Abstract

WAGNER, JESSICA LEIGH. A Parameter Analysis of a Physiologically Based Pharmacokinetic (PBPK) Model Describing the Movements of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin in the Mouse. (Under the direction of Alun Lloyd and Rory Conolly.)

Although its mechanism of action is not completely understood, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) is an unfortunate environmental contaminant. In order to understand the effects of TCDD better, we examine a physiologically based pharmacokinetic (PBPK) model that describes the movement of TCDD around the body of a mouse. The PBPK model examined in this work contains five tissue compartments: liver tissue, spleen tissue, adipose tissue, richly perfused tissue, and slowly perfused tissue. Each of these compartments has a subdivision to distinguish between the actual tissue and the blood contained in the tissue's vessels. The TCDD is introduced via the stomach and tracked as it travels around the body. The toxin can permeate each of the tissue types, but in the liver it also has the potential to bind to one of two proteins, CYP1A2 and the Ah receptor, or to undergo metabolism. An increased presence of the TCDD-AhR complex induces the liver tissue to produce more CYP1A2, which then affects the amount of TCDD molecules that are held in the liver tissue.

Unfortunately, the kinetic parameters that control the rates of binding and separation are very difficult to determine experimentally. In this work, the estimability of some of the parameters is analyzed through several different methods. Unfortunately, it is determined that it would require impossible experimental circumstances to have enough information to uniquely estimate every parameter in the model. The simulated model output is compared to an actual data set, but unfortunately they are drastically different. The analysis indicates that even if the data set were perfect, the values that were measured would only be able to provide enough information to estimate two out of twelve parameters. The model needs to be reparameterized or restructured in order for it to increase the scientific understanding of TCDD behavior. A dialog between the biologist who conducts experiments and the mathematician who conducts simulations can allow for a more efficient allocation of resources in future experimentation.

A Parameter Analysis of a Physiologically Based Pharmacokinetic (PBPK) Model Describing the Movements of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin in the Mouse

by
Jessica Leigh Wagner

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Biomathematics

Raleigh, North Carolina

2017

APPROVED BY:

Alun Lloyd Co-chair of Advisory Committee	Rory Conolly Co-chair of Advisory Committee

Rhyne Setzer

Dedication

In loving memory of my father, who always encouraged my scientific curiosity. I hope that you and Albert have become close friends on another metaphysical plane.

Biography

Jessica Wagner was born in 1987 in Pennsylvania, where she spent her childhood. She discovered a love for science somewhere in the midst of high school science fair projects. Following high school graduation, she enrolled in Gettysburg College to pursue a degree in Biochemistry and Molecular Biology. In her first year, a calculus class inspired her to reconsider her path. A lot of sweat and tears later, she became the first person to graduate from Gettysburg College with a dual degree in biochemistry and mathematics in 2009. Seeing the great potential for overlap of the two fields, Jessica sought a career path that would combine her specialties. Her search led her to enroll in the biomathematics graduate program at North Carolina State University.

Table of Contents

List of Tables

List of Figures

CHAPTER

1

INTRODUCTION

Dioxins are a class of organic chemicals which have a chemical structure consisting of four to eight chlorine atoms bound to varying locations on two benzene rings, which are connected to each other via oxygen atoms [39]. They are known to be highly toxic and cause a wide variety of health issues, but their exact molecular mechanism of action is still not completely understood. Unfortunately, this makes it much more difficult to develop any type of treatment. With most environmental contaminants, it is customary for experiments to be conducted on animals to determine the agent's toxicity, and then these values are scaled, typically by body mass, to assess the levels of human toxicity. This method, known as interspecies extrapolation, has obvious flaws, but it causes even more issues when the agent of interest is TCDD [23]. The tolerances observed in different species have been quite diverse [18], sometimes differing by 725 fold between humans and small mammals [23]. The body mass of an organism does not appear to correlate with the tolerance, considering that the lethal dose for guinea pigs appears to be approximately 1 μg/kg of body weight, compared to that of hamsters being approximately 1000 μg/kg [39]. By using PBPK modeling, we can simulate an organism's

internal environment using physiological parameter values specific to a species, which can provide more insight into the behaviors of a toxin.

A persistent organic pollutant (POP) is an organic chemical that remains stable in the environment for many years. Its resistance to degradation often creates a significant health risk to the human population, due to biomagnification in the food chain [33, 36]. In 2001, representatives from countries all over the world attended the Convention of Persistent Organic Pollutants in Stockholm, Sweden. The group realized the importance of minimizing these chemicals in the environment and compiled a list of twelve of the most dangerous POPs. Dioxins were on this list [33]. In the specific case of dioxins, their hydrophobic nature causes them to be absorbed readily into fat tissue. The combination of the storage in the fat tissue and the chemical stability causes the dioxins to remain in the body for a very long time [14]. The biological half-life of the most toxic dioxin molecule, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD), in humans has been estimated to be between seven and eleven years [39].

Dioxins are not produced intentionally or used in any mercantile capacities, but rather they are an unfortunate waste product of some commercial processes [1, 14, 39]. The incineration of chlorinated compounds, chlorine bleaching of pulp and paper products, chemical manufacturing of pesticides and herbicides all cause dioxins to be produced and released into the environment [14, 39]. Minute amounts of dioxin are found in nature due to some geographical events [39].

Humans are exposed to the dioxins almost entirely through the food chain [36]. All modern humans have some dioxins in their system, but the populations of more industrialized area typically have higher levels in their bodies [39]. The tendency of dioxins to be absorbed into fat tissues likely leads to increased levels of exposure in humans who eat animal products with high lipid content. Humans that maintain a vegan diet have been shown to have much lower dioxin levels that their omnivorous counterparts, adding more credence to this theory[36]. Unfortunately, the dioxins can also be transferred into breast milk. Nursing infants have been shown to be exposed to higher concentrations of dioxins than the rest of the population [36, 40]. This provides a very successful method for stored dioxin to leave the female adult, with one study estimating a 70% decrease in the TCDD levels of the mother over six months of lactation. As infants grow, they quickly increase their fat tissue volume, which helps to dilute

the TCDD concentration over time. By the time that they have reached puberty, the TCDD levels between breastfed and bottle-fed children are comparable[40]. The findings from these studies do not add insight into the molecular mechanics of the TCDD, but they do show that TCDD does not persist between generations.

While humans are not usually exposed to TCDD intentionally, there have been several events in history that have led to mass exposure. The largest of these is the infamous usage of the chemical known as Agent Orange during the Vietnam War. Agent Orange was designed to be an herbicide composed of two chemicals, 2,4-dichlorophenoxyacetic acid and 2,4,5-trichlorophenoxyacetic acid. The intention of the chemical was to defoliate areas that provided cover for enemy combatants, as well as to destroy crops [27, 32, 38, 40]. The production process of Agent Orange was imperfect, resulting in significant amounts of TCDD in the final product, although the exact amount is not possible to know [27, 32, 37, 38, 40]. Agent Orange was sprayed through various methods, including hand spraying, aircraft, and water vessels. This resulted in human exposure to TCDD not only via the food chain, but also directly through their skin and inhalation [32]. The exposure resulted in a large amount of birth defects [32]. Even many years after the war, Vietnam veterans experienced many different illnesses which were credited to TCDD, including several types of cancer [16]. The Vietnamese population living in the sprayed areas were found to have dioxin levels in their tissues much higher than other populations even after 20 years had passed [27]. This shows the persistence of this chemical and illustrates just how important it is to properly regulate it.

To provide an easier method to regulate dioxins, the concept of toxic equivalency (TEQ) was formulated. A TEQ value describes the toxicity of a sample as if it were only made of the most toxic dioxin, TCDD. TCDD is given a toxic equivalency factor (TEF) value of 1. All other dioxins are given TEF calculated by multiplying each dioxin's concentration by its respective TEF and adding all the resulting values together [36, 39, 40].

1.1 Biology of Dioxin

TCDD is a symmetric molecule (Fig 1.1), which makes it nonpolar and insoluble in water [2]. Thus, one would expect that it would gather in the adipose tissues almost exclusively.

However, evidence has shown that the liver can have a higher concentration compared to the adipose [5]. This is due to the interactions between liver proteins and the TCDD.

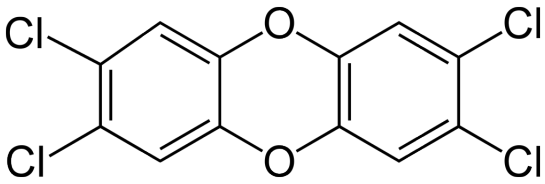


Figure 1.1 Molecular structure of TCDD.

When TCDD is introduced to a tissue, it must first travel across the membrane into the cell. The diffusion process is not entirely understood, but due to the large size of TCDD, it's unlikely to be by passive diffusion alone [14]. Once TCDD reaches the cytoplasm, it binds to the aryl hydrocarbon receptor (AhR). AhR is an important transcription factor in the cell, which has been found to affect cell cycle regulation, immune responses, and fatty acid metabolism. Also, the AhR assists in the metabolism and removal of foreign molecules [14].

TCDD binding to AhR triggers the beginning of a series of molecular events in the cell. The TCDD-AhR complex moves from the cytoplasm into the nucleus, where it impacts the activation of several genes. Unfortunately, this activity likely induces the carcinogenic effects of TCDD [19, 20]. While much of the details are still unknown, two of the most notable of these genes are known as CYP1A1 and CYP1A2, which are cytochrome P450 isozymes [46]. This group of enzymes is crucial to the proper metabolism and excretion of many different drugs, and is the root of many drug interactions, insensitivities, and dosage requirements [31]. TCDD binding to these two cytochrome P450 compounds (and probably others as well), causes more TCDD to be drawn into the liver and may be the reason that TCDD has been observed in higher levels in the liver tissues than in the adipose tissues, despite its very strong lipophilic nature [4, 5].

1.2 Previous modeling work on TCDD

For almost 30 years, scientists have been using PBPK models to describe the movement of dioxin. Table 1.1 highlights some of the differences between the models over time. The models differ in what proteins and what tissue types are included, as well as in how they handle the diffusion process of TCDD. However, they share the common goal of providing more insight into the molecular movements of dioxin around the body.

In 1988, an early model was designed by Leung with five compartments: blood, liver, adipose, slowly perfused, and richly perfused tissues (Figure 1.2)[25]. The only input comes from an injection into the peritoneal cavity, and the only output is via metabolism in the liver. TCDD in the blood compartment is found in two forms, only one of which could diffuse into the compartments. In the liver compartment, TCDD has the potential to bind to one of two different binding sites. In 1988, the knowledge of the molecular mechanics was limited, so these binding sites were vaguely described as a high affinity/low capacity receptor and a low affinity/high capacity receptor. Because the model only has a single blood compartment, these compartments are 'flow-limited'. In other words, the amount of toxin that gets into the tissues depends entirely on the rate of the blood flow through the tissues, because it is absorbed so quickly. Leung's original model was designed for mice [25], but they expanded the model to apply to rats in a second publication [26].

For many chemicals, the assumption that the compartments are 'flow-limited' is valid. However, in the case of TCDD, the molecule is relatively large and lipophilic, making it seem inappropriate in this situation. Andersen revised Leung's model, dividing each compartment into the actual tissue and the blood contained within the blood vessels of that tissue [6]. In doing so, the compartments were set up to be 'diffusion-limited', which assumes that instead of the blood flow, the rate of diffusion controls the concentration of the toxin in that compartment. While adding more biological realism to the mathematical model, this change also introduces more parameters, most of which are difficult to measure accurately with experimentation. One of these types of parameters is the permeability cross product, which considers the surface area in each compartment and will be discussed at length later. This model also focused on parameter values that were based on rat biology, not that of a mouse.

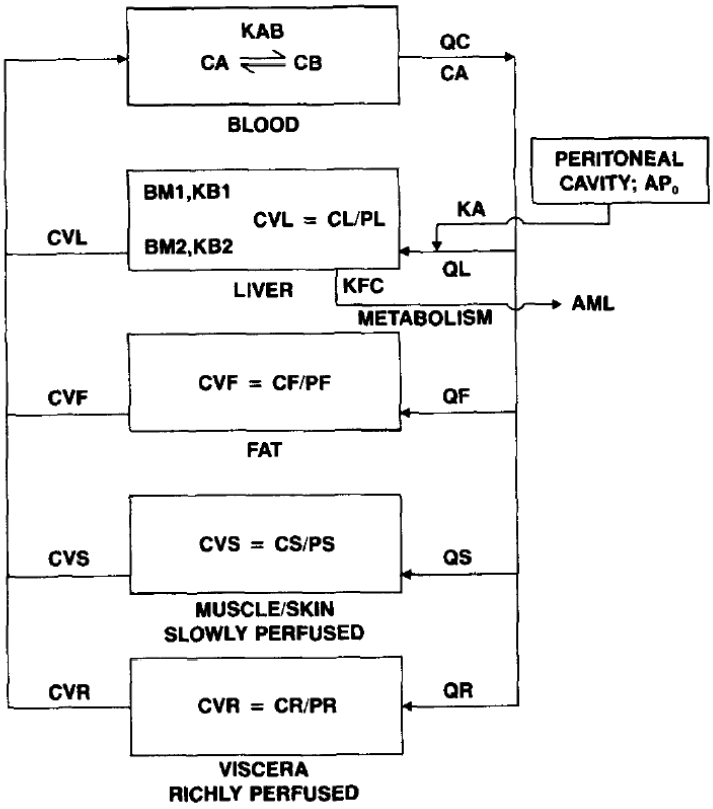


Figure 1.2 Visual representation of an early TCDD model. Reproduced from [25].

By the time Andersen's model was being formulated in 1993, more of the biology was understood. Once the TCDD gets into the liver tissue, it can potentially bind to the aryl hydrocarbon receptor (AhR). After this complex has formed, it binds to specific sections of DNA, which induces the cells to produce several different proteins. One of these proteins is CYP1A2, a cytochrome enzyme which can bind to the free TCDD in the liver. The TCDD cannot be metabolized while it is bound to either of these proteins [6].

A few years later, Wang modified the TCDD model further [47]. Fecal and urinary elimination were included, as well as the impact of the lymphatic system. Wang's work also divided the entire organism (which was a rat in their study) differently than the other models by considering the kidneys, lungs, spleen, and skin as separate compartments [47]. This study obtained its concentration data by using a radioactively tagged TCDD molecule. Unfortunately this leads to only one overall measurement, with no way to distinguish between free TCDD and bound TCDD concentrations.

The model and analysis that follow contain elements from each of these publications and expand on them. The original compartments that Leung's group described are still present, with the addition of the spleen, which has been examined separately as it was in Wang's work. The sub-compartments for tissue blood that Andersen designed have proven to be a valuable contribution and are included. Here we also examine the effects of the induction of CYP1A2, one of the TCDD binding proteins in the liver and scale it for a mouse. The mathematical predictions are compared to data points that have been obtained via a different experimental protocol involving mass spectrography instead of radiation level readings. The integrity of the model is investigated using sensitivity analysis and identifiability techniques.

Table 1.1 Information from older PBPK models examining TCDD.

Source	*Species*[a]	*Input(s)*	*Output(s)*	*Compartments*	*Diffusion*	*Hepatic Variables*
Leung, 1988 [25]	C57BL/6J mice DBA/2J mice	peritoneal cavity	metabolism (via liver)	blood liver fat slowly perfused richly perfused	flow limited	• a generalized, non-inducible, high affinity, low capacity, cytosolic receptor • a generalized, inducible, low affinity, high capacity, microsomal receptor
Leung, 1990a [26]	Sprague-Dawley rats	simulated oral gavage	metabolism (via liver)	blood liver fat slowly perfused richly perfused	flow limited	• a generalized, non-inducible, high affinity, low capacity, cytosolic receptor • a generalized, inducible, low affinity, high capacity, microsomal receptor
Leung, 1990b [24]	C57BL/6J mice (F)	peritoneal cavity	metabolism[b] (liver)	blood liver fat slowly perfused richly perfused skin	flow limited	• a generalized, non-inducible, high affinity, low capacity, cytosolic receptor • a generalized, inducible, low affinity, high capacity, microsomal receptor

[a] Gender unknown unless specified
[b] The metabolism process of this model includes an intermediate metabolite.

