypertensive parent	Yes	10(16.13)	37(22.16)	16(14.95)	0.278
Variables included from final visit					
Category	Value	Count (%)			
IHD	Yes	0(0.0)	9(5.39)	3(2.8)	0.130
Hypertension	Yes	1(1.61)	121(72.46)	44(41.12)	0.000
Dyslipidemia	Yes	21(33.87)	41(24.55)	27(25.23)	0.342
ECG Findings	Yes	5(8.06)	12(7.19)	4(3.74)	0.417
Hypothyroid	Yes	2(3.23)	11(6.59)	5(4.67)	0.562
Metformin	500mg	0(0.0)	4(2.4)	0(0.0)	0.050
	1000mg	18(29.03)	60(35.93)	23(21.5)	
	1500mg	18(29.03)	45(26.95)	30(28.04)	
	2000mg	26(41.94)	58(34.73)	54(50.47)	
Glimepiride	Not taking	6(9.68)	15(8.98)	5(4.67)	0.551
	1mg	0(0.0)	1(0.6)	1(0.93)	
	2mg	30(48.39)	79(47.31)	44(41.12)	
	4mg	26(41.94)	72(43.11)	57(53.27)	
Teneligliptin	Not taking	58(93.55)	153(91.62)	76(71.03)	0.000
	20mg	1(1.61)	6(3.59)	8(7.48)	
	40mg	3(4.84)	8(4.79)	23(21.5)	
Pioglitazone	Not taking	53(85.48)	156(93.41)	86(80.37)	0.015
	15mg	3(4.84)	4(2.4)	4(3.74)	
	30mg	6(9.68)	7(4.19)	17(15.89)	
Insulin	Yes	3(4.84)	8(4.79)	6(5.61)	0.952
Hydrochlorothiazide	Yes	0(0.0)	1(0.6)	1(0.93)	0.748
CCB	Yes	1(1.61)	53(31.74)	23(21.5)	0.000
ACE inhibitor	Yes	0(0.0)	74(44.31)	25(23.36)	0.000
Beta blockers	Yes	4(6.45)	7(4.19)	4(3.74)	0.692
ARBs	Yes	1(1.61)	26(15.57)	10(9.35)	0.009

Table 5.1: The post-clustering subject characteristics of the uncontrolled T2DM subgroups

Identified clusters	Cluster 1	Cluster 2	Cluster 3	p-value
<b>Variables</b>				
HbA1c (%)	7.74±0.31	7.46±0.19	8.41±0.28	0.014
Fasting Glucose(mg/dl)	172.29±20.01	133.14±5.69	180.38±13.29	0.002
30-min Glucose(mg/dl)	251.27±21.37	223.05±7.27	287.12±18.97	0.003
Fasting Insulin(uIU/ml)	4.71±0.69	6.68±0.75	6.05±0.89	0.360
30-min Insulin(uIU/ml)	20.35±4.67	20.79±2.46	17.28±1.94	0.597
<b>Indices of <math>\beta</math>-cell function</b>				
HOMA-B	52.29±32.70	43.96±7.32	26.86±4.22	0.107
Disposition Index (DI)	0.75±0.65	0.36±0.22	0.45±0.10	0.094
<b>Indices of Insulin Resistance</b>				
HOMA-IR	3.38±1.43	3.32±1.15	3.0±0.66	0.971
Matsuda Index	6.75±1.96	6.78±1.24	6.58±1.61	0.964
Insulinogenic Index	16.61± 0.63	11.26±0.54	10.77±0.54	0.460

Table 5.2: Summary statistics of the post-clustering features of all the T2DM clusters

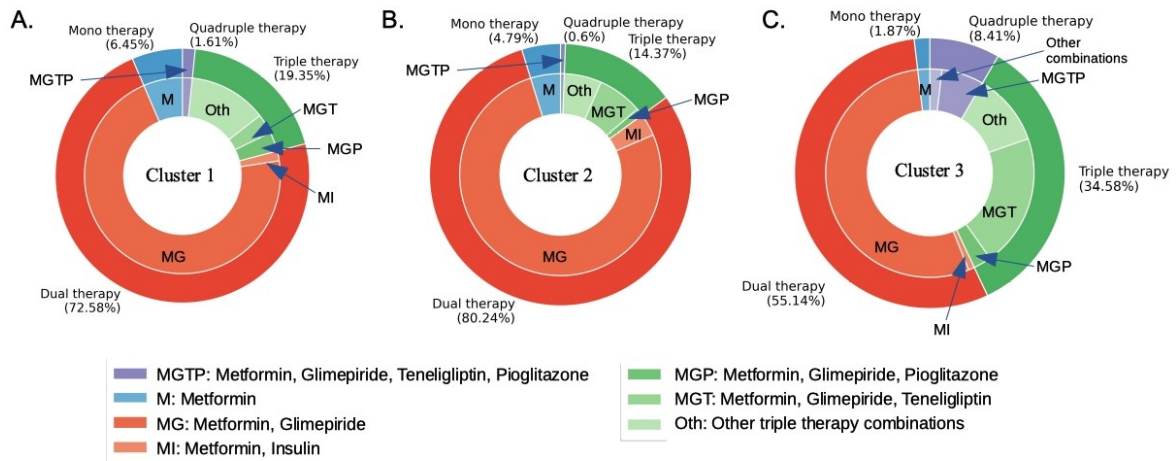


Figure 5.4: Anti-diabetic drug combinations and the percentage of patients receiving them in three clusters. Percentage of patients in Cluster 1 (A), Cluster 2 (B) and Cluster 3 (C) receiving monotherapy, dual therapy, triple therapy & quadruple therapy along with the specific anti-diabetic drug combinations for each therapy group.

0.74. Once the clusters are obtained (assigned as C1, C2 and C3), the clinical features are summarised for each cluster in Table 5.1.

### 5.3.1 Cluster characterisation

Cluster 1 (C1): C1 is the cluster with recent onset T2DM (diagnosed with T2DM within the last two months). None of the patients in this cluster presented with hypertension at the first visit, with only one patient from this group becoming hypertensive at the final visit. This cluster has the lowest average age (43.56 years) but the highest percentage of patients having T2DM parents (37.1%). Considering the duration of T2DM, the age of T2DM onset in this group is around 43 years. Distribution of patients from this cluster receiving monotherapy, dual

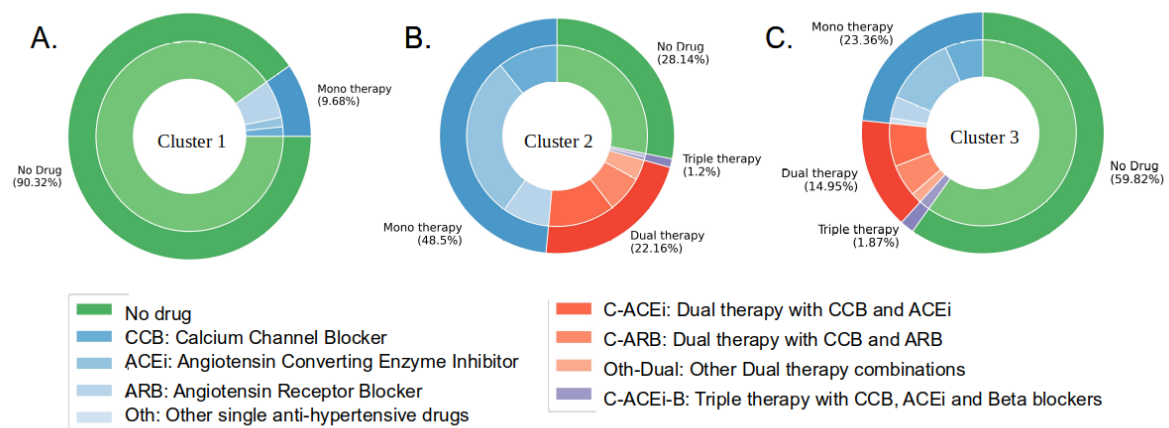


Figure 5.5: Anti-hypertensive drug combinations and the percentage of patients receiving them for each individual cluster. Percentage of patients in Cluster 1 (A), Cluster 2 (B) and Cluster 3 (C) receiving no medications, monotherapy, dual therapy & triple therapy along with the specific anti-diabetic drug combinations for each therapy group.

therapy, triple therapy and quadruple therapy for T2DM is found to be 6.45%, 72.58%, 19.35% and 1.61% respectively (Figure 5.4A).

Cluster 2 (C2): C2 has the highest average age (51 years), BMI ( $\geq 25$ ) and body fat percentage (20.2 %) with history of T2DM and hypertension for around 21 and 28 months, respectively. Considering the duration of both the diseases, this group has the onset of T2DM and hypertension between 48 and 50 years. Distribution of patients from this cluster receiving monotherapy, dual therapy, triple therapy and quadruple therapy for T2DM is found to be 4.79%, 80.24%, 14.37% and 0.6% respectively (Figure 5.4B). This group has the most patients (71.86%) receiving anti-hypertensive therapy in the final visit, among whom 48.5% received monotherapy, 22.16% received dual therapy, and 1.2% received triple therapy (Figure 5.5B).

Cluster 3 (C3): C3 is found to have a history of T2DM of over 6 years and no history of hypertension at initial presentation. With a mean age of 49 years, the age of T2DM onset for this group is around 43 years which is similar to that of C1. Distribution of patients from this cluster receiving monotherapy, dual therapy, triple therapy and quadruple therapy for T2DM is found to be 1.87%, 55.14%, 34.58% and 8.41% respectively (Figure 5.4C). Though the mean systolic blood pressure (SBP) and diastolic blood pressure (DBP) in this group is higher than that of C1 (SBP: 131.15 mm Hg in C3 vs 119.38 mm Hg in C1, DBP: 81.4 mm Hg in C3 vs 78.72 mm Hg in C1), the values at presentation did not reach the cut-off for hypertension. 41.12% of the patients from this cluster turned out to be hypertensive at the final visit. Distribution of patients receiving anti-hypertensive therapy in the final visit is: mono therapy received by 23.36%, dual therapy by 14.95% and triple therapy by 1.87% (Figure 5.5C).

### 5.3.2 Assessment of IR and $\beta$ -cell dysfunction within the clusters

To assess the degree of IR and  $\beta$ -cell dysfunction in all the clusters, 100 patients (17 from C1, 49 from C2 and 34 from C3) are randomly selected out of 339 uncontrolled T2DM patients and subjected to OGTT. These are the only patients for whom the requisite attributes are on record. C1 has the maximal  $\beta$ -cell function [HOMA-B = 52.29%, Disposition Index (DI) = 1.55] whereas C3 is found to have the worst revival with HOMA-B value of 26.7% and DI of 0.55. C2 has a moderate revival in  $\beta$ -cell function with HOMA-B value of 40.27% and DI of 0.77. Interestingly, both the IR indices are found to be similar for all the clusters (Table 5.2). Among the three clusters, C3 has the worst glycemic control (HbA1c = 8.41%) and C2 the best glycemic control (HbA1c = 7.46%) while C1 has a moderate glycemic control (HbA1c = 7.74%).

## 5.4 Discussion on uncontrolled T2DM clusters

Using an unsupervised clustering workflow on a clinical dataset of uncontrolled T2DM patients, we discover 3 subtypes of T2DM with respect to disease duration, presence of hypertension and severity of  $\beta$ -cell dysfunction. Interestingly, we find that the degree of IR remains similar across all the 3 clusters though  $\beta$ -cell function is different among them. Thus the ability of anti-diabetic drugs to recover the pancreatic  $\beta$ -cell function remains critical towards achieving glycemic control in T2DM. Our analysis reveals that Cluster C1 consisting of patients with recent onset T2DM with no history of hypertension represents an early onset of T2DM group where majority of the patients (72.58%) received dual therapy (Glimepiride and Metformin). As  $\beta$ -cell function remains relatively preserved in the earlier stage of T2DM [103, 104], this cluster has the best  $\beta$ -cell function among the three clusters. However, standard dual therapy could not help achieve the HbA1c target of 7% thereby suggesting the need for a more aggressive protocol for glycemic control. In contrast, Cluster 2 has the highest age and BMI, the two most important risk factors for T2DM [105, 106] and is the only group presenting with hypertension. Thus, Cluster C2 represent the classical form of obesity-associated T2DM. Interestingly, this cluster showed the best glycemic control (HbA1c = 7.46%) with 80.24% patients having received dual therapy (Glimepiride and Metformin). Cluster C3 has the longest history of T2DM with an age of onset similar to that of C1 (around 43 years) and no complaint of hypertension and displays the most severe defect in  $\beta$ -cell function. Though C3 has the maximum number of patients receiving triple therapy (34.58% in C3 vs 14.37% in C2 vs 19.35% in C1) and quadruple therapy (8.41% in C3 vs 0.6% in C2 vs 1.61% in C1), it shows the worst recovery in  $\beta$ -cell function and thereby has the worst glycemic control (HbA1c = 8.41%).

Our dataset including the anti-diabetic drug combinations for all patients also reveals a differential ability of the three clusters to revive the  $\beta$ -cell function. The cluster with recent

onset T2DM has a modest ability to increase  $\beta$ -cell function in response to anti-diabetic drugs. However, among the other two clusters with long-standing T2DM, the cluster with accompanying hypertension showed better glycemic control with dual anti-diabetic therapy whereas the other one with no clinical history of hypertension has the worst glycemic control even with triple anti-diabetic therapy.  $\beta$ -cell dysfunction progresses with time in untreated T2DM [103, 104]. As expected, C1 in our study with recent onset T2DM is found to have the best  $\beta$ -cell reserve compared to the other clusters. However, the difference in  $\beta$ -cell reserve between both the clusters with long-standing T2DM may be explained by their obesity status. Obesity has been shown to positively associate with  $\beta$ -cell function [107] with obese T2DM patients found to have relatively higher levels of fasting [98] and post-prandial insulin [108] at diagnosis compared to non-obese T2DM patients. The cluster exhibiting greater revival in  $\beta$ -cell function as well as better glycemic control is the most obese with the highest body fat percentage. Though both the clusters have similar degrees of IR, the obese T2DM cluster is found to have a greater  $\beta$ -cell reserve compared to the non-obese cluster.

The age of onset of T2DM is 43 years for both the recent onset and the long-standing non-obese cluster while it is 49 years for the obese cluster. Though T2DM occurs mainly in the fifth decade of life, studies have reported an early age of onset in the fourth decade of life among Asian Indians [109]. The similarity of the obese T2DM cluster to obesity-associated T2DM in terms of age of onset (fifth decade of life) indicates that T2DM manifests in this subgroup as a component of Metabolic Syndrome. Expectedly this subtype presents with the co-morbidity of hypertension. However, it is observed that achieving glycemic control is relatively easier in this subgroup. On the other hand, the long-standing non-obese subtype with no history of hypertension typically manifests in the early fourth decade of life and shows similarity to the Metabolically Unhealthy Non-obese (MUNO) Phenotype [110]. As seen in our study, achieving the glycemic target is difficult for this subtype even with triple and quadruple anti-diabetic therapy.

The strength of the study is defined by the regular clinical follow-up of the T2DM patients. However one important limitation of this study is its dependence on the rural population from a single clinic. This may be overcome by making the present study multi-centric and multi-ethnic.

## 5.5 Clinical significance and summary of T2DM subtypes

Uncontrolled T2DM comprises of three heterogeneous clusters with respect to duration of disease, co-morbid condition and  $\beta$ -cell function. While obesity-associated T2DM patients are relatively amenable to standard therapeutic regimen, achieving glycemic control in the early-onset non-obese T2DM becomes difficult with increase in disease duration. Stratification

of uncontrolled T2DM patients on the basis of patho-clinical features remains the first step towards precision management in T2DM. In addition, specific therapeutic strategies are to be devised for early and rapid glycemic control for non-obese T2DM patients.

In summary, the following conclusions can be drawn from our study- Cluster 2 and Cluster 3 respectively represent the obesity-associated and Asian-Indian diabetic variants whereas a considerable number of Cluster 1 patients might transit to Cluster 3 following significant damage of  $\beta$ -cells. Cluster 2 patients should be prioritised for the management of associated diseases while Cluster 1 and 3 must have a tighter glycemic control. Current therapy in the management of diabetes and its associated comorbidities is based on the classical obesity-associated T2DM. However, there is a need to revisit anti-diabetic therapy in the context of the Asian-Indian variant.

# Oral minimal model estimation

## Chapter outline

- Augmented oral minimal model with rate of appearance of glucose
- Parameter estimation of obese and non-obese T2DM and sensitivity analysis
- Distribution of parameters and statistical test
- Discussion on estimated parameters for obese and non-obese T2DM

Oral glucose tolerance test (OGTT) is the primary test used to diagnose type 2 diabetes mellitus (T2DM) in a clinical setting. Analysis of OGTT data using the oral minimal model (OMM) along with the rate of appearance of ingested glucose ( $R_a$ ) is performed to study differences in model parameters for control and T2DM groups. The differentiation of parameters of the model gives insight into the behaviour and physiology of T2DM. The model is also studied to find parameter differences among obese and non-obese T2DM subjects and the sensitive parameters are co-related to the known physiological findings. Sensitivity analysis is performed to understand changes in parameter values with model output and to support the findings, appropriate statistical tests are done. This seems to be the first preliminary application of the OMM with obesity as a distinguishing factor in understanding T2DM from estimated parameters of insulin-glucose model and relating the statistical differences in parameters to diabetes pathophysiology.

## 6.1 Overview of insulin-glucose regulatory system

Insulin glucose system (IG) is the biochemical regulatory system that helps to maintain a steady glucose level in the blood. Glucose is stored in liver cells in the form of glycogen. When the level of glucose falls in the blood (due to exercise or a long gap after the last meal),

pancreatic  $\alpha$ -cells secrete glucagon to release glucose (glycogenolysis) from the liver through the breakdown of glycogen until the glucose level rises to normal. When the level of glucose rises in the blood (after a meal), pancreatic  $\beta$ -cells secrete insulin to trigger the uptake of glucose by the peripheral tissue cells in the body, via the GLUT4 (Glucose transporter Type 4), until the glucose level falls back to normal. People suffering from diabetes mellitus (DM) need to regularly monitor their blood glucose level for hypoglycemia or hyperglycemia. OGTT is a commonly used test where fasting blood glucose is recorded and the subject is asked to ingest a certain amount of glucose dissolved in water and subsequently, glucose and insulin readings are taken at varying intervals [111]. It is a highly sensitive test used for screening and diagnosis of pre-diabetes and T2DM.

Machine learning (ML) models are used to identify and group data when adequate information about the data is not available. Features are extracted from the data in an abstract way in most of the ML models which makes it not easily explainable [112]. Clustering is an important ML technique employed to understand data through classification [113, 114, 115]. However, the factors leading to the classification may not be evident from the results. For example, data on diabetic patients may reveal distinct pathoclinical clusters [116], however, the causative factors may not be evident. Biological systems can be mathematically modelled using ordinary differential equations (ODEs) utilising expert domain knowledge. These models are typically parameterised, with each parameter representing a specific aspect of the system being studied. The values of these parameters are often estimated from available clinical data, particularly individual-specific information. For instance, the insulin-glucose regulatory system can be modelled using ODEs, with the model parameters being derived from data collected from subjects with T2DM as well as a control group. The pool of parameter sets can then be studied to understand the causative factors with respect to the underlying model. Furthermore, the model can suggest specific internal processes that may be difficult to observe directly. This can be a significant advantage if the analytical results are verified to be reliable. In the current work, as features are not adequate and the available ones are time-series, there is less scope for using abstract models and we focus on ODE models for the insulin-glucose system.

In the Indian context, a great majority of people suffering from diabetes are non-obese, having low body mass index (BMI) [117]. This work is an investigation on the usability of the OMM, with the rate of appearance of ingested glucose ( $R_a$ ), for studying differences in obese and non-obese T2DM. The OMM is based on the Bergman minimal model [13], where glucose analysis is done using an intravenous glucose tolerance test (IVGTT). The identification of subject-specific oral minimal models is carried out for the insulin-glucose system from the observed glucose and insulin readings in the OGTT data and the rate of appearance of ingested glucose from the OMM is inferred. The model parameters of the participating subjects are



studied towards suggesting differences in obese and non-obese diabetic physiology. A key consideration in this analysis is model identifiability. Both the minimal model based on IVGTT and the rate of glucose appearance (Model 2 as presented in [38]) are globally identifiable when both glucose and insulin levels are observed.

The problem with physiological models is the estimation of set of parameters both for an individual and a population. Some of the states in the model are not completely measurable, so estimating parameters becomes an important step. The unknown states are characterised by approximating the states and outputs that are measurable. The following sections describe the mathematical description of the oral minimal model, theory and methodology, followed by results and discussion.

## 6.2 Augmented oral minimal model

The OMM with rate of appearance of ingested glucose,  $R_a$ , is the computational model used in the work. This particular  $R_a$  model [38] is used as it is non-linear in accordance with the biphasic nature of gastric emptying of glucose. Initially the stomach contains the amount of ingested glucose. The gastric emptying rate then decreases to a minimum and rises back to  $k_{max}$  – this behaviour is also exhibited in the data used in this work. Also, the  $R_a$  model is modified by removing the solid phase glucose compartment from the stomach as it deals with a grinding rate parameter which is not relevant to our study as glucose solution is orally ingested. Equations 6.1, 6.2 and 6.3 form the modified  $R_a$  model and equations 6.4, 6.5 and 6.6 are same as explained in Section 2.4.

$$\frac{dq_{sto}}{dt} = -k_{empt} \cdot q_{sto}(t) + D \cdot \delta(t), \quad (6.1)$$

$$\frac{dq_{gut}}{dt} = -k_{abs} \cdot q_{gut}(t) + k_{empt} \cdot q_{sto}(t), \quad (6.2)$$

$$R_a(t) = f \cdot k_{abs} \cdot q_{gut}, \quad (6.3)$$

$$\frac{dG}{dt} = -p_1 \cdot (G - G_b) - X \cdot G + \frac{R_a(t)}{V}, \quad (6.4)$$

$$\frac{dX}{dt} = -p_2 \cdot X + p_3 \cdot (I - I_b), \quad (6.5)$$

$$\frac{dI}{dt} = \gamma \cdot (G - G_T)^+ - p_4 \cdot (I - I_b), \quad (6.6)$$

The amount of glucose in the stomach is described by  $q_{sto}$ ,  $\delta(t)$  is the impulse function,  $D$  is the amount of ingested glucose,  $q_{gut}$  is the glucose mass in the intestine,  $k_{empt}$  is the rate of gastric emptying,  $k_{abs}$  is the rate constant of intestinal absorption and  $f$  is the fraction of the

intestinal absorption which actually appears in plasma.

$$k_{\text{empt}}(q_{\text{sto}}) = k_{\text{min}} + \left( \frac{k_{\text{max}} - k_{\text{min}}}{2} \right) \cdot \{ \tanh(\alpha \cdot q_{\text{sto}} - b \cdot D) - \tanh(\beta \cdot (q_{\text{sto}} - c \cdot D)) + 2 \} \quad (6.7)$$

$$\alpha = \frac{5}{2 \cdot D \cdot (1 - b)}, \beta = \frac{5}{2 \cdot D \cdot c}$$

The gastric emptying rate,  $k_{\text{empt}}$  is a function of the amount of glucose in the stomach  $q_{\text{sto}}$  (Eqn 6.7). It equals  $k_{\text{max}}$  when the stomach contains the amount of the ingested glucose,  $D$ , then it decreases to a minimum value of  $k_{\text{min}}$ .  $b$  is the percentage of the dose for which  $k_{\text{empt}}$  decreases at  $(k_{\text{max}} - k_{\text{min}})/2$ .  $c$  is the percentage of the dose for which  $k_{\text{empt}}$  is back to  $(k_{\text{max}} - k_{\text{min}})/2$ .

### 6.3 Datasets used for parameter estimation

Two sets of data are used in the study. Dataset I consists of OGTT data of 300 individuals, of which 129 are normal subjects and others are type 2 diabetics. The outliers were removed and the final data has 116 normal subjects and 147 subjects with T2DM. These patients were asked to ingest 25g of glucose dissolved in 100 mL of water after fasting for 8-12 hours. Sample data from Dataset II are given in Table 6.1(a). FGLU indicates fasting glucose level.  $\text{GLU}\langle t \rangle$  and  $\text{INS}\langle t \rangle$  indicate glucose level and insulin level after time  $\langle t \rangle$  of orally ingesting glucose beyond fasting glucose measurement (at  $t = 0$ ), respectively. For example,  $\text{GLU}30$  indicates glucose level after 30 min. The measurements are taken at times 0, 30, 60, 90 and 120 minutes.

Dataset II consists of 40 T2DM patient data collected from a community clinic with 21 non-obese subjects ( $\text{BMI} \leq 25$ ) and 19 obese subjects ( $\text{BMI} > 25$ ) who underwent OGTT. The outliers were removed and the final data has 15 non-obese subjects with T2DM and 14 obese subjects with T2DM. These patients were asked to ingest 75g of glucose dissolved in 100 mL of water after fasting for 8-12 hours. The sample data from dataset-2 are shown in Table 6.1(b). The Dataset II consists of the following columns: Sex, Age, BMI, waist circumference (WC), Weight, Height and measurements of insulin and glucose at various time points. The column names are similar to Dataset-1 with only changes in time points. The measurements in Dataset II are taken at times 0, 15, 30, 45, 60, 90 and 120 minutes. All participants gave informed consent and the study was approved by the Institutional Human Ethics Committee of CSIR-IICB. For example,  $\text{INS}45$  indicates insulin level after 45 min.

### 6.4 Sensitivity indicators and parameter estimation

Sensitivity indicators, such as Parameter Importance Index and Collinearity Index, are statistical measures used to assess the sensitivity of a model's output to its input parameters. These

Code	FGLU	GLU15	GLU30	GLU60	GLU90	GLU120	FINS	INS15	INS30	INS60	INS90	INS120
1	202.5	237	279	313.02	384	385.02	90	200	220	170	300	300
2	195.48	263	314	385.92	419	419.04	18	13	21	34	42	38
3	120.96	154	232	311.04	364	383.94	7	21	30	54	69	70
4	148.50	194	249	347.04	376	334.08	40	33	90	150	190	140
5	256.50	299	320	417.96	455	466.02	6.5	13	24	36	31	38

(a) Sample from Dataset I

Code	SEX	AGE	BMI	WC	Weight	Height	FINS	INS15	INS120	FGLU	GLU15	GLU120
DM-24	F	48	23.8	82	47.4	141	0.64	0.47	6.38	167.80	192.50	400.69
DM-09	F	55	21.7	80	47.1	146	5.30	10.00	25.17	159.73	189.39	398.13
DM-28	F	42	24.2	86	51.5	146	2.42	3.42	31.06	145.86	146.23	245.98
DM-34	F	40	24.6	84	50.4	143	7.59	8.08	36.48	140.83	179.63	344.74
DM-06	F	29	20.8	75	45.9	147	7.24	10.43	36.63	154.75	178.68	309.04

(b) Sample from Dataset II

Table 6.1: Sample rows from Dataset I and Dataset II used for parameter estimation of OMM. Glucose units are in  $mg/dL$  and insulin units are in  $\mu U/mL$

indicators help identify which parameters have the most significant impact on the model variables and detect potential issues with parameter interactions. The parameter importance index  $\delta_l$  [118] can be calculated using the sensitivity of the  $l$ -th parameter  $p_l$  to the output vector as follows

$$\delta_l = \sqrt{\frac{1}{m} \cdot \sum_{i=1}^m S_{il}^2}, \quad (6.8)$$

where  $m$  is the number of parameters of the model. The collinearity index is calculated to analyse whether pairs of parameters have the same influence on the output vector. If two parameters have the same effect on the output vector, they are likely to be not identifiable. For this normalised sensitivity matrix is calculated as follows

$$S_{il}^N = \frac{S_{il}}{\delta_l \cdot \sqrt{m}}. \quad (6.9)$$

The submatrices are calculated as  $m \times l$  with first  $l$  columns of  $S_N$  with respect to sensitivities of first  $l$  parameters. Then the smallest eigenvalue  $\lambda_l$  of each  $l \times l$  matrix is computed and the index is defined as follows

$$\gamma_{col} = \frac{1}{\sqrt{l}}. \quad (6.10)$$

The parameters  $p_1$  to  $p_l$  are linearly dependent if the collinearity index is larger.

The model parameters in Section 6.2 need to be estimated using the datasets described in Section 6.3. A parameter optimisation problem is defined that minimises the error between outputs obtained from solving ODE and the observed data. The methods are implemented and coded in the Python framework using the Scipy packages `odeint()`, `minimize()` and `dist()` [119]. Parameter estimation runs were performed for each individual data. Different estimation techniques were tried out and the Nelder-Mead method [40] gave a better fit of the observed

data. A weighted error function  $E(n)$  is defined, which works with Nelder-Mead optimisation,

$$E(n) = \sum_{i=1}^m w_i \cdot \left( \frac{(y_i - y_{i_{model}}(n))}{\sigma_k} \right)^2, \quad (6.11)$$

where  $y_i$  are data points,  $y_{i_{model}}$  are simulated points and  $w_i$  are vector of weights associated with each time-point. In the OMM estimation, the initial time intervals are crucial and so higher weights are given to initial time intervals (15, 30 and 60 minutes) compared to the later ones.

## 6.5 Results from the oral minimal model using OGTT data

**Sensitivity indicators** The implication of doing sensitivity analysis is to understand the varying levels of sensitivity among the parameters. The less sensitive parameters are replaced with nominal values, but the overall fitting is not satisfactory. So, it is understood that the small variation in less sensitive parameters may be considered unimportant. However, these parameters cannot be ignored altogether as they may be significant for larger variation among any disease sub-categories. The range of these parameters may change over sub-groups. For less sensitive parameters, smaller variation is not considered as they vary significantly across subgroups. Morris screening algorithm is used to analyse the sensitivity of each individual data. The sensitivity order of parameters for all the groups remained mostly similar.

The parameter importance index (Eqn 6.8) and collinearity index (Eqn 6.9, 6.10) were computed for all the patients in each dataset and the order is similar for most of the individuals. The ordering of parameters based on the parameter importance index is shown in Figure 6.1(a).  $G_T$  is the parameter to which the output variable is most sensitive. The next set of important parameters are  $G_b, p_4, \gamma, f$  and  $p_1$ . The least important parameters are  $I_b$  and  $p_2$ . The collinearity index graph (Figure 6.1(b)) indicates that most of the parameter combinations close to the x-axis are not strongly collinear with each other. This observation is mostly similar across all individuals, with variations only among parameters that are less sensitive. The parameters which are collinear are found to be less sensitive to the output.

**Results of parameter estimation** The estimation considered 12 different parameters. The values of initial parameters are mainly from the literature and are used as required to match the unit of measurement. The initial values of insulin and glucose are taken at time point 0 (the time at which glucose is ingested) in the data and are not estimated. The model is estimated for parameter values for each of the groups separately. Simulation results of two random data samples from each of the control and type 2 diabetic groups are shown in Figure 6.2 using Dataset I. Sample results of two random subjects' data from obese and non-obese type 2 diabetic groups are shown in Figure 6.3 using Dataset II.

