ABSTRACT

SMITH, NIKKI SHAVON. A Comparison of Physiologically-Based Pharmacokinetic (PBPK) Models of Methyl-Tertiary Butyl Ether (MTBE). (Under the direction of Hien T. Tran and Marina V. Evans.)

Methyl-tertiary-butyl ether (MTBE) is a gasoline additive that gained widespread use after the passage of the Clean Air Act Amendments of 1990 designed to help reduce urban air pollution and toxic air emissions. Reports of water contamination and adverse health affects from inhalation exposure to MTBE also grew. We review MTBE carcinogenicity studies conducted in rodents and complete physiologically-based pharmacokinetic (PBPK) inhalation model comparisons, parameter estimation, and an assessment of design elements that help fit available human data.

PBPK models are mathematical constructs that consider different organs and tissues as compartments and make it possible to obtain a quantitative characterization of concentration-time profiles in the respective compartments. We specifically test a hypothesis whether complexity should be added to the lung compartment for substances with a large blood:air partition coefficient. Further, since rats exposed to MTBE exhibited renal tumors, we consider metabolism as a mode of action for that endpoint and we include a kidney metabolism pathway, which is generally not used in PBPK models, in our model. Finally, we use sensitivity analysis and model validation and comparison techniques to assess the lung complexity hypothesis, draw conclusions about the inclusion of the kidney metabolism pathway, and drive model design iterations by removing parameters, and thus, metabolism pathways that are not significant to the models' results.

Initial model analysis outcomes indicate that model simplicity prevails and adding lung complexity for the test case high blood:air partition coefficient substance (in this case MTBE's metabolite, tert-Butyl-alcohol) is not desirable. Additionally, we find that kidney metabolism parameters have a strong influence on each of the models' results. This significant influence suggests a meaningful relationship between—and further consideration for—kidney metabolism and the development of renal tumors in rodents.

A Comparison of Physiologically-Based Pharmacokinetic (PBPK) Models of Methyl-Tertiary Butyl Ether (MTBE)

by
Nikki Shavon Smith

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Applied Mathematics

Raleigh, North Carolina

2018

APPROVED BY:

Xiao-Biao Lin	Jesus Rodriguez
Hien T. Tran Co-chair of Advisory Committee	Marina V. Evans Co-chair of Advisory Committee

DEDICATION

To Manetta, who still says, "I want to be just like you when I grow up."
I hope I continue to make you proud.

To Tamika, whose support pushed me through many late nights:
Thank you.

To Mom, who didn't flinch when I said I was quitting work to return to school:
Your confidence was encouraging.

BIOGRAPHY

Nikki Shavon Smith was born to Mary and Milton Smith on January 31, some time ago, in Laurinburg, North Carolina.

While always a quiet child, Nikki was charismatic with an infectious smile, a true joy to be around. Country living kept her humble and active. However, her parents made it a point to introduce her to new and exciting things, letting her tag along with them to their college classes and labs. With a slew of cousins, aunts, and uncles nearby, there was always something to do, somewhere to go, and sports to play.

School always seemed to come naturally for Nikki so there was no surprise when she graduated high school with honors and attended North Carolina Agricultural and Technical State University on an academic scholarship. There, she obtained her Bachelor of Science in Electrical Engineering, graduating summa cum laude.

She later enrolled at the University of Tennessee, Knoxville, where she received her Master of Science in Mathematics. Upon graduation, she accepted a position with the Department of Defense in Charleston, SC, where she resided for several years until her passion for learning called her back to school.

She was accepted at North Carolina State University where she will receive her Doctor of Philosophy in Applied Mathematics under the generous guidance of Dr. Marina Evans and Dr. Hien Tran. In the meantime, she resides near Washington, DC, has resumed her post with the Department of Defense, and will collaborate further on PBPK models with her colleagues at the U.S. Food and Drug Administration.

ACKNOWLEDGMENTS

Before starting the applied mathematics program at North Carolina State University, I was admittedly afraid. I owe a wealth of gratitude to Dr. Stephen Campbell, who called personally to welcome me to the program and, although he didn't realize it, showed me that faculty are everyday people. That call calmed my fears and let me know that I would be OK.

At one point, though, I was far from OK and I was close to dismissal from the program. To Dr. Ernest Stitzinger, you will forever be my hero, the man who helped me turn things around. I can only offer finishing the program as my ultimate thanks.

I am thankful to Angelean Hendrix for her mentorship and giving me the tools I'd need to be a successful researcher. Every article I found was a reminder of my experience with you. My thanks also to Katie Schmidt and the 2014 NCSU Research Experience Undergraduate (REU) team for helping expand my PBPK knowledge and MATLAB skills.

To Joan Pharr, Rich Hammer, and Rui Hu: I would not have survived Analysis without you.

Courtney Waterman, you already know. :) The years after Numerical Analysis just weren't the same without you.

I have some of the biggest cheerleaders at home and abroad. I would be amazed at the level of support I have received while working toward this degree but I have known forever that I have amazing people in my life. Mom, Dad, Manetta, Tamika, Richelle, Ann, Linda, Teddy, Javier, Steve, my aunties and uncles. Each of you has been steady and got me through some tough times and nothing makes me happier than to celebrate my achievement with you.

Dr. Marina Evans and Dr. Hien Tran, I am not here without you. You have been my academic parents these past few years and I have grown so much under your guidance. Thank you for kindly nudging me in the right direction, for your patience, and for the laughs. I have and will continue to share how I could not have been more lucky to have you as advisors. Thank you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

Introduction

In 1990, the Clean Air Act Amendments (CAAA) required the addition of oxygenates in reformulated gasoline. Oxygenates added to gasoline make it burn cleaner and boost octane (a measure of gasoline's resistance to uncontrolled combustion). Methyl-tertiary butyl ether (MTBE) is one such oxygenate. Its primary use in the 1970s was to prevent engine "knocking" and performance loss. By the 1980s, since MTBE reduces carbon monoxide levels caused by automobile emissions, it was used to meet state and federal winter oxygenate requirements for areas that had not yet met air quality standards for carbon monoxide. During that same time, its usage rose as a result of the phase-out of leaded gasoline along with an increased demand for premium gasoline.

Further, because of MTBE's ease of production and low cost, its use was significantly boosted after the CAAA's passage. This widespread use led to an increase in groundwater contamination near underground storage tanks (USTs) across the country with the US Geological Survey (USGS) reporting in 1999 a 27% incidence of groundwater in urban areas contaminated with MTBE.

Years before the USGS report and groundwater crisis, though, people lodged health complaints related to MTBE exposure. These complaints included dizziness, nausea, disorientation, and eye irritation among others and were often the result of inhalation exposure at pumping stations, occupational exposure from regular handling of MTBE,

and other means.

In the years following the CAAA, a number of animal studies were conducted on exposure to MTBE. Select studies are summarized in Section 2.1. The studies cover oral (gavage) and inhalation exposure, both routes representative of drinking water contamination by leaking USTs and inhalation exposure at gas pumps and tanks. These studies led to conclusions about MTBE's carcinogenicity (animal and human) and its eventual ban from use in the US starting in 2006. The US still exports MTBE, however, and its use is quite prevalent in places like Asia, Latin America, and Europe [1, 29, 36] so its potential health and environmental risks persist.

The limited number of studies on humans focused on the kinetics of MTBE rather than its carcinogenicity. While more information is known about how MTBE affects animals' health (specifically rodents for this paper), information from those studies can be gathered for use in assessing the potential effects on humans. This information, coupled with the human kinetics, is the center of interest for the work going forward in this paper.

In this paper, physiologically-based pharmacokinetic (PBPK) models represent the complex biological processes related to the disposition (absorption, distribution, metabolism, and elimination [ADME]) of chemical compounds in the body. The models typically consider a body organ or tissue as a "compartment" and may be as simple or as complex as is required to adequately describe the ADME processes. This determination is made based on the definition and derivation of PBPK models discussed in Sections 4, 5, and 6.