Table 1.1 Continued.

Source	*Species*[a]	*Input(s)*	*Output(s)*	*Compartments*	*Diffusion*	*Hepatic Variables*
Andersen, 1993 [6]	Wistar rats (F)	subcutaneous injection	metabolism (via liver)	blood liver[b] fat[b] slowly perfused[b] richly perfused[b]	diffusion limited	• CYP1A1 • CYP1A2 • AhR • DNA
Andersen, 1997a [7]	Sprague-Dawley rats (F)	subcutaneous injection	metabolism (via liver)	blood fat[b] slowly perfused[b] richly perfused[b] liver[c]	diffusion limited	• a generic CYP protein • a generic receptor protein • DNA
Andersen, 1997b [5]	Sprague-Dawley rats	subcutaneous injection	metabolism (via liver)	blood fat[b] slowly perfused[b] richly perfused[b] liver[c]	flow limited	• CYP1A1 • CYP1A2 • AhR • 2 DNA binding sites

[a] Gender unknown unless specified
[b] This compartment contains a subcompartment which represents the blood volume contained in the tissue's blood vessels.
[c] This compartment contains five subcompartments, which represent different geographical regions of the liver with varied distances from the main blood source.

Table 1.1 Continued.

Source	*Species*[a]	*Input(s)*	*Output(s)*	*Compartments*	*Diffusion*	*Hepatic Variables*
Wang, 1997 [47]	Sprague-Dawley rats (F)	GI tract	metabolism (via liver) urine (via kidney)	blood lung spleen kidney[b] skin[b] fat[b] liver[b] rest[b]	varies by tissue type	• CYP1A2 • AhR • DNA • nonspecific proteins • lipid binding
Wang, 2000 [48]	Sprague-Dawley rats (F & M) Wistar rats (F) C57BL/6J mice (F)	GI tract	metabolism (via liver) urine (via kidney)	blood lung spleen kidney[b] skin[b] fat[b] liver[b] rest[b]	varies by tissue type	• CYP1A2 • AhR • DNA • nonspecific binding • lipid binding

[a] Gender unknown unless specified
[b] This compartment contains a subcompartment which represents the blood volume contained in the tissue's blood vessels.

CHAPTER

2

MODEL

2.1 Structure

The main structure of the model divides the mouse into five tissue compartments: richly perfused tissues, slowly perfused tissues, adipose tissue, liver tissue, and spleen tissue, each with its own sub-compartment representative of the blood contained inside that tissue's blood vessels. TCDD molecules have the ability to pass from the tissue blood to the tissue compartment itself and back again, but the blood flow of the arteries and veins only directly interacts with these tissue blood compartments. To mimic the natural blood flow through the portal vein of the liver, the spleen's output flows directly into the liver compartment instead of into the venous vessels. It is assumed that the TCDD being absorbed from the stomach enters into the liver blood before it travels around the body for the same reason (Figure 2.1).

The liver compartment is unique in its complexity. Like the other compartments, TCDD can travel between the tissue blood and the tissue. But once inside the tissue, the TCDD molecule can potentially bind to one of two proteins, CYP1A2 or AhR. The total concentration of AhR does not change over the course of the simulation, but AhR can exist in one of two forms, free and unbound or bound inside the TCDD-AhR complex. By contrast, the concentration of

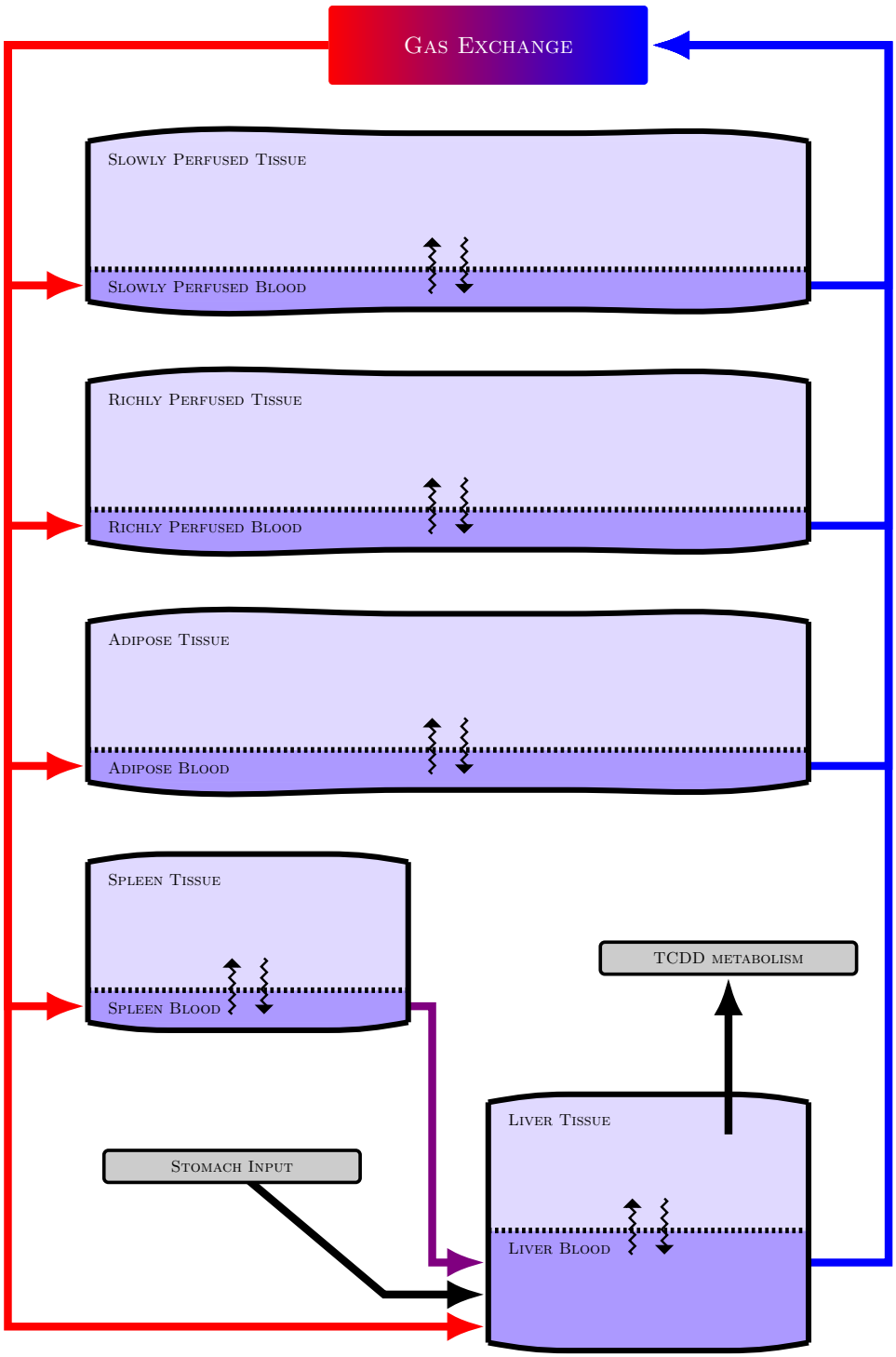


Figure 2.1 Organism level visualization of the model. Each arrow represents potential movement of TCDD across between different sections of the body. Red arrows represent freshly oxygenated blood coming from the lungs, and blue arrows represent venous blood leaving each compartment. The purple arrow from the spleen to the liver mimics the hepatic portal artery. Since lungs are not explicitly defined as a compartment and are instead lumped into the richly perfused compartment, the 'Gas Exchange' box is included to assist in the visualization of the model.

CYP1A2 is dynamic. Although a small amount of the enzyme is initially present, the presence of the TCDD-AhR complex stimulates the accumulation of more CYP1A2 enzyme in the tissue. Due to their large size, CYP1A2 and AhR do not travel across the cellular membrane into the bloodstream. The CYP1A2 enzyme has a first order degradation rate, so its concentration drops in the absence of the TCDD-AhR complex. The TCDD can be metabolized only in the liver, and only if the molecule is not bound to either enzyme (Figure 2.2).

2.2 Parameters

This model is based on a mostly closed system, with the only input being through the stomach and the only output through metabolic clearance in the liver compartment. Each of the compartments is assumed to be well mixed, with the tissue blood compartment being representative of blood contained in the tissue's capillaries and the tissue compartment being representative of the cells in the tissue. Due to the differences in composition and function of the different organs, most of these parameters have unique values for each of the tissue compartments.

2.2.1 Tissue Volumes/Blood Flow Rates

The fractional breakdown of the mouse's organs was found in literature. Using these values and the total body weight of the animal, it is possible to estimate the volume of each tissue type. Note that this also involves an assumption that the approximate density of the mouse tissue is 1 g/cm^3, which is required in order to convert the mass to a volume. This particular model assumes that the body weight of the animal stays constant over the course of the experiment.

Blood flow rates are treated in the same fashion as the tissue volumes, because the total blood flow through the body is divided into similar sections. When the body weight is constant, the secondary calculations required to determine these values only need to be calculated once and retained, but they would need to be recalculated on each time step if the body weight fluctuated over time. These equations and all parameter values can be seen in the appendix.

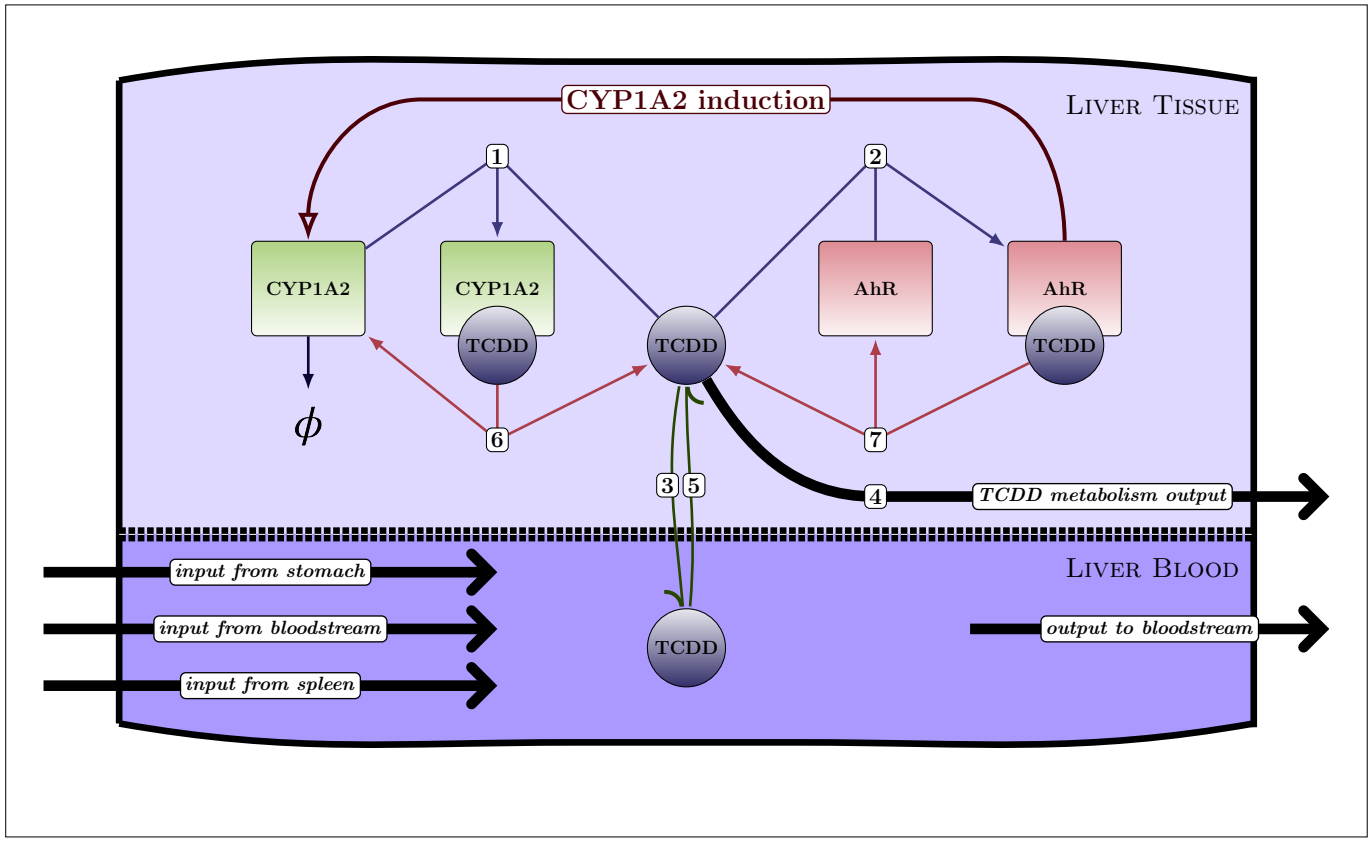


Figure 2.2 Liver level visualization of the model. As in every other compartment, the TCDD can diffuse between the tissue and its complementary blood compartment. However once a free TCDD molecule has diffused into the liver tissue, it has several potential fates, each of which are indicated in the figure: (1) Binding to a free AhR to form the TCDD-AhR complex; (2) Binding to a free CYP1A2 to form the TCDD-CYP complex; (3) Diffusing into liver blood compartment; (4) Being broken down and metabolized. Free TCDD enter the liver compartment via three methods: (5) Diffusing into the liver tissue compartment; (6) One of the resulting molecules when the TCDD-CYP complex separates; (7) One of the resulting molecules when the TCDD-AhR complex separates. Also note that the presence of the TCDD-AhR complex induces the production of more CYP1A2 enzyme which is simultaneously degrading, designated by ϕ.

2.2.2 *Partition Coefficients*

Partition coefficients are required to describe how the chemical agent in question interacts with each type of tissue. In this particular model, we assume that these values remain constant over time, but this is not true for every PBPK model. A partition coefficient represents the ratio between the tissue concentration and the concentration of the venous output from the compartment. There are many different methods to determine this value experimentally, but ideally the number would be measured when the system was at equilibrium. Tissue types that interact with the chemical agent via binding or metabolizing complicate this measurement because they prevent this equilibrium from occurring. Some literature also reverses the order of the ratio, so it is prudent to mention that the partition coefficients used in this model are tissue to blood partition coefficients [15]. Unlike the permeability coefficients, partition coefficients are not dependent on the surface area of blood vessels in the tissues or the rate of blood flow.

2.2.3 *Permeability Coefficients*

One of the major differences between this model and the one proposed by Leung deals with the way that diffusion into each tissue is handled. While the partition coefficients measure the overall tendency of a tissue to absorb TCDD, the permeability coefficients consider how quickly the blood flows through the tissues and the density of the blood vessels. The rate of absorption into the cells from the capillaries is dependent on the surface area available to interact with the tissue. This brings in the need for the permeability coefficient, which is also sometimes referred to as the diffusional tissue clearance or the permeability cross product in other literature. A new scaling constant parameter, J, is introduced here. Each individual blood flow must be scaled by J to calculate the permeability coefficient. For each tissue, the appropriate permeability coefficient thus equals JQ_i, where Q_i is the tissue specific blood flow rate. While the blood flow rate changes from one compartment to the next, the value for J remains equal for the entire system. The idea that J does not vary across tissue types assumes that subtle differences between the blood vessels of the organs will be accounted for in their different blood flow rates. While this may seem like a strong assumption, the alternative is

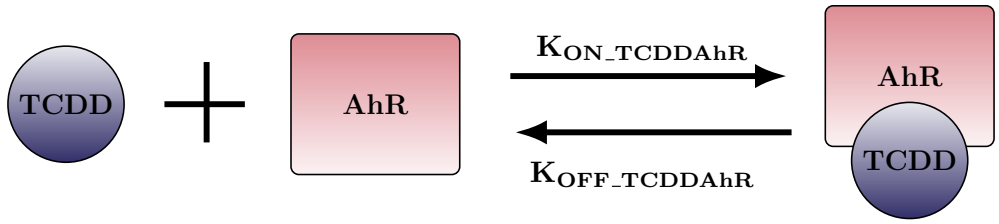


Figure 2.3 Visualization of the rate constants involved with the association process of the TCDD-AhR complex. The formation of the TCDD-CYP complex is analogous.

introducing five parameters, for which literature values are scarce or nonexistent, instead of only one.