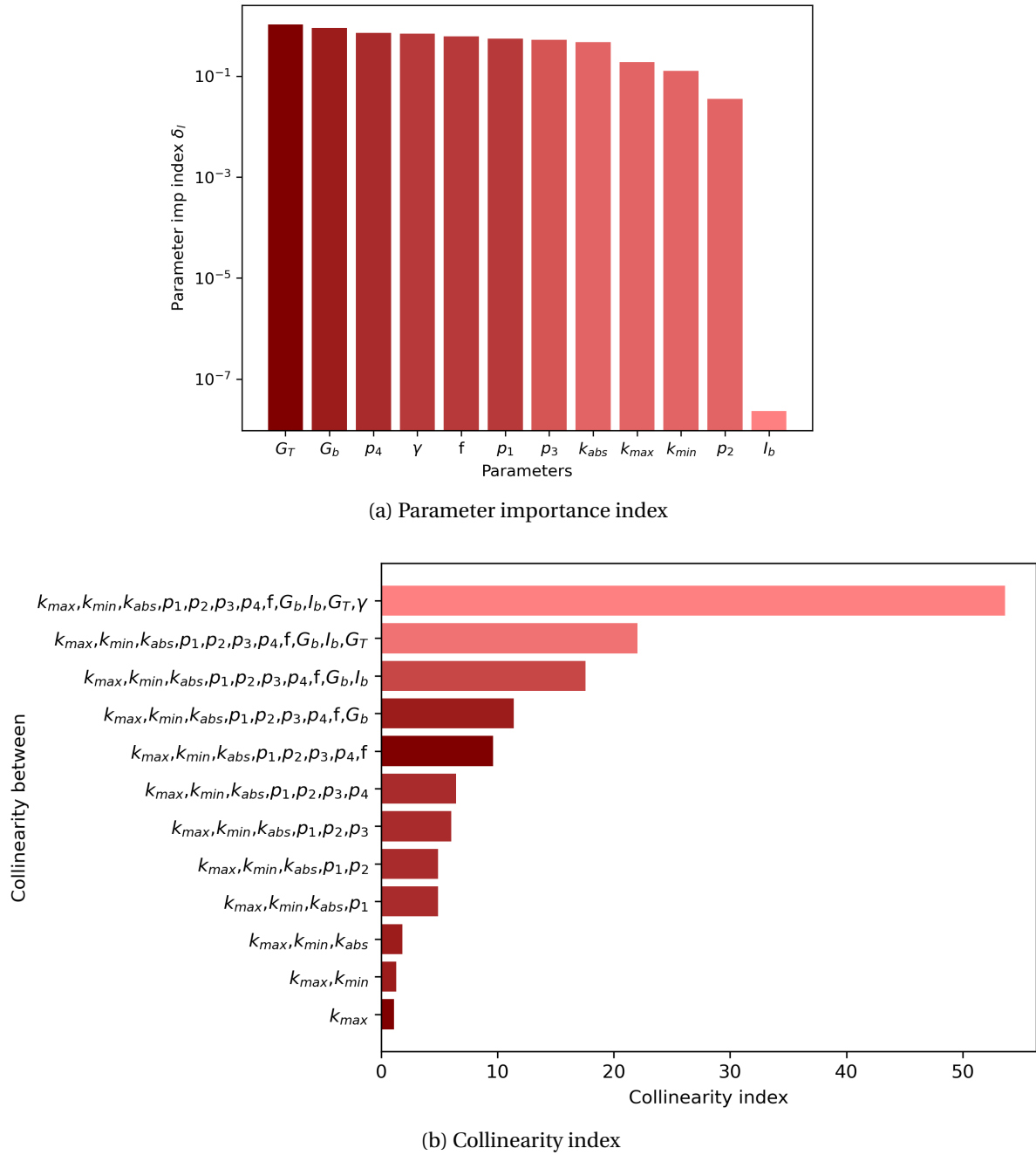
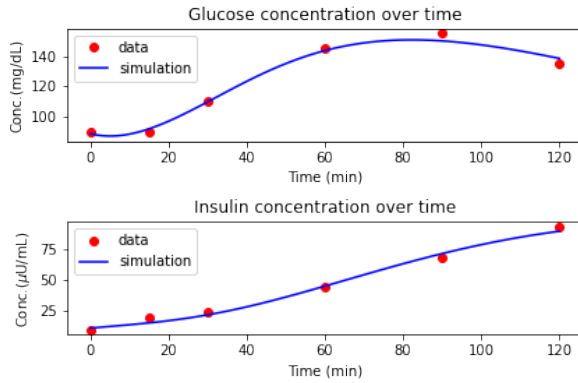
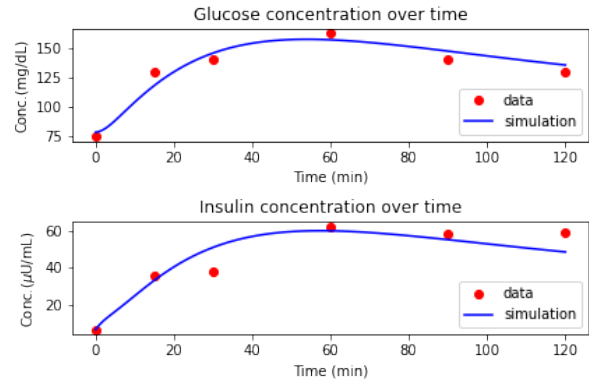


Figure 6.1: Parameter order in accordance with parameter importance index and collinearity index

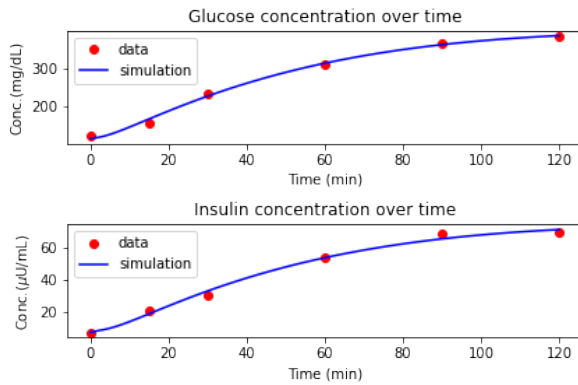
**Distribution of parameters and statistical test** The median values of parameters for control and diabetic groups are shown in Table 6.2. For obese diabetic and non-obese type 2 diabetic groups, the box plots are shown in Figure 6.4 and median values are shown in Table 6.3. The box plots of parameters which are distinctive are only shown. Each parameter has a different distribution and the statistical tests are chosen based on the corresponding distribution. For instance, in the control and diabetic group,  $k_{min}$  follows the Weibull distribution (Table 6.2)



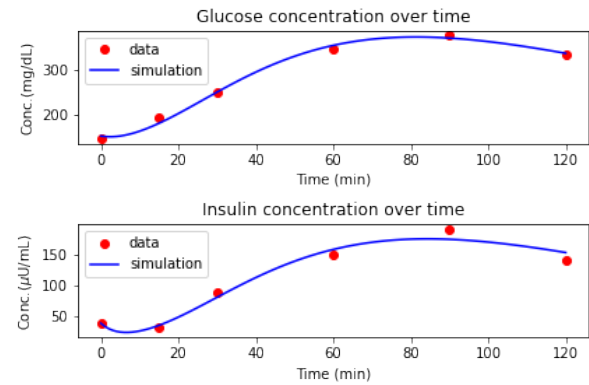
(a) Sample from control group (Subject 1)



(b) Sample from control group (Subject 2)



(c) Sample from diabetic group (Subject 1)



(d) Sample from diabetic group (Subject 2)

Figure 6.2: Sample simulation results for control and diabetic group

Parameter	Control group	Diabetic group	p-value	Distribution	Statistical test	Distinct feature
$k_{\max}$ [1/min]	0.0179	0.0269	0.000143	lognormal	Independent ttest	Yes
$k_{\min}$ [1/min]	0.00137	0.00751	7.49e-27	Weibull	Mann Whitney U test	Yes
$k_{\text{abs}}$ [1/min]	0.0289	0.0208	0.0901	lognormal	Independent ttest	No
$p_1$ [1/min]	0.0609	0.0208	2.45e-26	Weibull	Mann Whitney U test	Yes
$p_2$ [1/min]	0.0016	0.0016	0.579	Weibull	Mann Whitney U test	No
$p_3$ [(L/μmol)min <sup>-2</sup> ]	1.5e-07	1.5e-07	0.618	Weibull	Mann Whitney U test	No
$p_4$ [1/min]	0.154	0.863	5.61e-26	Weibull	Mann Whitney U test	Yes
$f$ [dimensionless]	0.756	0.891	0.000125	beta	Mann Whitney U test	Yes
$G_b$ [μmol/L]	4.76e+03	7.3e+03	1.06e-12	beta	Mann Whitney U test	Yes
$I_b$ [μmol/L]	2.22e-11	2.22e-11	0.232	Weibull	Mann Whitney U test	No
$G_T$ [μmol/L]	4.24e+03	3.57e+03	0.364	gamma (Γ)	Mann Whitney U test	No
$\gamma$ [min <sup>-2</sup> ]	2.8e-08	1.27e-08	2.15e-09	Weibull	Mann Whitney U test	Yes

Table 6.2: Comparison of parameters in control and diabetic group (Dataset I)

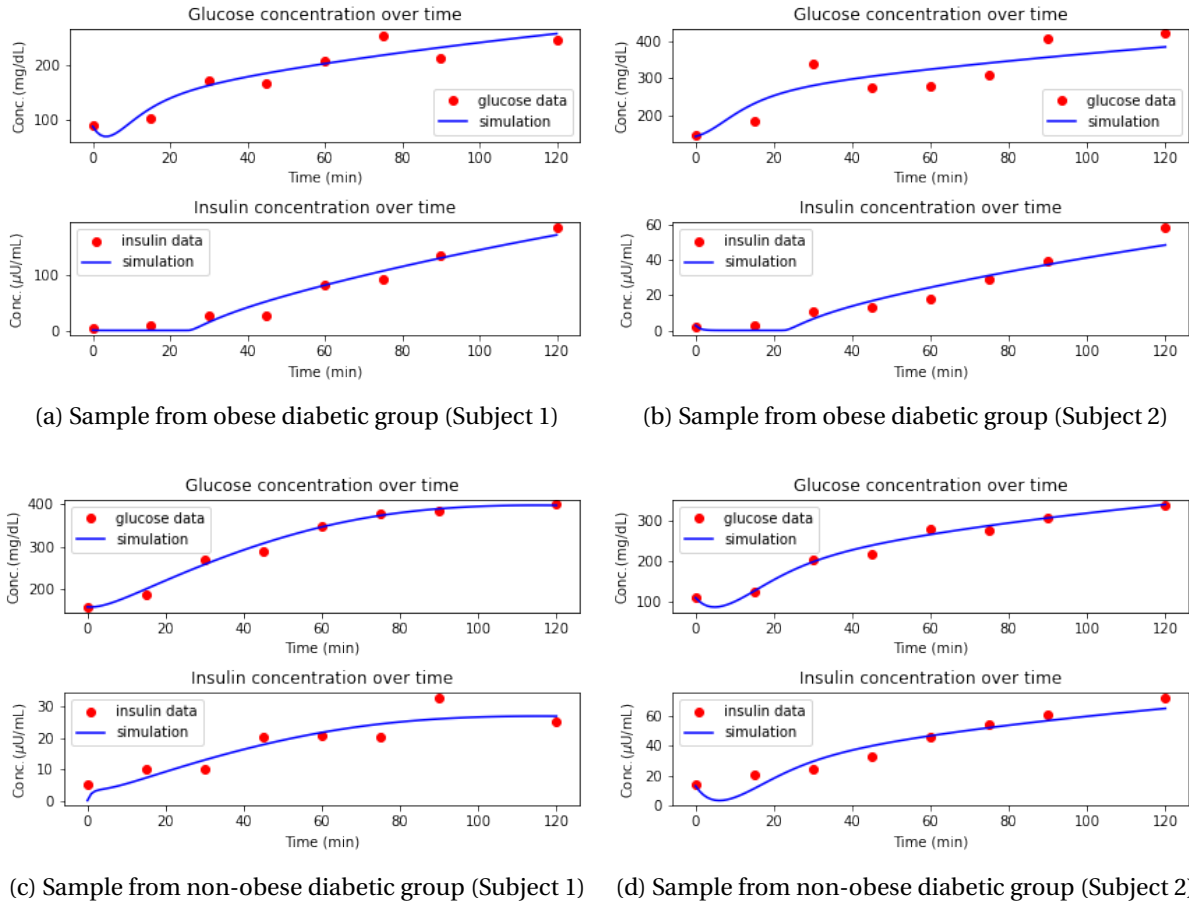


Figure 6.3: Sample simulation results for obese and non-obese diabetic group

Parameter	Obese diabetic (median)	Nonobese diabetic (median)	p-value	Distribution	Statistical test	Distinct feature
$k_{\max}$ [1/min]	0.0166	0.0182	0.561	gamma ( $\Gamma$ )	Mann Whitney U test	No
$k_{\min}$ [1/min]	0.00563	0.00252	0.000925	t	Mann Whitney U test	Yes
$k_{\text{abs}}$ [1/min]	0.00722	0.00141	1.16e-06	exponential	Mann Whitney U test	Yes
$p_1$ [1/min]	0.0565	0.15	0.000141	Weibull	Mann Whitney U test	Yes
$p_2$ [1/min]	0.00153	0.00808	0.451	Weibull	Mann Whitney U test	No
$p_3$ [(L/ $\mu\text{mol}$ ) $\text{min}^{-2}$ ]	2.48e-07	5.97e-08	0.186	Weibull	Mann Whitney U test	No
$p_4$ [1/min]	0.91	0.921	0.146	Weibull	Mann Whitney U test	No
f [dimensionless]	0.826	0.784	0.0259	Weibull	Mann Whitney U test	Yes
G <sub>b</sub> [ $\mu\text{mol/L}$ ]	7.31e+03	5.04e+03	0.0292	beta	Mann Whitney U test	Yes
I <sub>b</sub> [ $\mu\text{mol/L}$ ]	5.34e-12	5.01e-13	0.533	beta	Mann Whitney U test	No
$G_T$ [ $\mu\text{mol/L}$ ]	4.58e+03	9.27e+03	0.102	gamma ( $\Gamma$ )	Mann Whitney U test	No
$\gamma$ [ $\text{min}^{-2}$ ]	1.65e-08	3.54e-09	0.134	gamma ( $\Gamma$ )	Mann Whitney U test	No

Table 6.3: Comparison of obese diabetic and non-obese diabetic group (Dataset II)

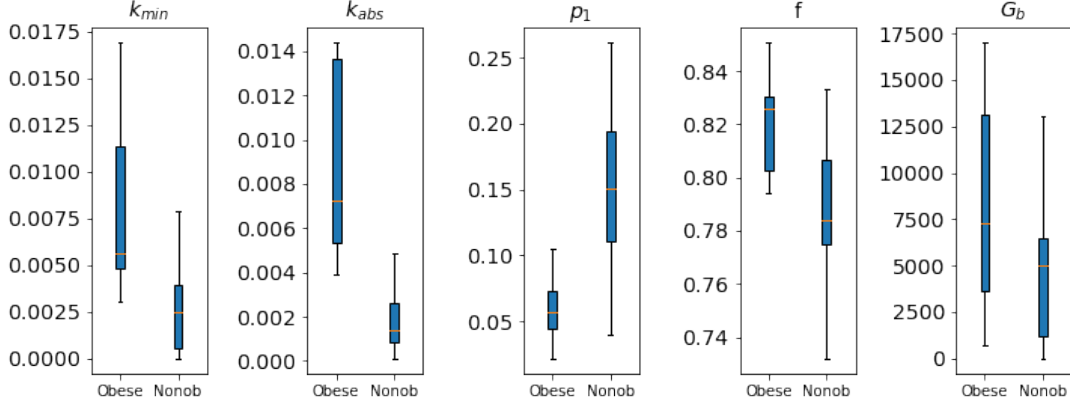


Figure 6.4: Box plots of parameters for obese and non-obese diabetic

and a common non-normal distribution test, known as the Mann Whitney U test, is used. Similarly, based on parameter distribution, appropriate statistical tests are chosen and the p-value is determined with significance level  $\alpha = 0.05$ . The distribution followed and tests used for the control and diabetic group using Dataset I are shown in Table 6.2 and for the obese and non-obese diabetic group using Dataset II are shown in Table 6.3.

## 6.6 Discussion on the findings

Our results emphasise the difference in parameter values in different T2DM groups. From Table 6.2 the control group differs from T2DM groups based on parameters  $k_{max}$ ,  $k_{min}$ ,  $f$ ,  $p_1$ ,  $G_b$ ,  $p_4$ , and  $\gamma$ . The model parameters conform with the previously published results for diabetic patients, which are described in detail in this section. The maximum and minimum amount of glucose emptying ( $k_{max}$  and  $k_{min}$ ) for the diabetic group is higher. In support of this observation, it can be argued that rapid gastric emptying is a frequent diabetic complication [120]. To improve the postprandial glycemic control in these patients, a slowing gastric emptying rationale may be considered. The fraction of intestinal absorption ( $f$ ), which actually appears in plasma, is higher for the diabetic group. In patients with T2DM, an increase of the Sodium/GLucose coTransporter 1 (SGLT1) protein and its mRNA in the enterocytes of the small intestine were found, which is involved in increased glucose absorption through the apical membrane [121]. The insulin-independent glucose uptake rate ( $p_1$ ), which also represents glucose effectiveness, is higher in the control group than in the type 2 diabetic group. This may be attributed to insulin resistance among the type 2 diabetic group. The basal level of glucose ( $G_b$ ) is higher in the diabetes group, which indicates pre-diabetes or diabetes, which is expected. As this parameter is less sensitive, the variation among subgroups is larger but less significant. Smaller changes in basal level of glucose do not change the model behaviour. The decay rate for insulin ( $p_4$ ) is significantly higher in the diabetic group.



The excess insulin produced is rapidly attenuated due to acute induction of mitochondrial superoxide production [122]. The rate of  $\beta$ -cells release of insulin ( $\gamma$ ) after the oral glucose intake is lower in the diabetic group. This may be due to a decrease in insulin sensitivity and secretion caused by a delay in glucose peak time [123].

The analysis of data using the model for obese and non-obese diabetic groups indicates that the parametric range is different for the two groups. From Table 6.3, it can be observed that the obese diabetic group differs from non-obese diabetic groups based on parameters  $k_{min}$ ,  $k_{abs}$ ,  $f$ ,  $p_1$  and  $G_b$ . The five estimated parameters are significant for obese and non-obese T2DM groups. The minimum level of gastric emptying ( $k_{min}$ ) is higher for obese, i.e., from that point, the emptying rate increases to  $k_{max}$ , not reducing any further. Obese subjects have a more rapid emptying rate for solids than non-obese subjects [124]. The absorption rate in the intestine ( $k_{abs}$ ) is higher for obese diabetics. The body surface area for obese subjects is larger and therefore, the absorption rate increases. The insulin-independent rate of glucose uptake ( $p_1$ ) is higher for non-obese diabetics. This observation is associated with adiposity with declining glucose tolerance. During increased insulin resistance, this mechanism helps preserve glucose uptake [125]. The fraction of intestinal absorption ( $f$ ) is higher for the obese group. This has the same explanation for why  $k_{abs}$  is higher for the obese diabetic group. The basal level of glucose ( $G_b$ ) is higher for the obese type 2 diabetic group. These significant changes in parameter values for obese and non-obese indicate a higher risk of diabetes as it causes both insulin resistance and  $\beta$ -cell dysfunction.

The differences in obese and non-obese diabetes groups have to be discussed further as it is being considered a novel category of diabetes. The degree of change and the direction of parameter values to increase or decrease diabetes risk is studied. The risk of T2DM increases for an obese group with an increase in the value of parameter  $k_{min}$  as compared to the non-obese group. The insulin-independent rate of glucose uptake ( $p_1$ ) is 2.6 times lower in the obese diabetic group, indicating a higher risk in obese diabetes. The absorption rate in the intestine ( $k_{abs}$ ) in the obese group is five times higher than in the non-obese group, indicating higher glucose levels in the non-obese diabetes group. This is also supported by a higher fraction of intestinal absorption ( $f$ ). The higher basal glucose level ( $G_b$ ) in the obese diabetic group indicates more insulin resistance. From the comparison, it can also be concluded that the obese group is more insulin-resistant than the non-obese group.

## 6.7 Conclusion

In this study, we adapted a published model to determine the significant changes in parameter values for healthy and T2DM groups and further among obese and non-obese T2DM groups. This is the first attempt to use the oral minimal model to co-relate the parameters for different

diabetic groups using OGTT data. The parameter values which are sensitive among groups are co-related to the known biological findings. The obese diabetic group are more insulin resistant, whereas the non-obese diabetic group are less insulin resistant. As non-obese T2DM is becoming more recognised, a model to identify the differences in insulin secretion and glucose absorption among these phenotypes is important. Using this model estimation, more relevant studies can be conducted, which pave the way to precision therapy in T2DM.

# Augmenting the insulin-glucose regulatory model with obesity indicators leptin and BMI

## Chapter outline

- The role of obesity indicators in insulin-glucose regulatory system
- Augmented oral minimal model including leptin and BMI
- Parameter estimation and sensitivity analysis of the augmented model
- Revision of indices derived from OGTT
- Results and validation of the augmented model

Leptin is a fat-derived hormone involved in satiety and body weight regulation. It also plays a critical regulatory role in the insulin-glucose regulatory system by modulating glucose metabolism and energy homeostasis. However, existing insulin-glucose models often fail to consider the impact of body weight indicators such as body mass index (BMI) and plasma leptin. To address this limitation, we propose augmenting the ordinary differential equations (ODEs) of the Oral Minimal Model (OMM) with an additional equation incorporating leptin as well as supplementary terms and parameters.

By estimating the model parameters, the model behaviour is aligned with the observed data of glucose, insulin and leptin for individuals with type 2 diabetes mellitus (T2DM). Based on model behaviour, revised indices formulated from Oral Glucose Tolerance Test (OGTT) data by including BMI and fasting leptin values correlate better with existing OGTT indices.

Additionally, parameter sensitivity analysis is performed to investigate the influence of the model parameters on the observed variables. Validation of the augmented model with clinical data (without leptin) demonstrates a superior fit to glucose and insulin data compared to the base model. This model emphasises the intricate associations between leptin, glucose and insulin concentrations with a potential for developing targeted interventions and therapies for T2DM. Notably, this work introduces the first ODE-based model incorporating leptin and BMI in the insulin-glucose pathway.

## 7.1 Association between obesity indicators and insulin-glucose dynamics in T2DM

T2DM occurs when the body becomes resistant to the effects of insulin and the pancreas fails to produce enough insulin to compensate for this resistance [126]. Excess body fat (obesity), particularly abdominal or visceral fat, is associated with insulin resistance (IR), a condition in which the cells of the body become less responsive to insulin. This reduced sensitivity to insulin leads to elevated blood glucose levels resulting in the characteristic symptoms and complications of T2DM [126]. Obesity-induced IR is followed by  $\beta$ -cell dysfunction leading to T2DM, however, the contribution of obesity and  $\beta$ -cell dysfunction in T2DM development vary significantly among the non-obese and obese population [98, 116].

Obesity is a major risk factor for the development of T2DM [127]. Leptin, a hormone secreted by fat cells, plays a crucial role in regulating body weight and energy balance. Leptin acts on the hypothalamus to suppress appetite and increase energy expenditure [128]. Additionally, leptin has various secondary functions in peripheral tissues, wherein it influences insulin level and glucose metabolism in the body. In the relationship analysis between leptin and body mass, serum leptin concentration is found to have a direct positive association with BMI [129].

There are several theoretical works [130, 131] that emphasise the role of leptin in IR and its importance in T2DM. However, to the best of our knowledge, there is a lack of mathematical models that adequately capture this behaviour. To overcome this limitation, a mathematical model including the role of leptin and BMI in the IG pathway is developed by augmenting the OMM proposed by Dallaman et al. [34, 38]. Our model aims to incorporate the regulatory influence of leptin on insulin-glucose dynamics, thereby improving the existing model fitting of glucose and insulin data and accurately capturing the leptin data. Consistent with the OMM approach, the current model uses clinical data derived from OGTT for parameter estimation and analysis.

OGTT is a frequently used diagnostic test in which the subject's fasting blood glucose levels are measured, followed by the ingestion of a specific amount of glucose dissolved in water

and subsequently, blood is collected at regular intervals to measure glucose and insulin levels at varying intervals [111]. The dataset used for modelling consists of data of individuals who have undergone a 2-hour OGTT. In addition to insulin and glucose, leptin concentrations are also taken at various time points for this study. OGTT indices (derived from OGTT data) are quantitative measures that provide information about the subject's insulin sensitivity,  $\beta$ -cell function and overall glucose regulation in response to a glucose load. The current work attempts to suggest modifications and enhancements to selected OGTT indices.

One of the crucial aspects of modelling is to understand how parameter values change with model output. The parameter variations affect the output of a model when dealing with critical data specific to biochemical processes. The interpretation of parameter values in these pathways is important to understand the relationship to their physiological properties. Sensitivity analysis on the augmented model is also conducted to explore the behaviour of model parameters, providing inference on the model's outcomes. Additionally, validation of the model is performed using a different dataset, as detailed in the subsequent sections. The present study introduces a novel approach to modelling the interaction between leptin and the insulin-glucose regulatory system, which has not been previously explored in detail.

## 7.2 Modelling mechanism

By integrating BMI and the adipokine leptin into the existing OMM, their associations with insulin and glucose concentrations can be modelled. Leptin is associated with insulin resistance and is positively correlated with BMI [132]. Leptin decreases insulin secretion to a certain extent [133] and also exhibits a negative correlation with fasting plasma glucose and postprandial glucose after two hours. Leptin also decreases hepatic glucose production, increases insulin sensitivity and decreases glucagon levels. This inhibitory action of leptin is modelled using enzyme kinetics in this work. Insulin, in turn, also plays a role in stimulating leptin production and secretion in the adipose tissue [130]. Impaired glucose metabolism and prolonged insulin resistance are also associated with a deficiency in leptin concentration [131].

The idea of parameter estimation is to find unknown parameters in a computational model which may describe the given biophysical process or phenomenon. These unknown parameters are estimated using experimental data which are collected from well-defined and standard conditions by minimising the error between the model simulations results and experimentally known data [134]. This way the behaviour of the model is captured effectively. The challenge is that no single optimal estimation technique exists for all models. Many different estimation methods have been developed so far to determine the best strategy for a given problem. The commonly used parameter estimation methods include maximum

likelihood estimation [39] and Nelder-Mead optimisation [40]. Nelder-Mead method is a numerical method in non-linear optimisation problems to find the minima of objective functions. Least square estimation is used in regression models and maximum likelihood estimation is used in statistical models [135]. In addition, evolutionary methods [136], namely genetic algorithms and particle swarm optimisation, are other efficient parameter estimation methods used in non-linear dynamic models.

**OGTT indices** The commonly used indices to access  $\beta$ -cell function and insulin resistance are briefly explained. HOMA-B (Homeostatic Model Assessment of  $\beta$ -Cell Function) is a mathematical model used to estimate  $\beta$ -cell function based on fasting glucose and fasting insulin levels, particularly in the context of insulin resistance and T2DM. The formula for calculating HOMA-B is  $\left(\frac{I_0 - 20}{G_0 - 3.5}\right)$ , where  $I_0$  is fasting insulin (in  $\mu U/mL$ ) and  $G_0$  is fasting glucose (in  $mmol/L$ ) values. HOMA2 is an improvement over the original HOMA model, as it considers glucose dynamics across population and different physiological states. Several software tools that accurately calculate HOMA2 values use advanced algorithms, as in the HOMA calculator [137] provided by the diabetes trials unit at the University of Oxford. HOMA-IR (Homeostatic Model Assessment of Insulin Resistance) is a widely used method for estimating insulin resistance, which is a key component of various metabolic disorders, including T2DM. The formula for calculating HOMA-IR is  $\left(\frac{I_0 \cdot G_0}{22.5}\right)$ , where  $I_0$  is in  $\mu U/mL$  and  $G_0$  is in  $mmol/L$ .

The insulin sensitivity index (ISI) is derived from the concept that insulin sensitivity can be assessed by evaluating the ability of insulin to regulate glucose levels. The formula for calculating ISI is  $\left(\frac{\Delta I_{0-30}}{\Delta G_{0-30}}\right)$ , where  $\Delta$  is the difference in fasting and 30 min values of insulin (I) and glucose (G). Higher ISI values indicate greater insulin sensitivity, meaning that the cells in the body are more responsive to insulin and effectively utilise glucose. The oral disposition index,  $DI_o$  [138] calculated as  $\left(\frac{\Delta I_{0-30}}{\Delta G_{0-30}}\right) \cdot \left(\frac{1}{I_0}\right)$ , where  $I_0$  is in  $pmol/L$  and  $G_0$  is in  $mmol/L$ , is commonly used as a measure of  $\beta$ -cell function. The *base* model used in this work is described in Section 6.2, Chapter 6. Our aim is to develop an underlying well-established model that includes the role of BMI and leptin by estimating the set of parameters that are significant.

### 7.3 Augmented model including leptin and BMI

Leptin level plays a major role in glucose homeostasis and is positively correlated with adiposity. In the presence of high blood glucose levels, leptin enhances tissue glucose uptake, independent of insulin and reduces hepatic glucose production [130]. Leptin has glucose-lowering effects via an insulin-independent mechanism which normalises hepatic glucose production and increases glucose uptake in peripheral tissues, including adipose tissue, muscle and heart [139]. Insulin promotes adipogenesis (formation of fat cells from pre-adipocytes)

and increases body fat mass [140]. It also stimulates the production and release of leptin, which acts centrally to decrease food intake and boost energy expenditure. Leptin, in turn, inhibits insulin secretion through both central and direct effects on  $\beta$ -cells. As plasma leptin level is directly linked to body fat mass, higher adiposity raises plasma leptin, thereby reducing insulin production and promoting a further increase in fat mass [141] (Figure 7.1). In this way, both leptin and insulin regulate each other, sharing the control of food intake and metabolism [130].