One concept under consideration is the complexity of the lung compartment. One viewpoint is to use a lung representation that considers alveolar dead space, diffusion between the pulmonary tissue and the upper respiratory tract (URT), and fractional chemical retention in the URT during gas exchange. This realization requires multiple lung sub-compartments. Another viewpoint is to consider a far simpler lung where the conducting airways carry the compound directly to the alveolar region where gas exchange occurs. The use of one lung representation over the other is postulated based upon the chemical's solubility in blood. This hypothesis is discussed further in Section 3.

A second concept under consideration, also discussed in Section 3, takes the information about the kidney endpoints in rodent studies and uses that information to guide PBPK model development. Specifically, we claim that the kidney tumors observed in rodents can be explained by the inclusion of a Michaelis-Menten metabolism pathway in the kidney.

With available data from inhalation studies with human subjects (Section 7), PBPK models are calibrated and validated using identifiability, estimability, and sensitivity analysis techniques described in Section 8. The results of these analyses are discussed in Section 9 and, finally, conclusions about the results and proposals for future work are presented in Section 10.

CHAPTER 2

Methyl-Tertiary-Butyl-Ether (MTBE)

Methyl-Tertiary-Butyl-Ether (MTBE, CAS 1634-04-4) is a fuel oxygenate used to make gasoline burn better and decrease carbon monoxide (CO) emissions [87]. The Clean Air Act Amendments (CAAA, 1990) required the use of oxygenated gas in areas that exceeded the federal CO air quality standard (i.e., had high levels of pollutants). Gas refiners chose MTBE over other oxygenates in reformulated gasoline (RFG) for its ease in meeting emission standards set forth by the CAAA; it was the most economical and the easiest to blend (e.g., versus ethanol) during the refining process.

Due to leaking underground storage tanks or spillage or other means, MTBE was found in groundwater and reservoirs [34]. These discoveries suggested its widespread use made it a potential large-scale hazard, in this case as a drinking water contaminant, since MTBE is highly soluble in water, not easily adsorbed to soil, and not easily biodegradable. In addition to drinking water exposure (oral), people could also be exposed by swimming or showering in MTBE-contaminated water (dermal exposure). Lastly, inhalation exposure may occur in a number of ways and is our primary focus.

The general public may come into contact with MTBE via inhalation while pumping gas or pouring it into containers or gas-fueled machinery. Individuals whose exposure routes included pumping gasoline, driving their cars, or working in gas stations reported having headaches, nausea, and dizziness. Those who worked with MTBE and had regular

exposure to it also complained of both eye and respiratory tract irritation. According to the Agency for Toxic Substances and Disease Registry, breathing small amounts of MTBE for short periods may cause nose and throat irritation [2]. These effects and MTBE's widespread use prompted a number of studies to be conducted to determine the immediate and long-term effects on rodents, humans, and the environment.

2.1 MTBE Studies in Rats and Mice

Subchronic and chronic exposure studies in laboratory animals have been conducted to assess the health hazard potential of MTBE on humans. The most common effect in test animals was to the nervous system. Other effects were kidney disease, larger-than-normal livers, liver tumors, testicular cancer, leukemia, lymphomas, and death [2]. While some of these effects were repeatable in independent studies, it is not readily known whether those effects have meaning for humans.

We examine some of the studies conducted on rats and mice and later outline their findings and human carcinogenic assessments.

2.1.1 Robinson, et al. (1990): Oral Gavage; Sprague-Dawley Rats

Oral gavage (tube feeding) studies were conducted over 14 and 90 days in male and female Sprague-Dawley rats [84]. Daily dose levels ranged from 0 to 1428 mg/kg body weight for the 14-day study and 0 to 1200 for the 90-day study.

No deaths were attributed to MTBE toxicity but diarrhea was common across all groups. Subjects at or above 1200 mg/kg dose levels experienced MTBE-induced anesthesia for approximately two hours.

In the 14-day study, males and females both experienced increased cholesterol while females experienced decreased blood-urea nitrogen (BUN) and creatinine. In the 90-day study, females experienced increased cholesterol and decreased BUN while males in the high dose range experienced decreased creatinine. Depending on the creatinine levels, a low BUN could indicate impaired kidney function.

Robinson et al. reported no significant pathophysiological effects where the dose was less than 1200 mg/kg. Males in the 1200 mg/kg MTBE group exhibited renal changes

compatible with $\alpha_{2\mu}$-globulin nephropathy (kidney disease/damage). Any kidney tumors are possibly attributed to the accumulation of the $\alpha_{2\mu}$-globulin protein.

2.1.2 Burleigh-Flayer (1992): Inhalation; CD-1 Mice

Over an 18-month study, 50 male and female CD-1 mice were exposed to MTBE at concentrations of 400, 3000, and 8000 parts per million (ppm) for six hours per day, five days per week. [26]

At the highest exposures, the study indicated a statistically significant increase in liver adenomas in female mice and liver carcinomas in male mice. These tumors, according to the Environmental Epidemiology Section of the NC Department of Environment, Health, and Natural Resources, provided evidence supporting MTBE carcinogenicity.

2.1.3 Chun (1992): Inhalation; Fischer-344 Rats

In this study, male and female Fischer-344 rats were exposed to the same concentrations as the CD-1 mice in the Burleigh-Flayer study: 400 (low-dose), 3000 (mid-dose), and 8000 (high-dose) ppm, for 104 weeks, 97 weeks, and 82 weeks, respectively [26, 28].

Early mortality was caused by toxicity effects in the high-dose group. In male rats, a statistically significant increase in interstitial cell testicular tumors were observed in mid- and high-dose groups. Additionally, there was a statistically significant increase in kidney tumors for the mid-dose group. Kidney tumors were also observed in the high-dose group but its statistical significance could not be determined because high morality rates led to early termination (at 82 weeks) for this group.

As in the Robinson study [84], kidney tumors are associated with the accumulation of $\alpha_{2\mu}$-globulin. Following an EPA review of the Chun bioassay and according to EPA criteria for renal toxicity [11], the protein accumulation did not appear to occur with MTBE, thus the kidney tumors provided evidence supporting MTBE carcinogenicity.

2.1.4 Belpoggi, et al. (1995, 1998): Oral Gavage; Sprague-Dawley Rats

Sixty male and 60 female rats were exposed to 250 mg/kg (low-dose) and 1000 mg/kg (high-dose) once daily, four times per week, for two years [16].

The rats did not develop renal tumors but there were increased observations of Leydig cell (testicular) tumors for the high-dose administration and increased incidences of leukemias and lyphomas in females for both low- and high-dose administrations. These responses provide evidence to support MTBE carcinogenicity.

Belpoggi, et al., repeated this study in 1998 administering MTBE by stomach tubes with results that strengthened the findings of the 1995 study [15].

2.1.5 Bird (1997): Inhalation; CD-1 mice and Fischer-344 Rats

In this study, male and female CD-1 mice were exposed to 400, 3000 or 8000 ppm MTBE for six hours per day, five days per week for 18 months while Fischer-344 rats were exposed for 24 months [17].

This study produced results of toxicity at inhalation exposure to 3000 and 8000 ppm MTBE. For CD-1 mice, chronic exposure to 8000 ppm MTBE resulted in increased incidence of hepatocellular (liver) adenomas. For the Fischer-344 male rats, exposure to 3000 and 8000 ppm MTBE resulted in increased incidences of renal tubular cell tumors and Leydig cell adenomas. These tumors were not observed in the female rats. Also, in the male Fischer rats exposed to both 3000 and 8000 ppm MTBE, there were incidences of increased testicular interstitial cell adenomas. Male and female CD-1 mice did not develop renal tumors following exposure.

2.1.6 Alpha-2μ Globulin

In some of the MTBE rodent studies, there were observations of kidney tumors in male rats: Bird et al., 1997; Chun et al., 1992; Robinson et al., 1990. The tumors found in rodent studies are not necessarily predictive of similar risks in human, particularly if it can be shown that a mode of action exhibited for the rodent does not operate in humans. In such case, the rodent response is not relevant for human cancer risk assessment [30].