2.2.4 Kinetic Parameters

While much of the movement in the model is due to the TCDD diffusing from one compartment to another, the liver compartment has more complex behavior (Figure 2.2). TCDD travels into the liver blood and tissue itself in the same manner as in the other compartments. However, once the free TCDD is in the liver tissue, it has several potential fates. One possibility is that the TCDD can bind to either the AhR or the CYP1A2 protein. The kinetic parameters that define the rates of these reactions, as well as their reverse counterparts, are difficult to measure accurately even experimentally, and there is some disagreement in the published literature values. To simplify the reaction rates, here we assume that the association and dissociation of the TCDD-AhR and TCDD-CYP complexes are driven by the law of mass action. This principle states that the rate of a chemical reaction is directly proportional to the product of the concentration of the reactants [30]. The rate of the reaction is calculated by using the concentrations and a proportionality constant (known as a rate constant). Because the complexes are simultaneously forming and separating, there are two rate constants involved in the calculations for each of the proteins.

The rate at which the complex forms can thus be described as $K_{ON\_TCDDAhR}C_{TCDD}C_{AhR}$, where C_i represents concentration of i and K represents the relevant rate constant. Similarly, the rate of dissociation is $K_{OFF\_TCDDAhR}C_{TCDD\_AhR}$ (Figure 2.3). The setup for the association and dissociation of the TCDD-CYP complex is analogous.

Other potential fates for the free TCDD in the liver tissue include traveling back across the membrane to the liver blood or undergoing metabolism. The chemical reactions involved in the metabolism are not explored here, but we assume that the TCDD can not be metabolized if it is bound to either CYP1A2 or AhR.

2.3 Equations

Each differential equation is a summation of inputs and outputs interacting with the appropriate compartment. The amount of TCDD in the compartment, designated by A_i, is measured in nanomoles. For clarity in their development, the equations in this section include concentrations, designated by C_i. In practice, each of these concentrations needs to be converted to an amount before computation to maintain the proper units. The equations listed in the appendix reflect this additional step.

The equations for the slowly perfused tissues, richly perfused tissues, adipose, and spleen compartments are of a similar form, so let us consider a generic tissue compartment, identified by the subscript 'T', and a generic tissue blood compartment, identified by the subscript 'TB'. There is an input from the tissue blood compartment, and a corresponding output to the tissue blood compartment. To simplify the system, we use the J term as a scaling factor to describe the permeability of the tissue when multiplied by the appropriate tissue's blood flow rate. When the TCDD leaves the tissue compartment, the partition coefficient comes into play to essentially describe how likely the toxin is to leave that tissue. Since the toxin is exiting the tissue and the partition coefficients in the model are defined as tissue/blood, the partition coefficient appears in the denominator of the term (Eqn 2.1).

$$\frac{dA_{\mathrm{T}}}{dt} = JQ_{\mathrm{T}}C_{\mathrm{TB}} - \frac{JQ_{\mathrm{T}}C_{\mathrm{T}}}{Pa_{\mathrm{T}}} \tag{2.1}$$

These two terms appear again in the generic tissue blood equation, with opposing signs due to the reverse directionality. The equation also includes a positive term that describes the input of the toxin from the arterial blood and a negative term that describes the concentration in the blood flow leaving the capillaries and entering the venous blood. In the case of the spleen, this blood enters the liver directly to mimic the hepatic artery blood flow (Eqn 2.2).

$$\frac{dA_{\mathrm{TB}}}{dt} = \frac{JQ_{\mathrm{T}}C_{\mathrm{T}}}{Pa_{\mathrm{T}}} - JQ_{\mathrm{T}}C_{\mathrm{TB}} + Q_{\mathrm{T}}C_{\mathrm{AB}} - Q_{\mathrm{T}}C_{\mathrm{TB}} \tag{2.2}$$

The liver blood equation is similar to the generic tissue blood equation, but it also includes terms for the TCDD input from the stomach and the input from the spleen. The liver tissue equation also contains all of the terms in the generic equation, but it requires the addition of several more terms due to the chemical reactions taking place in that compartment.

Most of the chemical reactions that take place in the liver are assumed to follow the law of mass action. This law states that the equation rates are based on the concentrations of the products of the reaction being multiplied together and then multiplied by a rate constant. Once rates for the reactions have been established, the differential equations become mass balance equations (Eqn 2.3). In these equations, the amount of free CYP1A2 and free AhR are represented by A_{CYP} and A_{AhR}, respectively. When the TCDD is bound to CYP1A2, it is referred to as the TCDD-CYP complex and appears in the equations as $A_{TCDDCYP}$. In similar fashion, AhR proteins which have a bound TCDD ligand are referred to as the TCDD-AhR complex, and the amount of the complex is $A_{TCDDAhR}$. Any free, unbound TCDD is accounted for in the C_{LT} term, which represents the concentration of TCDD in the liver tissue.

$$\begin{aligned}
\frac{dA_{\mathrm{TCDDAhR}}}{dt} &= K_{\mathrm{ON\ TCDDAhR}}C_{\mathrm{CYP}}C_{\mathrm{LT}} - K_{\mathrm{OFF\ TCDDAhR}}C_{\mathrm{TCDDAhR}} \\
\frac{dA_{\mathrm{AhR}}}{dt} &= -K_{\mathrm{ON\ TCDDAhR}}C_{\mathrm{AhR}}C_{\mathrm{LT}} + K_{\mathrm{OFF\ TCDDAhR}}C_{\mathrm{TCDDAhR}} \\
\frac{dA_{\mathrm{TCDDCYP}}}{dt} &= K_{\mathrm{ON\ TCDDCYP}}C_{\mathrm{CYP}}C_{\mathrm{LT}} - K_{\mathrm{OFF\ TCDDCYP}}C_{\mathrm{TCDDCYP}}
\end{aligned} \tag{2.3}$$

The exception to this assumption is the induction of the CYP1A2 enzyme production, which is stimulated by the presence of the TCDD-AhR complex (Eqn 2.4).

$$\begin{aligned}
\frac{dA_{\mathrm{CYP}}}{dt} = {} & K_{\mathrm{CYP\ BASE}} - K_{\mathrm{CYP\ DEG}}C_{\mathrm{CYP}} \cdots \\
& - K_{\mathrm{ON\ TCDDCYP}}C_{\mathrm{CYP}}C_{\mathrm{LT}} + K_{\mathrm{OFF\ TCDDCYP}}C_{\mathrm{TCDDCYP}} \cdots \\
& + (K_{\mathrm{CYP\ MAX}} - K_{\mathrm{CYP\ BASE}}) \left(\frac{C^N_{\mathrm{TCDDAhR}}}{K^N_{\mathrm{CYP\ D}} + C^N_{\mathrm{TCDDAhR}}} \right)
\end{aligned} \tag{2.4}$$

The equation for the amount of CYP1A2 protein in the liver includes a term for basal production, as well as a first order degradation term. This is required to balance the production and prevent an infinite production of the enzyme. The concentration of CYP1A2 is lowered by molecules binding to free TCDD and simultaneously increased by the TCDD-CYP complexes dissociating. The final term in Equation 2.4 represents the induction of the enzyme in the presence of the TCDD-AhR complex, used by Andersen in 1997 to describe this induction [5]. It takes the shape of a Hill function, with the N, also known as the Hill coefficient, changing the shape of the induction curve. The $K_{\text{CYP D}}$ parameter is representative of the microscopic dissociation.

Given the numerous additions and subtractions from the free TCDD concentration in the liver tissue observed in the above equations, the liver tissue equation requires many terms. One more term is necessary to describe the metabolism of TCDD. We assume that the metabolism is a first order reaction dependent on the concentration of free TCDD in the liver tissue (Eqn 2.5). This also assumes that only free, unbound TCDD is available to be metabolized. This metabolic term is also tracked as its own equation in the system to track the total amount of TCDD that has been metabolized.

$$\begin{aligned}
\frac{dA_{\text{LT}}}{dt} &= JQ_{\text{L}}C_{\text{LB}} - \frac{JQ_{\text{L}}C_{\text{LT}}}{Pa_{\text{L}}} - K_{\text{MET}}C_{\text{LT}} + \dots \\
&\quad K_{\text{OFF TCDDCYP}}C_{\text{TCDDCYP}} - K_{\text{ON TCDDCYP}}C_{\text{CYP}}C_{\text{LT}} + \dots \\
&\quad K_{\text{OFF TCDDAhR}}C_{\text{TCDDAhR}} - K_{\text{ON TCDDAhR}}C_{\text{AhR}}C_{\text{LT}} \\
\frac{dA_{\text{MET}}}{dt} &= K_{\text{MET}}C_{\text{LT}}
\end{aligned} \tag{2.5}$$

Now that all the equations for each compartment are defined, the circulatory system behavior must be considered. The venous blood acquires the blood flowing out of each of the blood tissue compartments, and then passes it to the arterial blood. The arterial blood pumps into each of these compartments on the next time step (Eqn 2.6). Note that due to the blood flow pattern, the rate of the blood flow exiting the liver is the combined rate of the input rates for both the spleen and the liver. This allows for the system to maintain a blood volume equilibrium in each compartment.

$$\begin{aligned}\frac{dA_{VB}}{dt} &= Q_S C_{SB} + Q_R C_{RB} + Q_F C_{FB} + (Q_L + Q_{SP}) C_{LB} - Q_{TOT} C_{VB} \\ \frac{dA_{AB}}{dt} &= Q_{TOT} C_{VB} - Q_S C_{AB} - Q_R C_{AB} - Q_F C_{AB} - Q_{SP} C_{AB} - Q_L C_{AB}\end{aligned} \tag{2.6}$$

In order to make any comparisons to real TCDD data values, an additional calculation is required. There is no evidence to suggest that the laboratory tests can distinguish free TCDD from the TCDD bound to an enzyme. Therefore the liver concentration value that is observed in experiments is assumed to be a sum of the concentrations of free TCDD, the TCDD-AhR complex, and the TCDD-CYP complex (Eqn 2.7).

$$C_{TOT\ TCDD}(t) = C_{LT}(t) + C_{TCDDAhR}(t) + C_{TCDDCYP}(t) \tag{2.7}$$

CHAPTER

3

MATHEMATICAL ANALYSIS

A common problem with biologically based mathematical models is that they tend to be complex. This complexity adds more parameters and more potential uncertainty into the model's predictions. Modelers often find themselves unable to determine unique values for every parameter, and the situation is made even worse in cases where data is sparse. Even in a situation where a laboratory has the means for extensive data collection, some values cannot be accurately measured experimentally, such as the concentration of an intermediate metabolite. The concepts of identifiability and estimability are very similar, but not quite equivalent. Identifiability aims to answer the question 'Can different values of model parameters lead to the same input-output behavior for the model?'. Comparatively, estimability asks 'Can all parameter values be estimated uniquely from the available experimental data?' [29]. Because the question of identifiability involves a more analytical approach, it becomes substantially more difficult in large models with complex equations. Almost every PBPK model contains too many parameters and state variables to make any analysis of structural identifiability practical, especially a global one. Even in cases where the parameters are shown to be identifiable, there is a possibility that their values cannot be estimated accurately given the data that is available.

Consider a general definition of a nonlinear, ordinary differential equation system described in Equation 3.1.

$$\begin{aligned} \frac{d\vec{x}(t)}{dt} &= f(\vec{x}(t), t, \vec{\theta}) \\ \vec{y} &= g(\vec{x}(t), t, \vec{\theta}) \end{aligned} \tag{3.1}$$

The model contains m state variables, represented by $\vec{x}(t)$, which vary with time, t. Besides time, the model requires an input of a vector of n parameters, $\vec{\theta}$. The model's predictions are described as $\vec{y}$, which is also a vector of length m. This system will be referenced throughout the techniques described in this section.

3.1 Structural Identifiability

A mathematical model is considered globally identifiable if all of its parameter values can be derived uniquely using a hypothetically perfect data set. Mathematically, this can be defined by Equation 3.2 holding true for any parameter vectors $\vec{\theta}_1$ and $\vec{\theta}_2$ in the appropriate parameter space.

$$f(\vec{x}, \vec{t}, \vec{\theta}_1) = f(\vec{x}, \vec{t}, \vec{\theta}_2) \iff \vec{\theta}_1 = \vec{\theta}_2 \tag{3.2}$$

It is possible for this relationship to only be true for some subset of the total parameter space, in which case the model is only locally identifiable [11, 29]. Especially for models that contain more than ten states and parameters, determining identifiability can prove to be quite complicated[34]. Many of the methods that have been developed rely on a linearized system, so the nonlinear nature of this model adds another dimension of difficulty [29]. If the analysis were successful, there is no guarantee that the identifiable parameters would even be estimable. However, there are reliable methods to implement that examine the practical identifiability of the model, also called the estimability, which we focus on in this work.

3.2 Estimability

Estimability analysis aims to determine if it is possible to uniquely estimate each parameter value of the model from the data available. If a parameter is found to be inestimable, there are several potential causes. First, the model may not be sensitive to that parameter. Secondly, there may be correlations in the model between two or more parameters. Such correlations can cause any changes made by a modified parameter value to be counteracted by a second parameter. This prevents any new information from being gleaned. Finally, a lack of sufficient data can cause parameter(s) to be inestimable. Any of these situations disrupts the ability to understand the relationships between the inputs and outputs of the model [29]. In this section, several different techniques were explored in an attempt to determine if the parameters in this TCDD model are estimable. Each of these methods attempts to estimate the impact that a small change in some parameter θ_j has on the ith state variable prediction, y_i. This can be expressed mathematically $s_{i,j} = \frac{\partial g_i}{\partial \theta_j}$.

There is a possibility for a model to be only partially estimable. The usual method of action is to remove inestimable parameters and set them at their literature values. The literature values that describe the general morphology of the organism, such as relative tissue volumes and body weight, are reliable. In certain situations, the structure of the model can be reconfigured in a way that allows for some parameters to be removed from the system entirely.

3.2.1 Qualitative Methods

Three methods are discussed here which estimate each $s_{i,j} = \frac{\partial g_i}{\partial \theta_j}$ over time in a qualitative manner. The results are viewed graphically in order to compare relative shapes and relative peak heights of each function.

3.2.1.1 Finite Difference

The simplest and easiest method to implement is the finite difference approximation. This method is based on the mathematical definition of a derivative (Eqn 3.3). For a general function, $p(x)$, the derivative is obtained by moving h units down the curve and comparing the two points. This stepsize h is decreased until it approaches zero [42].

$$\frac{dp(x)}{dt} = \lim_{h \to 0} \frac{p(x+h) - p(x)}{h} \tag{3.3}$$

Consider instead a more complicated differential equation system, such as the one described in Equation 3.1. The system is solved first with the input parameters at set values, and then again with a perturbation of size h added to a single parameter of interest identified as θ_j. Then the sensitivity information for θ_j can be estimated by the forward difference approximation (Eqn 3.4) [10, 28, 43, 44].

$$\frac{d\vec{x}(t)}{d\theta_j} = \frac{f(\vec{x}(t), t, \theta_j + h) - f(\vec{x}(t), t, \theta_j)}{h} \tag{3.4}$$

Modifications to this method also include the backward difference approximation (Eqn 3.5), where the perturbation is subtracted, and the central difference approximation (Eqn 3.6), which combines the two methods in an attempt to avoid bias [43].

$$\frac{d\vec{x}(t)}{d\theta_j} = \frac{f(\vec{x}(t), t, \theta_j) - f(\vec{x}(t), t, \theta_j - h)}{h} \tag{3.5}$$

$$\frac{d\vec{x}(t)}{d\theta_j} = \frac{f(\vec{x}(t), t, \theta_j + h) - f(\vec{x}(t), t, \theta_j - h)}{2h} \tag{3.6}$$

The finite difference method is the least code intensive, but it often ends up being slower and less accurate than other methods. At very small perturbation sizes, a significant amount of computational error occurs from the uncertainty originated from the subtraction operation in the numerator, but if the perturbation is not small enough, it will also cause inaccuracies [44].