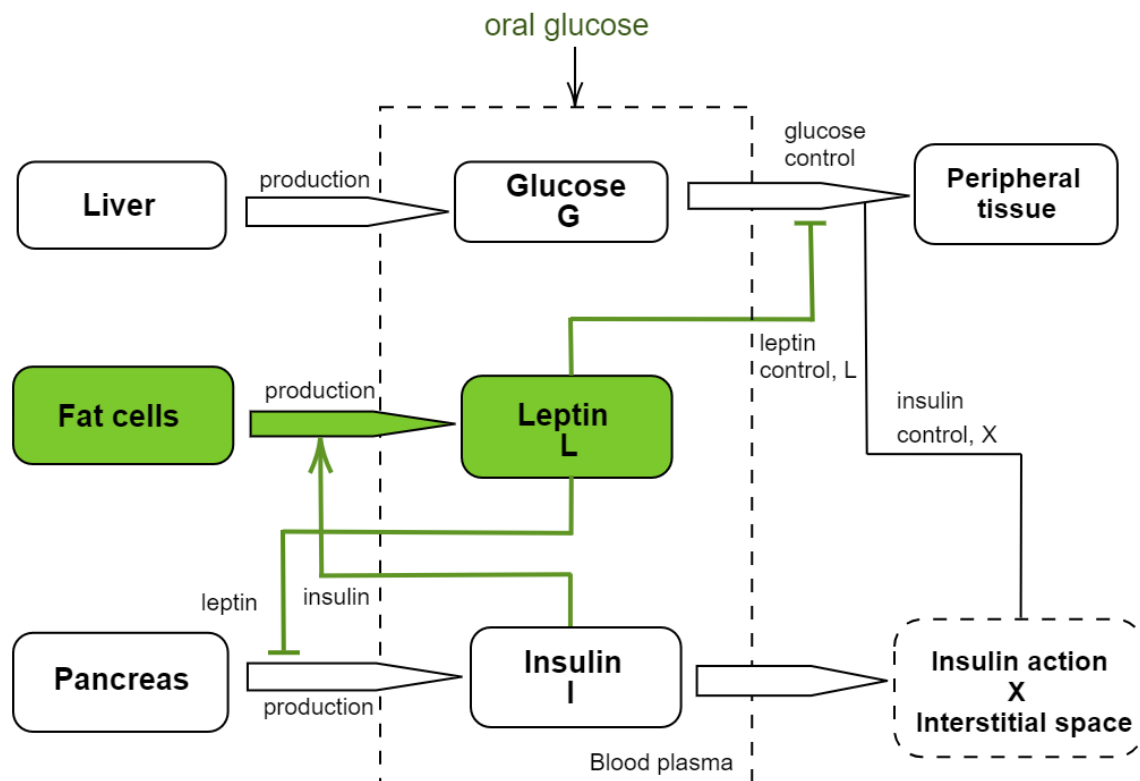


Figure 7.1: *Leptin* model: The augmentations to the model are visually highlighted using the colour green. Leptin has glucose-lowering effects independent of insulin. Insulin increases the production of leptin by adipose tissue (fat cells). Leptin feeds back to reduce insulin secretion.

BMI is a measure of body fat based on the weight and height of an adult. Adipose tissue produces leptin in quantities proportional to its mass and thereby regulates body weight. Leptin level also increases with BMI [142]. We have modelled the role of leptin and BMI in the IG regulatory system by augmenting the OMM as follows

- leptin-induced glucose absorption in peripheral tissues independent of insulin [128, 143],
- leptin-induced inhibition of insulin secretion [130],

- leptin production from adipose tissue modelled through BMI. This production can be dependent and independent of insulin [129, 130].

For accommodating leptin concentration, one additional differential equation (DE, Eqn 7.5) is added and additional terms ( $T_1 - T_6$ ) in the DEs for glucose and insulin concentrations (Eqns 7.1, 7.3) are added. Also, an expression (Eqn 7.4) for capturing the effect of BMI is introduced. The augmented model equations are given as follows

$$\frac{dG}{dt} = k_1 \cdot G_b - \left( k_1 + X + \underbrace{g_l \cdot L}_{T_1} \right) \cdot G + \frac{R_a(t)}{V}, \quad (7.1)$$

$$\frac{dX}{dt} = k_3 \cdot (I - I_b) - k_2 \cdot X, \quad (7.2)$$

$$\frac{dI}{dt} = \underbrace{\frac{\gamma}{(1 + k_6 \cdot L)}}_{T_2} \cdot (G - G_T)^+ - k_4 \cdot I, \quad (7.3)$$

$$k_{\text{BMI}} = \underbrace{\frac{1}{1 + e^{-\frac{\text{BMI}}{b_m}}}}_{T_3}, \quad (7.4)$$

$$\frac{dL}{dt} = k_{\text{BMI}} \cdot \left( \underbrace{k_b}_{T_4} + \underbrace{\frac{k_i \cdot I}{k_l + I}}_{T_5} \right) - \underbrace{k_5 \cdot L}_{T_6}, \quad (7.5)$$

where  $L$  is the leptin concentration in the blood,  $BMI$  is the body mass index of the subject,  $g_l$  denotes the ability of leptin to lower blood glucose levels,  $k_6$  denotes the inhibitory action of leptin on insulin,  $b_m$  is a scaling factor that regulates the effect of BMI,  $k_b$  is the BMI contribution factor to leptin production,  $k_i$  denotes leptin production based on insulin,  $k_l$  denotes the ability of insulin to stimulate leptin production and  $k_5$  is the leptin degradation rate.

The ability of leptin to independently lower blood glucose levels is modelled as  $T_1$  in Eqn 7.1. The effect is coupled with glucose as there exists a co-regulation between leptin and glucose via insulin. As leptin inhibits insulin production [133], this behaviour is modelled assuming reversible competitive inhibition [144] as  $T_2$  in Eqn 7.3. The role of BMI in leptin production is modelled empirically as a logistic function (Eqn 7.4) to incorporate the behaviour of leptin for the common BMI range of (18, 40); the parameters are determined through regression using Dataset 1. An additional equation for leptin concentration (Eqn 7.5) is added by introducing the effect of BMI and insulin in leptin production and a decay term to model the degradation of leptin over time.  $k_{\text{BMI}}$  is a function of BMI that regulates the effect of BMI in leptin concentration with a scaling parameter  $b_m$ . As leptin is mainly produced from body fat, this effect is introduced by the term  $k_{\text{BMI}} \cdot k_b$ ; as BMI increases, the effect of this term in leptin production also increases. The positive correlation of insulin in leptin concentration



Code	BMI	FINS	INS15	INS120	FGLU	GLU15	GLU120	FLEP	LEP15	LEP120
DM-06	28.5	0.68	2.22	9.48	289.57	313.60	485.08	36.03	27.54	57.01
DM-08	25.2	0.17	9.24	16.32	177.87	197.92	404.95	10.53	4.66	4.99
DM-28	26.4	1.93	2.98	58.35	145.17	182.76	420.99	30.24	11.75	13.33
DM-34	23.8	0.64	0.46	6.38	167.80	192.50	400.69	8.40	9.82	2.65

(a) Sample from Dataset-I

Code	FGLU	GLU15	GLU30	GLU90	GLU120	FINS	INS15	INS30	INS60	INS120
S1	202.5	237	279	384	385.02	90	200	220	300	300
S2	195.48	263	314	419	419.04	18	13	21	42	38
S3	120.96	154	232	364	383.94	7	21	30	69	70
S4	148.50	194	249	376	334.08	40	33	90	190	140
S5	256.50	299	320	455	466.02	6.5	13	24	31	38

(b) Sample from Dataset-II

Table 7.1: Sample rows from Dataset I and Dataset II of *leptin* model. Glucose units are in  $mg/dL$ , insulin units are in  $\mu U/mL$  and leptin units are in  $ng/mL$

is modelled following Michaelis–Menten (MM) kinetics (represented as  $T_5$ ). The term  $T_6$  models leptin degradation. The newly introduced model parameters and some of the existing parameters are estimated using available data. This model (Eqns 7.1-7.5) is referred to as *leptin* model in this chapter.

## 7.4 Parameter estimation

The parameters of the augmented model are estimated using the dataset described below. Seven new parameters -  $g_l, k_l, k_b, b_m, k_5, k_6$  and  $k_i$  are introduced in the augmented model. Model fitting is done to minimise the error between the observed data and simulated behaviour of the model.

**Dataset used for parameter estimation** Two sets of data are used in the study. Dataset-I consists of 40 rows of data related to subjects with T2DM, collected from a community-based primary health clinic located in the district of North 24 Parganas of West Bengal, India, who underwent OGTT. After pre-processing (removal of rows with missing values) the dataset consisted of data from 38 subjects. The subjects were asked to ingest 75g of glucose dissolved in 100mL of water after fasting for 8-12 hours. The sample rows from Dataset-I are shown in Table 7.1(a). The dataset consists of the following columns: BMI, measurements of insulin, glucose and leptin at various time points. FGLU indicates fasting glucose level.  $GLU\langle t \rangle$ ,  $INS\langle t \rangle$  and  $LEP\langle t \rangle$  indicate glucose, insulin and leptin levels after time  $\langle t \rangle$  of orally ingesting glucose beyond fasting glucose measurement (at  $t = 0$ ), respectively. For example, GLU30 indicates glucose level after 30 min and LEP45 indicates leptin level after 45 min. The measurements are taken at times: 0, 15, 30, 45, 60, 75, 90 and 120 minutes. All participants gave informed consent and the study was approved by Institutional Human Ethics Committee of CSIR-IICB.

Symbol	Description	Value	CV	Source	Unit
$g_l$	the ability of leptin to lower blood glucose levels	3.66	2.85	Estimated	$\frac{Lmin^{-1}}{\mu mol}$
$k_l$	ability of insulin to stimulate leptin production	4.99e-05	1.92	Estimated	$min^{-1}$
$k_5$	leptin degradation rate	0.03		Calculated	$min^{-1}$
$k_6$	the inhibitory action of leptin on insulin	749.74	2.80	Estimated	$L/\mu mol$
$k_b$	BMI contribution factor to leptin production	9.41e-05	0.89	Estimated	$min^{-1}$
$k_i$	leptin production based on insulin	5.15e-05	4.15	Calculated	$min^{-1}$
$b_m$	scaling factor that regulate effect of BMI	21		Calculated	constant
$k_1$	glucose effectiveness; the ability to promote glucose disposal and inhibit glucose production	0.027	0.73	Estimated	$min^{-1}$
$k_2$	the insulin-dependent increase in glucose uptake ability in tissue per unit of insulin concentration above $I_b$	0.110	2.89	Estimated	$min^{-1}$
$k_3$	scaling factor for the amplitude of insulin action	7.634	1.33	Estimated	$\frac{min^{-2}}{(\mu mol/L)^{-1}}$
$k_4$	decay rate for insulin in plasma	3.349	0.86	Estimated	$min^{-1}$
$\gamma$	the rate of the pancreatic $\beta$ -cells' release of insulin after the glucose injection and with glucose concentration above $D$	0.017e-6		Estimated from <i>base</i> model	$\mu mol/L$
$G_T$	the threshold value of glucose above which the pancreatic $\beta$ -cells release insulin	4974.59	0.55	Estimated	$\mu mol/L$
$G_b$	basal level of glucose	3000	0.26	Estimated from <i>base</i> model	$\mu mol/L$
$I_b$	basal level of insulin	2.221e-11		Estimated from <i>base</i> model	$\mu mol/L$

Table 7.2: Estimated parameters using the *leptin* model

Dataset-II consists of OGTT data of 129 individuals diagnosed with T2DM, collected from local pathology labs by a collaborator from a research institute. The subjects were asked to ingest 25g of glucose dissolved in 100 mL of water after fasting for 8-12 hours. Sample data from Dataset-II are given in Table 7.1(b). The column names are similar to Dataset-I with only changes in time points. The measurements in Dataset-II are taken at times: 0, 15, 30, 60, 90 and 120 minutes. These data were obtained from local pathology laboratories in anonymised form, ensuring the protection of patient confidentiality. The Dataset-I is used for the estimation of parameters in the leptin model and Dataset-II is used in the validation of

the model.

The condition which is characterised by abnormally high levels of insulin in the blood-stream (hyperinsulinemia) strongly influences leptin levels [129]. This is not accommodated in the present model as the available data used in the study only included subjects who were not affected by hyperinsulinemia.

**Estimation techniques** Parameter estimation is performed by solving an optimisation problem to minimise the squared error between outputs obtained from solving ODE and the observed data. The methods are implemented and coded in Python framework using the SciPy packages `odeint()` [119] and `lmfit()` [145]. Parameter estimation runs were performed for each individual subject and average parameter values were calculated from the estimated results. Different estimation techniques were tried out and Nelder-Mead optimisation [40] gave a better fit of observed data. The error function used in estimation is defined as follows:

$$E(\theta) = \sum_{i=1}^m (\hat{y}_i - y_i(\theta))^2 \quad (7.6)$$

where  $\hat{y}_i$  are observed data values and  $y_i$  are simulated values for a given parameter  $\theta$  at  $m$  time points.

When performing parameter estimation or fitting for the system of ODEs, it is common to impose bounds on the parameters to ensure that the estimated values are within a reasonable and physiologically meaningful range. Setting parameter bounds helps to constrain the optimisation process and prevents unrealistic or unbounded parameter estimates. This model considers parameter ranges reported in the relevant literature and physiological constraints while setting parameter bounds.

## 7.5 Revision of indices derived from OGTT

The concept of introducing a revised version of HOMA-B, HOMA2 and HOMA-IR, which incorporates BMI and fasting leptin, is a valuable approach to account for the influence of obesity on  $\beta$ -cell function and insulin resistance. The resulting indices can provide a more comprehensive assessment of insulin secretion and sensitivity. HOMA-B focuses only on fasting glucose and insulin values, while  $DI_o$  also considers post-prandial values. The revised indices for HOMA-B, HOMA2 and HOMA-IR are represented as functions of BMI and leptin,

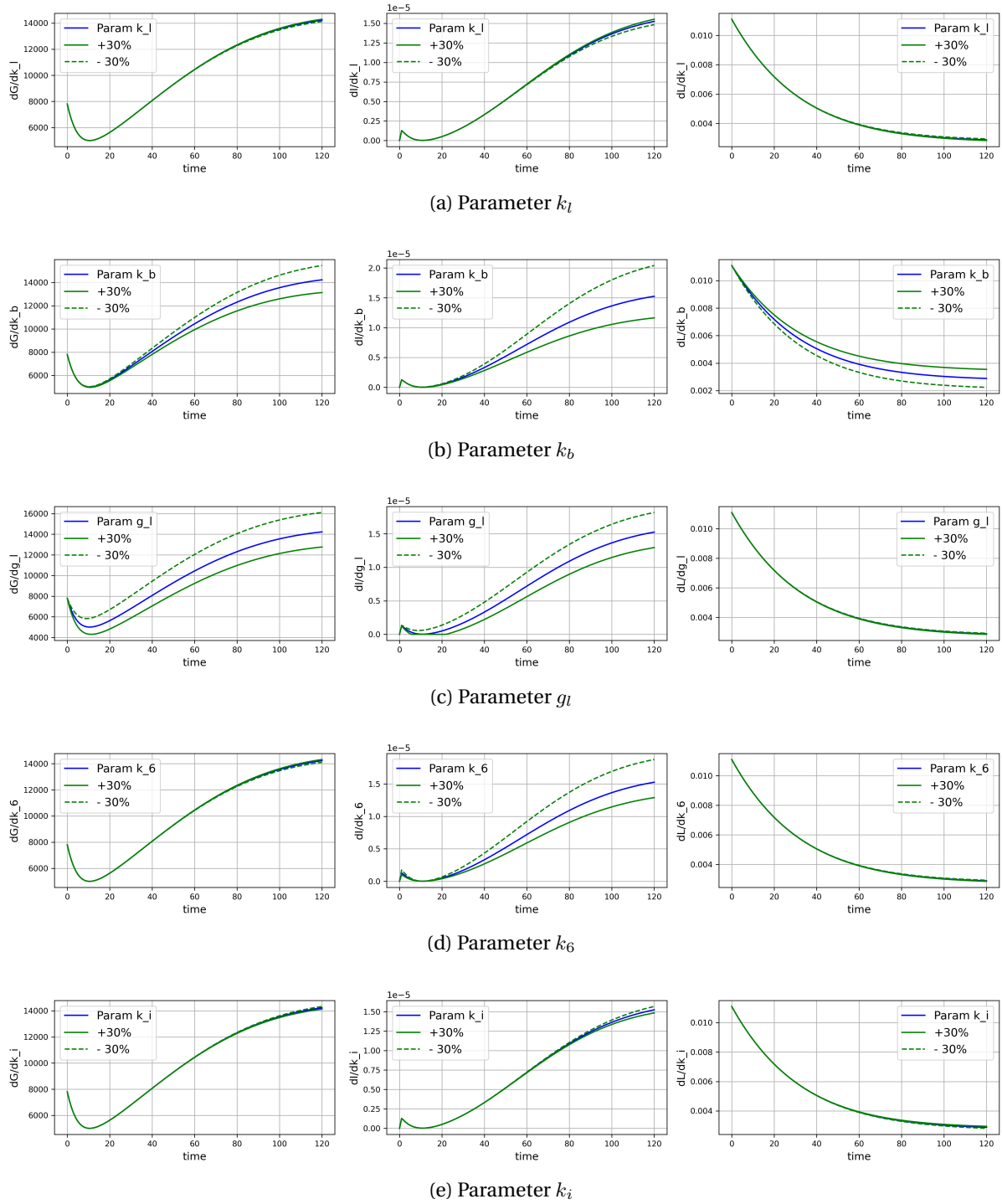


Figure 7.2: Sensitivity analysis of parameters with  $\pm 30\%$  variation wrt glucose, insulin and leptin

as shown in the equations

$$\text{HOMA-B (BMI)} = \text{HOMA-B} + 37 \cdot \text{BMI},$$

$$\text{HOMA2 (BMI)} = \text{HOMA2} + 24 \cdot \text{BMI},$$

$$\text{HOMA-IR (BMI)} = k_{\text{BMI}} \cdot \text{HOMA-IR},$$

$$\text{HOMA-IR (L}_0\text{)} = \log(\text{L}_0) \cdot \text{HOMA-IR},$$

$$\text{HOMA-IR (BMI, L}_0\text{)} = k_{\text{BMI}} \cdot \log(\text{L}_0) \cdot \text{HOMA-IR},$$

where  $L_0$  is fasting leptin in  $ng/mL$ . In order to establish a closer association with ISI, BMI is incorporated into HOMA-B and HOMA2 calculations by applying a scaling factor and the resulting values are analysed to determine their correlation with ISI. Through a systematic

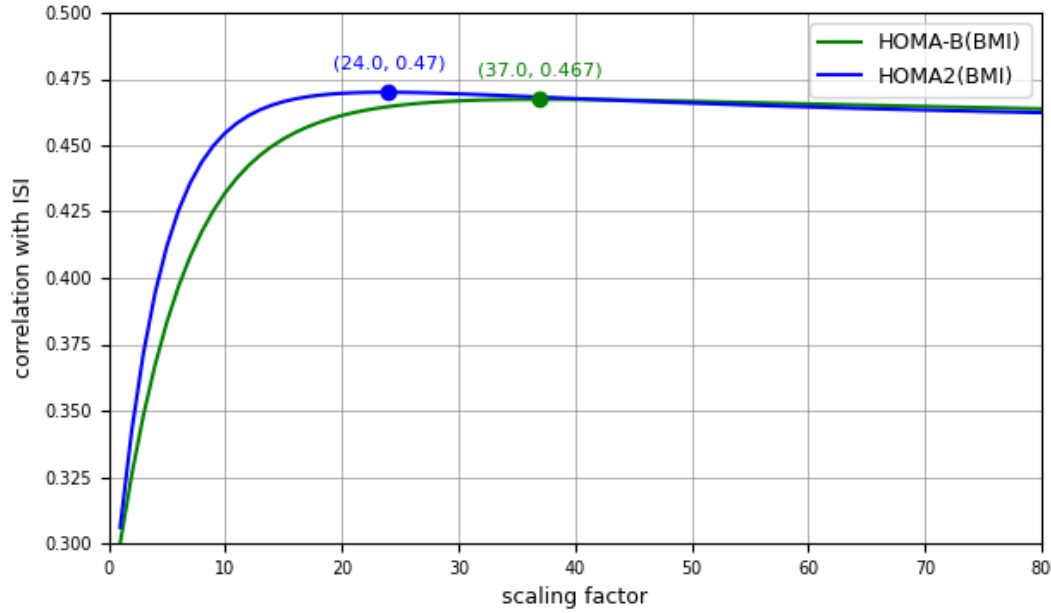


Figure 7.3: Variation of scaling factor used in revised HOMA-B and HOMA2 wrt correlation with ISI

evaluation of scaling factor values, the optimal scaling factor, which exhibits the highest correlation with the ISI, is identified (as shown in Figure 7.3). This iterative approach ensures that the scaling factor chosen is the minimum value that maximizes the required correlation. The HOMA-IR index is enhanced by incorporating  $k_{BMI}$  and  $L_0$  as separate variables, as well as by considering their combined effect. The correlation between the augmented HOMA-IR index and  $DI_0$  is also measured in each case to assess their relationship. By incorporating BMI and fasting leptin values into these indices, clinicians and researchers can gain insights into the combined effects of insulin secretion, insulin resistance and obesity on T2DM.

## 7.6 Results and validation

This section presents the parameter estimation results of the leptin model, along with analysis and validation of the model. The estimated parameters provide insights into the dynamics of glucose, insulin and leptin concentrations in IG pathway and validation with data ensures the reliability of *leptin* model.

**Results of parameter sensitivity analysis** The newly introduced parameters are sensitive to the observables, highlighting the positive aspects of the current model. Sensitivity analysis is

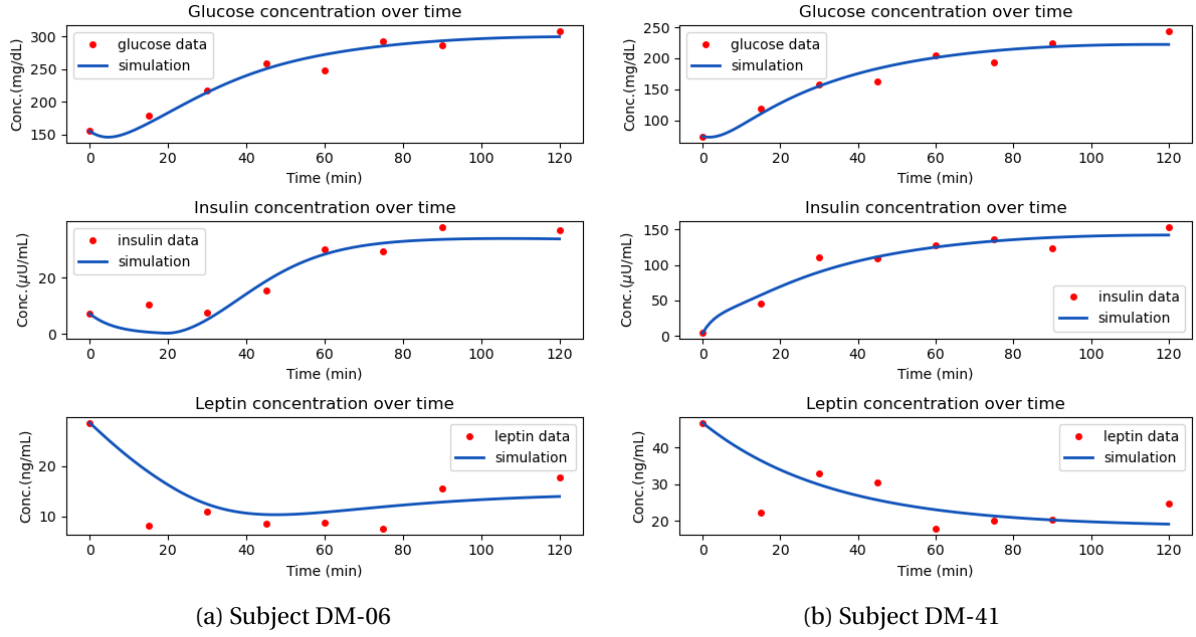


Figure 7.4: Simulation sample showing glucose, insulin and leptin concentrations for two subjects

done on five parameters  $k_l$ ,  $k_b$ ,  $g_l$ ,  $k_6$  and  $k_i$  in *leptin* model, considering a  $\pm 30\%$  variation with respect to glucose, insulin and leptin and the results are shown in Figure 7.2. The parameter  $g_l$  is highly sensitive to glucose and insulin whereas less sensitive to leptin which implies the ability of leptin to lower blood glucose levels impacts model behaviour significantly. This variation in  $g_l$  analysis helps in incorporating adequate ranges for  $g_l$ . Insulin concentration is affected by  $k_6$  and is moderately sensitive to the inhibitory action of leptin on insulin.  $k_b$  is highly sensitive to insulin, glucose and leptin, emphasising the role of BMI in the *leptin* model.  $k_l$  and  $k_i$  are related to the insulin action in the production of leptin and are less sensitive to insulin giving only a small boost in insulin concentration.

**Results of parameter estimation** The model is estimated for parameter values and the resulting simulation closely matches with the observed data. The parameter description, values, coefficient of variation (CV) and units are described in Table 7.2. Seven new parameters -  $g_l$ ,  $k_l$ ,  $k_b$ ,  $b_m$ ,  $k_5$ ,  $k_6$  and  $k_i$ , along with existing *base* model parameters are estimated using the Nelder-Mead optimisation method using `lmfit()` [145] in Python. The parameters of the  $R_a$  model for glucose absorption,  $k_{max}$ ,  $k_{min}$ ,  $k_{abs}$  and  $f$  are taken from literature [38]. Few of the parameters are estimated from the *base* model and few other parameters are calculated based on known relationships or processes. As BMI is usually in the range 18-32, the parameter  $b_m$  in  $k_{BMI}$  (Eqn 7.4) is fixed at 21 as the short range of the logistic function gives the effect of BMI adequately. This is the calculated average value of estimated  $b_m$  across all the subjects. The parameter  $k_5$ , which is the rate constant for leptin degradation, is found to have a half-life of

Indices	ISI	Indices	DI <sub>0</sub>
HOMA-B	0.299	HOMA-IR	0.707
HOMA2	0.306	HOMA-IR (L <sub>0</sub> )	0.718
HOMA-B (BMI)	0.467	HOMA-IR (BMI)	0.733
HOMA2 (BMI)	0.469	HOMA-IR (L <sub>0</sub> ,BMI)	0.735

(a)

(b)

Table 7.3: Correlation of indices: (a) Correlation of ISI to HOMA-B and HOMA2 and their revisions (b) Correlation of DI<sub>0</sub> to HOMA-IR and revised HOMA-IR

26 min [146] and the value is fixed at 0.03. The simulation results of two sample subjects are shown in Figure 7.4. The *leptin* model captures the observed time point measurements of leptin along with insulin and glucose. The model effectively captures the relationship between insulin and leptin, wherein an increase in insulin level leads to a decrease in leptin level. This agreement between the model predictions and observed behaviour highlights the ability of the model to represent the regulation between insulin and leptin adequately. CV is calculated to express the relative variability in the parameters. For instance, a CV value of  $k_b$  exhibits less variability and more precision, whereas  $k_l$  indicates a higher CV, which is attributed to the range of BMI in subjects (Table 7.2).

**Validation of the augmented model** To evaluate whether the developed *leptin* model is capable of capturing the dynamics of the insulin-glucose system, model validation is performed. The newly introduced parameter values are estimated from the augmented model and simulations are run for all the subjects. The BMI value is made a parameter for this simulation and the estimated average is 24.6. Other parameter values are taken from literature as given in Table 7.2. The simulations performed with the *leptin* model demonstrated a better fit to the observed data points when compared to the *base* model. Sample simulation results of the *base* model and *leptin* model using Dataset-II are shown in Figure 7.5. The simulated behaviour of leptin reveals a pattern that closely resembles the results previously observed for Dataset-I. To assess the statistical significance of the differences, a Chi-square goodness of fit test ( $\alpha = 0.05$ ) is performed on the simulated and observed data. The analysis revealed a statistically significant result of p-values with 74% good fits using the *leptin* model compared to 62% good fits using the *base* model. The *leptin* model is also validated using Dataset-I without considering leptin values and the fittings are found to be better. A sample simulation results of the *base* model and *leptin* model using Dataset-I are shown in Figure 7.6.

**Results of revised OGTT indices** The correlation values between indices are shown in Table 7.3. The inclusion of BMI in HOMA-B and HOMA2 exhibits a better correlation with ISI with coefficients of 0.467 and 0.469, respectively, suggesting a closer association with ISI. When

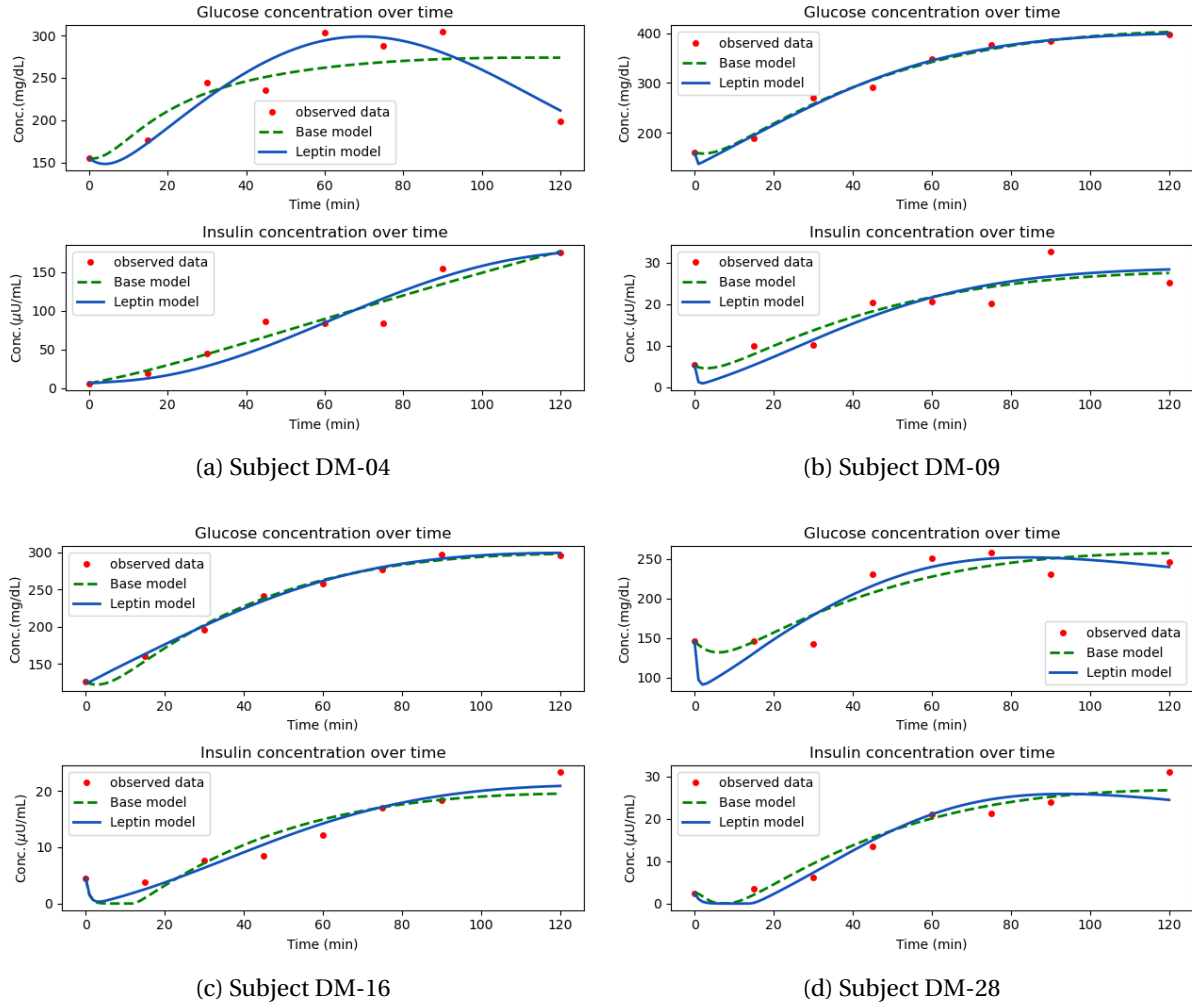


Figure 7.5: Simulation sample showing glucose and insulin (for 75g glucose) for *base* model and *leptin* model using Dataset I

examining the correlation between HOMA-IR and DI, the inclusion of BMI and leptin as separate variables, as well as their combined effect, resulted in slightly higher correlation values compared to the existing correlation between these factors. This suggests that BMI and leptin contribute to the relationship between HOMA-IR and DI and their combined impact further strengthens their association. These revised indices may be used clinically for better assessment of  $\beta$ -cell function and insulin resistance.