Chemically-induced $\alpha_{2\mu}$-globulin nephropathy is one mechanism for the development of kidney tumors that is not considered a predictor of carcinogenic risk to humans. This is because there is no production of an analogous protein in humans [35]. What was found for the rodents in the three studies listed above is that the increased kidney tumor response was not related to $\alpha_{2\mu}$-globulin accumulation. Additionally, there was no evidence to indicate that MTBE caused $\alpha_{2\mu}$-globulin accumulation.

Borghoff et al. established evidence that $\alpha_{2\mu}$-globulin exhibits an indirect connection to the formation of male rat kidney tumors from MTBE [22, 23, 75, 78]. Since there is no clear explanation for the disproportion between the accumulation of $\alpha_{2\mu}$-globulin and renal cell proliferation, we consider the possibility that male rat kidney tumors are induced by some other mode of action not specific to any species.

2.1.7 Overall Study Conclusions for Rats/Mice

At, or near, lethal doses of MTBE, animals experienced eye and mucous membrane irritation and central nervous system depression. At relatively high doses not typical for human environment exposure, the rat kidney was found to be a target of toxicity. Further, carcinogenic effects were observed in studies of doses administered via oral and inhalation routes.

Burleigh-Flayer et al. [26] reported increases in hepatocellular carcinomas in male CD-1 mice as well as adenomas and carcinomas in female mice. Chun et al. observed increases in tumors in male Fisher-344 rats including renal tumors in rats exposed via inhalation. And Belpoggi et al. results showed increases in tumor incidences for both male and female Sprague-Dawley rats. These incidences included male kidney tumors and female liver tumors.

Each study showed evidence of carcinogenic activity in the liver and kidney which led to inclusion of the liver (typically included) and the kidney (less often included) as target organs for model consideration in this paper. The kidney is further included as a target organ given the weak relationship between kidney tumor proliferation and $\alpha_{2\mu}$-globulin accumulation and the need to consider an alternate mode of action (e.g., metabolism) to explain the induced tumors.

2.2 MTBE Studies in Humans

Animal carcinogenicity studies for MTBE raised a number of concerns for the potential human health risk. A number of those discussed earlier confirmed carcinogenic effects in rodents.

Not surprisingly, far fewer studies have been conducted on humans than on rats/mice. Those that have been conducted primarily focused on the kinetics of MTBE and its blood concentrations in healthy volunteers [27, 42, 54, 77, 82, 94] but none of these studies

reported on chronic toxicity for MTBE. However, since the differences in reaction to MTBE between species is negligible, the extensive knowledge on toxicokinetic properties obtained from studies in rats/mice can be extended to humans [60].

2.3 Carcinogenicity and Classification

Data from a number of animal studies show that MTBE causes an abnormal proliferation of cells in several tissue sites independent of exposure route. The neoplasms occurred in both males and females across different species and rodent strains. The cancer effects in both rats and mice led to discussions on whether there was carcinogenic relevance to humans with varying conclusions and recommendations based on the evidence at the time:

- Not a carcinogen
  - The International Agency for Research on Cancer (IARC) evaluated MTBE as "not classifiable as to its carcinogenicity to humans" in 1999 [6, 37].
  - The National Toxicology Program (NTP) reviews but does not list MTBE in the 9th Report on Carcinogens (RoC), 1999 [66] with the motion to list defeated 6-5 with one abstention.
- Possibly a carcinogen
  - "The U.S. Environmental Protection Agency (USEPA) concludes that 'MTBE poses a potential for human carcinogenicity at high doses' based on animal data" but "these animal data 'do not support confident, quantitative estimation of risk at low exposure.'" [6, 90]
  - National Science and Technology Council (1996), Office of Science and Technology Policy (1997), USEPA (1997), and California Environmental Protection Agency (1999) "have published reports indicating that MTBE should be regarded as posing a potential carcinogenic risk to humans based on animal cancer data" [90].
  - Based on criteria established by national and international health agencies for classifying an agent in the absence of human data:

* IARC classifies MTBE as a "probable human carcinogen" (2006). [44]
* NTP classifies MTBE as "reasonably anticipated to be a human carcinogen" (2011). [67]
* USEPA finds MTBE "likely to be carcinogenic to humans" (2005). [93]

Based on the substantial evidence of carcinogenesis available from multiple studies showing carcinogenic effects at multiple sites across multiple species, the studies' conclusions indicate MTBE is likely to be a human carcinogen and human exposure to MTBE should be minimized.

2.4 Controversy and Bans

MTBE environmental concerns began with drivers feeling dizzy when filling their tanks with gasoline blended with MTBE. Coupled with reports of MTBE seeping into groundwater from underground storage tanks (USTs) and, to a lesser extent, pipelines, health concerns grew. These concerns were exacerbated by the oil industry's inability to stop the leaking storage tanks. Though MTBE met the criteria for the 1990 Clean Air Act and was relatively inexpensive to blend, its contamination of groundwater supplies led to legislation that would impose a voluntary nationwide phase-out by 2006, less than 10 years after its introduction and wide acceptance. [8, 43, 55, 92]

2.5 Alternatives To MTBE

MTBE's use is set to expand in the Middle East and Asia Pacific with China's increased market development leading to it projected as being the largest MTBE consuming market by 2020 [40]. This is in direct contrast to MTBE use in other countries. Though MTBE easily met emissions standards, its risk to the environment and study results led to legislation banning its use in the United States, Canada and other developed nations. Ethyl tert-butyl ether (ETBE, CAS 637-92-3) is one alternative gasoline additive. It has a higher relative cost than MTBE but it has lower water solubility and was possibly less of a threat to the environment than MTBE. Ethanol, Tertiary amyl methyl ether (TAME), and gasoline with alkylates [31, 73] were other such alternatives.

CHAPTER 3

Research Motivation

"Everything and anything is questionable and these questions make way for a research."
Dr. S.N. Sridhara

In this work, questions were boundless; however, we explore only three:

1. Is upper respiratory tract (URT) required to fit data?

   In PBPK models for volatile gases, it is often suggested that a model that includes the upper respiratory tract for the lung can provide more accurate data fits [47]. This is attributed to the washin-washout effect where some of an inhaled chemical is absorbed in the respiratory tract during inhalation and desorbed during exhalation. During this process, less of a chemical reaches the blood stream as compared to the amount when washin-washout is absent. When washin-washout is not considered, chemical concentrations in the organs can be predicted using the much simpler inert tube PBPK model. The inert tube serves as a passageway for the inhaled substance to the alveoli [47] and then to the blood where the substance is distributed to the organs.

   We examine the inconsistency when models include the upper respiratory tract or not. A hypothesis exists that the upper respiratory tract should be included to accurately explain the data when the blood:air partition coefficient is large

[47] or, more specifically, that an inert tube model is insufficient to describe the inhalation and exhalation kinetics for gases with high blood:air partition coefficients. Anderson et al. [7] state that gases that exchange in the airways "are very water soluble and have a high blood solubility ($P_{b:a} > 1000$)." They also conducted tests that show gases with $P_{b:a} < 10$ exchange predominantly in the alveoli, gases with $10 < P_{b:a} < 100$ exchange partially with the alveoli and partially with the airways, while gases with $P_{b:a} > 100$ exchange almost exclusively in the airways. The results of their tests indicate that gases with blood:air partition coefficients as low as 100 would require models that include the upper respiratory tract.

We use Methyl-Tertiary-Butyl-Ether (MTBE, CAS 1634-04-4), a fuel oxygenate used to make gasoline burn better and decrease carbon monoxide emissions, as a test case for the hypothesis. We seek to prove that inclusion of the upper respiratory tract is not required to estimate the parameters that adequately fit the observed data. [The water:air and blood:air partition coefficients are 15.2 and 17.7, respectively [48]. (*At steady-state, there's a direct relationship between water:air and blood:air partition coefficients.*)] We also use tert-Butyl alcohol (TBA, CAS 75-65-0), a metabolite of MTBE, which has a larger blood:air partition coefficient, $P_{b:a} = 462$, to test the hypothesis.