3.2.1.2 Complex Step Method

As mentioned above, the biggest downside to using the finite difference approximation method is the stepsize limitation. As the value decreases, floating point errors in the computer calculations become a serious problem, to the point that the estimation is no longer helpful. The complex step method is a viable alternative.

Consider a complex number, $z = x + \mathrm{i}y$, where x and y are real numbers. We input this value into a function with separate real and imaginary pieces (Eqn 3.7).

$$f(z) = u(x + \mathrm{i}y) + \mathrm{i}v(x + \mathrm{i}y) \tag{3.7}$$

In this function, u is a function describing the real portion and v is a function describing the imaginary piece. The developer of this method, Martins, et al, showed that this function is analytic and therefore differentiable in the complex plane [28]. Because of this, the Cauchy-Riemann equations are applicable, which means that equation 3.8 holds true.

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y} \\ \frac{\partial u}{\partial y} &= -\frac{\partial v}{\partial x} \end{aligned} \tag{3.8}$$

Using the definition of a derivative from earlier and combining it with this yields Eqn 3.9.

$$\frac{\partial u}{\partial x} = \lim_{h \to 0} \frac{v(x + \mathrm{i}(y + h)) - v(x + \mathrm{i}y)}{h} \tag{3.9}$$

Since we are dealing with PBPK models, we assume that the equations that we want to analyze are in the real plane. As a result, $y = 0$, $u(x) = f(x)$, and $v(x) = 0$ are all true statements, allowing us to simplify to Eqn 3.10.

$$\frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{\mathrm{Im}[f(x + \mathrm{i}h)]}{h} \tag{3.10}$$

Provided h is a small stepsize, we ultimately arrive at the complex step derivative approximation (Eqn 3.11).

$$\frac{\partial f}{\partial x} \approx \frac{\mathrm{Im}[f(x + \mathrm{i}h)]}{h} \tag{3.11}$$

This concept is applied to the general m-dimensional system described in Equation 3.1 with the notation in Equation 3.12.

$$\frac{d\vec{x}(t)}{d\theta_j} \approx \frac{\mathrm{Im}[f(\vec{x}(t), t, \theta_j + ih)]}{h} \tag{3.12}$$

Thus, if the parameter θ_j is perturbed on the imaginary plane rather than the real one, complex mathematics allows for a more precise solution. Note that this equation no longer contains a subtraction operation. Therefore, this method removes the error generated by subtractive cancellation typically generated by the finite difference approximation method. This also means that h can be taken as a much smaller value to increase the accuracy of the sensitivity analysis without a problem. The complex step method, while an improvement over the finite difference approximation, still has its own set of flaws. This method can be difficult to implement in cases where the programming language cannot process complex values, and it requires a lot of repetitive calculations.

3.2.1.3 Sensitivity Equations

The most complicated of the visual sensitivity methods discussed here involves the derivation of many differential equations. While the sensitivity equation method is difficult to implement initially, it provides explicit equations for the sensitivity analysis, preventing the introduction of more uncertainty through the arbitrary choice of the stepsize variable. This method is well documented [9, 13, 17, 41, 43, 44].

Considering a general model as described in Eqn 3.1, the goal of the sensitivity analysis is to examine the effect of different input parameters on the model's state variable output. As such, we take the derivative with respect to some parameter θ_j of both sides of the equation (Eqn 3.13). In this model, none of the parameters are dependent on time, which makes $\frac{\partial \theta_j}{\partial \theta_j} = 1$. After observing that $\frac{\partial t}{\partial \theta_j} = 0$, the equation is simplified. Also, the order of differentiation on the left hand side is modified.

$$
\begin{aligned}
\frac{d\vec{x}(t)}{dt} &= f(\vec{x}(t), t, \vec{\theta}) \\
\frac{d}{dt}\frac{\partial \vec{x}}{\partial \theta_j} &= \frac{\partial \vec{f}}{\partial \vec{x}}\frac{\partial \vec{x}}{\partial \theta_j} + \frac{\partial \vec{f}}{\partial \theta_j}\frac{\partial \theta_j}{\partial \theta_j} + \frac{\partial \vec{f}}{\partial t}\frac{\partial t}{\partial \theta_j} \\
\frac{d}{dt}\frac{\partial \vec{x}}{\partial \theta_j} &= \frac{\partial \vec{f}}{\partial \vec{x}}\frac{\partial \vec{x}}{\partial \theta_j} + \frac{\partial \vec{f}}{\partial \theta_j} \\
\dot{\vec{S}}_j &= \mathbf{J}\vec{S} + \vec{F}_j
\end{aligned}
\tag{3.13}
$$

In this setup, recall that $\vec{x}$ describes the state variables at time t, and $\vec{f}$ describes the differential equations of the state variables. The derivative of $\frac{\partial \vec{f}}{\partial \vec{x}}$ is usually referred to as the Jacobian matrix. The Jacobian matrix is a square matrix containing the derivative of the system equations with respect to each of the state variables and is defined as matrix **J**. The term $\frac{\partial \vec{f}}{\partial \theta_j}$ is called the parametric Jacobian, and defined as $\vec{F}_j$. Rather than a matrix, it is a column vector that contains the derivative of the system equations with respect to θ_j. After manipulation, $\vec{S}_j$ contains a set of m differential equations. These equations only represent the impact of one specific θ_j on each state variable x_i, so this process needs to be repeated n times, once for each θ_j. Fortunately, the Jacobian matrix, the most difficult piece to derive, will not change between different parameter examinations as long as the structure of the original system is not modified. This ultimately results in a collection of m original ODEs, plus mn sensitivity equations, for a total of $n(m+1)$ equations. An ODE solver inputs these equations with the original system to create a numerical simulation. Several methods exist for the actual generation of the sensitivity equations. Even a simple PBPK model can lead to very unpleasant partial derivatives, and manual derivation is rarely reasonable. In this work, the computing program Matlab, specifically the Symbolic Math Toolbox, was utilized to manipulate the symbolic expressions.

This method of deriving sensitivity equations, just like any other method, naturally has its drawbacks. The initial computation of the system, especially the Jacobian matrix, can be quite complicated and time consuming. Furthermore, in the case of this model, it will result in a stiff system, which not all ODE solvers are capable of manipulating. Running all of the sensitivity

equations and the original system concurrently is necessary, which is often computationally expensive and time consuming.

On the other hand, there are several benefits to this method. It requires a lot of initial effort, but once the system is established, the parameter space is much easier to modify in this method. Both the finite difference method and the complex step method require the system to be solved over and over again after each small parameter perturbation. Since there is no dependence between the equations generated with respect to θ_j and those generated with respect to θ_k, it is not necessary to include every single parameter in the system. Parameters with well established literature values can be fixed and not included in the analysis, with no ill effects. We took advantage of that in this work, in order to minimize the number of parameters in the analysis.

3.2.2 Quantitative Methods

While qualitative methods do provide insight into the model, it is important to also consider more stringent, quantitative methods. The ultimate goal of estimability analysis is to determine if there is adequate data available to uniquely estimate values for the parameters. Rather than the traditional route of comparing the model to one specific data set, here we explore the estimability of several hypothetical data sets. This approach allows for the experimental design of future research to be better structured, so that it contributes new information that can help make the model more concise.

In order to analyze different hypothetical data sets, we introduce several scenarios. Each scenario consists of two pieces of information: a set that contains the indexes of the state variables that were experimentally measured, SV, and a set that contains time points for which the data points are available, ST. The largest possible set for SV would contain every state variable, while the minimum set is a single state variable. Similarly, ST must contain at least one time point but at most contains every time point. A list of scenarios examined in this analysis is discussed at length in Chapter 4. It would be simple to expand this idea to consider a subset of parameters in each scenario as well, but since we have already fixed reliable parameters to literature values, this is not done here.

3.2.2.1 Full Sensitivity Matrix Development

The methods described here require the use of the sensitivity matrix, which is described in Equation 3.14.

$$\mathbb{S}\left(\vec{x}(t),t,\vec{\theta}\right)=\begin{pmatrix} S_1\left(\vec{x}(t),t,\vec{\theta}\right) \\ S_2\left(\vec{x}(t),t,\vec{\theta}\right) \\ \vdots \\ S_m\left(\vec{x}(t),t,\vec{\theta}\right) \end{pmatrix},$$

$$S_i\left(\vec{x}(t),t,\vec{\theta}\right)=\begin{pmatrix} s_{i,1}\left(\vec{x}(t_1),t_1,\vec{\theta}\right) & s_{i,2}\left(\vec{x}(t_1),t_1,\vec{\theta}\right) & \cdots & s_{i,n}\left(\vec{x}(t_1),t_1,\vec{\theta}\right) \\ s_{i,1}\left(\vec{x}(t_2),t_2,\vec{\theta}\right) & s_{i,2}\left(\vec{x}(t_2),t_2,\vec{\theta}\right) & \cdots & s_{i,n}\left(\vec{x}(t_2),t_2,\vec{\theta}\right) \\ \vdots & \vdots & \ddots & \vdots \\ s_{i,1}\left(\vec{x}(t_p),t_p,\vec{\theta}\right) & s_{i,2}\left(\vec{x}(t_p),t_p,\vec{\theta}\right) & \cdots & s_{i,n}\left(\vec{x}(t_p),t_p,\vec{\theta}\right) \end{pmatrix}, \tag{3.14}$$

$$s_{i,j}\left(\vec{x}(t_k),t_k,\vec{\theta}\right)=\frac{\partial y_i(t_k)}{\partial\theta_j}\frac{\theta_j}{\max\limits_{t\in\{t_1,\ldots,t_p\}} y_i}, \qquad \begin{matrix} i\in\{1,2,\ldots,m\} \\ j\in\{1,2,\ldots n\} \\ k\in\{1,2,\ldots p\} \end{matrix}$$

The full sensitivity matrix, $\mathbb{S}\left(\vec{x}(t),t,\vec{\theta}\right)$, consists of a vertical stack of smaller sensitivity matrices, each one of which corresponds to a different state variable in the model. These smaller matrices, $S_i\left(\vec{x}(t),t,\vec{\theta}\right)$, are dimension $p\times n$, where p is the number of time points and n is the number of parameters being examined. As such, the $\mathbb{S}\left(\vec{x}(t),t,\vec{\theta}\right)$ matrix has dimensions $mp\times n$. In order to normalize the matrix, each entry of the smaller matrices are scaled by a factor of $\frac{\theta_j}{\max y_i}$. Along with preventing division by zero, this choice of scaling has the added benefit of nondimensionalizing the matrix [8, 22, 35].

3.2.2.2 Normalized Sensitivity Coefficients

First, a relatively simple quantitative comparison can be accomplished by calculating a normalized sensitivity coefficient for each parameter,

$$C_j = \left\| s_{i,j} \right\|_2 = \left[\frac{1}{t_p} \int_0^{t_p} \left(\frac{\partial y_i(t)}{\partial \theta_j} \frac{\theta_j}{\max\limits_t y_i} \right)^2 \partial t \right]^{\frac{1}{2}}. \tag{3.15}$$

This equation essentially takes the mathematical norm of each column of the sensitivity matrix, and it allows for a relative ranking of the parameters' sensitivities [8, 22]. However, this calculation has no dependence on the type or amount of data that is available. Interpretation of these results is left to Chapter 4, while we explore other methods of estimability in this section.

3.2.2.3 Composing the Fisher Information Matrix

For each scenario, a customized sensitivity matrix must be fashioned, which excludes any information that does not pertain to the state variables and time points that the scenario specifies. As an example, let us consider the hypothetical experiment where the TCDD concentration levels were measured in both fat tissue and liver tissue over time. Measurements were recorded at one hour, two hours, and four hours after exposure. Its customized sensitivity matrix, which is designated $\mathbb{S}^*\left(\vec{x}(t), t, \vec{\theta}\right)$, would look like Equation 3.16.

$$\begin{aligned} \mathbb{S}^*\left(\vec{x}(t), t, \vec{\theta}\right) &= \begin{pmatrix} S_6^*\left(\vec{x}(t), t, \vec{\theta}\right) \\ S_{10}^*\left(\vec{x}(t), t, \vec{\theta}\right) \end{pmatrix} \\ &= \begin{pmatrix} s_{6,1}\left(\vec{x}(1), 1, \vec{\theta}\right) & s_{6,2}\left(\vec{x}(1), 1, \vec{\theta}\right) & \dots & s_{6,n^*}\left(\vec{x}(1), 1, \vec{\theta}\right) \\ s_{6,1}\left(\vec{x}(2), 2, \vec{\theta}\right) & s_{6,2}\left(\vec{x}(2), 2, \vec{\theta}\right) & \dots & s_{6,n^*}\left(\vec{x}(2), 2, \vec{\theta}\right) \\ s_{6,1}\left(\vec{x}(4), 4, \vec{\theta}\right) & s_{6,2}\left(\vec{x}(4), 4, \vec{\theta}\right) & \dots & s_{6,n^*}\left(\vec{x}(4), 4, \vec{\theta}\right) \\ s_{10,1}\left(\vec{x}(1), 1, \vec{\theta}\right) & s_{10,2}\left(\vec{x}(1), 1, \vec{\theta}\right) & \dots & s_{10,n^*}\left(\vec{x}(1), 1, \vec{\theta}\right) \\ s_{10,1}\left(\vec{x}(2), 2, \vec{\theta}\right) & s_{10,2}\left(\vec{x}(2), 2, \vec{\theta}\right) & \dots & s_{10,n^*}\left(\vec{x}(2), 2, \vec{\theta}\right) \\ s_{10,1}\left(\vec{x}(4), 4, \vec{\theta}\right) & s_{10,2}\left(\vec{x}(4), 4, \vec{\theta}\right) & \dots & s_{10,n^*}\left(\vec{x}(4), 4, \vec{\theta}\right) \end{pmatrix} \end{aligned} \tag{3.16}$$

The dimensions of $\mathbb{S}^*$ are $p^*m^* \times n^*$, where p^* is the modified number of time points, m^* is the modified number of state variables, and n^* is the modified number of parameters being analyzed. Several methods can be used to obtain the entries of $\mathbb{S}^*$, but in this work we use the differential algebra method from section 3.2.1.3.

In general, larger entries in this matrix are indicative of a strong influence of θ_j on the behavior of the system, but this is not always true. Similarly, small matrix entries imply that that θ_j may be inestimable. If any obvious linear dependencies are observed between the matrix columns, there may be a combination of parameters that are highly correlated. In such a case it is likely that only one of the group can be uniquely estimated unless the model is reparameterized.

Once the modified sensitivity matrix has been established, it is used to calculate the Fisher Information Matrix (FIM), via the formula

$$\text{FIM}^* = \left(\mathbb{S}^*\right)^{\mathrm{T}} \left(\mathbb{S}^*\right), \tag{3.17}$$

which has the dimensions $n^* \times n^*$. At this point, the rank of the FIM can provide some insight into the estimability of the system. If the rank is not equal to n^*, then at least one parameter in the set is not estimable. Depending on computational accuracy and model complexity, the rank may be difficult to determine, and it does not provide any insight into which of the parameters are contributing to the inestimable nature of the model. One option at this point is to reform the modified sensitivity matrix and recalculate the FIM for some subset of the original parameters examined. If the resulting FIM is nonsingular, then this may provide some insight into the model's estimability [34]. However, depending on the number of parameters that are being studied, the number of potential combinations to compare could escalate quickly.

3.2.2.4 Eigenvalue method

As with most algorithms, the eigenvalue method has been modified and adjusted often since its conception in 1989 [45]. Because it requires manual inspection of the eigenvalues and eigenvectors, the original algorithm does not lend itself for use with large PBPK models that contain many parameters. Here we use a modified version of the method, published in [34], which adds some automation into the procedure.

Before this algorithm can be implemented, the modified sensitivity matrix must be formulated as specified in the previous section. The eigenvalues and eigenvectors of the resulting FIM are calculated. The variable λ_j represents the jth eigenvalue of these calculations, and its corresponding eigenvector is $\vec{\gamma}_j$. We assume that the eigenvectors are organized so that $|\lambda_1| \leq |\lambda_2| \leq \ldots \leq |\lambda_{n^*}|$ and that the eigenvectors have been normalized. At the beginning of the procedure, we assume that all of the parameters are identifiable.