## 7.7 Discussion on augmented model

The objective of this work is to augment the existing mathematical model of IG regulatory system to accommodate the role of leptin and BMI. Parameters for the *leptin* model are estimated using the observed data of glucose, insulin and leptin for individuals with T2DM. The simulated results and the observed data have a statistically significant Chi-square fit of 74%. As leptin is involved in glucoregulatory actions, predicting its behaviour through the



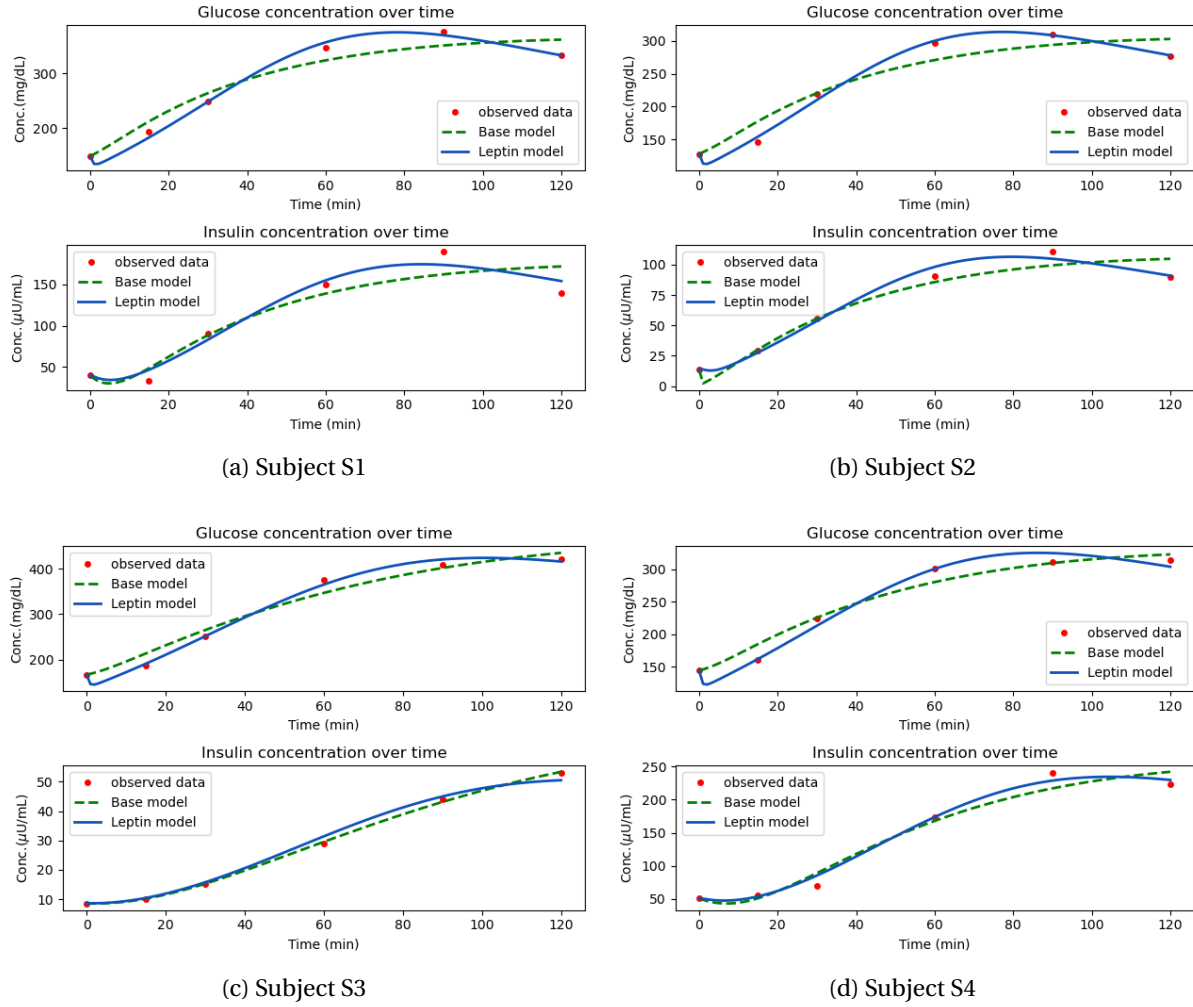


Figure 7.6: Simulation sample showing glucose and insulin (for 25g glucose) for *base* model and *leptin* model using Dataset II

*leptin* model can help develop targeted interventions and therapies for managing and treating T2DM, considering obesity-related issues. By influencing insulin secretion and sensitivity, leptin affects glucose metabolism and overall glycemic control in individuals with T2DM. Leptin is reported to have potent anti-diabetic actions independently of its effects on body weight and food intake [147]. These anti-diabetic effects of leptin can be explored and used in leptin therapy for the treatment of T2DM [147]. Further, long-term leptin-replacement therapy can significantly improve glycemic control and insulin sensitivity in patients with severe IR [148]. The new model can predict leptin behaviour in association with insulin and glucose concentrations using clinically significant parameters. A personalised treatment approach by integrating leptin behaviour may lead to more effective and targeted treatments for metabolic conditions.

The revised OGTT indices (HOMA-B (BMI), HOMA2 (BMI), HOMA-IR (BMI), HOMA-IR ( $L_0$ ), HOMA-IR (BMI,  $L_0$ )) proposed can aid in better insulin dosage estimation. The effect of

the estimated dosage may be better predicted using the *leptin* model. T2DM therapy largely depends on the degree of IR and insulin secretory dysfunction, the contribution of which is associated with obesity. Hence, the inclusion of obesity parameters in assessing IR and  $\beta$ -cell dysfunction is vital in the decision-making process of anti-diabetic therapy. The simple and usable indices proposed in this work will help decide a precision anti-diabetic therapy for non-obese T2DM patients and may open the door to other possibilities to build more accurate obesity-associated indices for personalised medicine. The claims in the work are based on the available data of 38 subjects with information on leptin concentration during OGTT. It would be desirable to perform the study on larger data to increase the confidence of the results. A potential future work may be to enhance the presented *leptin* model by incorporating other adipokines (such as adiponectin and adiponectin) that also play a role in the insulin-glucose regulatory system.

## 7.8 Conclusion

In this study, the OMM is augmented by incorporating the influence of BMI and leptin levels in the insulin-glucose regulatory system. The model is developed based on several theoretical studies on the association of leptin with insulin and glucose concentrations. The model parameters are estimated and analysed using the OGTT data collected in a clinical setting. The new model offers valuable insights into the workings of the insulin-glucose system, taking into account the obesity of the subject and uncovering related risk factors involved in T2DM. This seems to be the first attempt to model the role of obesity indicators, leptin and BMI in glucose regulation. The revised OGTT indices, with the inclusion of obesity parameters, hold the potential to aid in determining precise anti-diabetic therapies for T2DM. Simulation results from the augmented model with the identified parameters have a good match with the observed data of insulin, glucose and leptin. The results presented in this work are based on limited available data. It would be desirable to reinforce those by extending the study to a larger dataset.





## Conclusions

### Chapter outline

- Conclusions of present work
- Future work

This thesis presents computational techniques developed towards modelling, simulation and analysis primarily in the areas of *a)* physiologically based pharmacokinetic (PBPK) modelling and drug distribution in *in vitro* systems *b)* lumped parameter modelling of the cardiovascular system (CVS) *c)* clustering workflow to identify subtypes in type 2 diabetes mellitus (T2DM) *d)* insulin-glucose regulatory system *e)* parameter distinctions for obese and non-obese T2DM and *f)* the role of obesity indicators in the insulin-glucose regulatory system. The research presented in this thesis is the culmination of a comprehensive literature review and the development of various approaches on the aforementioned topics.

### 8.1 Conclusions of present work

The work on the five topics reported in this thesis, along with the methods, results and findings developed, are now summarised.

**Summary of mathematical modelling of drug distribution in *in vitro* systems** Drug development faces challenges in accurately predicting human responses, requiring effective preclinical testing strategies. The work on developing the digital liver-on-chip simulator (DigiLoCs) represents a significant stride towards a more accurate and efficient drug development process. DigiLoCs improve human clearance predictions through an integrated approach that translates *in vitro* findings to clinically relevant situations. We modelled drug depletion

kinetics using assay compound depletion data derived from liver cell-based organ-on-chip (OoC) to study drug depletion in the human liver by digital twinning.

Current methods oversimplify biological processes, lumping them into one clearance process and relying on minimal information. Our approach integrates three key components: biological, hardware, and physicochemical information. It distinguishes between active and passive processes by explicitly describing drug uptake, distribution and metabolism. Digital twinning links hardware architecture with advanced biological mapping, accurately capturing drug kinetics in intracellular space. Analysis of thirty-two compounds showed that our approach closely described on-chip kinetics, resolving systematic under-prediction issues compared to conventional methods. Notably, this improvement was achieved without additional biological experiments.

Using propranolol as a proof-of-concept, we demonstrated the superior predictive power of our approach in predicting human pharmacokinetics (PK). This workflow is easily adaptable to other compounds, making it a significant advancement in liver clearance prediction. By using ODE-constrained optimisation and incorporating detailed information on compound-specific characteristics and hardware-specific data, DigiLoCs enables precise predictions of clinical clearance. This work demonstrates the effectiveness of DigiLoCs in predicting *in vitro* liver clearance and successfully translating these findings to human PK. The DigiLoCs platform has far-reaching implications for pharmaceutical research, serving as a decision-support tool for estimating first-in-human doses, evaluating human PK and reducing reliance on animal experimentation. Moreover, the versatility of our approach enables its application across various physiological contexts, extending beyond liver metabolism to other organs, such as gut metabolism, blood-brain barrier models and placental barrier models. By integrating DigiLoCs into drug development workflows, the process can be streamlined to promote more sustainable and expedited clinical trials.

**Summary of lumped parameter modelling of the cardiovascular system (CVS)** In this work, the digital twinning approach is applied to model the human circulatory system using the hydraulic-electrical analogy. A comprehensive ODE model for the whole-body CVS is developed. The pressure and flow simulations of the comprehensive model help analyse the blood flow pattern in vessels at various places of interest. The modelling approach is initiated with the existing base model comprising heart chambers, systemic circulation and pulmonary circulation. The base model may be augmented with more compartments to attain a more refined model capable of detailed simulation of blood flow and pressures. In this work, augmentation with compartments for lungs, aorta, upper body, kidney, liver and lower body has been done to achieve the comprehensive model for a healthy individual. Most parameter values in the comprehensive model are obtained from the literature while others

are tuned manually to achieve physiologically reasonable behaviour of flows and pressures at various places of interest, as observed through model simulation. Notably, the addition of relevant compartments enables the simulation of pressure and flow in internal organs where observation is challenging.

Our comprehensive model simulations were validated with respect to the base model compartments and for additional compartments, with data available in the literature, ensuring consistency with established physiological values. We quantified the percentage of cardiac output distributed to each compartment and compared these values to known physiological ranges. To evaluate model convergence, we computed the RMSE across different numbers of cardiac cycles and showed that for a sufficiently large number of cycles (e.g., 1000+), both the flow difference and RMSE approached zero, indicating numerical stability and conservation. The results of the sensitivity analysis indicate that less sensitive parameters can be assigned nominal values. The more sensitive parameters can be determined optimally through parameter estimation. The broad and adaptable nature of our model allows for easy replication and modification, making it possible to simulate various disease states by selectively removing or adjusting specific compartments or variables. This flexibility is essential for capturing the interactions between multiple organs in disease states, such as those seen in sepsis, a multi organ dysfunction syndrome.

**Summary of clustering-based methods to identify subtypes of T2DM** This study investigates the heterogeneity of uncontrolled T2DM patients and identifies potential subtypes to aid in personalised treatment. We identified three distinct clusters by analysing electronic health records of 339 uncontrolled T2DM patients using Uniform Manifold Approximation and Projection (UMAP) and Density-Based Spatial Clustering of Applications with Noise (DBSCAN). These clusters differed significantly in disease duration, presence of co-morbidities, and  $\beta$ -cell function, without significant difference in insulin resistance. Notably, the clusters responded differently to treatment, with the second cluster showing the best glycemic control with dual anti-diabetic therapy, while the third cluster had the worst glycemic control despite receiving intensive therapy.

Our findings highlight the need for stratifying uncontrolled T2DM patients according to their patho-clinical characteristics to optimise treatment outcomes. Clusters 2 and 3 represent the obesity-associated and Asian-Indian diabetic variants, respectively. Many Cluster 1 patients may transition to Cluster 3 following substantial  $\beta$ -cell damage. Patients in Cluster 2 should be prioritised to manage associated diseases, while those in Cluster 1 and 3 require tighter glycemic control. Current diabetes treatment, which focuses on the classical obesity-associated T2DM, needs to be reevaluated in light of the Asian-Indian variant and its specific therapeutic needs. By recognising the heterogeneity within the population,

healthcare providers can move towards precision therapy, optimising treatment outcomes and improving patient care. This study lays the groundwork for further investigation into the clinical characteristics of uncontrolled T2DM subtypes. Additional research is needed to explore the causes of these differences and to develop personalised treatment strategies that account for the distinct pathophysiological profiles of each subtype. Ultimately, this precision medicine approach can potentially transform the management of T2DM.

### **Summary of parameter estimation of Oral Minimal Model in obese and non-obese T2DM**

This study successfully applied the Oral Minimal Model (OMM) to estimate parameters for obese and non-obese subjects with T2DM using Oral Glucose Tolerance Test (OGTT) data. The findings provide novel insights into the differences in insulin-glucose dynamics between obese and non-obese T2DM groups. The key findings of this study reveal that the obese diabetic group exhibited greater insulin resistance compared to the non-obese diabetic group. Statistical analysis of the estimated parameters revealed distinct distributions of model parameters between control and T2DM groups, as well as among obese and non-obese T2DM subjects. The implications of this study are multifaceted. Firstly, it demonstrates the potential of the OMM in understanding the pathophysiology of type 2 diabetes, particularly separating obese and non-obese T2DM. Secondly, the difference in identified parameters provides valuable information for personalised treatment strategies. Lastly, this preliminary application of the OMM with obesity as a distinguishing factor lays the groundwork for future research in diabetes management and treatment. Future research directions include further investigation to validate these findings in diverse populations. Additionally, integrating the OMM with other diagnostic tools may enhance T2DM diagnosis and treatment. This study contributes to the growing body of research on diabetes management, highlighting the importance of considering obesity in developing personalised treatment plans.

### **Summary of mathematical modelling of leptin in insulin-glucose regulatory system**

In this work, the OMM is augmented by incorporating the influence of body mass index (BMI) and leptin levels on the insulin-glucose regulatory system, providing a more comprehensive understanding of T2DM. Building on theoretical studies that established the association between leptin, insulin and glucose concentrations, our model offers insights into the interactions between these factors, particularly in obese individuals. The revised model, developed using ODEs, demonstrates a superior fit to clinical data on insulin, glucose and leptin levels (when available). Including obesity parameters in the OGTT indices holds promise for determining precise anti-diabetic therapies for T2DM. Notably, this study represents the first attempt to model the role of obesity indicators, leptin and BMI together, in glucose regulation by augmenting the OMM. The findings of this study underscore the importance of consid-



ering body weight indicators in developing effective treatments for T2DM. While the results are encouraging, they are based on limited available data. Future research should prioritise expanding the study to larger datasets to reinforce these findings. Nonetheless, this innovative approach lays the groundwork for targeted interventions and therapies, offering new avenues for improving the management and treatment of T2DM. The inclusion of obesity parameters in OGTT indices holds promise for determining precise anti-diabetic therapies. This study represents the first attempt to model the role of obesity indicators in insulin-glucose regulation, augmenting the ODEs of OMM.

## 8.2 Future work plan

The digital twin approach can be extended to other OoC devices. Measuring intracellular concentrations can help validate the three-compartment model. The physical-chemical properties of drugs can also be estimated with more observations from OoC experiments. The lumped parameter model of the cardiovascular system can be enhanced by integrating additional physiological subsystems such as CNS fluid flow, coronary circulation, and refined representations of the upper and lower body vasculature. Pharmacokinetics can be introduced to study drug response in the circulatory system. These extensions enable the development of a whole-body model of fluid transport that can simulate complex conditions like sepsis. Potential future applications of the comprehensive model include exploring the impact of comorbidities on cardiovascular function, optimising treatment protocols for specific patient populations, and investigating the mechanisms underlying cardiovascular disease. The developed leptin model can be validated using additional data, including observations of leptin levels. These models can also be extended by incorporating other adipokines, and their effect on the model can be studied and validated. The identified uncontrolled T2DM subtypes can be investigated further using larger cohorts to enable more informed decisions for personalised treatment strategies.



# Dissemination of the work

## Journals

- DigiLoCS: A leap forward in predictive organ-on-chip simulations. PloS One, 2025 Jan;20(1), e0314083. doi:10.1371/journal.pone.0314083. **Manoja Rajalakshmi Aravindakshan**, Chittaranjan Mandal, Alex Pothén and Christian Maass
- Parameter estimation for the oral minimal model and parameter distinctions between obese and non-obese type 2 diabetes. DNA and Cell Biology Reports 2024 5:1, 15-27. doi:10.1089/dcbr.2024.0018 **Manoja Rajalakshmi Aravindakshan**, Devleena Ghosh, Chittaranjan Mandal, K. V. Venkatesh, Jit Sarkar, Partha Chakrabarti and Sujay K. Maity
- Distinct patho-clinical clusters among uncontrolled type 2 diabetes patients: results from a prospective study in rural India. BMJ Open Diabetes Res Care. 2022 Feb;10(1), e002654. doi: 10.1136/bmjdr-2021-002654. PMID: 35144941. **Manoja Rajalakshmi Aravindakshan**, Sujay Krishna Maity, Avishek Paul, Partha Chakrabarti, Chittaranjan Mandal and Jit Sarkar

## Conferences/Workshops

- Oral presentation on "A whole-body lumped parameter model for cardiovascular simulation", 18<sup>th</sup> U.S. Association for Computational Mechanics 2025, July 20-24, Chicago, USA
- Talk on "Application of computational modeling towards digital twinning for human physiological systems" at OVDF Symposium, Dec 2024, Purdue University, USA
- Talk on "Mathematical modelling of drug distribution in liver-on-a-chip device". Data-driven mechanistic models of complex biomedical systems, School of Mathematics, University of Birmingham, UK. Dec 10 - 12, 2023.
- Oral presentation on "Parameter estimation for the oral minimal model and parameter distinctions between obese and non-obese type 2 diabetes", Bioinformatics and Computational Biology Conference, Dec 13-15 2022.

**Under revision**

- A closed-loop lumped parameter model for whole-body cardiovascular simulation.  
**Manoja Rajalakshmi Aravindakshan**, Brett A Meyers, Pavlos Vlachos, Chittaranjan Mandal and Alex Pothen

## About the author

Manoja Rajalakshmi Aravindakshan is enrolled as a PhD research scholar in the Department of Computer Science and Engineering, IIT Kharagpur, since January 2020. Before embarking on her doctoral journey, she worked as a Software Engineer at Global Embedon, Pune. She holds a B.Tech in Computer Science and Engineering from Government Engineering College, Palakkad. Her research interests include modelling drug distribution in organ-on-chip systems, lumped parameter modelling of the human cardiovascular system and insulin-glucose regulatory systems. Manoja has also contributed to research projects focused on adult metabolism models and signal interlocking of railway systems. Currently, she is a university graduate student in the School of Mechanical Engineering, Purdue University, IN, USA.

### Contact Information:

Manoja Rajalakshmi Aravindakshan  
Department of Computer Science and Engineering  
Indian Institute of Technology Kharagpur  
Kharagpur 721302, India  
Phone No: +91-8086285181  
Email: manojalax@gmail.com, manoja@iitkgp.ac.in

### Permanent Address:

Manoja Rajalakshmi Aravindakshan  
Chithira, Kottapuram PO  
Palakkad Dist  
Kerala 679513, India

### Research Interests:

Computational biology, Organ-on-chip, Cardiovascular system



# Bibliography

- [1] F Regazzoni and et al. Cardiac hemodynamics: normal values and clinical applications. *European Heart Journal*, 41(3):251–261, 2020. 45, 153
- [2] Deranged Physiology. Ventricular pressure-volume loops. <https://derangedphysiology.com/main/cicm-primary-exam/cardiovascular-system/Chapter-029/ventricular-pressure-volume-loops>, 2024. Accessed: 2025-06-01. 55, 56
- [3] Fatimah A. Alkhunaizi, Daniel Burkhoff, and Michael I. Brener. Right-sided mechanical circulatory support – a hemodynamic perspective. *Current Heart Failure Reports*, 19(5):265–275, 2022. © The Author(s) 2022 under exclusive licence to Springer Nature. 56
- [4] Patric Glynn, Sathya D Unudurthi, and Thomas J Hund. Mathematical modeling of physiological systems: an essential tool for discovery. *Life Sci.*, 111(1-2):1–5, August 2014. 1
- [5] Alexandre Vallée. Digital twin for healthcare systems. *Front. Digit. Health*, 5:1253050, September 2023. 1
- [6] Yubing Shi, Patricia Lawford, and Rodney Hose. Review of zero-d and 1-d models of blood flow in the cardiovascular system. *Biomed. Eng. Online*, 10(1):33, April 2011. 2, 5, 13, 41, 42
- [7] Anja Brunberg, Stefanie Heinke, Jan Spillner, Rüdiger Autschbach, Dirk Abel, and Steffen Leonhardt. Modeling and simulation of the cardiovascular system: a review of applications, methods, and potentials / modellierung und simulation des herz-kreislauf-systems: ein Überblick zu anwendungen, methoden und perspektiven. *Biomedizinische Technik/Biomedical Engineering*, 54(5):233–244, January 2009. 2, 42
- [8] James D Murray. *Mathematical Biology*. Interdisciplinary applied mathematics. Springer, New York, NY, 3 edition, January 2002. 3

- [9] Jingchen Feng, Stuart A Sevier, Bin Huang, Dongya Jia, and Herbert Levine. Modelling delayed processes in biological systems. *Physical Review E*, 94(3), September 2016. 3
- [10] Luca Docci, Nicolò Milani, Thomas Ramp, Andrea A Romeo, Patricio Godoy, Daniela Ortiz Franyuti, Stephan Krähenbühl, Michael Gertz, Aleksandra Galetin, Neil Parrott, and Stephen Fowler. Exploration and application of a liver-on-a-chip device in combination with modelling and simulation for quantitative drug metabolism studies. *Lab on a Chip*, 22(6):1187–1205, 2022. 5, 10, 20, 21, 22, 28, 29, 32, 137
- [11] N Tsamandouras, T Kostrzewski, C L Stokes, L G Griffith, D J Hughes, and M Cirit. Quantitative assessment of population variability in hepatic drug metabolism using a perfused three-dimensional human liver microphysiological system. *Journal of Pharmacology and Experimental Therapeutics*, 360(1):95–105, October 2016. 5, 10, 21, 22, 29, 32, 132
- [12] Ashutosh Dash, Karan Jain, Nirmalya Ghosh, and Amit Patra. Non-invasive detection of coronary artery disease from photoplethysmograph using lumped parameter modelling. *Biomedical Signal Processing and Control*, 77:103781, August 2022. 5, 42
- [13] Richard N Bergman. The minimal model of glucose regulation: a biography. In *Mathematical Modeling in Nutrition and the Health Sciences*, pages 1–19. Springer, 2003. 6, 14, 15, 74
- [14] Kajsa P Kanebratt, Annika Janefeldt, Liisa Vilén, Anna Vildhede, Kristin Samuelsson, Lucas Milton, Anders Björkbom, Marie Persson, Carina Leandersson, Tommy B Andersson, and Constanze Hilgendorf. Primary human hepatocyte spheroid model as a 3D *In Vitro* platform for metabolism studies. *Journal of Pharmaceutical Sciences*, 110(1):422–431, January 2021. 10, 21, 22, 140, 141, 142
- [15] Ia Hultman, Charlotta Vedin, Anna Abrahamsson, Susanne Winiwarter, and Malin Darnell. Use of H $\mu$ REL human coculture system for prediction of intrinsic clearance and metabolite formation for slowly metabolized compounds. *Mol. Pharm.*, 13(8):2796–2807, August 2016. 10
- [16] B Bonn, P Svanberg, A Janefeldt, I Hultman, and K Grime. Determination of human hepatocyte intrinsic clearance for slowly metabolised compounds: Comparison of a primary hepatocyte/stromal cell co-culture with plated primary hepatocytes and heparg. *Drug Metabolism and Disposition*, 44(4):527–533, February 2016. 10, 21, 22, 140, 141, 142, 143
- [17] Kenza Abouir, Caroline F Samer, Yvonne Gloor, Jules A Desmeules, and Youssef Daali. Reviewing data integrated for PBPK model development to predict metabolic drug-drug



- interactions: Shifting perspectives and emerging trends. *Front. Pharmacol.*, 12:708299, October 2021. [11](#)
- [18] André Dallmann, Juri Solodenko, Ibrahim Ince, and Thomas Eissing. Applied concepts in pbpk modeling: How to extend an open systems pharmacology model to the special population of pregnant women. *CPT: Pharmacometrics & Systems Pharmacology*, 7(7):419–431, April 2018. [11](#)
- [19] Sheila Annie Peters. *Physiologically based pharmacokinetic (PBPK) modeling and simulations: principles, methods, and applications in the pharmaceutical industry*. Wiley, Hoboken, NJ, 2011. [11](#), [133](#)
- [20] Attila A Seyhan. Lost in translation: the valley of death across preclinical and clinical divide – identification of problems and overcoming obstacles. *Transl. Med. Commun.*, 4(1), December 2019. [12](#)
- [21] Olivier J Wouters, Martin McKee, and Jeroen Luyten. Estimated research and development investment needed to bring a new medicine to market, 2009-2018. *JAMA*, 323(9):844–853, March 2020. [12](#)
- [22] Yuri Vassilevski, Maxim Olshanskii, Sergey Simakov, Andrey Kolobov, and Alexander Danilov. *Personalized computational hemodynamics: Models, methods, and applications for Vascular Surgery and Antitumor Therapy*. Academic Press, San Diego, CA, April 2020. [12](#)
- [23] Nico Westerhof, Jan-Willem Lankhaar, and Berend E. Westerhof. The arterial windkessel. *Medical & Biological Engineering & Computing*, 47(2):131–141, June 2008. [13](#), [42](#)
- [24] Louis Garber, Seyedvahid Khodaei, and Zahra Keshavarz-Motamed. The critical role of lumped parameter models in patient-specific cardiovascular simulations. *Archives of Computational Methods in Engineering*, 29(5):2977–3000, December 2021. [13](#), [42](#)
- [25] Xinzhou Xie, Minwen Zheng, Didi Wen, Yabing Li, and Songyun Xie. A new cfd based non-invasive method for functional diagnosis of coronary stenosis. *BioMedical Engineering OnLine*, 17(1), March 2018. [13](#)
- [26] S J Sherwin, L Formaggia, J Peiró, and V Franke. Computational modelling of 1D blood flow with variable mechanical properties and its application to the simulation of wave propagation in the human arterial system. *International Journal for Numerical Methods in Fluids*, 43(6–7):673–700, October 2003. [13](#)

- [27] Jing Wan, Brooke Steele, Sean A Spicer, Sven Strohband, Gonzalo R Feijó, Thomas J R Hughes, and Charles A Taylor. A one-dimensional finite element method for simulation-based medical planning for cardiovascular disease. *Computer Methods in Biomechanics and Biomedical Engineering*, 5(3):195–206, January 2002. 13
- [28] Kristen L Knutson, Karine Spiegel, Plamen Penev, and Eve Van Cauter. The metabolic consequences of sleep deprivation. *Sleep Medicine Reviews*, 11(3):163–178, June 2007. 14
- [29] Tobias A Weber and Andreas S Reichert. Impaired quality control of mitochondria: Aging from a new perspective. *Experimental Gerontology*, 45(7–8):503–511, August 2010. 14
- [30] W Bolie. Coefficients of normal blood glucose regulation. *J Appl Physiol.*, 16(5):783–8, 1961. 14
- [31] J Sturis, K S Polonsky, E Mosekilde, and E Van Cauter. Possible mechanisms underlying slow oscillations of human insulin secretion. *T. U. of Denmark*, page 160, 1991. 14
- [32] J Sturis, K S Polonsky, E Mosekilde, and E Van Cauter. Computer model for mechanisms underlying ultradian oscillations of insulin and glucose. *American Journal of Physiology - Endocrinology and Metabolism*, 260(5 23-5), 1991. 14, 15
- [33] Roman Hovorka, Valentina Canonico, Ludovic J Chassin, Ulrich Haueter, Massimo Massi-Benedetti, Marco Orsini Federici, Thomas R Pieber, Helga C Schaller, Lukas Schaupp, Thomas Vering, and Malgorzata E Wilinska. Nonlinear model predictive control of glucose concentration in subjects with type 1 diabetes. *Physiological Measurement*, 25(4):905–920, 2004. 14, 15
- [34] C D Man, A Caumo, and C Cobelli. The oral glucose minimal model: Estimation of insulin sensitivity from a meal test. *IEEE Transactions on Biomedical Engineering*, 49(5):419–429, 2002. 14, 15, 86
- [35] Richard N Bergman. Origins and history of the minimal model of glucose regulation. *Frontiers in Endocrinology*, 11, February 2021. 15
- [36] E D Lehmann and T Deutsch. A physiological model of glucose-insulin interaction in type 1 diabetes mellitus. *J. Biomed. Eng.*, 14(3):235–242, May 1992. 15
- [37] J D Elashoff, T J Reedy, and J H Meyer. Analysis of gastric emptying data. *Gastroenterology*, 83(6):1306–1312, December 1982. 15

- [38] C D Man, M Camilleri, and C Cobelli. A system model of oral glucose absorption: Validation on gold standard data. *IEEE Transactions on Biomedical Engineering*, 53(12):2472–2478, December 2006. 15, 75, 86, 96
- [39] Jérémie Guedj, Rodolphe Thiébaud, and Daniel Commenges. Maximum likelihood estimation in dynamical models of hiv. *Biometrics*, 63(4):1198–1206, 2007. 17, 88
- [40] Saša Singer and John Nelder. Nelder-mead algorithm. *Scholarpedia*, 4(7):2928, 2009. 17, 77, 88, 93
- [41] J A Nelder and R Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, January 1965. 17
- [42] Mami T Wentworth, Ralph C Smith, and H T Banks. Parameter selection and verification techniques based on global sensitivity analysis illustrated for an HIV model. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):266–297, January 2016. 18
- [43] Yunfei Chu and Juergen Hahn. Parameter set selection for estimation of nonlinear dynamic systems. *AIChE Journal*, 53:2858–2870, 2007. 18
- [44] Nora Franzen, Wim H van Harten, Valesca P Retèl, Peter Loskill, Janny van den Eijnden-van Raaij, and Maarten IJzerman. Impact of organ-on-a-chip technology on pharmaceutical R&D costs. *Drug Discovery Today*, 24(9):1720–1724, September 2019. 19, 20
- [45] Tinneke Denayer, Thomas Stöhr, and Maarten Van Roy. Animal models in translational medicine: Validation and prediction. *European Journal of Molecular & Clinical Medicine*, 2(1):5, August 2014. 19, 20
- [46] Kirstin Thelen, Katrin Coboeken, Stefan Willmann, Rolf Burghaus, Jennifer B Dressman, and Jörg Lippert. Evolution of a detailed physiological model to simulate the gastrointestinal transit and absorption process in humans, part 1: Oral solutions. *Journal of Pharmaceutical Sciences*, 100(12):5324–5345, December 2011. 19, 20, 26
- [47] 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 modelling in drug discovery and development: A pharmaceutical industry perspective. *Clinical Pharmacology & Therapeutics*, 97(3):247–262, January 2015. 19, 20, 21
- [48] R S Obach. Prediction of human clearance of twenty-nine drugs from hepatic microsomal intrinsic clearance data: An examination of *in vitro* half-life approach and nonspecific binding to microsomes. *Drug Metabolism and Disposition*, 27(11):1350–1359, November 1999. 19, 20