2. Does including kidney metabolism help explain renal carcinogenicity [in humans]?

   PBPK models do not often/always include metabolism in the kidney compartment. MTBE has been known to cause kidney damage in rodents (Bird, et al. [17], Chun et al. [28], Belpoggi et al. [15, 16], etc.). Humans and rodents have equivalent cytochrome P450 enzymes that metabolize MTBE [41], so metabolism is an important consideration even at low concentrations; damage can occur by different mechanisms at different doses. We look at a range of human exposure concentrations (1.7 - 75 ppm) and examine the relevance of kidney metabolism to the PBPK model. We include kidney metabolism with the expectation that its inclusion will have a compelling influence toward predicting MTBE's concentration levels using experimental data of venous blood and exhaled breath.

3. Could using sensitivity analysis drive model design?

   Several important decisions must be made when designing toxicokinetic and pharmacokinetic studies [83]: (i) animal species and strain, (ii) dose regimen, (iii) time

points for collection, (iv) tissues that will be collected, and (v) appropriate methods for sample preparation and storage.

From a mathematical modeling perspective, the analysis of a descriptive or predictive model that fits the data collected from a study may provide some information that could possibly drive the design of future kinetic studies. The factors that determine which tissues will be collected ((iv) above) also play a role in the design of the model used to describe the data. Blood data is always necessary for kinetic and model design. Sites of metabolism are typically included (liver, kidney, lungs) as well as target tissues (sites where specific effects have been observed). Properties of the chemical also drive the design. For example, if a chemical is known to collect in adipose tissue (fat), that tissue is an area of interest for both the kinetic and model design.

Once results of a mathematical model are found, a number of analysis techniques are performed (error, sensitivity, etc.). We discuss these techniques in detail later but focus momentarily on sensitivity analysis. In short, it is a process of investigating potential changes and errors in parameter values and assumptions of a model and the impact of these changes and errors on the conclusions that can be drawn from the model.

One such impact is model simplification: fixing inputs that have no effect on the output or removing redundant characteristics of the model structure. Sensitivity analysis can provide a quantitative assessment of which parameter values and/or state variables (or tissue sites under consideration), if any, can be eliminated from the current model. If it is determined that a state variable does not affect the outcome of a model, the corresponding tissue site could potentially be excluded from collection in later kinetic studies design.

CHAPTER 4

Physiologically-Based Pharmacokinetics (PBPK): Establishing Baseline Concepts

Pharmacokinetics (PK) is the study of the movement of a chemical compound or drug within the body from the time of its administration to its elimination. It quantitatively describes the time-course behavior of chemicals and the processes influencing absorption, distribution, metabolism, and elimination (ADME) of a compound in the body. Early PK arose as a way to make comparisons, evaluations and predictions of chemical disposition based on a set of parameters used to describe pharmacokinetic behavior. These parameters then allow us to understand the factors that affect ADME and thus determine the concentration of the compound throughout various organs of the body.

In early PK modeling, time-course curves were analyzed by estimating a small number of parameters. All ADME processes were treated as first-order rates, that is, the rate values were directly proportional to a chemical's concentration. These models also assumed a data-based compartmental model structure that later came into question due to their ability to account for (i) "saturation of elimination pathways, and (ii) the possibility that blood flow, rather than metabolic capacity, might limit the clearance of a chemical." [57] When an elimination pathway becomes saturated, then the PK model is no longer first-order and exact solutions to the model questions become more difficult to obtain.

For our disposition predictions to be as sound as possible, the pharmacokinetic models

we build should be as biologically realistic as possible while also realizing the simplest formulation of the model. Models that include biologically relevant parameters are called physiologically-based pharmacokinetic (PBPK) models. These parameters are typically compound- or species-specific. Species-specific values include organ volumes and blood flow rates and allow for extrapolation to other species whereas compound-specific values include lipophilicity, partitioning and metabolic clearance.

4.1 Absorption, Distribution, Metabolism, and Excretion

Absorption, distribution, metabolism, and excretion are known collectively as ADME processes. The concentrations of chemical substances in the blood and tissues and the length of time that it remains in the blood/tissue are dependent upon these processes. [83, 85]

Absorption. Once a substance is administered, either by inhalation, oral ingestion, intravenous injection, etc., it is transferred into the bloodstream. The rate and extent of this absorption process is dependent upon the route of administration and the physiochemical properties of the substance, for example, whether the compound is lipophilic ("fat loving"), volatile vs. nonvolatile, hydrophilic ("water loving"), soluble vs. insoluble, etc.

Distribution is the process of carrying a compound throughout the body once it is absorbed. Once in the blood, the compounds are taken up into the organs and tissues but blood flow to each of these is different and we have the least vitally important tissues (e.g. fat) receiving the least amount of blood whereas the greatest amount of blood flows to the more important organs (e.g., brain and liver).

Metabolism. Physiochemical properties of a compound further determine whether the compound will have a tendency to undergo metabolic activation. These biochemical reactions typically occur in the liver but can also occur in other tissues like the kidneys or gastrointestinal tract. The enzymes that metabolize the compounds fall into two major metabolism phases:

- Phase I: Oxidation-reduction reactions. These reactions are often catalyzed by the cytochrome P450 enzyme system. The amount of available enzyme dictates the

rate of metabolism. The reaction rate is accelerated by the presence of enzymes and the rate of conversion will continue for as long as the enzyme is present. When the enzyme is used up, metabolism slows until the enzyme becomes available again. This rate dependence upon the amount of available enzyme makes this metabolism case a *saturable* process.

- Phase II: Conjugation reactions. The compound is joined with another substance such as glutathione, glucuronic acid or sulfate, making it more water soluble and thus easier for it to be excreted in the urine or feces.

Metabolism is a particularly important concept for our PBPK modeling and we turn our attention later specifically to the Michaelis-Menten model of enzyme kinetics.

Elimination is the complete removal of a chemical substance from the body. The physiochemical properties mentioned earlier also determine how quickly a compound will be cleared from the body. In the case of repeated exposures to a substance, the body must remove the substance or concentrations of the substance will rise with each exposure. The primary routes for elimination are urinary excretion, fecal elimination, and exhalation. A compound may be eliminated unchanged (through the kidneys, exhalation, or into bile) or may be metabolized (e.g., by the liver clearing toxins from the body). If a substance is not made more water soluble (via a metabolism process), it will be reabsorbed into the bloodstream and recycled through the body.

Overall, the ADME process starts with absorption of a substance into the bloodstream. The substance is then distributed to the tissues (or other fluids) or passed again to the bloodstream. Those particles of the substance in the bloodstream that are available to undergo metabolism (in the kidney, liver or other tissues) do so while the remaining particles may be cycled again through the body where, ultimately, the substance and its metabolites are eliminated from the body (through metabolism or excretion in urine, exhaled breath, etc.)

4.2 Michaelis-Menten Kinetics

For all our models, we rely upon enzyme kinetics to describe the metabolism processes. The general idea is that the body wishes to convert a substrate (S), the substance being observed in a chemical reaction, into a product (P), the substance formed in a chemical

reaction. The substrate and enzyme (E), the substance (catalyst) added to the system to cause a reaction, form an enzyme-substrate complex (ES) which can then form a product plus the initial enzyme. The reactions are represented by Equation (4.1)

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} P + E, \tag{4.1}$$

where k_1 is the rate in which the enzyme and substrate react to form the complex; k_{-1} is the rate in which the complex converts back to the initial substrate and enzyme; and k_2 is the rate in which the complex forms the product.

In 1913, Leonor Michaelis and Maud Menten [49, 61] published their work on a mechanism for the catalysis of chemical reactions in biological systems. They described the mechanism of enzyme-catalyzed reactions and gave a relationship between the reaction rates and the concentrations of enzyme and substrate. First, define the production rate of product, P, at a given substrate concentration, S, as

$$v_0 = v(S) = \frac{dP}{dt} = \frac{V_{max}S}{K_m + S}. \tag{4.2}$$

This is the well-known Michalis-Menten kinetics equation where V_{max} is the maximum production rate of P (or the maximum rate of elimination or the maximum velocity of reaction) and K_m is Michaelis-Menten constant, the affinity of the enzyme for the substrate, defined as

$$K_m = \frac{k_{-1} + k_2}{k_1}, \tag{4.3}$$

the sum of the reaction rates *from* the enzyme-substrate complex divided by the reaction rate *to* the enzyme-substrate complex. K_m also satisfies $v(K_m) = V_{max}/2$. That is, when the substrate concentration is equal to K_m, the reaction velocity is one-half of V_{max}. Conversely, when the reaction velocity is one-half V_{max}, we know that the substrate concentration is equal to K_m. (See Appendix C for the full Michaelis-Menten equation derivation.)