This algorithm is based on the concept that a singular FIM indicates a nonidentifiable model, because a singular matrix does not have full rank. One of the properties of a singular matrix is a determinant that equals zero, which guarantees at least one eigenvalue to also be zero. However, many computer calculations have taken place between the point of the original model simulation and the determination of the eigenvalues and eigenvectors. This may have potentially added floating point errors to a model that already contains some uncertainty from estimated parameter values. As such, the likelihood of $|\lambda_1|$ equaling precisely zero is very low. The algorithm introduces a variable, ϵ, which is a small and positive value. The value assigned to this variable is somewhat arbitrary and does require some manual inspection. Any eigenvalues $\left|\lambda_j\right| < \epsilon$ are assumed to be essentially equal to zero and treated as such in the rest of the algorithm (Table 3.1).

Table 3.1 Steps involved with the eigenvalue method, as described by [29].

Prior to beginning the algorithm:
Obtain $\lvert\lambda_1\rvert \leq \lvert\lambda_2\rvert \leq \ldots \leq \lvert\lambda_{n^*}\rvert$, (the ordered list of eigenvalues of the FIM).
Obtain $\vec{\gamma}_j, j \in \{1, ..., n^*\}$, (the corresponding eigenvectors of the FIM).
Set $I = \{1, ..., n^*\}$, (the indexes of the parameters that are identifiable).
Set $U = \{\varnothing\}$, (the indexes of the parameters that are not identifiable).
Choose ϵ value
Set $n_p = n^*$.
1. If $\lvert\lambda_1\rvert < \epsilon$, this indicates that there is an unidentifiable parameter in I.
2. The component of $\vec{\gamma}_j$ which has the largest absolute magnitude indicates the least identifiable parameter in the model. Select k such that $\{\lvert\gamma_1(1,k)\rvert = \max(\lvert\gamma_1(1,1)\rvert, \lvert\gamma_1(1,2)\rvert, ..., \lvert\gamma_1(1,n_p)\rvert\}$. Remove k from I, and add it to U. Set n_p equal to the updated length of I.
3. Modify $\mathbb{S}^*$ to remove the information corresponding to the kth parameter. 4. Repeat until $\lvert\lambda_1\rvert \geq \epsilon$. Any elements that are removed from I are done so in order from least to most identifiable.

CHAPTER

4

INTERPRETATION

In order to simplify the analysis of this model, parameters with reliable literature values were not included. This leaves twelve parameters, mostly related to the kinetics of the chemical reactions in the liver tissue compartment, which make up

$$\begin{aligned}\vec{\theta} = [&J, K_{MET}, K_{ABS}, K_{OFF\ TCDDCYP}, K_{D\ CYP}, K_{OFF\ TCDDAhR}, \dots \\ &K_{D\ AhR}, K_{CYP\ BASE}, K_{CYP\ DEG}, K_{CYP\ MAX}, N, K_{CYP\ D}].\end{aligned} \tag{4.1}$$

4.1 Visual Inspection

The qualitative method we examine is the differential sensitivity equations, detailed in section 3.2.1.3. This method results in a large system of equations, each which represent $\frac{\partial x_i}{\partial \theta_j}$, for $j \in \{1, ..., n\}$ and $i \in \{1, ..., m\}$. In the case of this model, which contains 18 state variables and 12 parameters, 216 sensitivity equations are produced. For clarity, they are divided into 18 graphs, each one representing the impact of the parameters on a single state variable. Several of the graphs have sharp peaks, which are a product of the time scale of the integration.

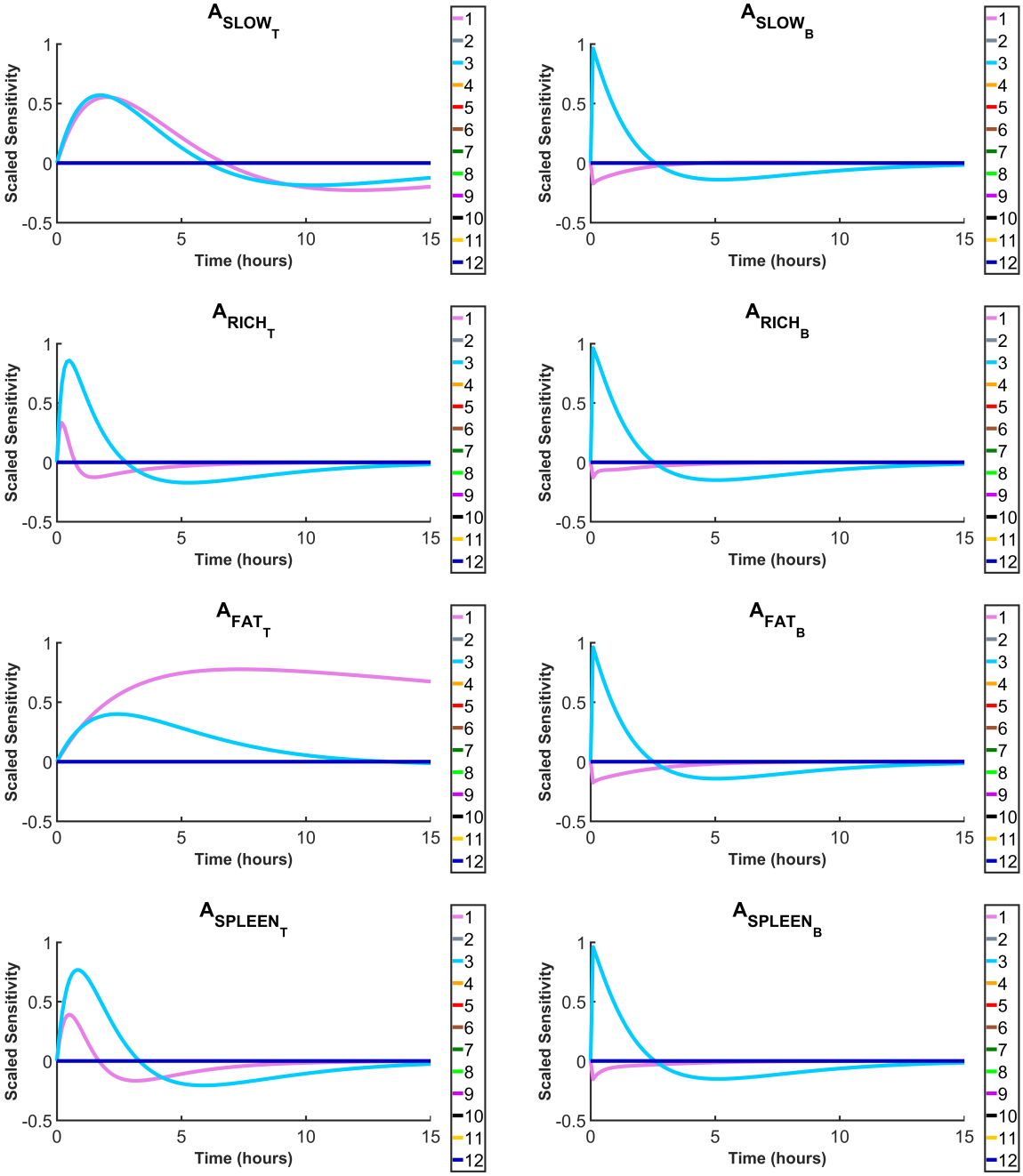


Figure 4.1 Differential sensitivity equations over time. These equations represent $\frac{\partial x_i}{\partial \theta_j}$, for $j \in \{1, ..., 12\}$ and $i \in \{2, ..., 9\}$. The $\vec{\theta} = [J, K_{MET}, K_{ABS}, K_{OFF\ TCDDCYP}, K_{D\ CYP}, K_{OFF\ TCDDAhR}, K_{D\ AhR}, K_{CYP\ BASE}, K_{CYP\ DEG}, K_{CYP\ MAX}, N, K_{CYP\ D}]$. The x-axis has been truncated to assist in readability.

The less complex compartments produce very similar results to each other (Figure 4.1). Each of the tissue compartments shows a high impact of θ_1 and θ_3, but all other sensitivities are approximately zero. The shape of each curve differs slightly, most likely because of the differences in partition coefficients. The corresponding tissue blood compartments also only show sensitivity to these two parameters. Since θ_1 represents J, its impact is not unexpected. This parameter is involved in the rate of permeability for each tissue compartment. The other important parameter, θ_3, represents K_{ABS}, the absorption rate of the TCDD from the stomach into the bloodstream. The value of K_{ABS} has a large impact on the concentration in the blood compartments, which in turn impacts the tissue concentrations. None of these compartments appear to be sensitive to any other parameters, given the very small sensitivity equation values.

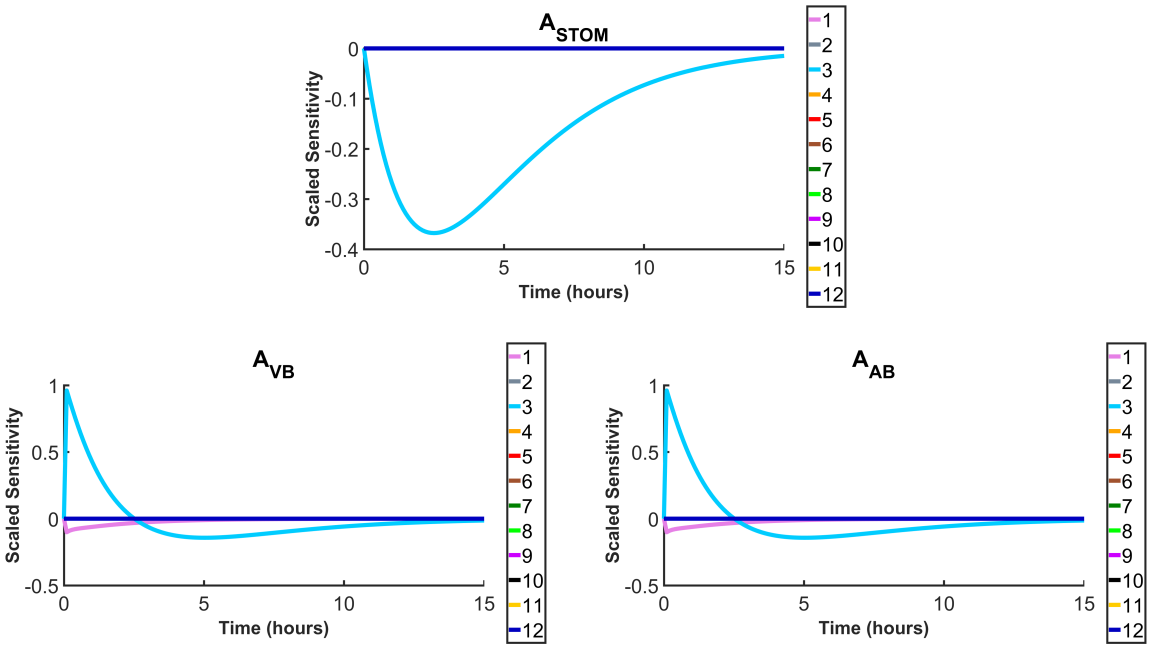


Figure 4.2 Differential sensitivity equations over time. These equations represent $\frac{\partial x_i}{\partial \theta_j}$, for $j \in \{1, ..., 12\}$ and $i \in \{1, 16, 17\}$. The $\vec{\theta} = [J, K_{MET}, K_{ABS}, K_{OFF\ TCDDCYP}, K_{D\ CYP}, K_{OFF\ TCDDAhR}, K_{D\ AhR}, K_{CYP\ BASE}, K_{CYP\ DEG}, K_{CYP\ MAX}, N, K_{CYP\ D}]$. The x-axis has been truncated to assist in readability.

The state variables which demonstrate the introduction of the TCDD into the system are the amount left in the stomach, the amount in the arterial blood, and the amount in the venous blood (Figure 4.2). The strong impact that θ_3 has on the stomach is expected, given that this parameter represents the absorption rate. That same parameter also makes a strong appearance in both of the blood compartments. The θ_1 parameter, which controls the rate of permeability of the dioxin into the tissue compartments, has influence on the bloodstream concentration.

Next we examine the sensitivities for the state variables that are located in the liver (Figure 4.3). The sensitivities of the liver tissue compartment behave similarly to those of the other tissue compartments, with a strong influence by θ_1 and θ_3. In the liver, there is also an influence by θ_2, which represents the metabolic clearance rate of the TCDD. The sensitivities that appear to have the largest magnitudes are those that impact the A_{CYP} and $A_{TCDDCYP}$ variables, but these also level out the most quickly. Recall that the differential equation for A_{CYP} is the most intricate, and it contains most of the parameters which are being examined in this identifiability analysis. The other enzyme pair, A_{AhR} and $A_{TCDDAhR}$, does not undergo induction and is thus less complicated. These state variables are impacted by θ_1 and θ_3, which is a common theme throughout the system, but in this case, θ_7 also has a strong influence. This parameter represents the dissociation constant of the AhR enzyme, and it only appears in the two corresponding state variable equations. Finally, it can be seen that θ_1 and θ_3 have a strong impact on the amount of TCDD that has been metabolized. This finding is unexpected, because the only parameter that appears in this state variable's differential equation is θ_2, the rate of metabolic clearance. This sensitivity graph seems to indicate that the ability of the TCDD to permeate each of the tissues and the rate at which it enters the system is more important than the actual rate of clearance.

The other method that we examine here is the ranking of the normalized sensitivity coefficients, described in section 3.2.2.2. Unfortunately, it can be seen in Figure 4.4 that uniquely estimating the different parameters will be difficult for this model. Large vertical gaps in this graph indicate the potential for division between the identifiable and nonidentifiable parameters. Seven of the parameters appear to have almost identical sensitivity coefficients, with a large gap above and below them. Upon closer inspection, five of these parameters are

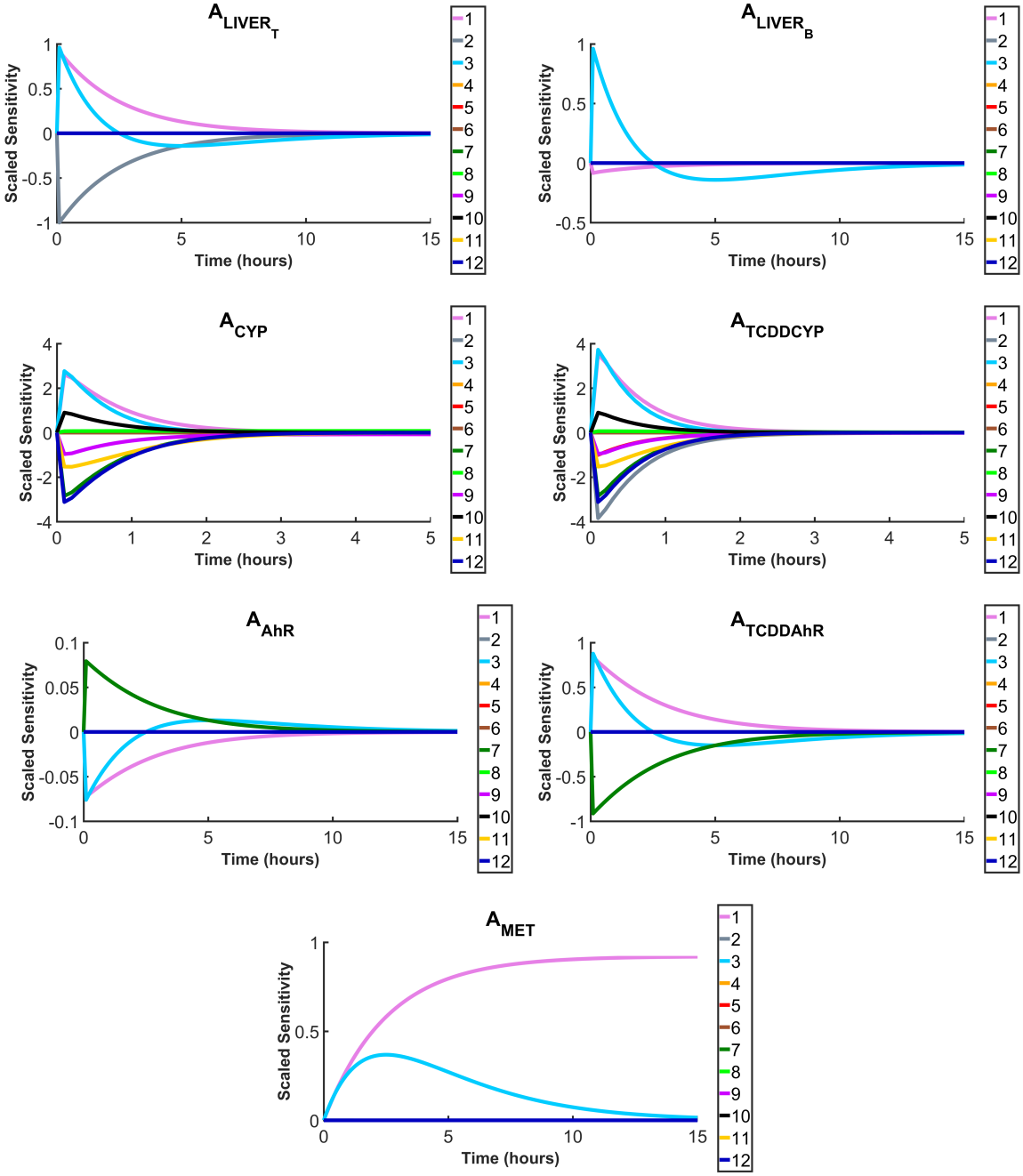


Figure 4.3 Differential sensitivity equations over time. These equations represent $\frac{\partial x_i}{\partial \theta_j}$, for $j \in \{1, ..., 12\}$ and $i \in \{10, .., 15, 18\}$. The $\vec{\theta} = [J, K_{MET}, K_{ABS}, K_{OFF\ TCDDCYP}, K_{D\ CYP}, K_{OFF\ TCDDAhR}, K_{D\ AhR}, K_{CYP\ BASE}, K_{CYP\ DEG}, K_{CYP\ MAX}, N, K_{CYP\ D}]$. The x-axis has been truncated to assist in readability.

involved only in the differential equation for A_{CYP}, so it logically follows that they would have similar sensitivity coefficients.