- [49] Hayley S Brown, Michael Griffin, and J Brian Houston. Evaluation of cryopreserved human hepatocytes as an alternative *in vitro* system to microsomes for the prediction of metabolic clearance. *Drug Metabolism and Disposition*, 35(2):293–301, November 2006. 19, 20
- [50] Micaela B Reddy, Kelly D Mccarley, and Annette L Bunge. Physiologically relevant one-compartment pharmacokinetic models for skin. 2. comparison of models when combined with a systemic pharmacokinetic model. *Journal of Pharmaceutical Sciences*, 87(4):482–490, April 1998. 19, 20
- [51] Yukiko Murata, Sibylle Neuhoﬀ, Amin Rostami-Hodjegan, Hiroyuki Takita, Zubida M Al-Majdoub, and Kayode Ogungbenro. *In vitro* to *In vivo* extrapolation linked to physiologically based pharmacokinetic models for assessing the brain drug disposition. *The AAPS Journal*, 24(1), January 2022. 19, 21
- [52] David Hallifax, Joanne A Foster, and J Brian Houston. Prediction of human metabolic clearance from *in vitro* systems: Retrospective analysis and prospective view. *Pharmaceutical Research*, 27(10):2150–2161, July 2010. 20, 21
- [53] Tanvi Shroﬀ, Kehinde Aina, Christian Maass, Madalena Cipriano, Joeri Lambrecht, Frank Tacke, Alexander Mosig, and Peter Loskill. Studying metabolism with multi-organ chips: new tools for disease modelling, pharmacokinetics and pharmacodynamics. *Open Biology*, 12(3), March 2022. 20
- [54] European Medicines Agency. Guideline on the investigation of drug interactions. [https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-investigation-drug-interactions-revision-1\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-investigation-drug-interactions-revision-1_en.pdf), 2012. Accessed: 22-3-2024. 20, 30
- [55] Christian Maass, Cynthia L Stokes, Linda G Griﬃth, and Murat Cirit. Multi-functional scaling methodology for translational pharmacokinetic and pharmacodynamic applications using integrated microphysiological systems (mps). *Integrative Biology*, 9(4):290–302, 2017. 20
- [56] Stephen Fowler, Wen Li Kelly Chen, David B Duignan, Anshul Gupta, Niresh Hariparsad, Jane R Kenny, W George Lai, Jennifer Liras, Jonathan A Phillips, and Jinping Gan. Microphysiological systems for adme-related applications: current status and recommendations for system development and characterisation. *Lab on a Chip*, 20(3):446–467, 2020. 20
- [57] Christian Maass, Nathan B Sorensen, Jonathan Himmelfarb, Edward J Kelly, Cynthia L Stokes, and Murat Cirit. Translational assessment of drug-induced proximal tubule

- injury using a kidney microphysiological system. *CPT: Pharmacometrics & Systems Pharmacology*, 8(5):316–325, April 2019. 20
- [58] Anna Herland, Ben M Maoz, Debarun Das, Mahadevabharath R Somayaji, Rachelle Prantil-Baun, Richard Novak, Michael Crouce, Tessa Huffstater, Sauveur S F Jeanty, Miles Ingram, Angeliki Chalkiadaki, David Benson Chou, Susan Marquez, Aaron Delahanty, Sasan Jalili-Firoozinezhad, Yuka Milton, Alexandra Sontheimer-Phelps, Ben Swenor, Oren Levy, Kevin K Parker, Andrzej Przekwas, and Donald E Ingber. Quantitative prediction of human pharmacokinetic responses to drugs via fluidically coupled vascularised organ chips. *Nature Biomedical Engineering*, 4(4):421–436, January 2020. 20
- [59] Shiny Amala Priya Rajan, Jason Sherfey, Shivam Ohri, Lauren Nichols, J Tyler Smith, Paarth Parekh, Eugene P Kadar, Frances Clark, Billy T George, Lauren Gregory, David Tess, James R Gosset, Jennifer Liras, Emily Geishecker, R Scott Obach, and Murat Cirit. A novel milli-fluidic liver tissue chip with continuous recirculation for predictive pharmacokinetics applications. *The AAPS Journal*, 25(6), October 2023. 21, 22, 28, 139, 144
- [60] Collin D Edington, Wen Li Kelly Chen, Emily Geishecker, Timothy Kassis, Luis R Soenksen, Brij M Bhushan, Duncan Freake, Jared Kirschner, Christian Maass, Nikolaos Tsamandouras, Jorge Valdez, Christi D Cook, Tom Parent, Stephen Snyder, Jiajie Yu, Emily Suter, Michael Shockley, Jason Velazquez, Jeremy J Velazquez, Linda Stockdale, Julia P Papps, Iris Lee, Nicholas Vann, Mario Gamboa, Matthew E LaBarge, Zhe Zhong, Xin Wang, Laurie A Boyer, Douglas A Lauffenburger, Rebecca L Carrier, Catherine Communal, Steven R Tannenbaum, Cynthia L Stokes, David J Hughes, Gaurav Rohatgi, David L Trumper, Murat Cirit, and Linda G Griffith. Interconnected microphysiological systems for quantitative biology and pharmacology studies. *Scientific Reports*, 8(1), March 2018. 25
- [61] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. 25
- [62] Patrick Poulin and Sami Haddad. Hepatocyte composition-based model as a mechanistic tool for predicting the cell suspension: Aqueous phase partition coefficient of drugs in *in vitro* metabolic studies. *Journal of Pharmaceutical Sciences*, 102(8):2806–2818, August 2013. 27
- [63] Rupert P Austin, Patrick Barton, Sarfraz Mohamed, and Robert J Riley. The binding of drugs to hepatocytes and its relationship to physiochemical properties. *Drug Metabolism and Disposition*, 33(3):419–425, December 2004. 27

- [64] Armanda V Rodrigues, Graça Alexandre-Pires, Ana Valério-Bolas, David Santos-Mateus, Mariana Rafael-Fernandes, Maria A Pereira, Dário Ligeiro, Telmo Nunes, Raquel Alves-Azevedo, Marcos Santos, Isabel Pereira da Fonseca, and Gabriela Santos-Gomes. 3d-hepatocyte culture applied to parasitology: Immune activation of canine hepatic spheroids exposed to leishmania infantum. *Biomedicines*, 8(12):628, December 2020. 28
- [65] Sangeeta N Bhatia and Donald E Ingber. Microfluidic organs-on-chips. *Nat. Biotechnol.*, 32(8):760–772, August 2014. 27
- [66] André Mateus, Pär Matsson, and Per Artursson. Rapid measurement of intracellular unbound drug concentrations. *Molecular Pharmaceutics*, 10(6):2467–2478, May 2013. 29
- [67] Hadley Wickham. *Ggplot2. Use R!* Springer International Publishing, Cham, Switzerland, 2 edition, June 2016. 30
- [68] 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, October 2016. 30
- [69] Lars Borgström, Carl-Gustav Johansson, Hans Larsson, and Rune Lenander. Pharmacokinetics of propranolol. *Journal of Pharmacokinetics and Biopharmaceutics*, 9(4):419–429, August 1981. 37, 135
- [70] Isidor Kokalari, Theodor Karaja, and Maria Guerrisi. Review on lumped parameter method for modeling the blood flow in systemic arteries. *Journal of Biomedical Science and Engineering*, 06(01):92–99, 2013. 42
- [71] Luca Formaggia, Alfio Quarteroni, and Alessandro Veneziani. *Cardiovascular Mathematics: Modeling and simulation of the circulatory system*, volume 1. Springer Science & Business Media, 2010. 42
- [72] Wilco Kroon, Wouter Huberts, Marielle Bosboom, and Frans van de Vosse. A numerical method of reduced complexity for simulating vascular hemodynamics using coupled 0d lumped and 1d wave propagation models. *Computational and Mathematical Methods in Medicine*, 2012:1–10, 2012. 42
- [73] S. V. Frolov, S. V. Sindeev, V. A. Lischouk, D. S. H. Gazizova, D. Liepsch, and A. Balasso. A lumped parameter model of cardiovascular system with pulsating heart for diagnostic studies. *Journal of Mechanics in Medicine and Biology*, 17(03):1750056, May 2017. 42

- [74] Eun-Ok Jung and Wan-Ho Lee. Lumped parameter models of cardiovascular circulation in normal and arrhythmia cases. *J. Korean Math. Soc.*, 43(4):885–897, July 2006. 42
- [75] Francesco Regazzoni, Matteo Salvador, Pasquale Claudio Africa, Marco Fedele, Luca Dede', and Alfio Quarteroni. A cardiac electromechanics model coupled with a lumped parameters model for closed-loop blood circulation. part i: model derivation. *arXiv preprint arXiv:2011.15040*, 2020. 42, 43, 44, 52
- [76] R. Mukkamala, A. T. Reisner, H. M. Hojman, R. G. Mark, and R. J. Cohen. Continuous cardiac output monitoring by peripheral blood pressure waveform analysis. *IEEE Transactions on Biomedical Engineering*, 53(3):459–467, March 2006. 42
- [77] Abdul Jalil Rufaihah, Ching Kit Chen, Choon Hwai Yap, and Citra N. Z. Mattar. Mending a broken heart: In vitro, in vivo and in silico models of congenital heart disease. *Disease Models & Mechanisms*, 14(3), March 2021. 42
- [78] Elham Kayvanpour, Tommaso Mansi, Farbod Sedaghat-Hamedani, Ali Amr, Dominik Neumann, Bogdan Georgescu, Philipp Seegerer, Ali Kamen, Jan Haas, Karen S. Frese, Maria Irawati, Emil Wirsz, Vanessa King, Sebastian Buss, Derliz Mereles, Edgar Zitron, Andreas Keller, Hugo A. Katus, Dorin Comaniciu, and Benjamin Meder. Towards personalized cardiology: Multi-scale modeling of the failing heart. *PLOS ONE*, 10(7):e0134869, July 2015. 42
- [79] Pablo Blanco and Raúl Feijóo. A 3D-1D-0D computational model for the entire cardiovascular system. September 2009. 44
- [80] Patrick Kidger. *On Neural Differential Equations*. PhD thesis, University of Oxford, 2021. 47
- [81] International Diabetes Federation. *IDF Diabetes Atlas - 2019*. 2019. 61
- [82] Ana Beatriz Valverde Mendes, João Antônio Saraiva Fittipaldi, Raimundo Celestino Silva Neves, Antônio Roberto Chacra, and Edson Duarte Moreira. Prevalence and correlates of inadequate glycemic control: Results from a nationwide survey in 6,671 adults with diabetes in Brazil. *Acta Diabetologica*, 47(2):137–145, 2010. 61
- [83] Afsana Afroz, Liaquat Ali, Md Nazmul Karim, Mohammed J Alramadan, Khurshid Alam, Dianna J Magliano, and Baki Billah. glycemic Control for People with Type 2 Diabetes Mellitus in Bangladesh - An urgent need for optimization of management plan. *Scientific Reports*, 9(1):1–10, dec 2019. 61



- [84] Surendra S. Borgharkar and Soma S. Das. Real-world evidence of glycemic control among patients with type 2 diabetes mellitus in India: The TIGHT study. *BMJ Open Diabetes Research and Care*, 7(1):1–8, 2019. 61
- [85] Steven E Kahn, Mark E Cooper, and Stefano Del Prato. Pathophysiology and treatment of type 2 diabetes: Perspectives on the past, present and future. *Lancet*, 383(9922):1068–1083, 2014. 61
- [86] American Diabetes Association. Standards of medical care in diabetes 2016. *Diabetes Care*, 39:Supplement 1, 2016. 61, 64
- [87] American Diabetes Association. Standards of Medical Care in Diabetes — 2018. 41(January), 2018. 61
- [88] John M Lachin, Saul Genuth, David M Nathan, Bernard Zinman, and Brandy N Rutledge. Effect of glycemic exposure on the risk of microvascular complications in the diabetes control and complications trial-revisited. *Diabetes*, 57(4):995–1001, apr 2008. 61
- [89] Emma Ahlqvist, Petter Storm, Annemari Käräjämäki, Mats Martinell, Mozhgan Dorkhan, Annelie Carlsson, Petter Vikman, Rashmi B. Prasad, Dina Mansour Aly, Peter Almgren, Ylva Wessman, Nael Shaat, Peter Spégl, Hindrik Mulder, Eero Lindholm, Olle Melander, Ola Hansson, Ulf Malmqvist, Åke Lernmark, Kaj Lahti, Tom Forsén, Tiinamaija Tuomi, Anders H. Rosengren, and Leif Groop. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *The Lancet Diabetes and Endocrinology*, 6(5):361–369, may 2018. 62, 63
- [90] Ranjit Mohan Anjana, Viswanathan Baskar, Anand Thakarakkattil Narayanan Nair, Saravanan Jebarani, Moneeza Kalhan Siddiqui, Rajendra Pradeepa, Ranjit Unnikrishnan, Colin Palmer, Ewan Pearson, and Viswanathan Mohan. Novel subgroups of type 2 diabetes and their association with microvascular outcomes in an Asian Indian population: A data-driven cluster analysis: The INSPIRED study. *BMJ Open Diabetes Research and Care*, 8(1), aug 2020. 62, 63
- [91] Wayne W Daniel. Prevalence and co-prevalence of comorbidities among patients with type 2 diabetes mellitus. *Biostatistics: A foundation for analysis in the health sciences*, (New York: John Wiley & Sons 6th ed.), 1995. 62
- [92] Shreemathi S Mayya, Ashma D Monteiro, and Sachit Ganapathy. Types of biological variables. *Journal of Thoracic Disease*, 9(6):1730–1733, 2017. 62



- [93] Steven E Kahn, Rebecca L Hull, and Kristina M Utzschneider. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature*, 444(7121):840–846, dec 2006. 63
- [94] Robert J Smith, David M Nathan, Silva A Arslanian, Leif Groop, Robert A Rizza, and Jerome I Rotter. Individualizing therapies in type 2 diabetes mellitus based on patient characteristics: What we know and what we need to know. In *Journal of Clinical Endocrinology and Metabolism*, volume 95, pages 1566–1574. Endocrine Society, 2010. 63
- [95] Mark I McCarthy. Painting a new picture of personalised medicine for diabetes, 2017. 63
- [96] Hugo Fitipaldi, Mark I McCarthy, Jose C Florez, and Paul W Franks. A global overview of precision medicine in type 2 diabetes. *Diabetes*, 67(10):1911–1922, oct 2018. 63
- [97] Anna L Gloyn and Daniel J Drucker. Precision medicine in the management of type 2 diabetes. *The Lancet Diabetes and Endocrinology*, 6(11):891–900, nov 2018. 63
- [98] Jit Sarkar, Sujay Krishna Maity, Abhishek Sen, Titli Nargis, Dipika Ray, and Partha Chakrabarti. Impaired compensatory hyperinsulinemia among nonobese type 2 diabetes patients: a cross-sectional study. *Therapeutic Advances in Endocrinology and Metabolism*, 10:204201881988902, jan 2019. 63, 71, 86
- [99] Leland McInnes, John Healy, Nathaniel Saul, and Lukas Großberger. UMAP: Uniform manifold approximation and projection. *J. Open Source Softw.*, 3(29):861, September 2018. 65
- [100] Yiqing Song, Joann E Manson, Lesley Tinker, Barbara V Howard, Lewis H Kuller, Lauren Nathan, Nader Rifai, and Simin Liu. Insulin sensitivity and insulin secretion determined by homeostasis model assessment and risk of diabetes in a multiethnic cohort of women: the women’s health initiative observational study. *Diabetes Care*, 30(7):1747–1752, July 2007. 65, 66
- [101] Ralph A DeFronzo and Masafumi Matsuda. Reduced time points to calculate the composite index. *Diabetes Care*, 33(7):e93, July 2010. 65
- [102] Kristina M Utzschneider, Ronald L Prigeon, Mirjam V Faulenbach, Jenny Tong, Darcy B Carr, Edward J Boyko, Donna L Leonetti, Marguerite J McNeely, Wilfred Y Fujimoto, and Steven E Kahn. Oral disposition index predicts the development of future diabetes above and beyond fasting and 2-h glucose levels. *Diabetes Care*, 32(2):335–341, February 2009. 66

- [103] Michael Stumvoll, Barry J Goldstein, and Timon W van Haeften. Type 2 diabetes: principles of pathogenesis and therapy. *Lancet*, 365(9467):1333–1346, 2005. 70, 71
- [104] Bernardo L Wajchenberg. Beta-cell failure in diabetes and preservation by clinical treatment. *Endocr. Rev.*, 28(2):187–218, April 2007. 70, 71
- [105] Parvez Hossain, Bisher Kavar, and Meguid El Nahas. Obesity and diabetes in the developing world—a growing challenge. *N. Engl. J. Med.*, 356(3):213–215, January 2007. 70
- [106] Vincenzo De Tata. Age-related impairment of pancreatic beta-cell function: pathophysiological and cellular mechanisms. *Front. Endocrinol. (Lausanne)*, 5:138, September 2014. 70
- [107] Ele Ferrannini, Stefania Camastra, Amalia Gastaldelli, Anna Maria Sironi, Andrea Natali, Elza Muscelli, Geltrude Mingrone, and Andrea Mari. Beta-cell function in obesity: Effects of weight loss. *Diabetes*, 53 Suppl 3:S26–33, December 2004. 71
- [108] Hiromi Iwahashi, Yuki Yoshi Okauchi, Miwa Ryo, Midori Noguchi, Sachiko Morita, Ken Kishida, Kohei Okita, Tetsuya Ohira, Tohru Funahashi, Tadashi Nakamura, Akihisa Imagawa, and Iichiro Shimomura. Insulin-secretion capacity in normal glucose tolerance, impaired glucose tolerance, and diabetes in obese and non-obese Japanese patients. *J. Diabetes Investig.*, 3(3):271–275, June 2012. 71
- [109] Ranjit Mohan Anjana, Mohan Deepa, and Rajendra Pradeepa et. al. Prevalence of diabetes and prediabetes in 15 states of india: results from the ICMR–INDIAB population-based cross-sectional study. *The Lancet Diabetes & Endocrinology*, 5(8):585–596, aug 2017. 71
- [110] Unjali P Gujral, Mary Beth Weber, Lisa R Staimez, and K M Venkat Narayan. Diabetes among non-overweight individuals: An emerging public health challenge. *Curr. Diab. Rep.*, 18(8):60, July 2018. 71
- [111] U.S. Department of Health and Human Services. Diabetes & prediabetes tests. <https://www.niddk.nih.gov/health-information/professionals/clinical-tools-patient-management/diabetes/diabetes-prediabetes?dkrd=hisce0124>, 2020. 74, 87
- [112] Nadia Burkart and Marco F Huber. A survey on the explainability of supervised machine learning. *J. Artif. Intell. Res.*, 70:245–317, January 2021. 74

- [113] Md Rezaul Karim, Oya Beyan, Achille Zappa, Ivan G Costa, Dietrich Rebholz-Schuhmann, Michael Cochez, and Stefan Decker. Deep learning-based clustering approaches for bioinformatics. *Briefings in Bioinformatics*, 22(1):393–415, feb 2020. 74
- [114] Wei Tse Li, Jiayan Ma, Neil Shende, Grant Castaneda, Jaideep Chakladar, Joseph C Tsai, Lauren Apostol, Christine O Honda, Jingyue Xu, Lindsay M Wong, Tianyi Zhang, Abby Lee, Aditi Gnanasekar, Thomas K. Honda, Selena Z Kuo, Michael Andrew Yu, Eric Y Chang, Mahadevan Raj Rajasekaran, and Weg M Ongkeko. Using machine learning of clinical data to diagnose COVID-19: a systematic review and meta-analysis. *BMC Medical Informatics and Decision Making*, 20(1), sep 2020. 74
- [115] M Ambigavathi and D Sridharan. *Analysis of Clustering Algorithms in Machine Learning for Healthcare Data*. Springer Singapore, 2020. 74
- [116] Manoja Rajalakshmi Aravindakshan, Sujay Krishna Maity, Avishek Paul, Partha Chakrabarti, Chittaranjan Mandal, and Jit Sarkar. Distinct pathoclinical clusters among patients with uncontrolled type 2 diabetes: results from a prospective study in rural india. *BMJ Open Diabetes Research & Care*, 10(1):e002654, February 2022. 74, 86
- [117] Anoop Misra. Ethnic-specific criteria for the classification of body mass index: a perspective for asian indians and american diabetes association position statement. *Diabetes technology & therapeutics*, 17(9):667–671, 2015. 74
- [118] Roland Brun, Peter Reichert, and Hans R Künsch. Practical identifiability analysis of large environmental simulation models. *Water Resources Research*, 37(4):1015–1030, 2001. 77
- [119] Pauli Virtanen, Gommers et. al., and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. 77, 93
- [120] Raj K Goyal, Vivian Cristofaro, and Maryrose P Sullivan. Rapid gastric emptying in diabetes mellitus: Pathophysiology and clinical importance. *Journal of Diabetes and its Complications*, 33(11):107414, November 2019. 82
- [121] Lyudmila V Gromova, Serguei O Fetissov, and Andrey A Gruzdkov. Mechanisms of glucose absorption in the small intestine in health and metabolic diseases and their role in appetite regulation. *Nutrients*, 13(7):2474, July 2021. 82
- [122] Kyle L Hoehn, Adam B Salmon, Cordula Hohnen-Behrens, Nigel Turner, Andrew J Hoy, Ghassan J Maghzal, Roland Stocker, Holly Van Remmen, Edward W Kraegen, Greg J Cooney, Arlan R Richardson, and David E James. Insulin resistance is a cellular

- antioxidant defense mechanism. *Proceedings of the National Academy of Sciences*, 106(42):17787–17792, October 2009. [83](#)
- [123] Xinlei Wang, Xiaoqin Zhao, Ranran Zhou, Yunjuan Gu, Xiaohui Zhu, Zhuqi Tang, Xinlu Yuan, Wei Chen, Rongping Zhang, Chen Qian, and Shiwei Cui. Delay in glucose peak time during the oral glucose tolerance test as an indicator of insulin resistance and insulin secretion in type 2 diabetes patients. *Journal of Diabetes Investigation*, 9(6):1288–1295, April 2018. [83](#)
- [124] Richard A Wright, Sam Krinsky, Carla Fleeman, Juan Trujillo, and Edward Teague. Gastric emptying and obesity. *Gastroenterology*, 84(4):747–751, April 1983. [83](#)
- [125] Reiner Jumpertz, Marie S Thearle, Joy C Bunt, and Jonathan Krakoff. Assessment of non-insulin-mediated glucose uptake: association with body fat and glycemic status. *Metabolism*, 59(10):1396–1401, 2010. [83](#)
- [126] Unai Galicia-Garcia et al. Pathophysiology of type 2 diabetes mellitus. *Int. J. Mol. Sci.*, 21(17):6275, August 2020. [86](#)
- [127] Ann Smith Barnes. The epidemic of obesity and diabetes: trends and treatments. *Tex. Heart Inst. J.*, 38(2):142–144, 2011. [86](#)
- [128] Thomas H Meek and Gregory J Morton. Leptin, diabetes, and the brain. *Indian J. Endocrinol. Metab.*, 16(Suppl 3):S534–42, December 2012. [86](#), [89](#)
- [129] Maria do Carmo Martins, Luís Lima Faleiro, and Aidil Fonseca. Relationship between leptin and body mass and metabolic syndrome in an adult population. *Rev. Port. Cardiol. (Engl. Ed.)*, 31(11):711–719, November 2012. [86](#), [90](#), [93](#)
- [130] Gilberto Paz-Filho, Claudio Mastronardi, Ma-Li Wong, and Julio Licinio. Leptin therapy, insulin sensitivity, and glucose homeostasis. *Indian J Endocrinol Metab.*, page S549–S555, 2012. [86](#), [87](#), [88](#), [89](#), [90](#)
- [131] Jonathan P German et al. Leptin deficiency causes insulin resistance induced by uncontrolled diabetes. *Diabetes*, 59(7):1626–1634, July 2010. [86](#), [87](#)
- [132] Amita Yadav, Pramila Jyoti, S K Jain, and Jayashree Bhattacharjee. Correlation of adiponectin and leptin with insulin resistance: a pilot study in healthy north indian population. *Indian J. Clin. Biochem.*, 26(2):193–196, April 2011. [87](#)
- [133] P G Cammisotto and M Bendayan. Leptin secretion by white adipose tissue and gastric mucosa. *Histol. Histopathol.*, 22(2):199–210, February 2007. [87](#), [90](#)

- [134] Klaus Schittkowski. Parameter estimation in dynamic systems. In *Progress in Optimization*, pages 183–204. Springer, 2000. 87
- [135] E García-Portugués. *Notes for Predictive Modeling*. 2022. Version 5.9.10. ISBN 978-84-09-29679-8. 88
- [136] Olcay Akman and Elsa Schaefer. An evolutionary computing approach for parameter estimation investigation of a model for cholera. *J. Biol. Dyn.*, 9(1):147–158, 2015. 88
- [137] Radcliffe Department of Medicine. Homa calculator. <https://www.rdm.ox.ac.uk/about/our-clinical-facilities-and-mrc-units/DTU/software/homa>, 2023. Accessed: 22-3-2022. 88
- [138] Kristina M Utzschneider et al. Oral disposition index predicts the development of future diabetes above and beyond fasting and 2-h glucose levels. *Diabetes Care*, 32(2):335–341, February 2009. 88
- [139] Anna M D'souza, Ursula H Neumann, Maria M Glavas, and Timothy J Kieffer. The glucoregulatory actions of leptin. *Molecular Metabolism*, 6(9):1052–1065, September 2017. 88
- [140] Angelo Cignarelli, Valentina Annamaria Genchi, Sebastio Perrini, Annalisa Natalicchio, Luigi Laviola, and Francesco Giorgino. Insulin and insulin receptors in adipose tissue development. *Int. J. Mol. Sci.*, 20(3):759, February 2019. 89
- [141] R M Hoffman, R C Boston, D Stefanovski, D S Kronfeld, and P A Harris. Obesity and diet affect glucose dynamics and insulin sensitivity in thoroughbred geldings. *J. Anim. Sci.*, 81(9):2333–2342, September 2003. 89
- [142] Rubina Faisal Paul, Mukhtiar Hassan, Hassan Shehzad Nazar, Saima Gillani, Naeema Afzal, and Iftikhar Qayyum. Effect of body mass index on serum leptin levels. *J. Ayub Med. Coll. Abbottabad*, 23(3):40–43, July 2011. 89
- [143] Jonathan P German et al. Leptin activates a novel CNS mechanism for insulin-independent normalization of severe diabetic hyperglycemia. *Endocrinology*, 152(2):394–404, 2011. 89
- [144] John Dawson. *BIOC\*2580: Introduction to Biochemistry*. Pressbooks, 2021. Biochemistry. 90
- [145] Matt Newville et al. lmfit/lmfit-py: 1.2.1. May 2023. 93, 96
- [146] Volker Morath et al. PASylation of murine leptin leads to extended plasma half-life and enhanced in vivo efficacy. *Mol. Pharm.*, 12(5):1431–1442, May 2015. 97

- [147] Roberto Coppari and Christian Bjørbæk. Leptin revisited: its mechanism of action and potential for treating diabetes. *Nat. Rev. Drug Discov.*, 11(9):692–708, September 2012. 99
- [148] Gilberto Paz-Filho et al. Changes in insulin sensitivity during leptin replacement therapy in leptin-deficient patients. *Am. J. Physiol. Endocrinol. Metab.*, 295(6):E1401–8, December 2008. 99
- [149] European Medicines Agency. Ich s7a safety pharmacology studies for human pharmaceuticals. <https://www.ema.europa.eu/en/ich-s7a-safety-pharmacology-studies-human-pharmaceuticals>, 2018. Accessed: 2024-03-24. 133
- [150] European Medicines Agency. Guideline on the investigation of bioequivalence. [https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-investigation-bioequivalence-rev1\\_en.pdf](https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-investigation-bioequivalence-rev1_en.pdf), 2010. Accessed: 2024-03-24. 133
- [151] U.S. Food and Drug Administration. Physiologically based pharmacokinetic analyses — format and content guidance for industry. <https://www.fda.gov/media/101469>, 2019. Accessed: 2024-03-24. 133
- [152] Kirstin Thelen, Katrin Coboeken, Stefan Willmann, Jennifer B Dressman, and Jörg Lippert. Evolution of a detailed physiological model to simulate the gastrointestinal transit and absorption process in humans, Part II: Extension to describe performance of solid dosage forms. *Journal of Pharmaceutical Sciences*, 101(3):1267–1280, March 2012. 133
- [153] Paulo Paixão, Luís F Gouveia, and José A G Morais. Prediction of the *in vitro* permeability determined in caco-2 cells by using artificial neural networks. *European Journal of Pharmaceutical Sciences*, 41(1):107–117, September 2010. 135
- [154] Yi Zheng, Leslie Z Benet, Hideaki Okochi, and Xijing Chen. ph dependent but not p-gp dependent bidirectional transport study of s-propranolol: The importance of passive diffusion. *Pharmaceutical Research*, February 2015. 135
- [155] Radiopaedia contributors. Atrial volume, 2025. Accessed: 2025-06-01. 153
- [156] Luigi P. Badano, Theodore J. Koliass, Denisa Muraru, and et al. Left atrial volumes and function by three-dimensional echocardiography: Reference values, accuracy, reproducibility, and comparison with two-dimensional echocardiographic measurements. *Circulation: Cardiovascular Imaging*, 9(7):e004229, 2016. 153

- [157] N O Fowler. Cardiac function: the right atrium and ventricle. *Circulation*, 62(2):I37–I43, 1980. 153
- [158] Harvey Feigenbaum, William F. Armstrong, and Thomas Ryan. Feigenbaum's echocardiography. *Lippincott Williams & Wilkins*, 2019. 153
- [159] N O Fowler. The normal right atrial pressure pulse in man. *Circulation*, 8:1–8, 1953. 153
- [160] J Widrich and et al. Pulmonary vein anatomy. *StatPearls*, 2024. 153
- [161] E Erdal and et al. Pulmonary artery pressure: reference values and clinical significance. *Turkish Journal of Medical Sciences*, 50(4):1012–1020, 2020. 153
- [162] John E Hall and Michael E Hall. *Guyton and Hall Textbook of Medical Physiology*. Guyton Physiology. Elsevier - Health Sciences Division, Philadelphia, PA, 14 edition, June 2020. 153, 154
- [163] David B Mount. Cerebral perfusion pressure. *UpToDate*, 2024. 153
- [164] Ethan Kung, Giancarlo Pennati, Francesco Migliavacca, Tain-Yen Hsia, Richard Figliola, Alison Marsden, and Alessandro Giardini. A simulation protocol for exercise physiology in fontan patients using a closed loop lumped-parameter model. *Journal of Biomechanical Engineering*, 136(8), June 2014. 153
- [165] David A. Rubenstein, Wei Yin, and Mary D. Frame. *Flow through the kidney*, pages 473–513. Elsevier, 2022. 153
- [166] Atul Malhotra and David P. White. Obstructive sleep apnoea. *The Lancet*, 360(9328):237–245, 2002. 153
- [167] Naim Alkhouri, Charles Winans, and Vera F. Hupertz. *Portal Hypertension*, pages 829–839.e4. Elsevier, 2011. 153
- [168] T Markou and et al. Normal values of central venous pressure in adults. *Journal of Cardiothoracic and Vascular Anesthesia*, 18(5):575–579, 2004. 153
- [169] John E Hall and Michael E Hall. *Guyton and Hall Textbook of Medical Physiology*. Elsevier, Philadelphia, PA, 15 edition, 2021. 154
- [170] Matthew N Levy, Bruce M Koeppen, and Bruce A Stanton. *Berne and Levy Physiology*. Elsevier, Philadelphia, PA, 8 edition, 2022. 154
- [171] Laleh Zarrinkoob, Kim Ambarki, Anders Wåhlin, Richard Birgander, Anders Eklund, and Jan Malm. Blood flow distribution in cerebral arteries. *J Cereb Blood Flow Metab*, 35(4):648–654, 2015. 154

- 
- [172] G. Ahlborg and M. Jensen-Urstad. Arm blood flow at rest and during arm exercise. *Journal of Applied Physiology*, 70(2):928–933, February 1991. 154
- [173] M.M. Madani and E. Golts. *Cardiovascular Anatomy*. Elsevier, 2014. 154
- [174] Christy K. Holland, Janis M. Brown, Leslie M. Scoutt, and Kenneth J.W. Taylor. Lower extremity volumetric arterial blood flow in normal subjects. *Ultrasound in Medicine & Biology*, 24(8):1079–1086, October 1998. 154



## Supplementary information for Chapter 3

The primary objective of this study is to develop an innovative digital twin framework that integrates microphysiological systems (MPS) and organ-on-chip (OoC) data into advanced computational models of biology to enhance the accuracy of predicting clinical clearances. The resulting tool, DigiLoCs, is a digital liver-on-chip simulator capable of effectively capturing the complexities of on-chip biology. It incorporates information on various biological processes such as clearance, permeability, and partitioning, along with hardware-specific details from the studied *in vitro* systems and compound-specific data.