Figure 4.1 shows the graphical relationship between the production velocity/rate and the substrate concentration. As the substrate increases, the production velocity approaches, but never exceeds, V_{max}. Thus V_{max} is the limiting value where all enzyme molecules form an enzyme-substrate complex and any additional substrate added to the reaction has no effect on the rate of reaction. In other words, V_{max} is the value where

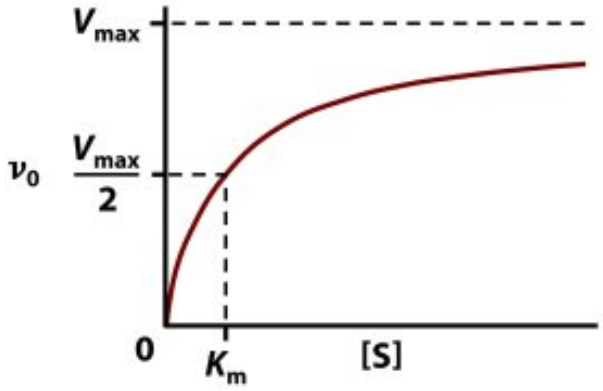


Figure 4.1: Michaelis-Menten reaction velocity vs. substrate concentration

the reaction process is using all the enzyme all the time but there is no increase in the reaction rate beyond V_{max}.

The kinetics of a model are described by "orders" which are given in terms of the concentration since this is the main factor for kinetics. A zero-order process has an elimination rate which is independent of the concentration, that is, the elimination of the substance is a constant quantity of the substance per unit time, thus the rate of change of mass of substance per unit time is given by:

$$\frac{dA}{dt} = k_0.$$

A first-order process exhibits elimination of a constant fraction of the substance present per unit time, that is, the process elimination is proportional to the substance concentration and is given by

$$\frac{dA}{dt} = k_1 CV.$$

A second-order process is proportional to two concentrations giving

$$\frac{dA}{dt} = k_2 C_1 C_2 V,$$

where V is the volume of the compartment with concentration C, C_1, or C_2 above.

Nonlinear kinetics arises when saturation occurs in one of the pharmacokinetic mechanisms (e.g., hepatic metabolism). Saturation occurs when a first-order process at low

concentrations transitions to a zero-order process at high concentrations [83]. During saturable elimination, the elimination rate tends to reach a maximal value when the concentration exceeds a certain value. The kinetics for each of the models represented for this research is described by Michaelis-Menten kinetics. We saw by Equation (4.2) that the Michaelis-Menten rate of metabolism is given by

$$Metab = \frac{V_{max}C}{K_m + C},$$

where V_{max} is the maximum rate of elimination and K_m is the enzyme affinity, the concentration of substance that gives an elimination rate of one-half V_{max}. When the concentration, C, is much larger than the affinity constant, K_m ($C \gg K_m$), saturation of the enzymes occurs, the rate of elimination approaches a fixed value equal to V_{max}, and the elimination becomes a zero-order process thereby satisfying the properties of a saturable process.

CHAPTER 5

Physiologically-Based Pharmacokinetic (PBPK) Modeling

Models that are good at prediction are referred to as "models of data" (versus "models of systems") by DiStefano and Landaw (1984) [20, 32, 52]. These models are empirical in nature, require "few assumptions about the data generating mechanism," and are useful when the data is not accompanied by knowledge of the underlying processes from which it is generated. Allometric scaling, a scaling based on body size, is one example of this type of model.

Models of systems, on the other hand, are based on physical and physiological principles and should include as many systems features as the data will support. This is the basis of physiologically based pharmacokinetic (PBPK) models, characterized in part by the transport into and out of tissues being modeled as a function of blood flow and permeability between the blood and tissue.

A PBPK approach to modeling is based on the use of a physiologically realistic representation of the biological system. This representation is a quantitative description of the ADME of a compound and includes information about the biochemical and physiological interactions that occur. The ADME determinants include blood flow rates, rates of absorption, tissue-to-blood partition coefficients, diffusion across membranes, and tissue volumes. They are one of either physiological, physiochemical or biochemical type. These parameters are also classified as either species- or drug-specific. Replacing the

species-specific data, allows for extrapolation to different species or to human.

Extrapolations play an important role in the application of PBPK modeling. They convert from one exposure route of administration to another, from high dose to low dose, and from one population to another. When PBPK models are based on animal data, one clear extrapolation would be from animals to humans. These extrapolations are understandably necessary when we have studies that indicate evidence of toxicity of a substance in one environment and we wish to predict the conditions that could lead to similar outcomes in another environment.

PBPK models, due to their ability to increase toxicity testing efficiency, significantly boost the accuracy of extrapolations. By using biologically relevant model parameters, PBPK models also reduce uncertainty in high-to-low dose extrapolations.

5.1 Principles of PBPK Modeling

Models represent the system under study [20]. Modeling is a process that requires iteration. Our initial structure was determined by identifying the important kinetic processes driving the chemical behavior (in this case, metabolism), target sites of interest (e.g., endpoints for toxicity), the availability of parameter data obtained from the literature, and most importantly, the availability of experimental data (the outputs of the system). Model parameters are fitted to the measured data, maintaining realistic biochemical, physiological, and biological properties. For example, partition coefficients and chemical concentrations are always positive values. The model is then tested against the data and, oftentimes, an updated model is developed.

Best practices in PBPK model development are always to start with the simplest model possible, including only compartments of interest, for example, those for which experimental data is available (e.g., blood plasma) or those which are endpoints for toxicity (e.g., liver). While this might not always be possible, model designers should keep in mind that the number of input parameters increases as more model compartments are added [83]. Further, these parameters must be estimated from experimental data, thus the availability of data is a limiting factor in the model's design.

Experimental data aside, first, consider tissue grouping. How many compartments are needed? Can certain organs be grouped into a larger compartment if they are not identified as data or target compartments? Yes. A minimal PBPK model might

have rapidly perfused tissues (heart, kidney, liver, brain, etc.) grouped into a single compartment and slowly perfused tissues (muscle, fat, skin, etc.) grouped into another. This "lumping" can be achieved when tissues are pharmacokinetically and toxicologically similar. In contrast, tissues which are pharmacokinetically and toxicologically dissimilar or distinct must be separated. [80]

Next, we consider the processes controlling kinetics and other compartments that contribute to the distribution, elimination, and toxicity of a substance. For example, is distribution flow-limited or diffusion-limited? In a flow-limited system, the amount of substance reaching a tissue is dependent upon the blood flow rate to that tissue. The amount can be increased if the blood flow rate is increased. Further, equilibrium for this compartment can be reached, an important property in the definition of partition coefficients (discussed in Section 5.3). In a diffusion-limited system, the amount of substance entering a tissue is dependent upon the tissue's barrier and its rate of diffusion (movement of a substance from a region of high concentration to a region of low concentration). Thus the amount crossing the tissue's membrane can not be increased with blood flow. The models presented in this paper are all flow-limited systems.

Finally, what are the uptake routes? Do we have dermal exposure to the substance? Is it inhaled? Ingested? Injected? Each of these exposure routes indicates the requirement of an associated compartment: skin, lung, gastrointestinal (GI) tract, and/or blood, respectively. The exposure routes represent the perturbation to the system. The associated amounts along with duration are the inputs to the system.

When all these considerations have been made, it is important to also maintain mass balance for the system. That is, the sum of the tissue blood flows must equal the total cardiac output: $\sum_i Q_i = Q_{total}$. If a compartment is split from a larger compartment (e.g, separating kidney from the rapidly perfused tissue compartment), that compartment's blood flow and volume must also be separated from the total for the larger compartment.