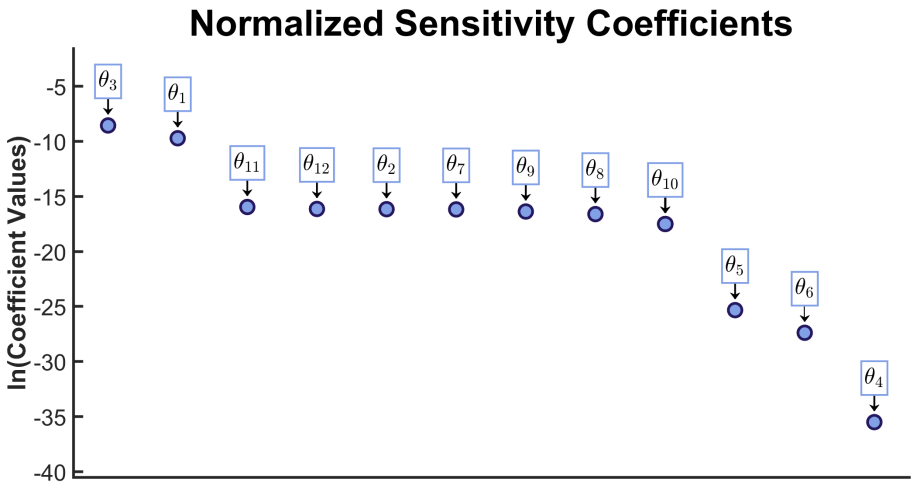


Figure 4.4 Normalized sensitivity coefficients. These values are calculated as in Equation 3.15, and arranged from largest to smallest. Each coefficient is labeled with its corresponding parameter. Note that the y-axis is on the natural log scale. The $\vec{\theta} = [J, K_{MET}, K_{ABS}, K_{OFF\ TCDDCYP}, K_{D\ CYP}, K_{OFF\ TCDDAhR}, K_{D\ AhR}, K_{CYP\ BASE}, K_{CYP\ DEG}, K_{CYP\ MAX}, N, K_{CYP\ D}]$.

Recall that these methods have not taken into consideration the data that is available. Even before considering the more involved eigenvalue method, the chances that all of the parameters in this model are estimable appears to be low.

4.2 Eigenvalue Method

The eigenvalue method, described in section 3.2.2.4, takes a different approach than the other methods examined in this work, because it considers the amount and type of data points that are available.

4.2.1 Scenarios

As a modeler trying to calibrate a PBPK model, more data is not always helpful unless it answers different questions. One of the ways that PBPK models can help to provide scientific

insight is by aiding in the design of an experiment (Table 4.1). In this analysis, we explore whether eight different, hypothetical experimental designs would be able to provide enough information to accurately estimate the parameters of this model, using the techniques outlined in the previous chapter.

There are potentially an infinite number of experimental designs, but any laboratory work inevitably faces constraints. This model is considering the toxin as it flows around the body of the mouse, which contains a very small amount of tissue. The experimental protocol which measures the concentration of the TCDD requires at least a volume of tissue similar to that of the mouse's spleen. Since the liver and spleen are both explicitly defined in the model and large enough for measurement, they are good candidates for a helpful experimental design. Although more complicated to measure, adipose tissue also has the potential to add information. The compartments in the model which represent the richly and slowly perfused tissues are more abstract and thus are more difficult to measure concretely in the lab. With the current technology, the concentration of TCDD in the blood cannot be tested, which is an unfortunate limitation. There is also no way to distinguish the TCDD that is free from the TCDD that is bound to another molecule. The hypothetical experimental scenarios were designed with these restrictions in mind.

The simulation from this model is compared to a set of data in section 4.3. That experiment provided data was available for the eleven time points in scenario (a), and they were only available for the liver tissue. Scenario (b) illustrates an experiment that was similarly designed, but which measured the concentrations in each of the tissues. The experimental procedure which measures the concentration of TCDD unfortunately cannot be performed with blood. Rather than focus on more of the body, scenario (c) takes the opposite approach and examines more of the detail in the liver tissue. At this point, there is no process to distinguish between the bound and free concentrations of TCDD in the liver. Thus, both scenarios (b) and (c) are not practical for real experimentation, but they are included for comparison. Scenarios (d) and (e) explore the merits of including measurements from one other tissue type, adipose or spleen respectively, along with those for the liver.

The timing of the experimental data could also have an effect on how much information that it can provide to the mathematical model. To determine the effect of more data points,

Table 4.1 List of the hypothetical experimental scenarios explored. SV represents the set of state variables which are measured, and ST represents the time points at which data are available. Scenario (a) represents a documented experiment [12, 21]. Scenarios (b) and (c) are not possible, but they are provided for comparison. Scenarios (d) and (e) involve multiple tissues. Scenarios (f) and (g) involve different time lines. Scenario (h) combines scenario (e) and (f).

	Scenario Description
(a)	**only liver tissue** $SV = \{x_{10}\}$ $ST = \{2, 4, 8, 12, 18, 24, 72, 168, 672, 2016, 4032\}$
(b)	**all of the tissues, no blood** $SV = \{x_2, x_4, x_6, x_8, x_{10}\}$ $ST = \{2, 4, 8, 12, 18, 24, 72, 168, 672, 2016, 4032\}$
(c)	**three liver enzyme measurements** $SV = \{x_{10}, x_{13}, x_{15}\}$ $ST = \{2, 4, 8, 12, 18, 24, 72, 168, 672, 2016, 4032\}$
(d)	**liver and fat tissues** $SV = \{x_6, x_{10}\}$ $ST = \{2, 4, 8, 12, 18, 24, 72, 168, 672, 2016, 4032\}$
(e)	**liver and spleen tissues** $SV = \{x_8, x_{10}\}$ $ST = \{2, 4, 8, 12, 18, 24, 72, 168, 672, 2016, 4032\}$
(f)	**twice as many data points** $SV = \{x_{10}\}$ $ST = \{1, 2, 3, 4, 8, 12, 16, 18, 24, 48, 72, 120, 168, 240, ...$ $...336, 504, 672, 1344, 2016, 2688, 3360, 4032\}$
(g)	**same number of data points, but earlier** $SV = \{x_{10}\}$ $ST = \{1, 2, 3, 4, 8, 12, 16, 18, 24, 48, 72\}$
(h)	**spleen, liver, and twice the data points** $SV = \{x_8, x_{10}\}$ $ST = \{1, 2, 3, 4, 8, 12, 16, 18, 24, 48, 72, 120, 168, 240, ...$ $...336, 504, 672, 1344, 2016, 2688, 3360, 4032\}$

scenario (f) contains twice as many time points as the other data sets. Because this system seems to approach equilibrium quickly, the system appears to be much more dynamic at the beginning of the experiment. Scenario (g) takes this into account, because it still contains eleven data points, but they all take place closer to the beginning of the experiment. Finally, scenario (h) represents a combined effort of twice as many data points along with both spleen and tissue measurements.

4.2.2 Results

The results of the eigenvalue method are observed in Table 4.2. The method's findings are also expressed visually in Figure 4.5. Each ϵ value was calculated by examining the largest vertical gap in the graph.

There are variations among the findings, but all except scenario (c) have a very similar outcome. Only a couple of the parameters in each of the scenarios are deemed to be estimable by the algorithm. The parameters θ_1 and θ_3 are seen in several of these cases, which is to be expected after their strong presence in the sensitivity equations of section 4.1. However, θ_2 also appears frequently, which is somewhat surprising. This parameter only appears in the equation of the liver tissue compartment, and there is so much activity in that compartment that it was unexpected for the parameter to be estimable.

Note that scenario (c) has more parameters that are estimable, but still several that are not. Scenario (c) consists of a hypothetical set of data that will likely never exist, given the technological difficulties of distinguishing between bound and free TCDD. Yet even in that ''perfect'' scenario, the data would not allow for a completely estimable model.

Even though this calculation of ϵ is somewhat arbitrary, the rank of the FIM is a firm calculation. It is evident that none of the scenarios generate a FIM that is anywhere near full rank, although the idealized scenario (c) comes the closest. Examining the values of the eigenvalues in Figure 4.5 strengthens the idea that estimability is a problem. Even the largest eigenvalues for each scenario are incredibly small, often below 10^{-10}. For scenarios (b), (c), and (d), the rank of the FIM matches the number of parameters that were determined to be identifiable by the eigenvalue method. For the other scenarios, the length of I is less than the

Table 4.2 Results from the eigenvalue method.
Scenario lettering corresponds to full descriptions found in Table 4.1. For each scenario, the U and I sets are displayed. The elements of U are ordered from less to more estimable, but the elements of I have no specific ordering. The parameter vector is defined in Equation 4.1.

Scenario	*rank*(FIM*)	U *(non identifiable)*	I *(identifiable)*
(a)	4	$\{\theta_6, \theta_4, \theta_9, \theta_{10}, \theta_8, \theta_5, \theta_{11}, \theta_{12}, \theta_7, \theta_1\}$	$\{\theta_2, \theta_3\}$
(b)	3	$\{\theta_4, \theta_6, \theta_8, \theta_{10}, \theta_9, \theta_5, \theta_{11}, \theta_{12}, \theta_7\}$	$\{\theta_1, \theta_2, \theta_3\}$
(c)	7	$\{\theta_4, \theta_9, \theta_{10}, \theta_8, \theta_6\}$	$\{\theta_1, \theta_2, \theta_3, \theta_5, \theta_7, \theta_{11}, \theta_{12}\}$
(d)	3	$\{\theta_6, \theta_9, \theta_4, \theta_{10}, \theta_8, \theta_5, \theta_{11}, \theta_{12}, \theta_7\}$	$\{\theta_1, \theta_2, \theta_3\}$
(e)	4	$\{\theta_6, \theta_8, \theta_4, \theta_5, \theta_{10}, \theta_9, \theta_{11}, \theta_{12}, \theta_7\}$	$\{\theta_1, \theta_2, \theta_3\}$
(f)	5	$\{\theta_6, \theta_4, \theta_9, \theta_{10}, \theta_8, \theta_5, \theta_{11}, \theta_7, \theta_{12}, \theta_1\}$	$\{\theta_2, \theta_3\}$
(g)	5	$\{\theta_4, \theta_6, \theta_8, \theta_5, \theta_9, \theta_{10}, \theta_{11}, \theta_7, \theta_{12}, \theta_1\}$	$\{\theta_2, \theta_3\}$
(h)	5	$\{\theta_6, \theta_4, \theta_8, \theta_5, \theta_{10}, \theta_9, \theta_{11}, \theta_7, \theta_{12}\}$	$\{\theta_1, \theta_2, \theta_3\}$

rank. This is not surprising, recalling the similarity of the normalized sensitivity coefficients of seven of the parameters.

Although their positioning changes slightly between scenarios, it is common for θ_4 and θ_6 to be quickly removed as the least identifiable parameter. This is an unfortunate finding, because these parameter represent the rate at which the liver enzymes release the TCDD ligand, and it is a value that is likely impossible to determine experimentally.

4.3 Comparison to Available Data

Data points were obtained from [12, 21] for comparison to the output of this model (Table 4.3). In these experiments, female C57BL/6 mice were given a bolus dose of dioxin at time 0. At each of the time points indicated, mice were sacrificed, and the concentration of TCDD in the liver was recorded.

Unfortunately, these data points are orders of magnitude different than the simulated results of this mathematical model (Figure 4.6). The maximum value of TCDD measured in the experiment was 664 nM, observed after 3 days. The model simulation peaks at approximately

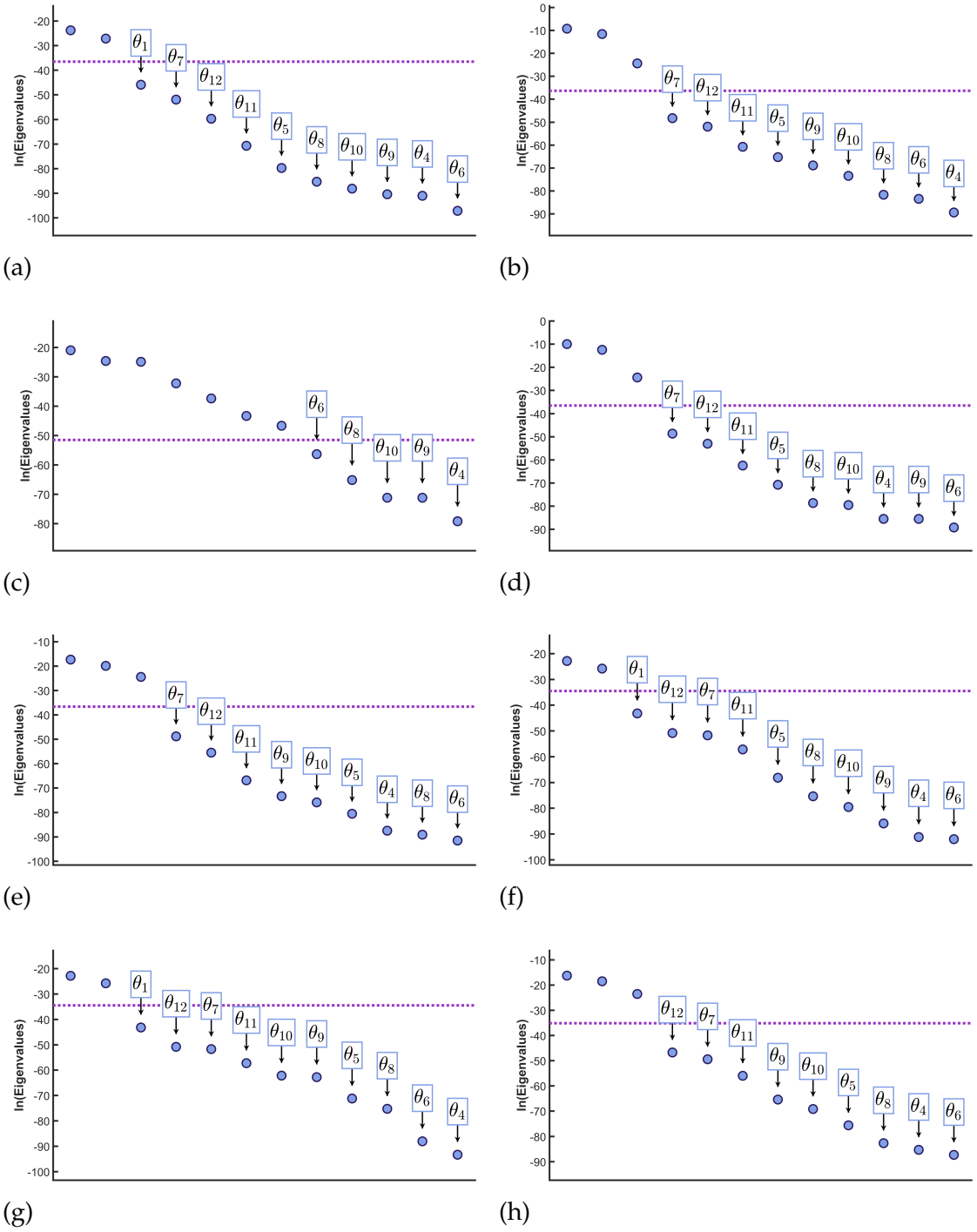
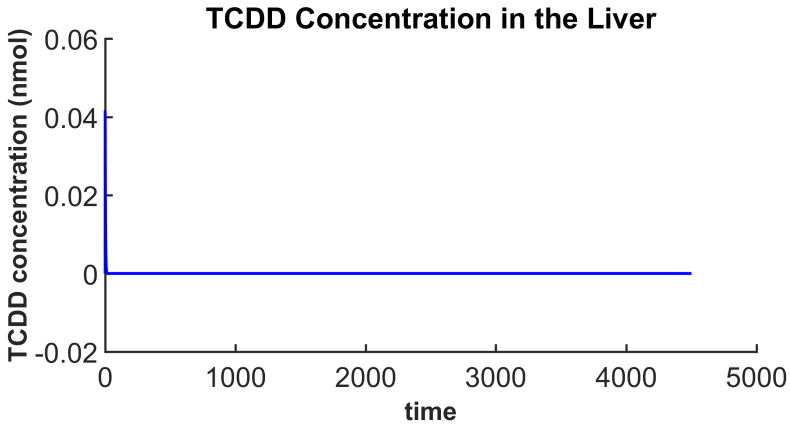