DigiLoCs enable the differentiation between active biological processes like metabolism and passive ones such as permeability and partitioning, which contrasts with existing approaches where passive processes are often lumped together into a single clearance process. The tool's performance was evaluated using drug depletion kinetics data of 32 compounds from literature covering commercially available liver-on-chips and 3D spheroids. Results indicate that DigiLoCs significantly outperform current prediction approaches. Furthermore, a proof-of-concept study involving propranolol demonstrated the tool's potential in predicting human pharmacokinetics (PK) more accurately compared to state-of-the-art and literature-based approaches.

### A.1 Simulation of drug depletion data

The digital twin-based model simulation of on-chip kinetics of all the drugs involved in the study are shown in Figures A.1, A.2, A.3 and A.4. The drug-specific information is shown in Table A.1 and the analysis of estimated parameters is shown in Table A.2. The correlation plots

for different *in vitro* systems along with average fold error (AFE) are shown in Figures A.5, A.6 and A.7.

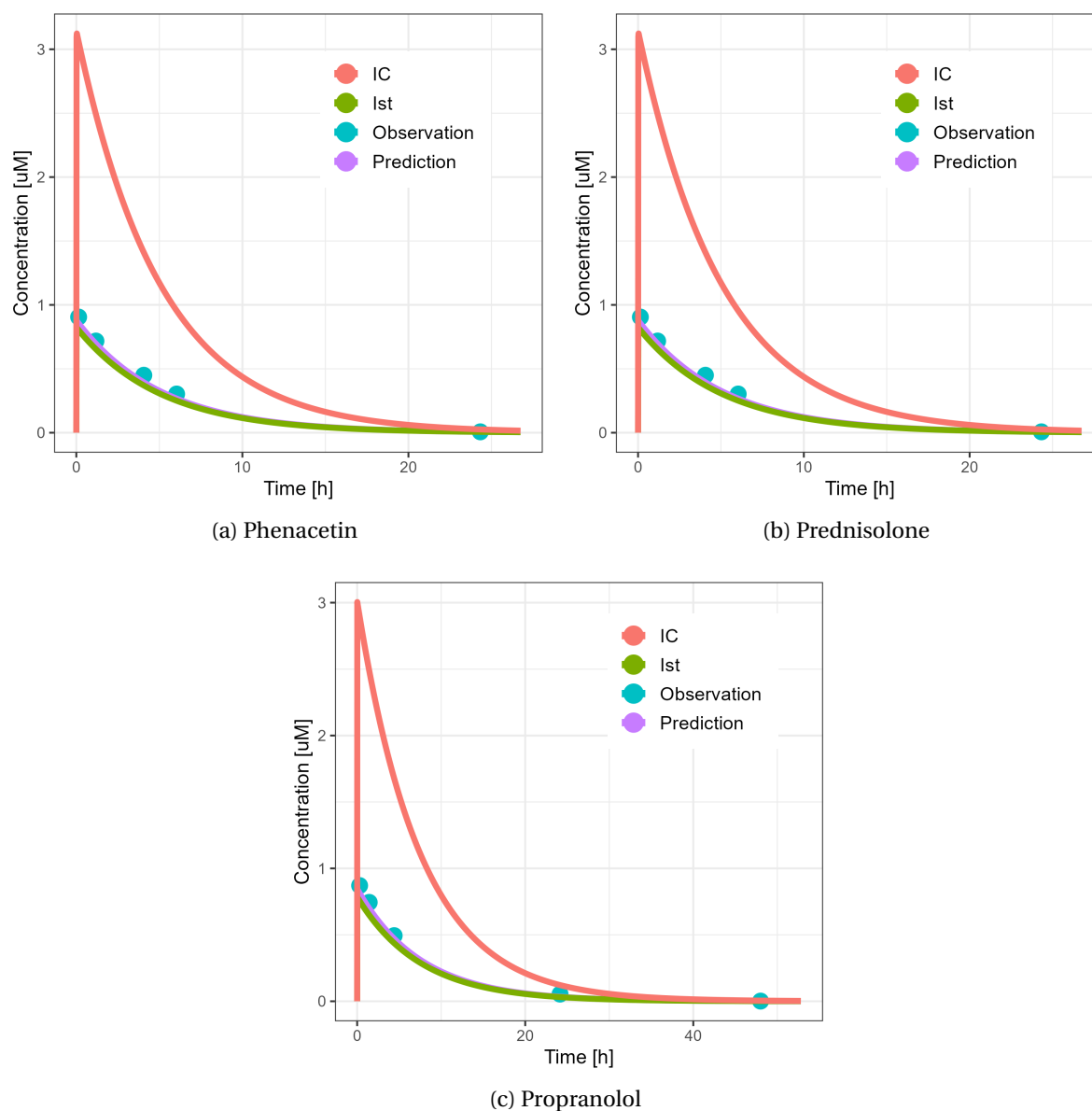


Figure A.1: Digital twin-based model simulation of on-chip kinetics after fitting parameters for three compounds from Tsamandourous et al. [11]; IC = intracellular, Ist = interstitium.

## A.2 Human PBPK modelling of propranolol

The open systems pharmacology suite (OSPS) (PK-Sim and MoBi) provides functionality for the full range of PBPK applications from physicochemistry-based prediction models for all relevant preclinical and clinical species to elaborated simulations for drug-drug interactions and special populations. The suite comprises functionalities for parameter identification, model qualification, and automated reporting as well as interfaces to R (The R Project for Sta-

Names	Lipophilicity	Molecular weight [g/mol]	pKa	Species	Unbound fraction plasma	Unbound fraction blood	Blood-to-plasma ratio (Rbp)	Unbound fraction media
Diclofenac	4.36	318.14	4.00	Acid	0.01	0.018	0.55	0.045
Lorazepam	3.10	321.16	10.61	Neutral	0.11	0.11	1.00	0.80
Midazolam	4.32	325.77	-	Neutral	0.03	0.06	0.55	0.10
Naloxone	1.30	327.38	10.07	Neutral	0.56	0.46	1.22	0.87
Oxazepam	2.45	286.72	10.61	Neutral	0.05	0.04	1.10	0.60
Posaconazole	4.57	700.79	-	Neutral	0.02	0.02	1.00	0.37
Quinidine	3.17	324.42	13.89	Base	0.13	0.15	0.87	0.63
Tolbutamide	1.78	270.35	4.33	Acid	0.02	0.03	0.75	0.45
Zidovudine	-0.20	267.25	9.96	Neutral	0.74	0.75	0.99	0.98
Propranolol	2.58	259.35	14.9	Base	-	0.14	-	0.95
Prednisolone	1.56	360.45	12.59	Neutral	-	0.10	-	0.87
Phenacetin	2.04	179.22	-	Neutral	-	0.6	-	1.00
Imipramine	2.30	280.40	9.20	Base	0.03	0.13	0.23	0.59
Clozapine	3.627	326.80	3.70	Base	-	0.07	-	0.35
Paracetamol	0.29	151.16	9.50	Neutral	-	0.04	-	0.93
Irbesartan	1.30	428.50	7.40	Zwitter	-	0.72	-	0.52
Ketoprofen	-0.15	254.80	4.45	Acid	-	0.01	-	0.87
Disopyramide	3.10	339.50	10.42	Base	-	0.16	-	0.83
(S)-Warfarin	2.74	308.30	5.00	Neutral	-	0.01	-	0.95
Carvedilol	3.10	406.50	8.10	Base	-	0.03	-	0.43
Imipramine	2.30	280.40	9.20	Base	0.03	0.13	0.23	0.92
Disopyramide	3.10	339.50	10.42	Base	-	0.16	-	0.98
(S)-Warfarin	2.74	308.30	5.00	Neutral	-	0.01	-	0.95
Carvedilol	3.10	406.50	8.10	Base	-	0.03	-	0.43
Midazolam	4.32	325.77	-	Neutral	0.03	0.10	0.55	0.49
Dextromethorphan	3.50	271.00	9.85	Base	0.46	0.38	1.20	0.79
Diclofenac	4.36	318.14	4.00	Acid	0.01	0.01	0.55	0.04
Propranolol	2.58	259.35	14.9	Base	-	0.14	-	0.98
Raloxifene	4.60	473.00	9.00	Base	0.02	0.02	0.78	0.09
Tolbutamide	1.78	270.35	4.33	Acid	0.02	0.05	0.75	0.60
Verapamil	4.00	454.00	9.68	Base	0.19	0.28	0.67	0.35
Zidovudine	-0.20	267.25	9.96	Neutral	0.74	0.75	0.99	0.88

Table A.1: Drug-specific information of 32 compounds involved in the study

tistical Computing, [www.R-project.org](http://www.R-project.org)) and PK-Sim release 11.3 (<http://www.open-systems-pharmacology.org/>). The methodological framework and the systematic presentation of findings adhere to the criteria stipulated by the EMA, FDA and OECD in their guidelines for PBPK M&S reporting [149, 150, 151].

In a comprehensive PBPK model (illustrated in Figure A.8), there is a detailed depiction of key organs involved in the drug's absorption, distribution, metabolism and excretion [152]. These organs typically include the heart, lungs, brain, stomach, spleen, pancreas, intestines, liver, kidneys, gonads, thymus, adipose tissue, muscles, bones and skin. The model considers each organ's physiological/pharmacological role, volume [19], composition, surface area and protein expression levels. These tissues are interconnected through arterial and venous blood compartments, each characterised by unique attributes such as blood flow rate, lymph flow rate, volume, vascular, interstitial and intracellular fractions, tissue-partition coefficients and permeability.

Thus, PBPK models provide an intricate structural representation of physiological mech-

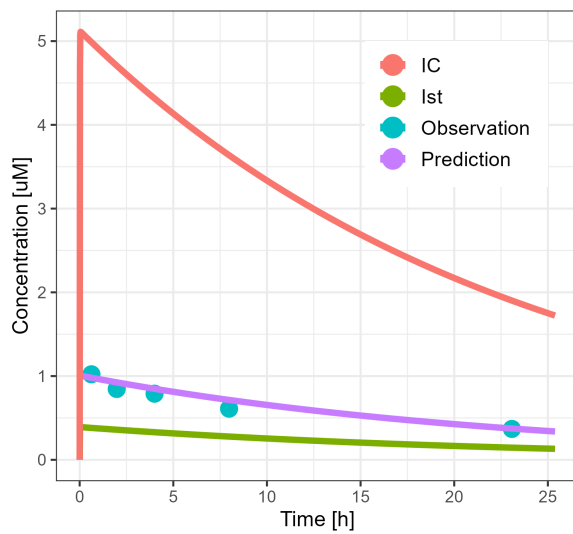
Names	$\frac{CL_{obs}}{ml/min/kg}$	$\frac{CL_{conv}}{ml/min/kg}$	$\frac{CL_{pred}}{ml/min/kg}$	$\frac{CL_{pred}}{CL_{obs}}$	$\frac{CL_{conv}}{CL_{obs}}$	On-chip clearance [ $ml/min$ ]	Surface area [ $cm^2$ ]	Study
Diclofenac	6.83	3.62	6.78	0.99	0.53	2.55E-03	2.90E-03	Docci
Lorazepam	1.43	0.69	1.36	0.95	0.48	5.48E-04	1.26E-05	Docci
Midazolam	11.28	12.67	11.08	0.98	1.12	2.20E-03	8.90E-04	Docci
Naloxone	16.89	16.30	19.45	1.15	0.96	1.11E-02	9.50E-04	Docci
Oxazepam	1.30	1.41	1.30	0.99	1.08	1.05E-03	7.16E-05	Docci
Posaconazole	2.24	0.42	0.81	0.36	0.19	5.67E-04	1.34E-05	Docci
Quinidine	6.33	3.79	5.68	0.90	0.60	1.65E-03	3.60E-05	Docci
Tolbutamide	0.21	0.10	0.36	1.73	0.47	1.02E-04	1.30E-04	Docci
Zidovudine	14.58	3.27	14.18	0.97	0.22	9.90E-04	6.80E-06	Docci
Propranolol	10.80	1.55	13.06	1.21	0.14	2.75E-03	5.34E-05	Tsamandouras
Prednisolone	2.02	0.25	2.83	1.40	0.12	3.58E-04	2.60E-02	Tsamandouras
Phenacetin	19.50	9.33	13.08	0.67	0.48	1.48E-03	2.37E-02	Tsamandouras
Imipramine	13.81	2.12	17.83	1.29	0.15	3.10E-04	2.70E-01	Hultman
Clozapine	8.41	1.33	8.47	1.01	0.16	2.90E-04	2.20E-01	Hultman
Paracetamol	0.40	0.23	0.52	1.31	0.57	7.50E-06	7.20E-02	Hultman
Irbesartan	19.28	2.86	18.44	0.96	0.15	3.89E-04	1.00E-01	Hultman
Ketoprofen	1.28	0.01	2.58	2.01	0.01	4.11E-04	9.00E-01	Hultman
Disopyramide	0.90	0.49	0.75	0.83	0.55	1.00E-05	1.40E-04	Hultman
(S)-Warfarin	0.06	0.002	0.06	0.94	0.03	4.80E-05	2.47E-01	Hultman
Carvedilol	8.70	4.05	10.95	1.26	0.47	1.51E-04	1.04E-02	Bonn
Imipramine	13.81	4.35	14.49	1.05	0.32	1.00E-06	2.90E-02	Kanebratt
Disopyramide	0.90	1.53	0.81	0.90	1.70	7.90E-06	1.00E-01	Kanebratt
(S)-Warfarin	0.06	0.09	0.03	0.56	1.47	1.00E-06	5.30E-03	Kanebratt
Carvedilol	8.70	4.05	6.82	0.78	0.47	1.00E-06	2.00E-02	Kanebratt
Midazolam	5.30	2.87	4.61	0.87	0.54	4.80E-04	1.59E-05	Rajan
Dextromethorphan	18.40	2.68	18.77	1.02	0.15	4.24E-05	2.12E-03	Rajan
Diclofenac	3.50	3.02	3.38	0.96	0.86	8.00E-05	2.50E-03	Rajan
Propranolol	10.80	5.00	8.06	0.75	0.46	1.90E-04	3.80E-03	Rajan
Raloxifene	9.33	13.8	13.29	1.42	1.48	1.50E-05	2.29E-03	Rajan
Tolbutamide	0.21	0.24	0.23	1.08	1.14	1.10E-04	2.88E-03	Rajan
Verapamil	17.60	7.85	13.69	0.78	0.45	8.50E-05	1.60E-03	Rajan
Zidovudine	14.58	4.20	15.18	1.04	0.29	2.20E-04	7.80E-03	Rajan

Table A.2: Overview of observed, conventional, on-chip and predicted clearance values of the investigated 32 drugs across different *in vitro* systems. CL is clearance, obs is observation, conv is conventional, pred is prediction

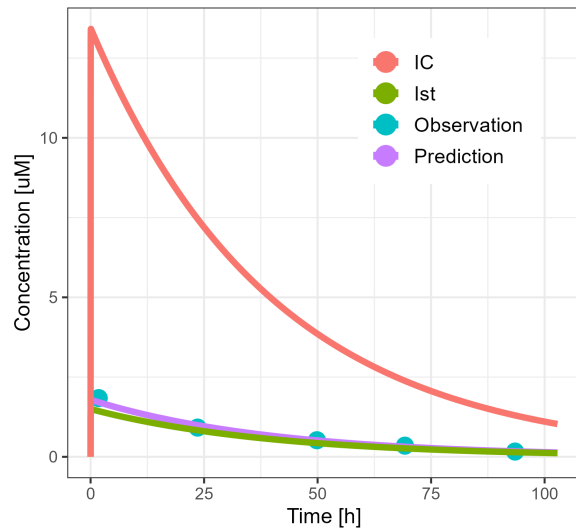
anisms, with most parameters derived from existing knowledge bases or from rigorously validated equations. This allows for a clear differentiation between pre-established organism parameters and drug-specific parameters. Despite the complexity of PBPK models, often involving hundreds of ordinary differential equations, the number of independent parameters for a new drug is typically limited (generally less than five per drug). This is made possible by the extensive use of existing physiological data. Similar to distribution models, these drug-specific parameters are typically consistent across various species or administration methods.

Drug characteristics like lipophilicity, solubility and molecular weight are independent of the organism's physiology. Conversely, drug-biological properties, such as unbound drug fraction in plasma or the tissue-plasma partition coefficient, are unique to the drug but also influenced by the drug's interaction with the biological system. By integrating these

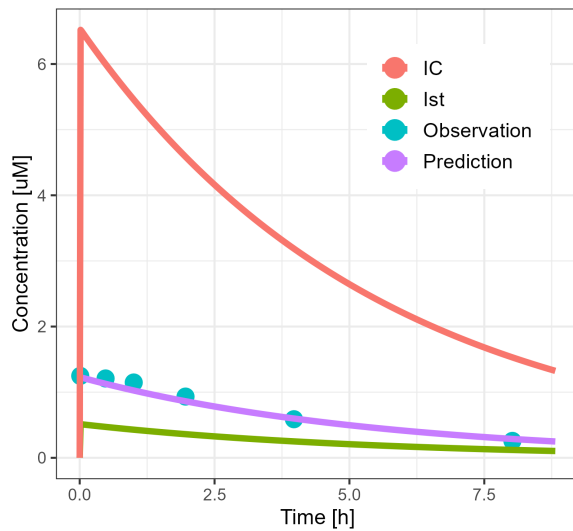
drug characteristics with the organism's anatomical and physiological traits, it becomes feasible to estimate parameters for passive processes that govern drug distribution in the body, such as membrane permeation. Additionally, data on the administration protocol and formulation attributes are crucial for configuring a PBPK simulation. The model can also incorporate temporal factors like gallbladder emptying or meal consumption, allowing for the assessment of their effects on drug PK. Properties for propranolol, including absorption after oral administration, clearance and human plasma PK, were taken from literature [153, 69, 154].



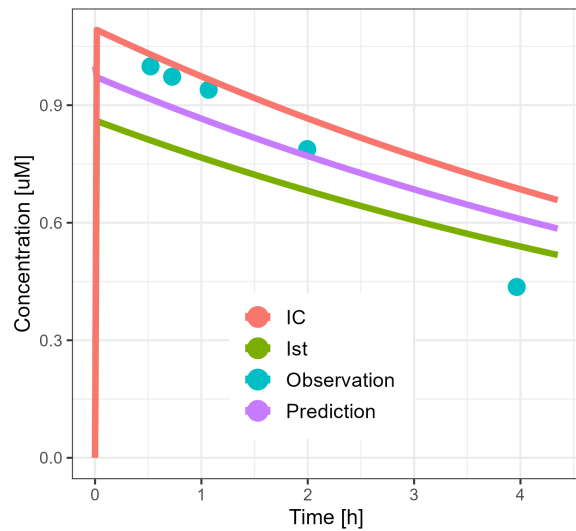
(a) Diclofenac



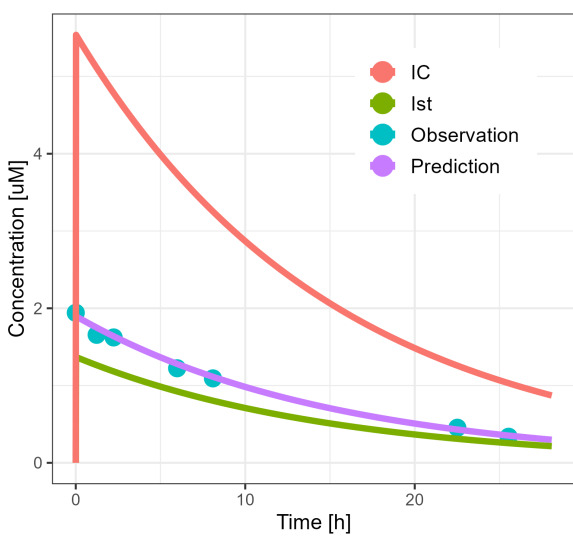
(b) Lorazepam



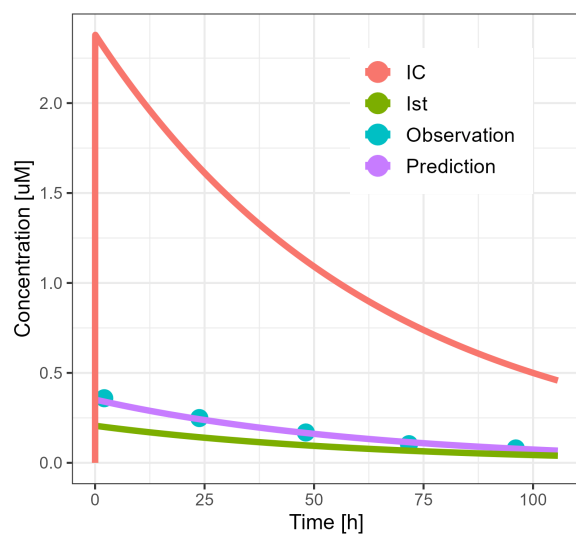
(c) Midazolam



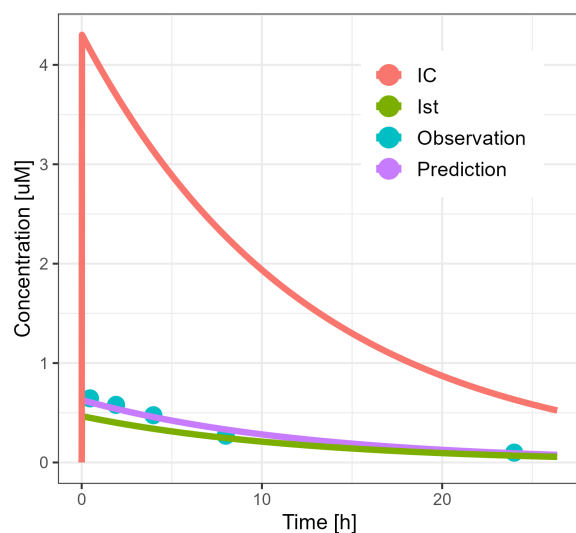
(d) Naloxone



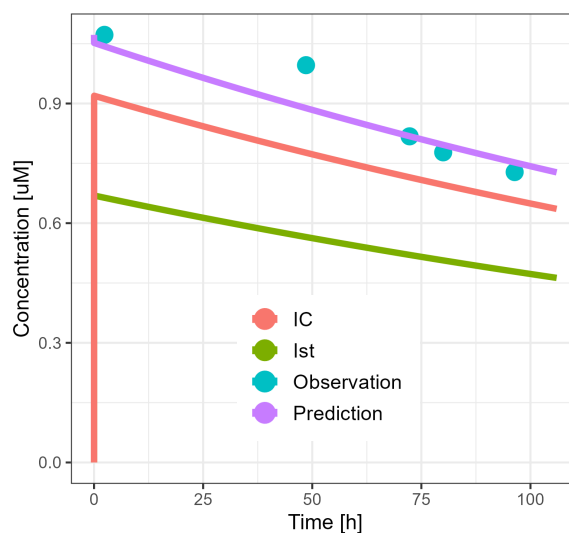
(e) Oxazepam



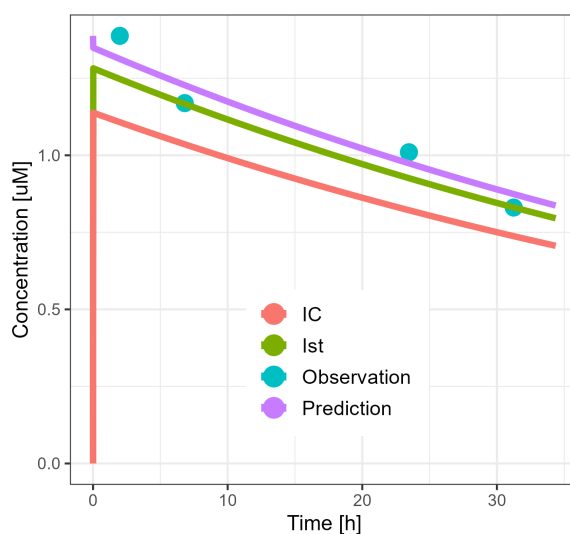
(f) Posaconazole



(g) Quinidine

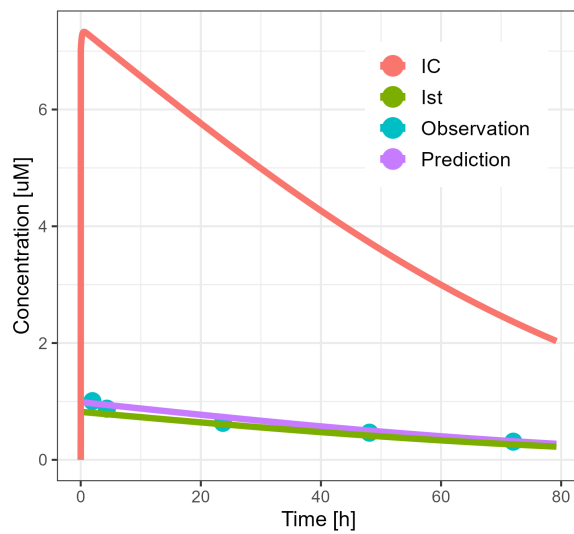


(h) Tolbutamide

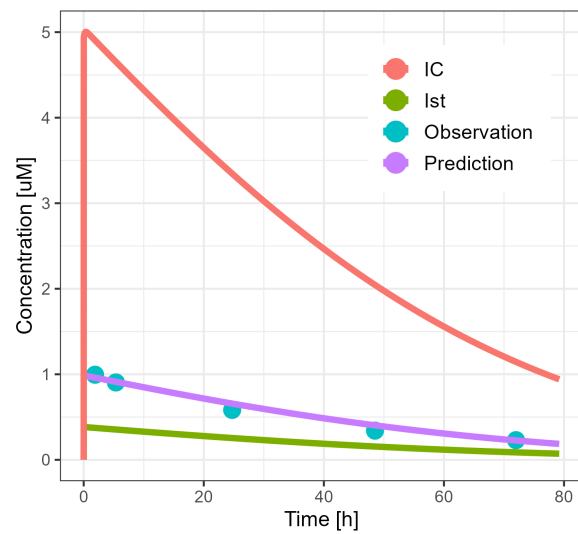


(i) Zidovudine

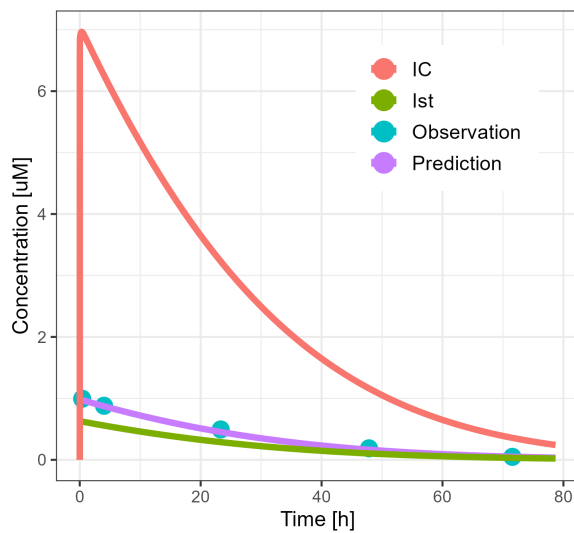
Figure A.2: Digital twin-based model simulation of on-chip kinetics after fitting parameters for nine compounds from Docci et al. [10]; IC = intracellular, Ist = interstitium.