5.2 Gas Exchange in the Lung: the Wash-in Wash-out effect

Washin-washout is a phenomenon where some of a chemical is retained in the upper airways during respiration [80]. It is the process where a chemical is diffused into the lung's mucous layer during both inhalation and exhalation. During inhalation (washin),

less of the chemical participates in gas exchange with the alveoli, thus less of the chemical reaches the blood stream to be distributed to the rest of the body. Upon exhalation, some fractional component of the chemical is desorbed from the respiratory airway (washout) and removed in expired air.

When washin-washout is not considered, chemical concentrations in the organs can be predicted using the much simpler inert tube PBPK model. The inert tube serves as a conducting passageway to transport the inhaled substance to the alveoli [47] and then to the blood where the substance is distributed to the organs. PBPK models for highly soluble substances that treat the airways as an inert tube have been shown to underestimate the exhaled values anticipated [50].

The washin-washout effect is most dramatic when the blood:air (or water:air) partition coefficient is high. A high partition coefficient ($P_{b:a} > 1000$ [7]) indicates a highly water soluble substance and is therefore more likely to be absorbed into the mucous lining of the lung and upper airways whereas low blood-soluble gases ($P_{b:a} < 10$) exchange via inert tube in the alveoli.

5.3 Partition Coefficients

Key parameters in any PBPK model are the tissue-to-blood partition coefficients. They measure the steady-state chemical concentration between two compartments. Modeling pharmacokinetics requires detailed information about the partition coefficients for each organ. The most-used partition coefficients for our model are the tissue:blood coefficients given by the ratio of the tissue concentration to the blood concentration

$$P_{t:b} = \frac{C_t}{C_b}.$$

The greater the partition coefficient, the greater the affinity of the chemical for the tissue compared to the blood. For example, if $P_{t:b} = 5$, then the affinity of the chemical toward the tissue is five times greater than its affinity toward blood. We represent the concentration leaving a tissue, C_{v_t}, and entering the venous blood compartment as

$$C_{v_t} = \frac{C_t}{P_{t:b}},$$

thus, the greater the partition coefficient, the smaller the tissue's contribution to the

venous blood compartment.

Since partition coefficients are a steady-state measure, the coefficients for some relationships may be determined from other known quantities. For example, the blood:air partition coefficient of a compound can be computed from the ratio of tissue:air and tissue:blood partition coefficients. (We make use of this when determining the lung:air partition coefficient for our models.)

5.4 Assumptions and Constraints

PBPK models, in general, follow some basic rules and assumptions. The first assumption is that each organ can be modeled as a single well-stirred compartment. Further, we assume the blood-tissue exchange is flow limited, that is, the transportation of a substance into one tissue depends on its blood supply. Under this type of transport, equilibrium between blood and tissue could be reached instantly. Thus, substance distribution is assumed to occur instantaneously and homogeneously and the tissue concentration and the venous concentration exiting the tissue quickly achieve a steady-state ratio. Finally, elimination is typically assumed to occur based on either a linear or a saturable Michaelis-Menten process.

The following constraints must be adhered to for a PBPK model to be complete. First, the sum of blood flow rates to all tissues making up the model compartments should add up to the cardiac output. The weights of individual tissues making up the compartments should be less than or equal to the body weight and the sum of the weights of these tissues should equal total body weight. Last, the mass balance of the compound of interest should be maintained. That is, the total amount of a compound and its metabolites in the body at any time and the amount of compound and its metabolites eliminated by that time should add up to the initial dose administered [72].

5.5 Parameters and Estimation

For each metabolism pathway, we estimate Michaelis-Menten parameters, V_{max} (the maximum enzymatic reaction rate achieved by the kinetic system) and K_m (the measure of the affinity of the enzyme for the substance and the substrate concentration at which the reaction rate is half of V_{max}).

MATLAB (The MathWorks, Inc.) was used for simulation with `fminsearch` used for optimization. Where initial "guesses" are required, we use Blancato's published values [18]. Blancato, however, does not specify V_{max} or K_m parameter values for kidney metabolism. In this case, the author uses an arbitrary initial value for both parameters.

Data fitting and parameter estimation are done via a two-stage method (Section 8.2). That is, we estimate the model parameters for each subject in the data set then compute the mean of the estimates which is then used as the common predictor set for all data.

CHAPTER 6

Model Development/Derivation

The structure of a model is determined by the question one is attempting to answer and by the availability of data to validate the model. Not all data are sufficient for validation but a useful model is one that is mathematically and biologically realistic, can be validated, and is used to make a decision. On the path to its usefulness, a model must have an appropriate level of detail but must also be parsimonious, that is, be as simple as possible. This requires a delicate balance since increased detail decreases simplicity.

The general process of creating a mathematical model is to analyze the problem; formulate a model; fit the model to experimental data; validate the model; reformulate, refit, and revalidate the model as often as is required; then interpret and communicate the results [20].

At the forefront of model development is making assumptions about the processes needed to determine the form of the model to fit the data. The general process laid out above assumes data have already been collected and that the data is fit for use across candidate models. If this is not the case and new data is needed to fit the model, then a typical recommendation is to design a new experiment to collect new data; however, this is not always feasible since data collection can be time- and cost-intensive. We continue with the assumption that the data is sufficient for candidate models.

Once assumptions have been deemed valid (and data is presumed accurate and

unbiased), we next identify and formulate a candidate model. The model is then evaluated and verified. Is the model built right? Are the input parameters correctly represented? Is the logical structure correct? This step is followed by fitting the model and estimating unknown parameters.

Next is validating the model. Was the correct model built? The answer to this is achieved by comparing the model to the actual system behavior (i.e., comparing the model to reality) and using the differences between them to modify and improve the model. This process of formulating, fitting, and validating the model is repeated until its accuracy (the differences between the model output and actual system behavior) is acceptable.

The final step is to communicate the results, interpretations, and conclusions of the model. Its usefulness rests on its fit for purpose. For example, Does the model predict the "important" aspects of the data? Does it guide experiment design? Once the model and its results are accepted, any number of events may cause the model to be revisited. New data may be obtained or basic assumptions may be shown later to be incorrect. These examples point to the dynamic nature of models: They can—and should—change as new information is obtained.

6.1 PBPK Models

The most common methods used for pharmacokinetic analysis estimation are noncompartmental analysis (NCA), classical compartmental analysis, and physiologically-based pharmacokinetic models. Noncompartmental analysis involves calculations that are model-independent and requires fewer assumptions than model-based approaches. However, the results of NCA calculations cannot be used to extrapolate to other exposure routes or conditions [83]. Classical compartmental models overcome this limitation and allow data generated from one exposure condition to simulate a different exposure condition. For example, predictions on the behavior of a repeated exposure event can be made using a compartmental model with parameters obtained from data from a single exposure event. While these models can extrapolate across exposure conditions, they cannot be used to extrapolate to a different exposure route or to a different species since the parameters and the compartments of these models have no physiological meaning.

Physiologically-based pharmacokinetic models are classified as lumped systems. That

is, the dependent variables are a function of time alone versus distributed[1] systems where the dependent variables are functions of time and one or more other variables. In this system, processes are lumped together based on time, location or a combination of the two.

6.2 PBPK Model Parameters

A number of parameters are required when considering a PBPK model. Specifically, we include the use of physiochemical tissue:blood partition coefficients, $P_{t:b}$, and a blood:air partition coefficient, $P_{b:a}$. These parameters are species-, substance-, and tissue-specific. The Michaelis-Menten kinetics parameters, V_{max} and K_m, are dependent upon the metabolic pathway(s) primarily involved in biotransformation.

Other parameters in a PBPK model are purely physiological as well as species-specific. These include body weights, BW/bw; tissue weights/volumes, vol_t, typically given as fractions of body weight; and tissue blood flow rates, Q_i, typically given as fractions of total cardiac output, Q_{tot}, measured in liters per hour (L/hr).

The physiological parameters are well-documented in the literature. These values were almost exclusively obtained from Brown, et al. [25]. We note that body weights are often provided in units of mass but tissue values are often represented as volumes. These terms appear to be interchangeable in most literature but a major assumption for PBPK models is that tissues exhibit "unit density." Thus, one kilogram of mass is equivalent to one liter of volume for tissues.