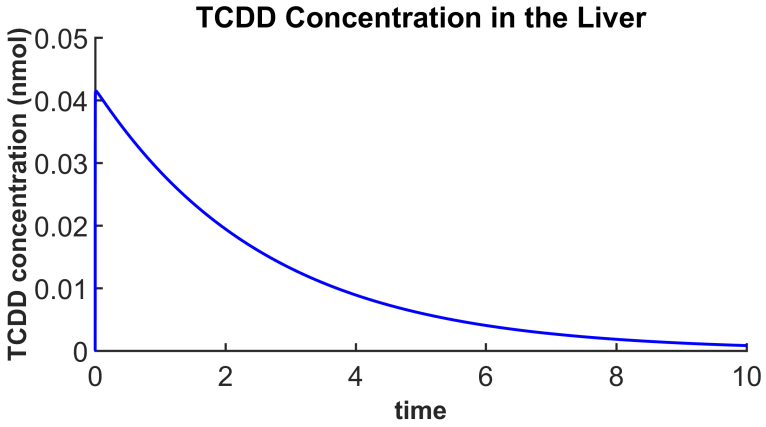
Figure 4.5 Visualization of the eigenvalues associated with the modified FIM during the eigenvalue method analysis. Each of the subplots corresponds to a unique scenario, identified in Table 4.1. The parameter which was deemed non identifiable and removed in that step of the algorithm is indicated about the eigenvalue. The ϵ cutoff value is displayed with a dotted horizontal line. Note that the vertical axis is log scaled.

Table 4.3 TCDD concentration data from [12, 21]. These concentrations were measured in liver tissue of female, C57BL/J mice, using mass spectrography techniques.

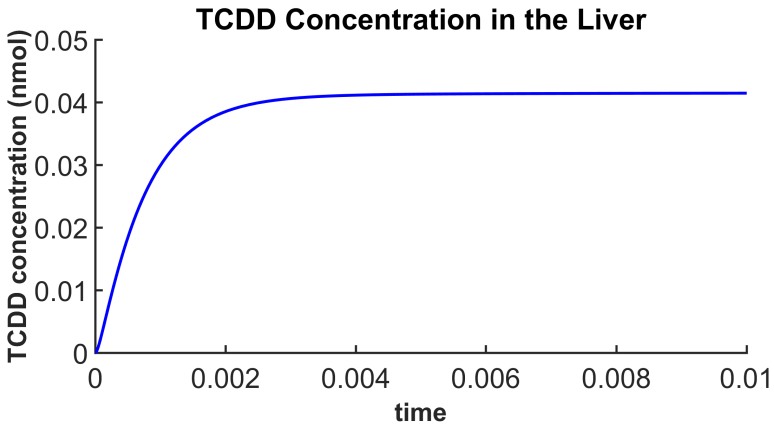
Time (hours)	*Concentration (nM)*
2	193.37
4	322.46
8	477.23
12	419.46
18	520.70
24	593.17
72	663.56
168	271.84
672	41.532
2016	1.7184
4032	0.041304



(a)



(b)



(c)

Figure 4.6 Simulated liver tissue concentration of TCDD over time. Multiple timelines are shown for easier visual interpretation.

0.04 nM, and the system has completely cleared the TCDD in less than 12 hours. These differences illustrate that drastic errors exist in the parameter space that was used. However, as the mathematical analysis in this work illustrates, it is not possible to uniquely estimate all of the parameters.

4.4 Concluding Thoughts

TCDD is potentially a very dangerous chemical, but its behavior inside the organism is far from being understood. Mathematical models are often very helpful in providing scientific insight, but unfortunately the model that we analyze in this work cannot provide extra information.

Approximately 12 of the parameters in the model do not have reliable literature values. The results in this work illustrate that it is not possible to uniquely estimate these parameters given the current structure of the model and the current scientific capabilities to measure TCDD. Five of these parameters are involved solely in the equation that describes the concentration of the CYP1A2 enzyme. It is likely that this equation could be reparameterized, but it would be difficult to do so in such a way that still allowed for each parameter to have a biological representation. Another option would be to simplify the model structure by removing one or more terms from the equation. However, this model was specifically formulated with the intention of a more complex CYP1A2 enzyme in mind, so this would defeat its purpose.

These findings are specific to the exact model described in this work, but it is possible that similar PBPK models based on TCDD have made errors in this area as well. Before the findings from such a model can be truly accepted, it would be prudent for a similar identifiability analysis to be conducted on its model structure to ensure that the results are valid. Otherwise, a model may be fit with a set of parameters that closely resemble a data set, without being sure that it in fact the only set of parameters that would generate that outcome.

Bibliography

[1] NTP (National Toxicology Program). 2016. Report on Carcinogens, Fourteenth Edition.; Research Triangle Park, NC: U.S. Department of Health and Human Services, Public Health Service. <http://ntp.niehs.nih.gov/go/roc14/>.

[2] National Center for Biotechnology Information. PubChem Compound Database, CID=15625.

[3] SEER Training Modules, *Anatomy & Physiology.* U.S. National Institutes of Health, National Cancer Institute. Accessed 2016 <https://training.seer.cancer.gov/>.

[4] Abraham, K. et al. "Pharmacokinetics and biological activity of 2,3,7,8-tetrachlorodibenzo-ρ-dioxin". *Archives of Toxicology* **62** (1988), pp. 359–368.

[5] Andersen, M. E. et al. "Regional hepatic CYP1A1 and CYP1A2 induction with 2,3,7,8-tetrachlorodibenzo-ρ-dioxin evaluated with a multicompartment geometric model of hepatic zonation." *Toxicology and applied pharmacology* **144**.1 (1997), pp. 145–55.

[6] Andersen, M. et al. "Modeling Receptor-Mediated Processes with Dioxin: Implications for Pharmacokinetics and Risk Assessment". *Risk Analysis* **13** (1993), pp. 25–36.

[7] Andersen, M. E. et al. "A multicompartment geometric model of the liver in relation to regional induction of cytochrome P450s". *Toxicology and applied pharmacology* **144**.1 (1997), pp. 135–144.

[8] Arthur, J. G. et al. "Feasibility of parameter estimation in hepatitis C viral dynamics models". *Journal of Inverse and Ill-posed Problems* (2016).

[9] Banks, H. T. & Tran, H. T. *Mathematical and Experimental Modeling of Physical and Biological Processes*. CRC Press, 2009.

[10] Beck, J. & Arnold, K. *Parameter estimation in engineering and science*. John Wiley & Sons, 1977.

[11] Bellman, R. & Åström, K. J. "On structural identifiability". *Mathematical biosciences* **7** (1970), pp. 329–339.

[12] Boverhof, D. R. et al. "Temporal and dose-dependent hepatic gene expression patterns in mice provide new insights into TCDD-mediated hepatotoxicity." *Toxicological sciences : an official journal of the Society of Toxicology* **85**.2 (2005), pp. 1048–63.

[13] Capaldi, A. et al. "Parameter estimation and uncertainty quantification for an epidemic model." *Mathematical biosciences and engineering : MBE* **9**.3 (2012), pp. 553–76.

[14] Casalegno, M et al. "Hydrophobic aggregation and collective absorption of dioxin into lipid membranes: insights from atomistic simulations". *Physical Chemistry Chemical Physics* **17**.4 (2015), pp. 2344–2348.

[15] Chen, H. S. & Gross, J. F. "Estimation of tissue-to-plasma partition coefficients used in physiological pharmacokinetic models." *Journal of pharmacokinetics and biopharmaceutics* **7** (1979), pp. 117–125.

[16] David, P. "Dioxin lawsuits: Agent Orange in the courts". *Nature* **304** (1983), p. 6.

[17] Eslami, M. *Theory of Sensitivity in Dynamic Systems: an introduction*. 1st. Springer Berlin Heidelberg, 1994.

[18] Gasiewicz, T. A. & Rucci, G. "Cytosolic receptor for 2,3,7,8-tetrachlorodibenzo-p-dioxin. Evidence for a homologous nature among various mammalian species." *Molecular pharmacology* **26**.1 (1984), pp. 90–8.

[19] Knutson, J. C. & Poland, A. "Response of murine epidermis to 2,3,7,8-tetrachlorodibenzo-p-dioxin: Interaction of the Ah and hr loci". *Cell* **30**.1 (1982), pp. 225–234.

[20] Kohn, M. C. et al. "Physiological modeling of a proposed mechanism of enzyme induction by TCDD". *Toxicology* **162**.3 (2001), pp. 193–208.

[21] Kopec, A. et al. "Toxicogenomic evaluation of long-term hepatic effects of tcdd in immature, ovariectomized C57BL/6 mice". *Toxicological Sciences* **135**.2 (2013), pp. 465–75.

[22] Lankford, G. B. "Optimization, Modeling, and Control : Applications to Klystron Designing and Hepatitis C Virus Dynamics" (2016).

[23] Lawrence, G. S. & Gobas, F. A.P. C. "A pharmacokinetic analysis of interspecies extrapolation in dioxin risk assessment". *Chemosphere* **35**.3 (1997), pp. 427–452.

[24] Leung, H. W. et al. "Pharmacokinetics of $[^{125}I]$-2-iodo-3,7,8-trichlorodibenzo-p-dioxin in mice: analysis with a physiological modeling approach." *Toxicology and applied pharmacology* **103**.3 (1990), pp. 411–419.

[25] Leung, H.-w. et al. "A physiologically based pharmacokinetic model for 2,3,7,8-tetrachlorodibenzo-ρ-dioxin C57BL/6J and DBA/2J mice". *Toxicology Letters* **42** (1988), pp. 15–28.

[26] Leung, H.-W. et al. "A physiological pharmacokinetic description of the tissue distribution and enzyme-inducing properties of 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin in the rat". *Toxicology and applied pharmacology* **103**.3 (1990), pp. 399–410.

[27] Li, Q et al. "High level of dioxin-TEQ in tissue is associated with Agent Orange exposure but not with biochemical recurrence after radical prostatectomy". *Prostate cancer and prostatic diseases* **16**.4 (2013), pp. 376–381.

[28] Martins, J et al. "An automated method for sensitivity analysis using complex variables". *AIAA paper* (2000).

[29] McLean, K. A. P. & Mcauley, K. B. "Mathematical modelling of chemical processes-obtaining the best model predictions and parameter estimates using identifiability and estimability procedures". *Canadian Journal of Chemical Engineering* **90**.2 (2012), pp. 351–366.

[30] Nelson, D. & Cox, M. *Lehninger principles of biochemistry*. 2005.

[31] Ogu, C. C. & Maxa, J. L. "Drug interactions due to cytochrome P450." *Proceedings (Baylor University. Medical Center)* **13**.4 (2000), pp. 421–423.

[32] Palmer, M. G. "The Case of Agent Orange: International Perspectives and an Homage to Victims". *Contemporary Southeast Asia: A Journal of International and Strategic Affairs* **29**.1 (2007), pp. 172–195.

[33] Perkins, E. "2001 YEARBOOK: HAZARDOUS MATERIALS AND ENERGY: The Stockholm Convention on Persistent Organic Pollutants: A Step Toward the Vision of Rachel Carson". *COLO. J. INT'L ENVTL. L. & POL'Y* **2001** (2001), pp. 191–215.

[34] Quaiser, T. & Mönnigmann, M. "Systematic identifiability testing for unambiguous mechanistic modeling–application to JAK-STAT, MAP kinase, and NF-κB signaling pathway models." *BMC systems biology* **3** (2009), p. 50.

[35] Sawyer, M. E. "Compartmentalizing the Sunlight Vitamin: Physiologically Based Pharmacokinetic Modeling and Vitamin D" (2013).

[36] Schecter, A. et al. "Intake of dioxins and related compounds from food in the US population". *Journal of toxicology and environmental health Part A* **63**.1 (2001), pp. 1–18.

[37] Schecter, A. et al. "Are Vietnamese food exports contaminated with dioxin from Agent Orange?" *Journal of Toxicology and Environmental Health Part A* **66**.15-16 (2003), pp. 1391–1404.

[38] Schecter, A. et al. "Dioxin, Dibenzofuran, and Polychlorinated Biphenyl (PCB) Levels in Food from Agent Orange–sprayed and Nonsprayed Areas of Laos". *Journal ofToxicology and Environmental Health, Part A* **66**.22 (2003), pp. 2165–2186.

[39] Schecter, A. et al. "Dioxins: an overview". *Environmental research* **101**.3 (2006), pp. 419–428.

[40] Scialli, A. R. et al. "Agent Orange Exposure and 2, 3, 7, 8-Tetrachlorodibenzo-p-Dioxin (TCDD) in Human Milk". *Birth Defects Research Part B: Developmental and Reproductive Toxicology* **104**.3 (2015), pp. 129–139.

[41] Smith, R. *Uncertainty quantification: theory, implementation, and applications*. Vol. 12. SIAM, 2013.

[42] Stewart, J. *Calculus: Early Transcendentals*. 5th. Belmont, CA: Brooks/Cole, 2003.

[43] Tortorelli, D. a. & Michaleris, P. "Design sensitivity analysis: Overview and review". *Inverse Problems in Science and Engineering* **1**.1 (1994), pp. 71–105.

[44] Turanyi, T. & Rabitz, H. "Local Methods". *Sensitivity Analysis*. Ed. by Saltelli, A. et al. West Sussex, England: John Wiley & Sons, 2000. Chap. 5, pp. 81–99.

[45] Vajda, S et al. "Qualitative and quantitative identifiability analysis of nonlinear chemical kinetic models". *Chemical Engineering Communications* **83** (1989), pp. 191–219.

[46] Walker, N. J. et al. "Characterization of the Dose-Response of CYP1B1 , CYP1A1 , and CYP1A2 in the Liver of Female Sprague-Dawley Rats Following Chronic Exposure to 2,3,7,8-Tetrachlorodibenzo-ρ-dioxin". **286** (1999), pp. 279–286.

[47] Wang, X et al. "Determination of parameters responsible for pharmacokinetic behavior of TCDD in female Sprague-Dawley rats." *Toxicology and applied pharmacology* **147** (1997), pp. 151–68.