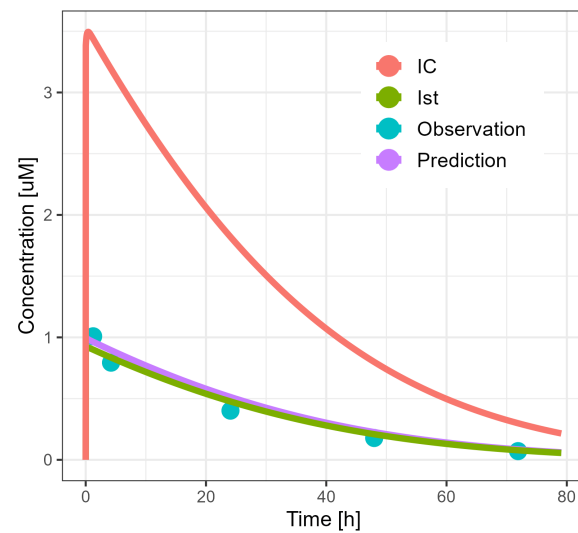
(a) Dextromethorphan



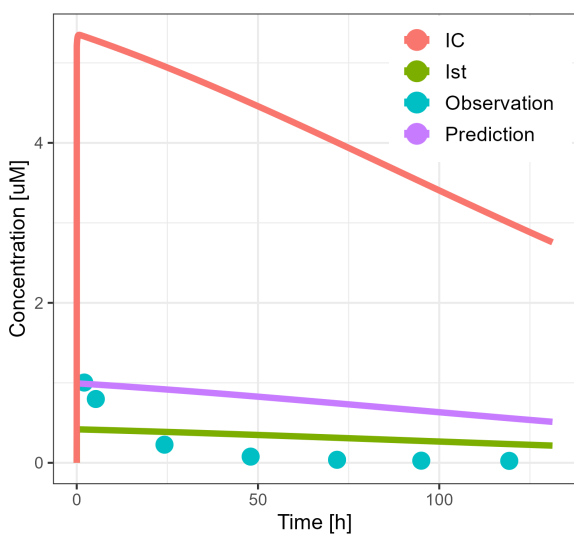
(b) Diclofenac



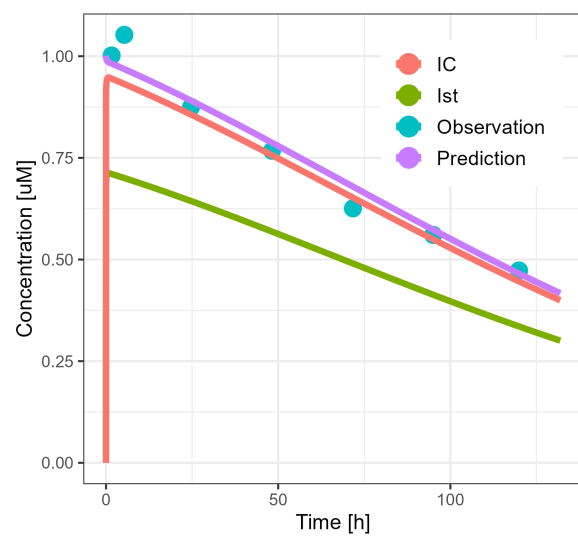
(c) Midazolam



(d) Propranolol



(e) Raloxifene



(f) Tolbutamide



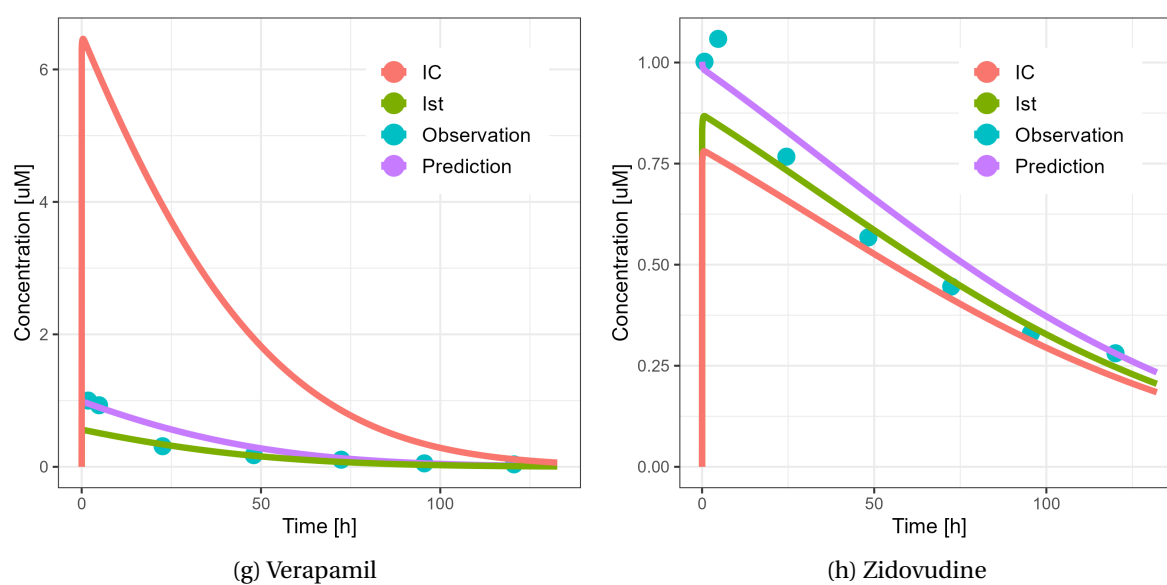
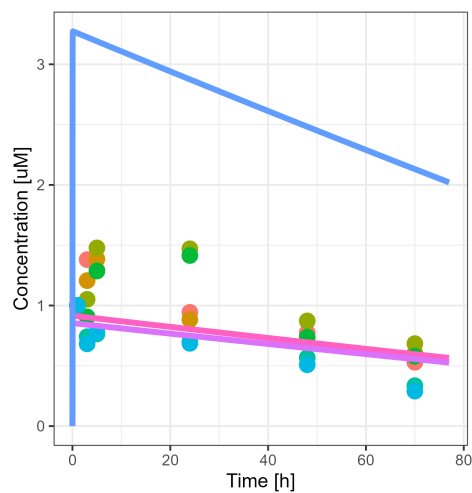
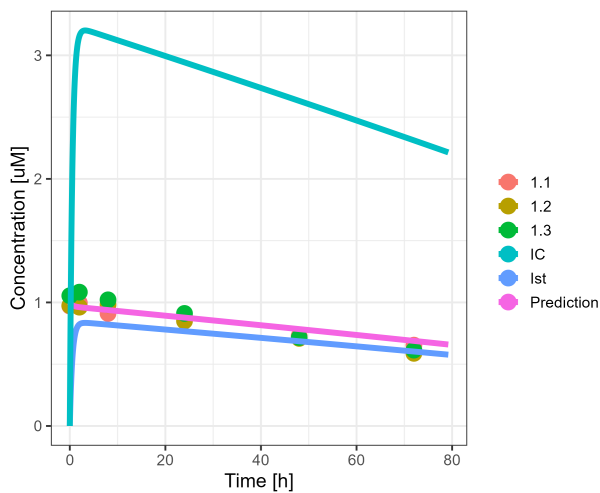


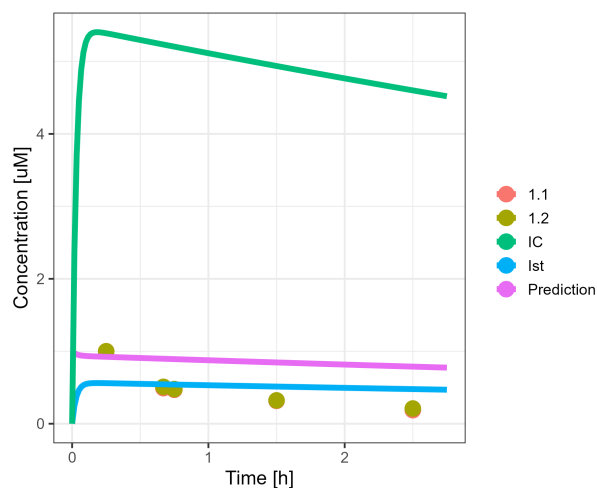
Figure A.3: Digital twin-based model simulation of on-chip kinetics after fitting parameters for eight compounds from Rajan et al. [59]; IC = intracellular, Ist = interstitium.



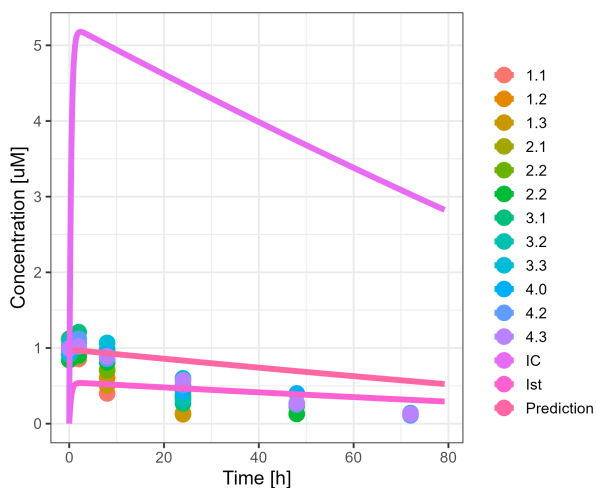
(a) (S)-warfarin [16]



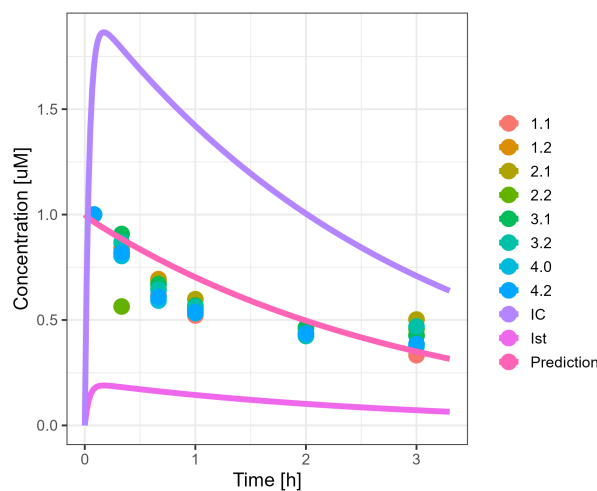
(b) (S)-warfarin [14]



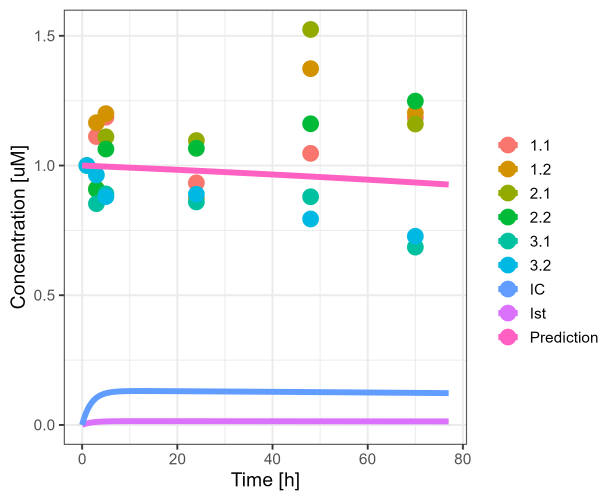
(c) Carvedilol [16]



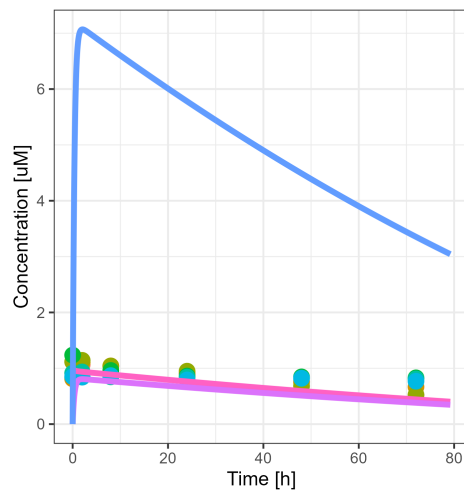
(d) Carvedilol [14]



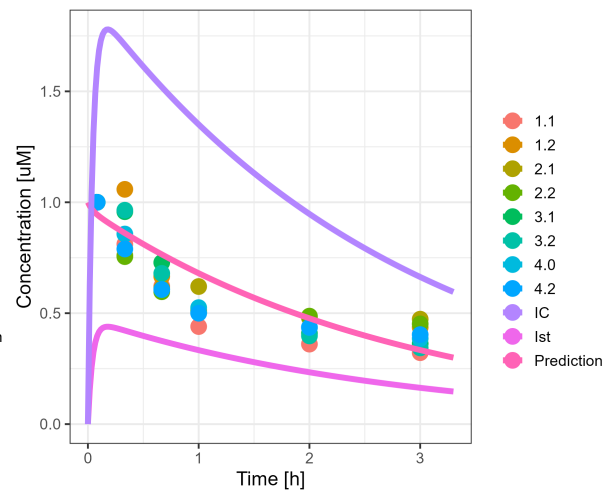
(e) Clozapine [16]



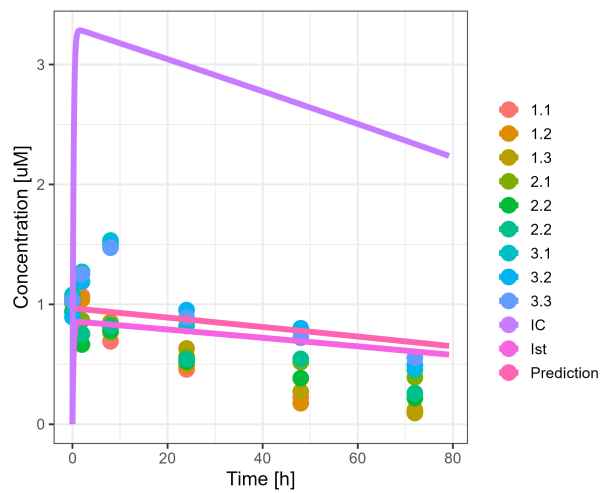
(f) Disopyramide [16]



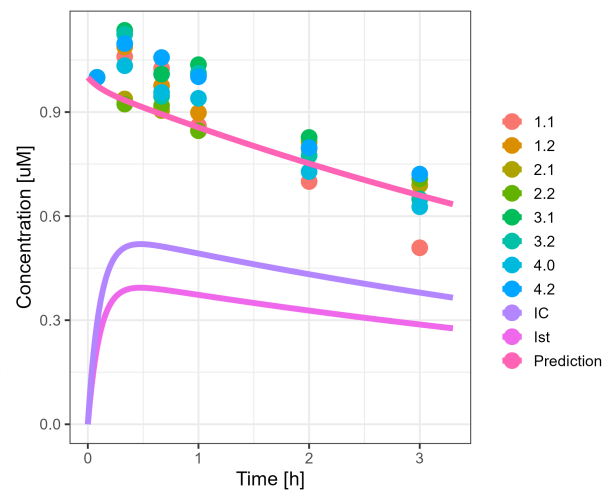
(g) Disopyramide [14]



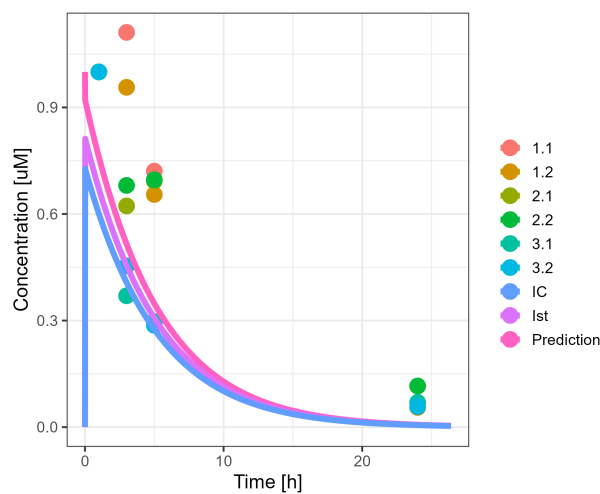
(h) Imipramine [16]



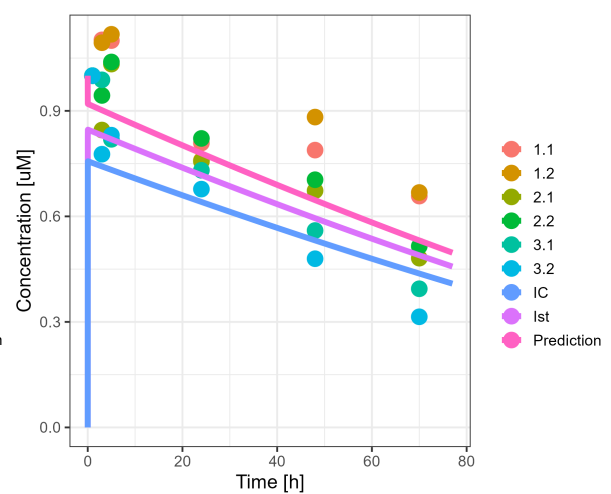
(i) Imipramine [14]



(j) Irbesartan [16]



(k) Ketoprofen [16]



(l) Paracetamol [16]

Figure A.4: Digital twin-based model simulation of on-chip kinetics after fitting parameters for 3D spheroids (Kanebratt et al. [14], Bonn et al. [16]); IC = intracellular, Ist = interstitium. Numbers indicate individual biological replicates.

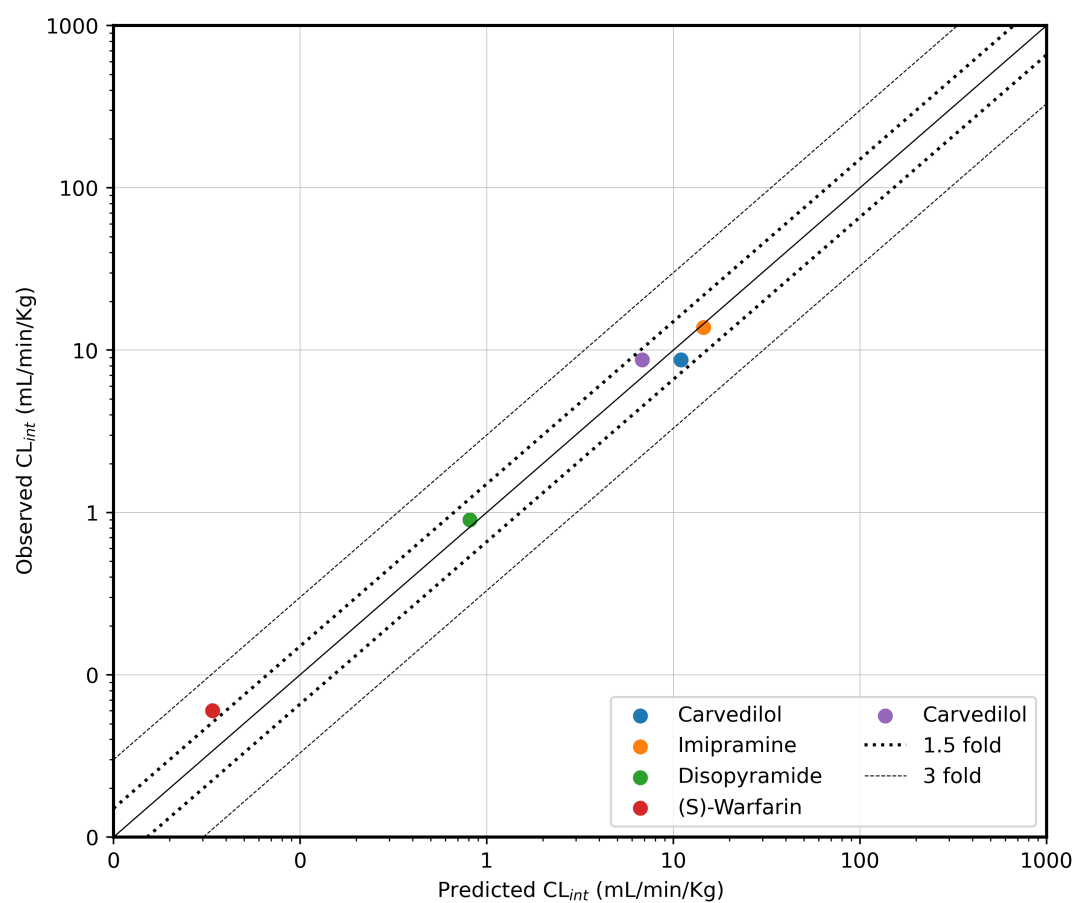


Figure A.5: Correlation between observed and predicted *in vivo* intrinsic clearance ( $CL_{int}$ ) using the three-compartment model for four drugs from Kanebratt et.al [14] and Carvedilol from Bonn et al. [16] (AFE = 0.879). The solid line shows the line of unity.

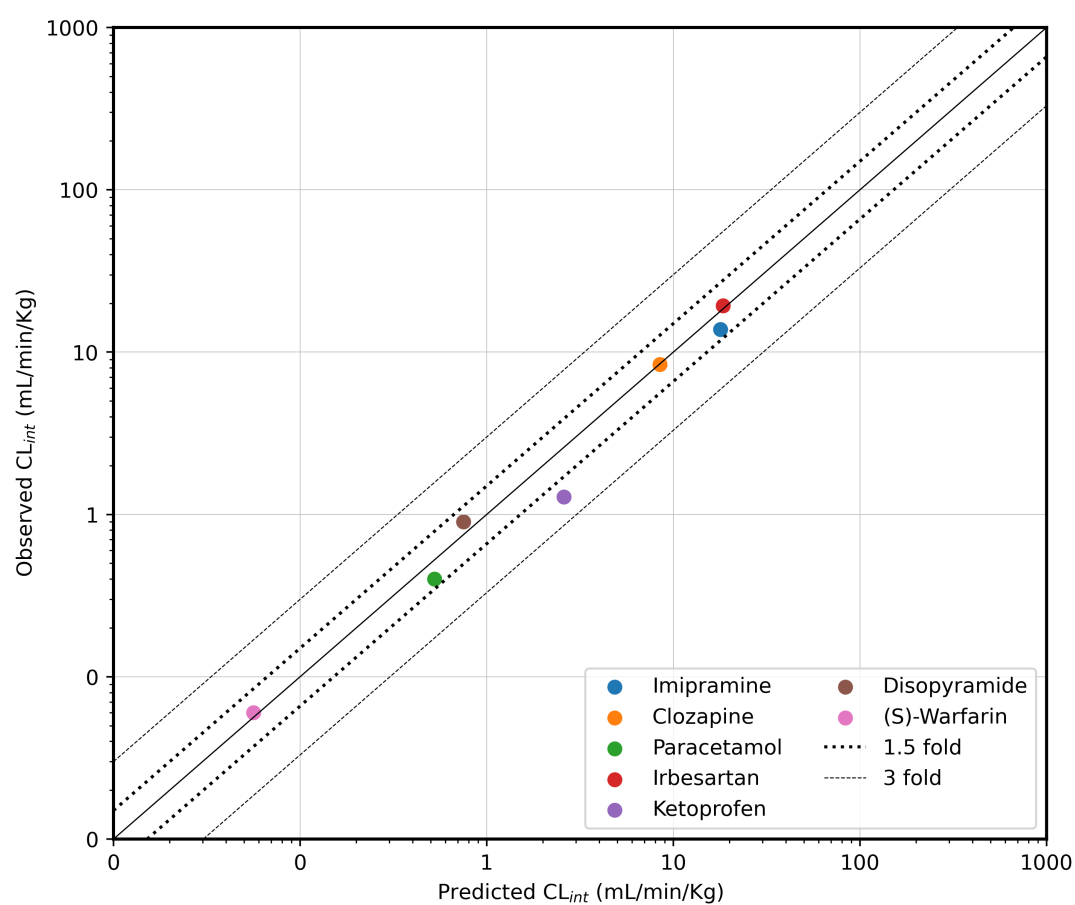


Figure A.6: Correlation between observed and predicted *in vivo* intrinsic clearance ( $CL_{int}$ ) using the three-compartment model for seven drugs from Hultman et.al [16] (AFE= 1.143). The solid line shows the line of unity.

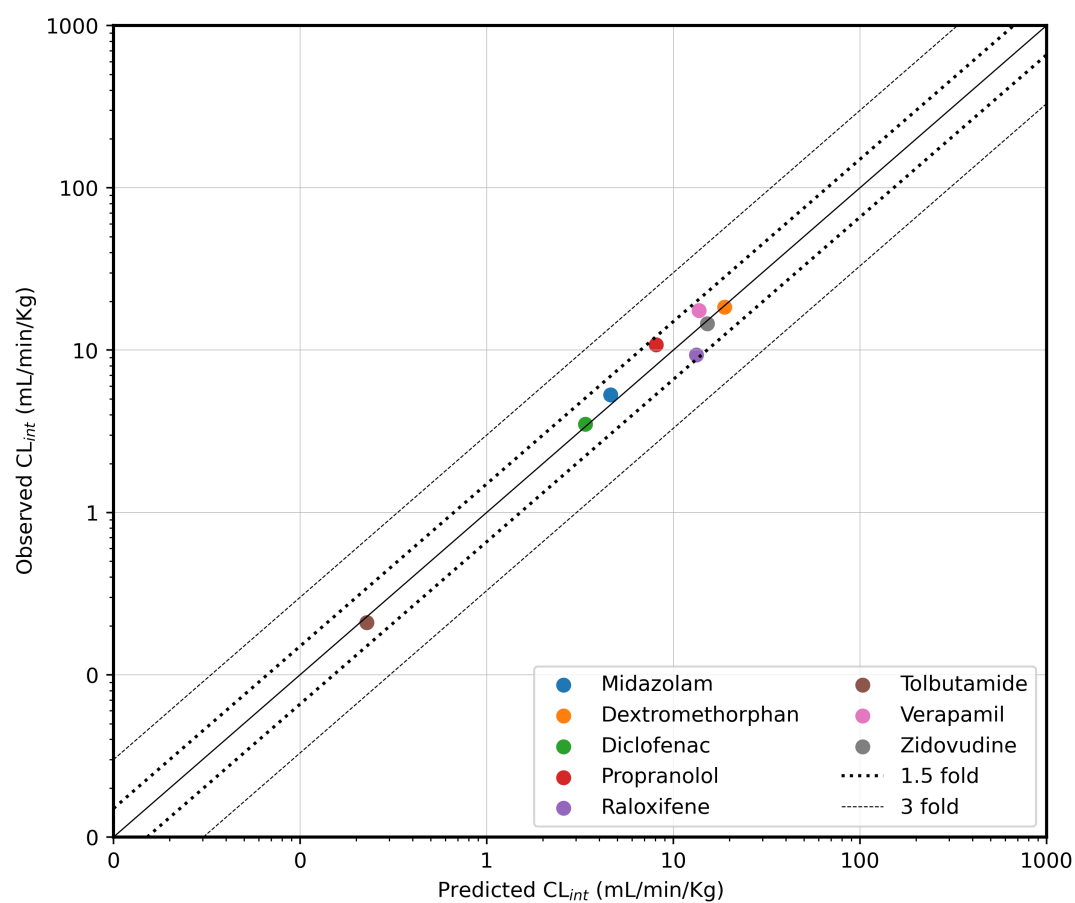


Figure A.7: Correlation between observed and predicted *in vivo* intrinsic clearance ( $CL_{int}$ ) using three-compartment ODE liver chip for eight drugs from Rajan et. al [59] (AFE=0.972). The solid line shows the line of unity.

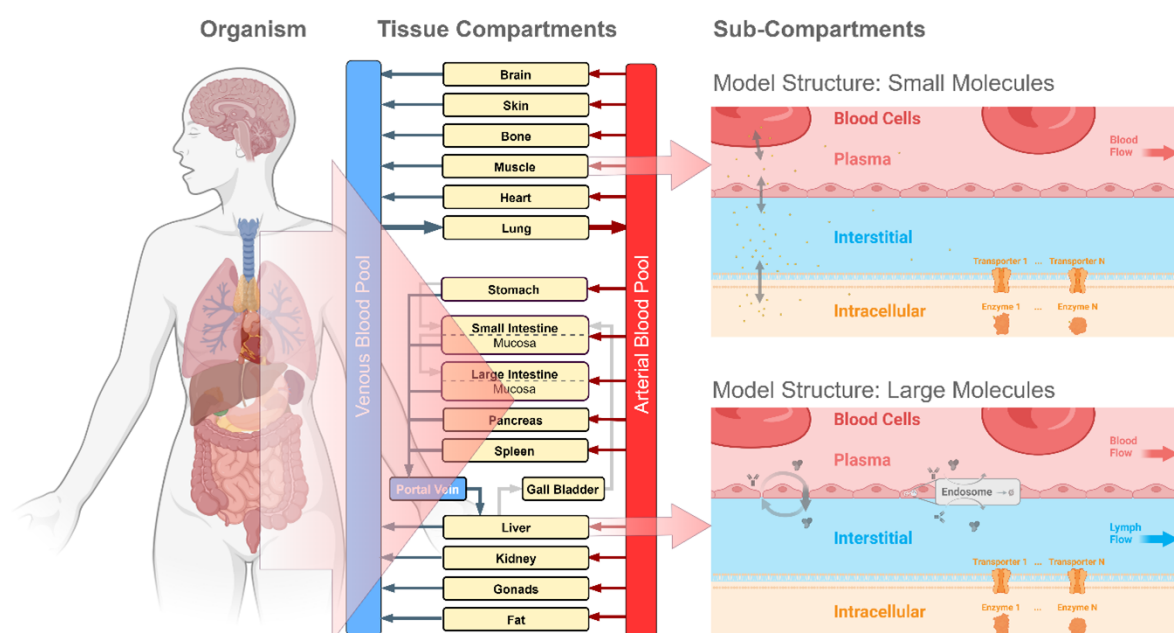


Figure A.8: Representation of the generic structure of a whole-body PBPK model





# Supplementary information for Chapter 4

## B.1 Compartments and ODEs

### Pulmonary circulation

$$\begin{aligned}
 C_{Ar}^{Pul} \frac{dP_{Ar}^{Pul}}{dt} &= Q_{PV}(t) - Q_{Ar}^{Pul}(t), \\
 L_{Ar}^{Pul} \frac{dQ_{Ar}^{Pul}}{dt} &= -R_{Ar}^{Pul} \cdot Q_{Ar}^{Pul}(t) + P_{Ar}^{Pul}(t) - P_{Lung}(t), \\
 C_{Lung} \frac{dP_{Lung}}{dt} &= Q_{Ar}^{Pul}(t) - Q_{Ar}^{LPul}(t) - Q_{Ar}^{RPul}(t), \\
 C_{LL} \frac{dP_{LL}}{dt} &= Q_{Ar}^{LPul}(t) - Q_{LL}(t), \\
 L_{Ar}^{LPul} \frac{dQ_{Ar}^{LPul}}{dt} &= -R_{Ar}^{LPul} \cdot Q_{Ar}^{LPul}(t) + P_{Lung}(t) - P_{LL}(t), \\
 L_{LL} \frac{dQ_{LL}}{dt} &= -R_{LL} \cdot Q_{LL}(t) + P_{LL}(t) - P_{Ven}^{Pul}(t), \\
 C_{RL} \frac{dP_{RL}}{dt} &= Q_{Ar}^{RPul}(t) - Q_{RL}(t), \\
 L_{Ar}^{RPul} \frac{dQ_{Ar}^{RPul}}{dt} &= -R_{Ar}^{RPul} \cdot Q_{Ar}^{RPul}(t) + P_{Lung}(t) - P_{RL}(t), \\
 L_{RL} \frac{dQ_{RL}}{dt} &= -R_{RL} \cdot Q_{RL}(t) + P_{RL}(t) - P_{Ven}^{Pul}(t), \\
 C_{Ven}^{Pul} \frac{dP_{Ven}^{Pul}}{dt} &= Q_{RL}(t) + Q_{LL}(t) - Q_{Ven}^{Pul}(t), \\
 L_{Ven}^{Pul} \frac{dQ_{Ven}^{Pul}}{dt} &= -R_{Ven}^{Pul} \cdot Q_{Ven}^{Pul}(t) + P_{Ven}^{Pul}(t) - P_{LA}(t).
 \end{aligned}$$

### Ascending aorta and aortic arch

$$\begin{aligned}
C_{\text{AscA}} \frac{dP_{\text{AscA}}}{dt} &= Q_{\text{AV}}(t) - Q_{\text{AscA}}(t), \\
L_{\text{AscA}} \frac{dQ_{\text{AscA}}}{dt} &= -R_{\text{AscA}} \cdot Q_{\text{AscA}}(t) + P_{\text{AscA}} - P_{\text{AArc}}(t), \\
C_{\text{AArc}} \frac{dP_{\text{AArc}}}{dt} &= Q_{\text{AscA}}(t) - Q_{\text{Sub}}(t) - Q_{\text{CCar}}(t) - Q_{\text{DscA}}(t).
\end{aligned}$$