The physiochemical values for MTBE and TBA were obtained from various sources (e.g., [63, 64]). See Tables B.1 and B.2 for values and descriptions.

6.3 PBPK Model Equations

Section 5.1 covers the principles and structure of a PBPK model. We now turn our attention to building the PBPK model by obtaining deterministic model equations that represent the biological properties of the model. We discussed earlier that target tissues and other representative compartments should be identified as well as routes of exposure.

[1]Distributed systems require solutions to partial differential equations and are rarely seen in the pharmacokinetics field.

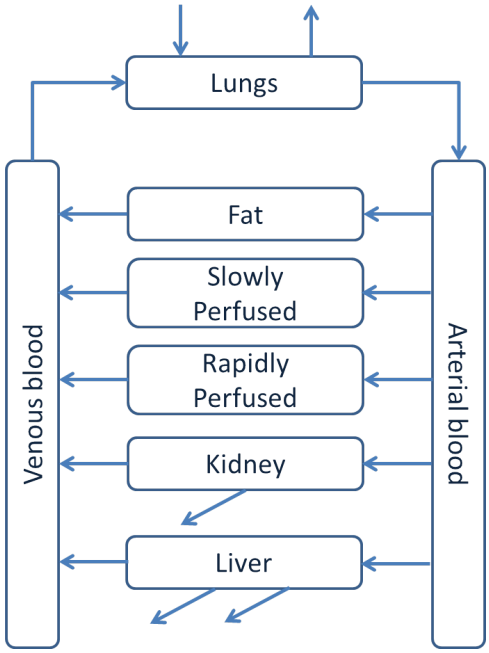


Figure 6.1: Eight-compartment PBPK model

We incorporate appropriate physiological properties of the system such as tissue volume, blood flows, and clearance pathways for the substance and its metabolites. Finally, we include relevant biochemical constants, those specific to the substance under study, namely, tissue affinities or partition coefficients.

In practice, model design should include only compartments that add significantly to the model. Here, we start with a relatively simple model (Figure 6.1) that includes arterial and venous blood compartments, the connective conduits for other compartments; non-metabolizing compartments (adipose/fat, rapidly perfused tissues, slowly perfused tissues); a metabolizing compartment (liver, kidney); and the lung as an intake (via inhalation) compartment.

Each compartment is described with a mass-balance differential equation based on Fick's law of perfusion, a process in which the rate a substance moves across a tissue membrane is dependent upon the blood flow to that tissue. Recall, this is the perfusion-limited process briefly described in Section 5.1. This process assumes that the compartments are well-stirred. It further assumes a rapid equilibrium of tissue along with steady-state conditions. In fact, we established that the tissue:blood partition coefficients are defined as the steady-state ratio of concentration of a substance in a compartment to concentration

of the substance in the blood leaving the compartment.

6.3.1 Non-elimination Compartment

We define the rate of change for the amount of substance,

$$\frac{dA}{dt},$$

in a compartment as "what goes in" minus "what goes out." To write this as a differential equation, we start with adipose as a representative compartment (Figure 6.2). The rate of substance (mg/hr) entering the adipose tissue ("what goes in") is

$$Q_{fat}C_{art},$$

the product of the concentration of substance leaving the arterial compartment, C_{art} (mg/L), and blood flow rate to the adipose/fat compartment Q_{fat} (L/hr). The rate of

Figure 6.2: Non-elimination compartment

substance (mg/hr) exiting the adipose tissue ("what goes out") is

$$Q_{fat}C^*_{fat} = Q_{fat}C_{v_{fat}},$$

where $C_{v_{fat}}$ is the concentration of substance leaving the adipose compartment entering the venous blood compartment. This concentration is not simply the concentration of substance inside the adipose compartment, but is the concentration inside the compartment divided by the unitless tissue:blood partition coefficient for the compartment. Thus, the

concentration exiting the adipose compartment is

$$Q_{fat}\frac{C_{fat}}{P_{fat:blood}}. \tag{6.1}$$

Putting everything together, we have

$$\begin{aligned}
\frac{dA_{fat}}{dt} &= Q_{fat}C_{art} - Q_{fat}C_{v_{fat}} && (6.2)\\
&= Q_{fat}C_{art} - Q_{fat}\frac{C_{fat}}{P_{fat:blood}} && (6.3)\\
&= Q_{fat}\left(C_{art} - \frac{C_{fat}}{P_{fat:blood}}\right), && (6.4)
\end{aligned}$$

where:

- Q_{fat} is the blood flow rate to/from the adipose compartment (L/hr),
- C_{fat} is the concentration of substance inside the adipose compartment (mg/L), and
- $C_{v_{fat}}$ is the concentration of substance in venous blood leaving the adipose compartment (mg/L).

Equation (6.4), then, is the general form of a PBPK equation for a non-elimination compartment (adipose tissue, rapidly perfused tissue, slowly perfused tissue, etc.).

Notice we have an inconsistency in variables amount (A_{fat}) and concentration (C_{fat}). This is a simple remedy by noting that concentration is defined as amount per volume, thus we have

$$C_{fat} = \frac{A_{fat}}{V_{fat}} \implies A_{fat} = C_{fat}V_{fat}, \tag{6.5}$$

where V_{fat} (L) is a constant and Equation (6.4) becomes

$$\begin{aligned}
\frac{d(C_{fat}V_{fat})}{dt} &= Q_{fat}\left(C_{art} - \frac{C_{fat}}{P_{fat:blood}}\right) && (6.6)\\
V_{fat}\frac{dC_{fat}}{dt} &= Q_{fat}\left(C_{art} - \frac{C_{fat}}{P_{fat:blood}}\right) && (6.7)\\
\frac{dC_{fat}}{dt} &= \frac{Q_{fat}}{V_{fat}}\left(C_{art} - \frac{C_{fat}}{P_{fat:blood}}\right), && (6.8)
\end{aligned}$$

where this final form of the equation is written in terms of concentrations.

6.3.2 Elimination Compartment

To derive the equations for an elimination compartment (Figure 6.3), liver in this case, we follow similar steps as in Section 6.3.2. The "what goes in" (mg/hr) is again

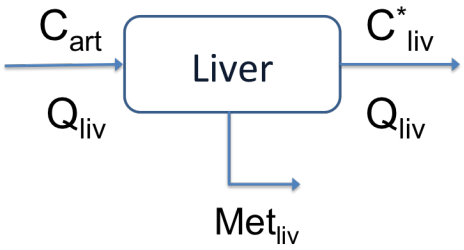


Figure 6.3: Elimination compartment

$$Q_{liv}C_{art}.$$

The "what goes out" (mg/hr) starts with

$$Q_{liv}C^*_{liv} = Q_{liv}C_{v_{liv}},$$

but must also consider an elimination process. The elimination process used exclusively in this paper is Michaelis-Menten kinetics. (See Appendix C for derivation details.) The Michaelis-Menten rate of metabolism, Met_{liv} (mg/hr), is given as

$$Met_{liv} = \frac{V_{max}C_{v_{liv}}}{K_m + C_{v_{liv}}}, \tag{6.9}$$

where:

- V_{max} is the maximum enzymatic reaction rate achieved by the kinetic system (mg/hr),
- K_m is the substrate concentration at which half the enzyme's active sites are occupied by substrate and is also the substrate concentration at which the reaction rate is

half of V_{max} (mg/L); and

- $C_{v_{liv}}$ is the concentration of substance in venous blood leaving the liver compartment (mg/L).

Putting it all together, we have

$$\frac{dA_{liv}}{dt} = Q_{liv}C_{art} - Q_{liv}C_{v_{liv}} - Met_{liv} \tag{6.10}$$

$$= Q_{liv}C_{art} - Q_{liv}C_{v_L} - \frac{V_{max}C_{v_L}}{K_m + C_{v_L}} \tag{6.11}$$

$$= Q_{liv}\left(C_{art} - C_{v_L}\right) - \frac{V_{max}C_{v_L}}{K_m + C_{v_L}}, \tag{6.12}$$

where C_{v_L} is similarly defined as in Equation (6.1) where we divide instead by the unitless liver:blood partition coefficient and consistency in variables is obtained similarly as in Equation (6.5). If there were other routes of elimination, these would be included as "what goes out."