[48] Wang, X. et al. "Extrapolation of a PBPK model for dioxins across dosage regimen, gender, strain, and species". *Toxicological Sciences* **56** (2000), pp. 49–60.

Appendix

APPENDIX

A

FULL MODEL DETAILS

This section provides the minute details of the mathematical model that was explored in this work. The equations for each of the state variables (Table A.1) are detailed in Equation A.1. This system contains many parameters of different types. The information in Table A.3 relates to the overall morphology of the mouse's body and its tissue makeup. Some of these parameters need to undergo secondary calculations before they are helpful for the model, which are detailed in Table A.2. Table A.6 includes the parameters which are kinetic in nature. These variables also make up the $\vec{\theta}$ vector used in the analysis of this model's identifiability and estimability. The partition coefficients are listed separately in Table A.4, and the initial conditions for the system are in Table A.5.

Table A.1 Descriptions of each state variable, along with their position in the $\vec{x}$ vector.

Index	***Symbol***	***Description***
x_1	A_{STOM}	amount of TCDD left in the stomach
x_2	A_{ST}	amount of TCDD in the slowly perfused tissue
x_3	A_{SB}	amount of TCDD in the blood of the slowly perfused tissue
x_4	A_{RT}	amount of TCDD in the richly perfused tissue
x_5	A_{RB}	amount of TCDD in the blood of the richly perfused tissue
x_6	A_{FT}	amount of TCDD in the adipose tissue
x_7	A_{FB}	amount of TCDD in the blood of the adipose tissue
x_8	A_{SPT}	amount of TCDD in the spleen tissue
x_9	A_{SPB}	amount of TCDD in blood of the the spleen tissue
x_{10}	A_{LT}	amount of TCDD in the liver tissue
x_{11}	A_{LB}	amount of TCDD in the blood of the liver tissue
x_{12}	A_{CYP}	amount of unbound CYP1A2 enzyme in the liver tissue
x_{13}	$A_{TCDDCYP}$	amount of TCDD bound to the CYP1A2 (in the liver tissue)
x_{14}	A_{AhR}	amount of unbound Ah receptor in the liver tissue
x_{15}	$A_{TCDDAhR}$	amount of TCDD bound to the Ah receptor (in the liver tissue)
x_{16}	A_{VB}	amount of TCDD in the venous blood
x_{17}	A_{AB}	amount of TCDD in the arterial blood
x_{18}	A_{MET}	amount of TCDD that has been metabolized

The full system of differential equations are listed here in Equation A.1. Note that these equations differ slightly from those described earlier in the text. Each concentration was replaced by the corresponding amount divided by the volume of the compartment. This substitution allows for the state variables to be directly observable in the equations.

$$
\begin{aligned}
(1)\quad & \frac{dA_{STOM}}{dt} = -K_{ABS}A_{STOM} \\
(2)\quad & \frac{dA_{ST}}{dt} = \frac{JQ_SA_{SB}}{V_{SB}} - \frac{JQ_SA_{ST}}{V_{ST}Pa_S} \\
(3)\quad & \frac{dA_{SB}}{dt} = \frac{Q_SA_{AB}}{V_{AB}} - \frac{Q_SA_{SB}}{V_{SB}} - \frac{JQ_SA_{SB}}{V_{SB}} + \frac{JQ_SA_{ST}}{V_{ST}Pa_S} \\
(4)\quad & \frac{dA_{RT}}{dt} = \frac{JQ_RA_{RB}}{V_{RB}} - \frac{JQ_RA_{RT}}{V_{RT}Pa_R} \\
(5)\quad & \frac{dA_{RB}}{dt} = \frac{Q_RA_{AB}}{V_{AB}} - \frac{Q_RA_{RB}}{V_{RB}} - \frac{JQ_RA_{RB}}{V_{RB}} + \frac{JQ_RA_{RT}}{V_{RT}Pa_R} \\
(6)\quad & \frac{dA_{FT}}{dt} = \frac{JQ_FA_{FB}}{V_{FB}} - \frac{JQ_FA_{FT}}{V_{FT}Pa_F} \\
(7)\quad & \frac{dA_{FB}}{dt} = \frac{Q_FA_{AB}}{V_{AB}} - \frac{Q_FA_{FB}}{V_{FB}} - \frac{JQ_FA_{FB}}{V_{FB}} + \frac{JQ_FA_{FT}}{V_{FT}Pa_F} \\
(8)\quad & \frac{dA_{SPT}}{dt} = \frac{JQ_{SP}A_{SPB}}{V_{SPB}} - \frac{JQ_{SP}A_{SPT}}{V_{SPT}Pa_{SP}} \\
(9)\quad & \frac{dA_{SPB}}{dt} = \frac{Q_{SP}A_{AB}}{V_{AB}} - \frac{Q_{SP}A_{SPB}}{V_{SPB}} - \frac{JQ_{SP}A_{SPB}}{V_{SPB}} + \frac{JQ_{SPT}A_{SPT}}{V_{SPT}Pa_{SP}} \\
(10)\quad & \frac{dA_{LT}}{dt} = \frac{JQ_LA_{LB}}{V_{LB}} - \frac{JQ_LA_{LT}}{V_{LT}Pa_L} - \frac{K_{MET}A_{LT}}{V_{LT}} \ldots \\
& \quad + \frac{K_{OFF\_TCDDCYP}A_{TCDDCYP}}{V_{LT}} - \frac{K_{ON\_TCDDCYP}A_{CYP}A_{LT}}{V_{LT}^2} \ldots \\
& \quad + \frac{K_{OFF\_TCDDAhR}A_{TCDDAhR}}{V_{LT}} - \frac{K_{ON\_TCDDAhR}A_{AhR}A_{L_T}}{V_{LT}^2}
\end{aligned}
\tag{A.1}
$$

$$(11)\quad \frac{dA_{LB}}{dt} = \frac{Q_L A_{AB}}{V_{AB}} + \frac{Q_{SP} A_{SPB}}{V_{SPB}} + \frac{JQ_L A_{LT}}{V_{LT} Pa_L} \ldots - \frac{JQ_L A_{LB}}{V_{LB}} - \frac{(Q_L + Q_{SP}) A_{LB}}{V_{LB}} + K_{ABS} A_{STOM}$$

$$(12)\quad \frac{dA_{CYP}}{dt} = K_{CYP\_BASE} - \frac{K_{CYP\_DEG} A_{CYP}}{V_{LT}} \ldots - \frac{K_{ON\_TCDDCYP} A_{CYP} A_{LT}}{V_{LT}^2} + \frac{K_{OFF\_TCDDCYP} A_{TCDDCYP}}{V_{LT}} \ldots + (K_{CYP\_MAX} - K_{CYP\_BASE}) \frac{\frac{A_{TCDDAhR}}{V_{LT}}^N}{\frac{A_{TCDDAhR}}{V_{LT}}^N + K_{CYP\_D}^N}$$

$$(13)\quad \frac{dA_{TCDDCYP}}{dt} = \frac{K_{ON\_TCDDCYP} A_{CYP} A_{LT}}{V_{LT}^2} - \frac{K_{OFF\_TCDDCYP} A_{TCDDCYP}}{V_{LT}}$$

$$(14)\quad \frac{dA_{AhR}}{dt} = -\frac{K_{ON\_TCDDAhR} A_{AhR} A_{LT}}{V_{LT}^2} + \frac{K_{OFF\_TCDDAhR} A_{TCDDAhR}}{V_{LT}}$$

$$(15)\quad \frac{dA_{TCDDAhR}}{dt} = \frac{K_{ON\_TCDDAhR} A_{AhR} A_{LT}}{V_{LT}^2} - \frac{K_{OFF\_TCDDAhR} A_{TCDDAhR}}{V_{LT}}$$

$$(16)\quad \frac{dA_{VB}}{dt} = \frac{Q_S A_{SB}}{V_{SB}} + \frac{Q_R A_{RB}}{V_{RB}} + \frac{Q_F A_{FB}}{V_{FB}} + \frac{(Q_L + Q_{SP}) A_{LB}}{V_{LB}} - \frac{Q_{TOT} A_{VB}}{V_{VB}}$$

$$(17)\quad \frac{dA_{AB}}{dt} = -\frac{Q_S A_{AB}}{V_{AB}} - \frac{Q_R A_{AB}}{V_{AB}} - \frac{Q_F A_{AB}}{V_{AB}} \ldots - \frac{Q_L A_{AB}}{V_{AB}} - \frac{Q_{SP} A_{AB}}{V_{AB}} + \frac{Q_{TOT} A_{VB}}{V_{VB}}$$

$$(18)\quad \frac{dA_{MET}}{dt} = \frac{K_{MET} A_{LT}}{V_{LT}}$$

Table A.2 Parameters which require extra computation, which is accomplished via these equations.

Symbol	***Description***	***Units***	***Calculation***
Q_S	blood flow rate to slowly perfused tissues	L/hr	$Q_{TOT} \cdot QF_S$
Q_R	blood flow rate to richly perfused tissues	L/hr	$Q_{TOT} \cdot QF_R$
Q_F	blood flow rate to adipose tissues	L/hr	$Q_{TOT} \cdot QF_F$
Q_{SP}	blood flow rate to the spleen	L/hr	$Q_{TOT} \cdot QF_{SP}$
Q_L	blood flow rate to the liver	L/hr	$Q_{TOT} \cdot QF_L$
V_{SB}	volume of slowly perfused tissue blood	L	$BW \cdot VF_S \cdot VF_{SB}$
V_{ST}	volume of slowly perfused tissue	L	$BW \cdot VF_S$
V_{RB}	volume of richly perfused tissue blood	L	$BW \cdot VF_R \cdot VF_{RB}$
V_{RT}	volume of richly perfused tissue	L	$BW \cdot VF_R$
V_{FB}	volume of adipose tissue blood	L	$BW \cdot VF_F \cdot VF_{FB}$
V_{FT}	volume of adipose tissue	L	$BW \cdot VF_F$
V_{SPB}	volume of spleen tissue blood	L	$BW \cdot VF_{SP} \cdot VF_{SPB}$
V_{SPT}	volume of spleen tissue	L	$BW \cdot VF_{SP}$
V_{LB}	volume of liver tissue blood	L	$BW \cdot VF_L \cdot VF_{LB}$
V_{LT}	volume of liver tissue	L	$BW \cdot VF_L$
V_B	volume of total blood	L	$BW \cdot VF_B - V_{SB} - V_{RB}...$ $-V_{FB} - V_{SPB} - V_{LB}$
V_{AB}	volume of arterial blood	L	$V_B \cdot VF_{AB}$
V_{VB}	volume of venous blood	L	$V_B - V_{AB}$

Table A.3 Parameter values related to animal morphology. Unless otherwise noted, values are based on estimates for female C57BL/6J mice.

Symbol	*Description*	*Units*	*Value*	*Source*
BW	total body weight	kg	0.025	[21, 48]
Q_{TOT}	total blood flow rate	L/hr	1.045	[24]
QF_S	fractional blood flow to slowly perfused tissues	unitless	0.17	[24]
QF_R	fractional blood flow to richly perfused tissues	unitless	0.485	[24]
QF_F	fractional blood flow to adipose tissue	unitless	0.07	[24]
QF_{SP}	fractional blood flow to spleen	unitless	0.025	estimated[a]
QF_L	fractional blood flow to liver	unitless	0.25	[24]
VF_{ST}	volume fraction of slowly perfused tissues to organism	unitless	0.761	[24, 25]
VF_{RT}	volume fraction of richly perfused tissues to organism	unitless	0.035	[24, 25]
VF_{FT}	volume fraction of adipose tissue to organism	unitless	0.059	[24, 25]
VF_{SPT}	volume fraction of spleen tissue to organism	unitless	0.005	estimated[a]
VF_{LT}	volume fraction of liver tissue to organism	unitless	0.05	[24, 25]
VF_{SB}	volume fraction of slowly perfused tissue blood to slowly perfused tissues	unitless	0.05	[6][b]
VF_{RB}	volume fraction of richly perfused tissue blood to richly perfused tissues	unitless	0.01	[6][b]
VF_{FB}	volume fraction of adipose tissue blood to adipose tissue	unitless	0.05	[6][b]
VF_{SPB}	volume fraction of spleen tissue blood to spleen	unitless	0.01	estimated[c]
VF_{LB}	volume fraction of liver tissue blood to liver	unitless	0.01	[6][b]
VF_B	volume fraction of total blood to organism	unitless	0.05	[24]
VF_{AB}	volume fraction of arterial blood to total blood	unitless	0.3	[3][d]

[a] Due to the mass of the mouse spleen being approximately one tenth of the mass of the liver, the fractional volume of the spleen and the fractional blood flow to it was assumed to be one tenth the value of the liver.
[b] These values are based on data from female Wistar rats, not mice.
[c] Due to the similarities between the liver and spleen organs, it was assumed that the spleen's blood vessel density was approximately the same.
[d] Human estimate.

Table A.4 List of partition coefficients. All values are based on female C57BL/J mice

Symbol	***Description***	***Units***	***Value***	***Source***
Pa_S	partition coefficient of slowly perfused tissues/blood	unitless	3	[24]
Pa_R	partition coefficient of richly perfused tissues/blood	unitless	10	[24]
Pa_F	partition coefficient of adipose tissue/blood	unitless	300	[24]
Pa_{SP}	partition coefficient of spleen tissue/blood	unitless	10	estimated[a]
Pa_L	partition coefficient of liver tissue/blood	unitless	10	[24]

[a] Assumed to have the same chemical attribute as the liver.

Table A.5 Initial conditions for the state variables. Sources are based on female C57BL/J mice. All other concentrations are equal to zero at the start of the simulation, due to the assumed lack of any other exposure to TCDD.

Symbol	***Description***	***Units***	***Value***	***Source***
$C_{STOM}(0)$	initial concentration of TCDD in the stomach	nmoles/L	2.3292	[21][a]
C_{AhR}	concentration of AhR	nmoles/L	0.35	[48]
$C_{CYP}(0)$	initial concentration of CYP	nmoles/L	1.5	[48]

[a] Value based on the experimental protocol in [21], which calls for a bolus dose of 30μg/kg of body weight.

Table A.6 Kinetic parameters involved in the chemical reactions.

Index	***Symbol***	***Description***	***Units***	***Value***	***Source***
θ_1	J	permeability coefficient multiplier for all tissues	unitless	0.08	[48]
θ_2	K_{MET}	first order metabolism rate of TCDD in liver	L/hr	$\frac{2.5}{BW^{0.3}}$	[5]
θ_3	K_{ABS}	absorption rate of TCDD from stomach	L/hr	0.4	[48]
θ_4	$K_{OFF\ TCDDCYP}$	rate constant for separation of TCDDCYP complex	L/hr	35	estimated[a]
θ_5	$K_{D\ CYP}$	dissociation constant of CYP1A2 (K_{OFF}/K_{ON})	nmoles/L	35	[48]
θ_6	$K_{OFF\ TCDDAhR}$	rate constant for separation of TCDDAhR complex	L/hr	0.1	estimated[a]
θ_7	$K_{D\ AhR}$	dissociation constant of AhR (K_{OFF}/K_{ON})	nmoles/L	0.1	[48]
θ_8	$K_{CYP\ BASE}$	basal production rate of CYP1A2	nmoles/g of liver/hr	0.004	[5]
θ_9	$K_{CYP\ DEG}$	degradation rate of CYP1A2	1/hr	0.04	[5]
θ_{10}	$K_{CYP\ MAX}$	maximum production rate of CYP1A2	nmoles/g of liver/hr	0.4	[5]
θ_{11}	N	Hill coefficient for CYP induction	unitless	4	[5][b]
θ_{12}	$K_{CYP\ D}$	Hill binding constant	pmoles/L	50	[6][c]

[a] The dissociation constant of an enzyme has more biological relevance than either K_{OFF} or K_{ON} separately. The values of $K_{ON\ CYP}$ and $K_{ON\ AhR}$ were both set to 1 for simplicity. This allows the K_{OFF} values for each enzyme to be equal to its corresponding dissociation constant.
[b] Value based on Sprague-Dawley rats.
[c] Value based on Wistar rats.