### Upper body- subclavian, cerebral, facial, upper limbs and superior vena cava

$$\begin{aligned}
L_{\text{Sub}} \frac{dQ_{\text{Sub}}}{dt} &= -R_{\text{Sub}} \cdot Q_{\text{Sub}}(t) + P_{\text{AArc}}(t) - P_{\text{ULimb}}(t), \\
C_{\text{CCar}} \frac{dP_{\text{CCar}}}{dt} &= Q_{\text{CCar}}(t) - Q_{\text{ICar}}(t) - Q_{\text{ECar}}(t), \\
L_{\text{CCar}} \frac{dQ_{\text{CCar}}}{dt} &= -R_{\text{CCar}} \cdot Q_{\text{CCar}}(t) + P_{\text{AArc}}(t) - P_{\text{CCar}}, \\
L_{\text{ICar}} \frac{dQ_{\text{ICar}}}{dt} &= -R_{\text{ICar}} \cdot Q_{\text{ICar}}(t) + P_{\text{CCar}}(t) - P_{\text{Cer}}, \\
C_{\text{Cer}} \frac{dP_{\text{Cer}}}{dt} &= Q_{\text{ICar}}(t) - Q_{\text{Cer}}(t), \\
L_{\text{Cer}} \frac{dQ_{\text{Cer}}}{dt} &= -R_{\text{Cer}} \cdot Q_{\text{Cer}}(t) - R_{\text{CerT}} \cdot Q_{\text{Cer}}(t) + P_{\text{Cer}}(t) - P_{\text{SVC}}, \\
[1.5ex] L_{\text{ECar}} \frac{dQ_{\text{ECar}}}{dt} &= -R_{\text{ECar}} \cdot Q_{\text{ECar}}(t) + P_{\text{CCar}}(t) - P_{\text{Fac}}, \\
C_{\text{Fac}} \frac{dP_{\text{Fac}}}{dt} &= Q_{\text{ECar}}(t) - Q_{\text{Fac}}(t), \\
L_{\text{Fac}} \frac{dQ_{\text{Fac}}}{dt} &= -R_{\text{Fac}} \cdot Q_{\text{Fac}}(t) - R_{\text{FacT}} \cdot Q_{\text{Fac}}(t) + P_{\text{Fac}}(t) - P_{\text{SVC}}, \\
[1.5ex] C_{\text{ULimb}} \frac{dP_{\text{ULimb}}}{dt} &= Q_{\text{Sub}}(t) - Q_{\text{ULimb}}(t), \\
L_{\text{ULimb}} \frac{dQ_{\text{ULimb}}}{dt} &= -R_{\text{ULimb}} \cdot Q_{\text{ULimb}}(t) + P_{\text{ULimb}}(t) - P_{\text{SVC}}(t), \\
C_{\text{SVC}} \frac{dP_{\text{SVC}}}{dt} &= Q_{\text{ULimb}}(t) + Q_{\text{Cer}}(t) - Q_{\text{SVC}}(t), \\
L_{\text{SVC}} \frac{dQ_{\text{SVC}}}{dt} &= -R_{\text{SVC}} \cdot Q_{\text{SVC}}(t) + P_{\text{SVC}}(t) - P_{\text{RA}}(t).
\end{aligned}$$

### Descending aorta, abdominal aorta, renal circulation, lower limbs and inferior vena cava

$$\begin{aligned}
C_{\text{DscA}} \frac{dP_{\text{DscA}}}{dt} &= Q_{\text{DscA}}(t) - Q_{\text{Ar}}^{\text{Ren}}(t) - Q_{\text{AbdA}}(t), \\
L_{\text{DscA}} \frac{dQ_{\text{DscA}}}{dt} &= -R_{\text{DscA}} \cdot Q_{\text{DscA}}(t) + P_{\text{AArc}}(t) - P_{\text{DscA}}(t), \\
C_{\text{AbdA}} \frac{dP_{\text{AbdA}}}{dt} &= Q_{\text{AbdA}}(t) - Q_{\text{Cel}}(t) - Q_{\text{Mes}}(t) - Q_{\text{Iliac}}(t), \\
L_{\text{AbdA}} \frac{dQ_{\text{AbdA}}}{dt} &= -R_{\text{AbdA}} \cdot Q_{\text{AbdA}}(t) + P_{\text{DscA}}(t) - P_{\text{AbdA}}(t),
\end{aligned}$$

$$\begin{aligned}
L_{Ar}^{Ren} \frac{dQ_{Ar}^{Ren}}{dt} &= -R_{Ar}^{Ren} \cdot Q_{Ar}^{Ren}(t) + P_{DscA}(t) - P_{Ren}(t), \\
C_{Ar}^{Ren} \frac{dP_{Ar}^{Ren}}{dt} &= Q_{AbdA1}(t) - Q_{Ar}^{Ren}(t), \\
L_{Ar}^{Ren} \frac{dQ_{Ar}^{Ren}}{dt} &= -R_{Ar}^{Ren} \cdot Q_{Ar}^{Ren}(t) + P_{Ren}(t) - P_{IVC}(t), \\
C_{LLimb} \frac{dP_{LLimb}}{dt} &= Q_{Iliac}(t) - Q_{LLimb}(t), \\
L_{Iliac} \frac{dQ_{Iliac}}{dt} &= -R_{Iliac} \cdot Q_{Iliac}(t) + P_{AbdA}(t) - P_{LLimb}(t), \\
L_{LLimb} \frac{dQ_{LLimb}}{dt} &= -R_{LLimb} \cdot Q_{LLimb}(t) + P_{LLimb}(t) - P_{IVC}(t), \\
C_{IVC} \frac{dP_{IVC}}{dt} &= Q_{LLimb}(t) + Q_{Hep}(t) + Q_{Ren}(t) + Q_{Ven}^{Por}(t) - Q_{IVC}(t), \\
L_{IVC} \frac{dQ_{IVC}}{dt} &= -R_{IVC} \cdot Q_{IVC}(t) + P_{IVC}(t) - P_{RA}(t).
\end{aligned}$$

### Celiac, spleen, mesenteric, hepatic and portal vein

$$\begin{aligned}
L_{Cel} \frac{dQ_{Cel}}{dt} &= -R_{Cel} \cdot Q_{Cel}(t) + P_{AbdA}(t) - P_{Cel}, \\
C_{Spl} \frac{dP_{Spl}}{dt} &= Q_{Ar}^{Spl}(t) - Q_{Spl}(t), \\
L_{Ar}^{Spl} \frac{dQ_{Ar}^{Spl}}{dt} &= -R_{Ar}^{Spl} \cdot Q_{Ar}^{Spl}(t) + P_{Cel}(t) - P_{Spl}(t), \\
L_{Spl} \frac{dQ_{Spl}}{dt} &= -R_{Spl} \cdot Q_{Spl}(t) + P_{Spl}(t) - P_{Ven}^{Por}, \\
C_{Mes} \frac{dP_{Mes}}{dt} &= Q_{Ar}^{Mes}(t) - Q_{Mes}(t), \\
L_{Ar}^{Mes} \frac{dQ_{Ar}^{Mes}}{dt} &= -R_{Ar}^{Mes} \cdot Q_{Ar}^{Mes}(t) + P_{AbdA}(t) - P_{Ar}^{Mes}, \\
L_{Mes} \frac{dQ_{Mes}}{dt} &= -R_{Mes} \cdot Q_{Mes}(t) + P_{Mes}(t) - P_{Ven}^{Por}, \\
C_{Hep} \frac{dP_{Hep}}{dt} &= Q_{Ar}^{Hep}(t) + Q_{Ven}^{Por}(t) - Q_{Hep}, \\
L_{Ar}^{Hep} \frac{dQ_{Ar}^{Hep}}{dt} &= -R_{Ar}^{Hep} \cdot Q_{Ar}^{Hep}(t) + P_{Cel} - P_{Hep}(t), \\
L_{Hep} \frac{dQ_{Hep}}{dt} &= -R_{Hep} \cdot Q_{Hep}(t) + P_{Hep} - P_{IVC}(t), \\
C_{Ven}^{Por} \frac{dP_{Ven}^{Por}}{dt} &= Q_{Spl}(t) + Q_{Mes}(t) - Q_{Ven}^{Por}(t), \\
L_{Ven}^{Por} \frac{dQ_{Ven}^{Por}}{dt} &= -R_{Ven}^{Por} \cdot Q_{Ven}^{Por}(t) + P_{Ven}^{Por} - P_{IVC}(t).
\end{aligned}$$

### B.2 Supplementary figures and tables

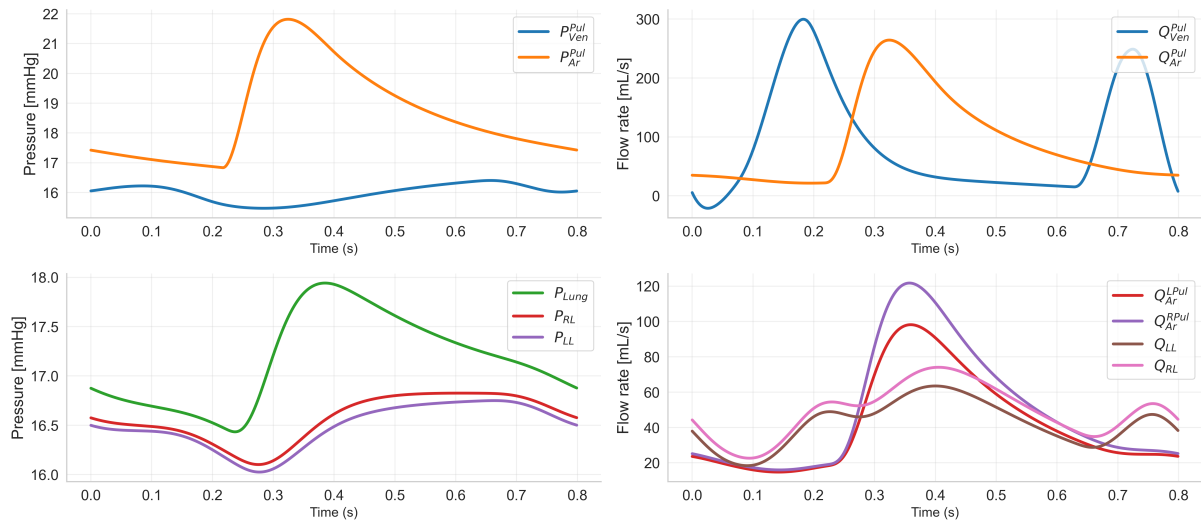


Figure B1: Simulated pressure and flow rate for pulmonary circulation (left and right lung, pulmonary arteries and veins).

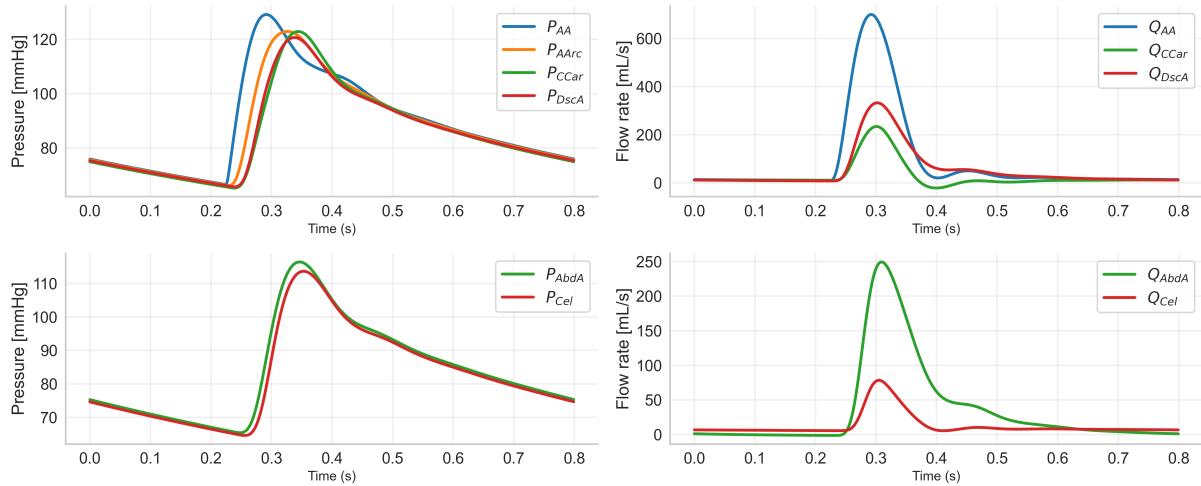


Figure B2: Simulated pressure and flow rate for major arteries.

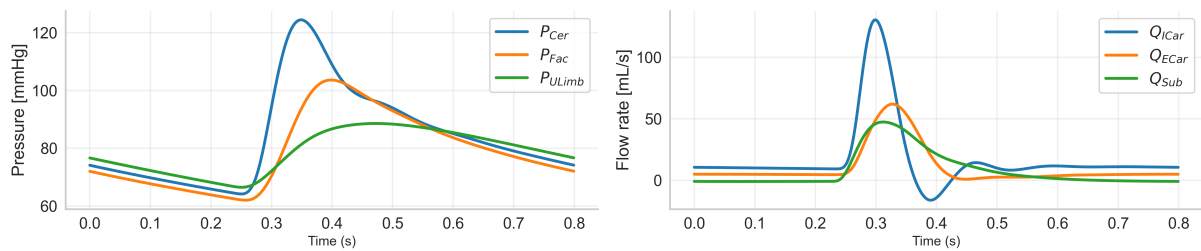


Figure B3: Simulated pressure and flow rate for upper body organs.

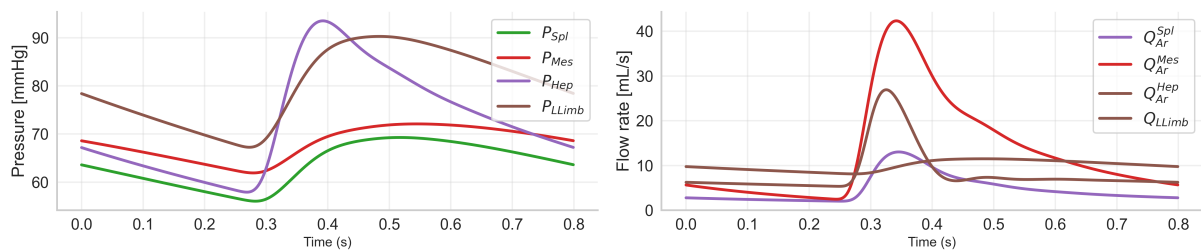


Figure B4: Simulated pressure and flow rate for abdominal organs.

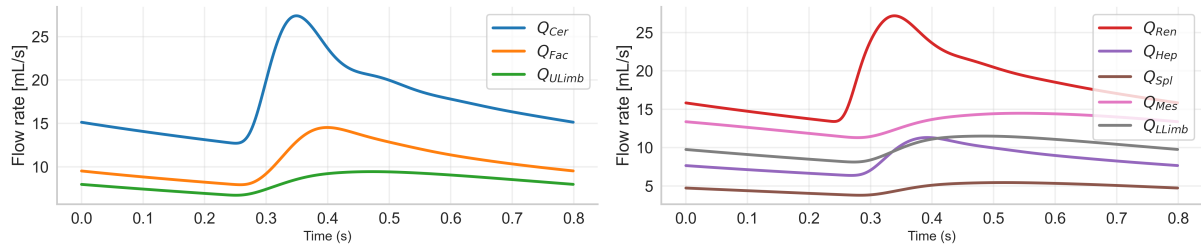


Figure B5: Simulated pressure and flow rate for abdominal organs.

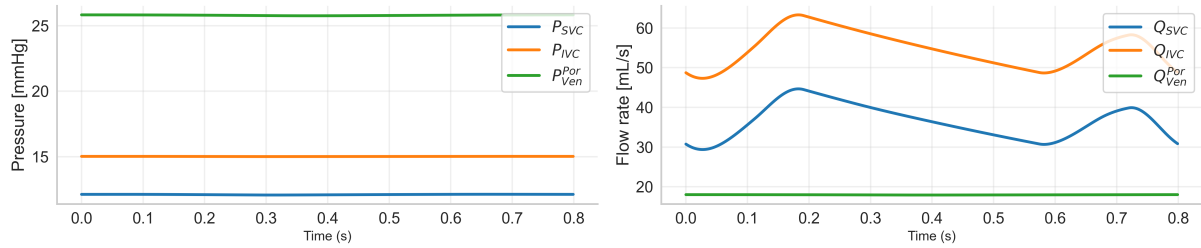


Figure B6: Simulated pressure and flow rate in major veins.

Variable	Initial value	Variable	Initial value	Variable	Initial value
$V_{LA}$	80	$P_{CCar}$	40	$P_{Cer}$	60
$V_{LV}$	110	$Q_{CCar}$	0.5	$Q_{Cer}$	0.5
$V_{RA}$	80	$Q_{ICar}$	0.5	$P_{Spl}$	30
$V_{RV}$	110	$Q_{ECar}$	0.5	$Q_{Spl1}$	0.5
$P_{Ar}^{Pul}$	20	$P_{Cel}$	50	$Q_{Spl2}$	0.5
$P_{Ven}^{Pul}$	15	$Q_{Cel}$	0.5	$P_{Hep}$	30
$Q_{Ar}^{Pul}$	5.0	$P_{Ren}$	15	$Q_{Hep1}$	0.5
$Q_{Ven}^{Pul}$	5.0	$Q_{Ren}$	0.5	$Q_{Hep2}$	0.5
$P_{Lung}$	5	$Q_{Ar}^{Ren}$	0.5	$P_{Mes}$	30
$P_{LL}$	15	$P_{DscA}$	80	$Q_{Mes1}$	0.5
$Q_{LL1}$	0.1	$Q_{DscA}$	1.0	$Q_{Mes2}$	0.5
$Q_{LL2}$	0.1	$P_{AbdA}$	70	$P_{sys}$	10
$P_{RL}$	15	$Q_{AbdA}$	1.0	$Q_{Psys}$	0.5
$Q_{RL1}$	0.1	$P_{ULimb}$	40	$P_{LLimb}$	30
$Q_{RL2}$	0.1	$Q_{Iliac}$	0.5	$Q_{LLimb}$	0.5
$P_{AscA}$	120	$Q_{ULimb}$	0.5	$P_{Ven}^{Por}$	10
$Q_{AscA}$	5.0	$P_{Fac}$	40	$Q_{Ven}^{Por}$	0.5
$P_{AArc}$	100	$Q_{Fac}$	0.5		

Table B.1: Initial values of variables. P, Q and V represent blood pressure ( $mmHg/s$ ), flow ( $mL/s$ ) and volume ( $mL$ ).

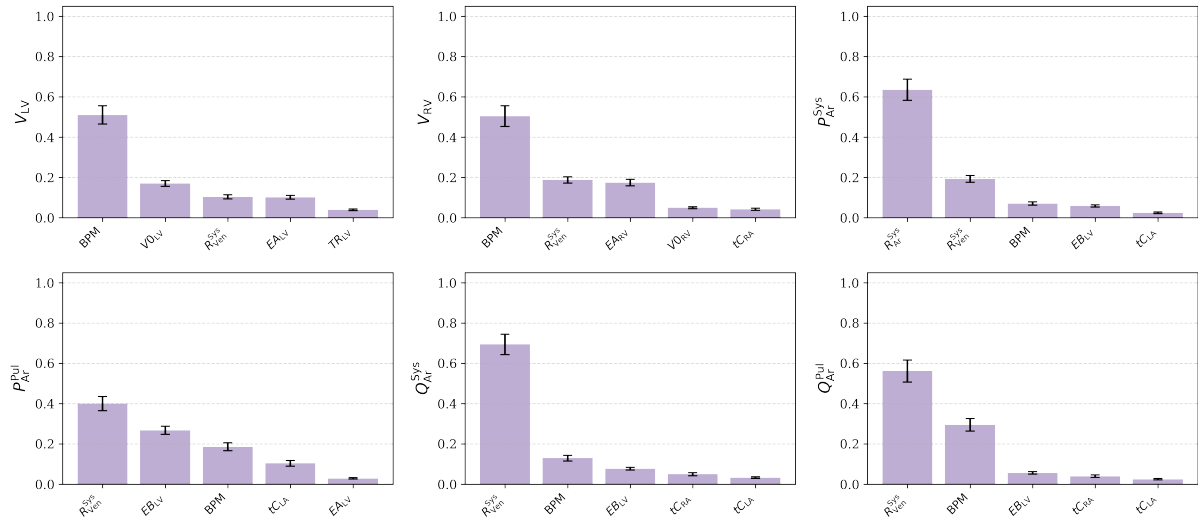


Figure B7: Pressure, volume and flow rates during a heartbeat in heart chambers

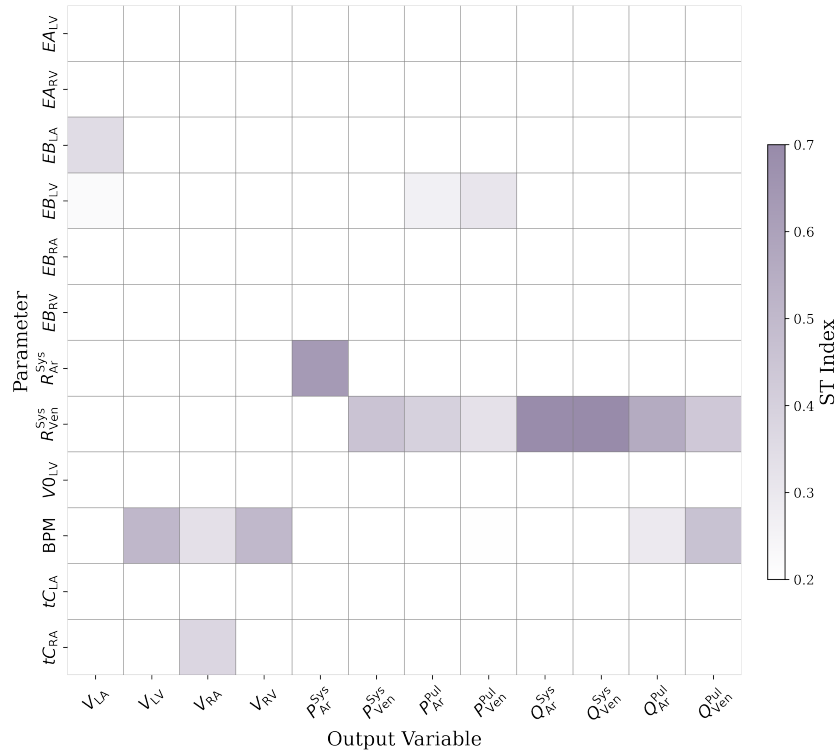


Figure B8: Total Sobol indices for all model parameters and output variables for base model. Darker shades indicate higher sensitivity.

Variables	Description	Simulated range	Min reference range	Max reference range	Status
$V_{LA}$	Left atrium	41–90	Men: 31 Women: 28 [155]	Men: 112 Women: 100 [155, 156]	Acceptable
$V_{LV}$	Left ventricle	41–114	Men: 62 Women: 46 [155]	Men: 155 Women: 106 [155]	Acceptable
$V_{RA}$	Right atrium	52–97	Men: 24 Women: 24 [155]	Men: 105 Women: 81 [155]	Satisfactory
$V_{RV}$	Right ventricle	46–118	Men: 74 Women: 58 [155]	Men: 202 Women: 150 [155]	Acceptable
$P_{LA}$	Left atrium	10–17	2 [157]	12 [157, 158]	Satisfactory
$P_{LV}$	Left ventricle	9–134	3 [158]	120 [158]	Acceptable
$P_{RA}$	Right atrium	4–7	0 [158]	7 [158, 159]	Acceptable
$P_{RV}$	Right ventricle	3–26	0 [158]	30 [158, 1]	Acceptable
$P_{Ven}^{Pul}$	Pulmonary veins	15–16	5 [160]	12 [160, 161]	Satisfactory
$P_{Ar}^{Pul}$	Pulmonary artery	17–22	8 [1]	25 [1, 161]	Satisfactory
$P_{AscA}$	Ascending aorta	67–129	80 [162]	120 [162]	Acceptable
$P_{AArc}$	Aortic arch	66–123	80 [162]	120 [162]	Acceptable
$P_{CCar}$	Common carotid	65–123	70 [162]	120 [162]	Acceptable
$P_{Cer}$	Cerebral artery	64–124	60 (MAP) [163]	80 (MAP) [163]	Acceptable
$P_{Fac}$	Facial circulation	62–104	–	–	Unavailable
$P_{ULimb}$	Upper limb	67–89	80 [164]	120 [164]	Satisfactory
$P_{DscA}$	Descending aorta	66–121	80 [162]	120 [162]	Acceptable
$P_{AbdA}$	Abdominal aorta	65–116	80 [162]	120 [162]	Acceptable
$P_{Ren}$	Renal arterial	64–118	80 [165, 162]	100 [165, 162]	Acceptable
$P_{Cel}$	Celiac artery	65–114	80 [162]	120 [162]	Acceptable
$P_{Mes}$	Mesenteric artery	62–72	–	–	Unavailable
$P_{Spl}$	Splanchnic artery	56–69	–	–	Unavailable
$P_{Hep}$	Hepatic artery	58 - 94	–	–	Unavailable
$P_{LLimb}$	Lower limb	67–90	80 [166]	120 [166]	Satisfactory
$P_{Ven}^{Por}$	Portal vein	26–26	5 [167]	15 [167]	Satisfactory
$P_{IVC}$	Inferior vena cava	15 - 15	2 [168]	14 [168]	Satisfactory
$P_{SVC}$	Superior vena cava	12 - 12	2 [164]	13 [164]	Satisfactory

Table B.2: Hemodynamics for whole-body cardiovascular model. Simulated and reference ranges for volumes (mL) and pressures (mmHg) in all model compartments.

Variables	Description	Simulated mean (mL/s)	Mean flow reference (mL/s)	Status
$Q_{Ven}^{Pul}$	Pulmonary veins	91.0	80–100 [169, 170]	Acceptable
$Q_{Ar}^{Pul}$	Pulmonary artery	91.0	80–100 [169, 170]	Acceptable
$Q_{AV}$	Aortic valve	91.0	80–100 [169, 170]	Acceptable
$Q_{PV}$	Pulmonary valve	91.0	80–100 [169, 170]	Acceptable
$Q_{TV}$	Tricuspid valve	91.0	80–100 [169, 170]	Acceptable
$Q_{MV}$	Mitral valve	91.0	80–100 [169, 170]	Acceptable
$Q_{Ren}$	Renal circulation	18.0	16–22 [170, 162]	Acceptable
$Q_{AscA}$	Ascending aorta	91.0	80–100 [169, 170]	Acceptable
$Q_{CCar}$	Common carotid	28.0	16–26 [171, 170]	Satisfactory
$Q_{ICar}$	Internal carotid	18.0	12–16 [171]	Satisfactory
$Q_{ECar}$	External carotid	11.0	4–8 [171]	Satisfactory
$Q_{Cer}$	Cerebral circulation	18.0	12–15 [171, 170]	Satisfactory
$Q_{Fac}$	Facial circulation	11.0	2–4 [171]	Satisfactory
$Q_{ULimb}$	Upper limb	8.0	6–10 [172, 170]	Satisfactory
$Q_{IVC}$	Inferior vena cava	55.0	40–60 [169, 170]	Acceptable
$Q_{SVC}$	Superior vena cava	36.0	15–25 [169, 170]	Satisfactory
$Q_{DscA}$	Descending aorta	55.0	50–70 [169]	Acceptable
$Q_{AbdA}$	Abdominal aorta	36.0	30–50 [169, 170]	Acceptable
$Q_{Cel}$	Celiac artery	13.0	7–12 [170]	Satisfactory
$Q_{Mes}$	Mesenteric circulation	13.0	10–15 [170]	Acceptable
$Q_{Spl}$	Spleen, pancreas, gastric	5.0	3–5 [162]	Acceptable
$Q_{Ven}^{Por}$	Portal vein	18.0	18–25 [170, 173]	Acceptable
$Q_{Hep}$	Hepatic circulation	8.0	6.7–11.7 [170]	Acceptable
$Q_{LLimb}$	Lower limb	10.0	8–10 [174, 170]	Acceptable

Table B.3: Hemodynamics for whole-body cardiovascular model: Simulated and reference mean flows in all compartments. The range of flow in the carotid arteries, the upper limb, and the lower limb is multiplied by two.



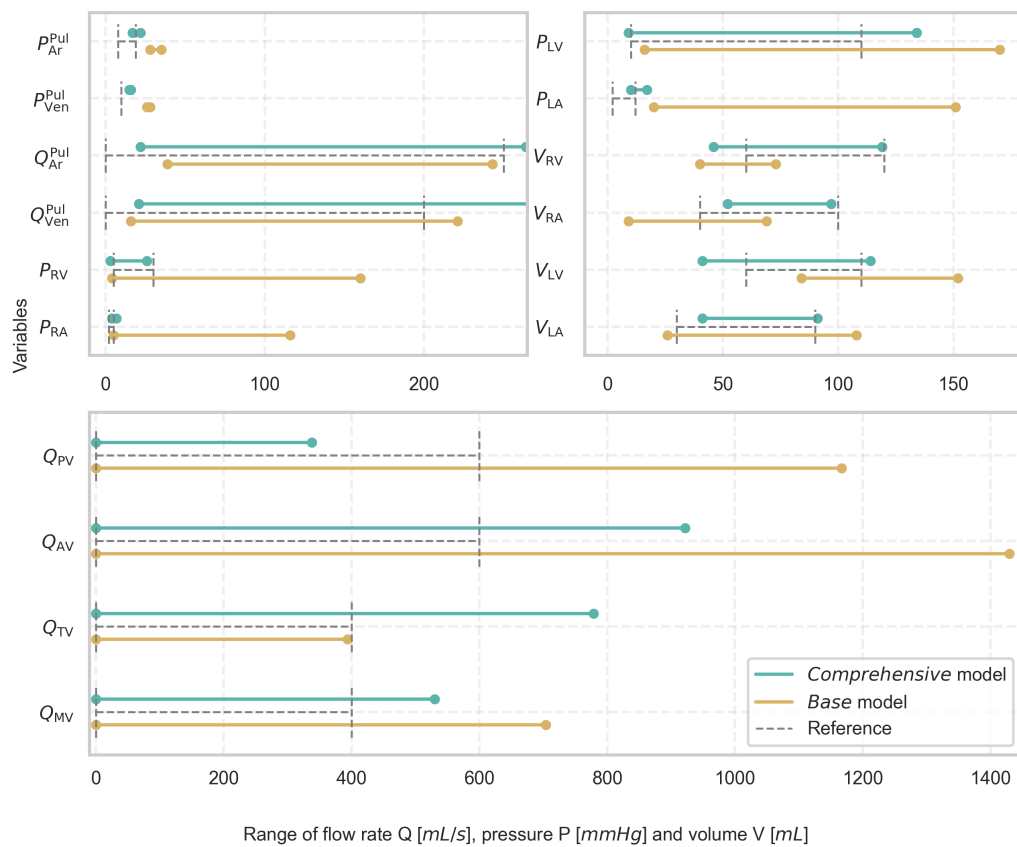


Figure B9: Range of variables with the reference (Table B.2,B.3) for both base and comprehensive model. The base model overshoots the reference range for most parameters, including  $Q_{PV}$ ,  $Q_{AV}$  and  $Q_{MV}$ .