6.3.3 Intake Compartment

Routes of exposure come in various forms: inhalation, oral/ingestion, dermal, injection (e.g., intravenous), and others. These routes dictate the inclusion of the compartments which are directly affected by the exposure. For example, one would include a skin compartment for dermal exposure or include the gastrointestinal tract for oral exposure. For our model, we have only inhalation to consider and thus include the lung in our model.

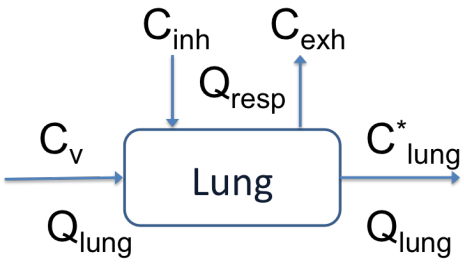


Figure 6.4: Intake compartment

In Figure 6.4, the "what goes in" are the concentration of the substance being inhaled with a ventilation/respiration rate of Q_{resp},

$$Q_{resp}C_{inh},$$

and the concentration of the substance in the venous blood compartment carried by the lung's blood flow rate, Q_{lung} (which is equal to the total cardiac output rate, Q_{tot})

$$Q_{lung}C_v.$$

"What goes out" is the concentration of substance exhaled again with a ventilation rate of Q_{resp},

$$Q_{resp}C_{exh},$$

and the concentration of substance entering the arterial blood compartment

$$Q_{lung}C^*_{lung}.$$

As we saw with other tissue compartments, the concentration leaving the lung is based on the partition coefficient for that tissue, thus we have

$$Q_{lung}\frac{C_{lung}}{P_{lung:blood}}.$$

We turn our attention the exhaled concentration, C_{exh}. It is again the concentration of substance leaving the lung but rather than having blood on the other side of the tissue membrane, we have air. Thus, we require a lung:air or blood:air partition coefficient (Sections 5.3 and 6.2).

Putting it all together, the equation representing the lung compartment is

$$\frac{dA_{lung}}{dt} = Q_{resp}C_{inh} + Q_{lung}C_v - Q_{lung}\frac{C_{lung}}{P_{lung:blood}} - Q_{resp}\frac{C_{lung}}{P_{lung:air}}. \tag{6.13}$$

The uptake in the lung and distribution to the arterial compartment equilibrates rapidly [5, 12] and one may see the equation derivation for this exchange written for the arterial

compartment, rather than the lung, as follows:

$$\frac{dA_{art}}{dt} = Q_{resp}C_{inh} + Q_{tot}C_v - Q_{tot}C_{art} - Q_{resp}\frac{C_{art}}{P_{blood:air}}. \tag{6.14}$$

Assuming a rapid equilibrium in the arterial compartment means there is no change over time in the amount of substance in the arterial blood, thus $\frac{dA_{art}}{dt} = 0$ and one may compute the concentration of the arterial blood compartment, C_{art}, algebraically

$$C_{art} = \frac{Q_{resp}C_{inh} + Q_{tot}C_v}{Q_{tot} + \frac{Q_{resp}}{P_{blood:air}}}. \tag{6.15}$$

This form of the equation can be found in many PBPK articles with inhalation as a route of exposure.

CHAPTER 7

PBPK Data

7.1 Cain, 1.7 ppm

The subjects for this study consisted of two males and two females, aged 18-26 years. Volunteers were exposed to 1.7 ppm of substance in air for 1 hour. Six venous blood samples per subject were collected during exposure (at t = 2, 5, 10, 20, 30, and 60 minutes). Seven venous blood samples were collected over 1.5 hours post-exposure (at t_{post} = 2, 5, 10, 20, 40, 60, and 90 minutes or t = 62, 65, 70, 80, 100, 120, and 150 minutes). The exposures were carried out in two controlled environmental chambers measuring 18.5m^3 each.

Measurements were performed by gas chromatography/mass spectroscopy (GC/MS) using purge and trap extraction of MTBE and TBA from blood.

7.2 Pleil, 3 ppm

The subjects for this study consisted of fourteen males, aged 20-30. Volunteers were exposed to 3 ppm (10.8 mg/m^3) of substance in air for 1 hour. Venous blood samples were collected over 24 hours before, during and after exposure (one baseline measurement and others at t = 5, 15, 30, 45, 60, 65, 75, 90, 120, 180, 240, 360, and 1440 minutes).

Exhaled breath samples were collected from seven of the 14 subjects and were obtained at the same time as blood samples. The exposures were carried out in a converted body plethysmograph measuring 1.75m^3.

As with the Cain study, measurements were performed by GC/MS using purge and trap extraction of MTBE and TBA from blood. Breath analysis was performed via GC/MS using protocols derived from EPA Method TO-14 which tests ambient air for toxic organic compounds.

CHAPTER 8

Model Analysis

8.1 Parameter Estimation

A classical inverse problem is one given by the system of equations,

$$f(\mathbf{x}; \mathbf{q}) = \mathbf{d}, \tag{8.1}$$

where a vector of system parameters, $\mathbf{q}$, characterizing the model is estimated given data, $\mathbf{d}$; f is the forward function relating $\mathbf{q}$ and $\mathbf{d}$; and $\mathbf{x}$ is the vector of state variables for the system [10, 62, 88]. Forward modeling allows us to make predictions on the results of measurements on some observable parameters (i.e., find $\mathbf{d}$ given $\mathbf{q}$) whereas parameter estimation is the process of using empirical data to approximate model parameter values (i.e., find $\mathbf{q}$ given $\mathbf{d}$). It is the basis for solving inverse problems, assuming a solution exists.

One of the most common techniques for parameter estimation is the least squares method. It is an iterative process requiring an initial set of values that leads to estimates close to the real values of the parameters. We apply a criteria function based on the differences between the observed values and the predicted values obtained from using the estimated parameters in the model. These differences are called *residuals* and are

discussed further in Section 8.3.1 but are defined here as

$$\boldsymbol{e} = \mathbf{d} - f(\mathbf{x}; \tilde{\mathbf{q}}), \tag{8.2}$$

where $\tilde{\mathbf{q}}$ is the vector of estimated parameters and $f(\mathbf{x}; \tilde{\mathbf{q}})$ is the vector of predicted values that ultimately lead to an optimum $\tilde{\mathbf{q}}$. At each step, we seek to find a $\tilde{\mathbf{q}}$ that minimizes the distance between the observed values and the predicted values, $\|\boldsymbol{e}\|^2 = \|\mathbf{d} - f(\mathbf{x}; \tilde{\mathbf{q}})\|^2$. Thus, the least squares objective (or cost) function is defined as

$$Cost = \sum_i \left[d_i - f(x_i; \tilde{\mathbf{q}})\right]^2 \tag{8.3}$$

and this is the quantity we wish to minimize. In practice, we use a variation of this equation,

$$Cost = \sum_i \left[\log d_i - \log f(x_i; \tilde{\mathbf{q}})\right]^2, \tag{8.4}$$

to reduce the dispersion of the data set and reduce the effect of extreme values.

We generally assume that there are no constraints on the model's parameter values, $\mathbf{q}$; they can range from $-\infty$ to ∞. While this may be valid for some parameters, it is not true for our pharmacokinetic parameters. For example, all rate constants (Q, V_{max}, etc.) are positive values. Additionally, the Michaelis-Menten affinity constant, K_m, is a positive value. Therefore, these constraints must be considered in the data-fitting process. We address this in practice by using MATLAB's constrained optimization option for `fminsearch`, limiting all applicable estimated parameters to $[0, \infty)$ and, in a few cases, limiting some parameter values to those that are biologically plausible, $[0, q_{max}]$.

8.2 Data Fitting Via Two-Stage Method

In pharmacokinetic studies, data are often obtained from multiple subjects. Individual data can be used in a model to obtain parameter estimates for each subject. Repeatedly applying the model will yield a vector of parameters for each subject. The logical next step is to determine statistical properties about the collection of parameters to make a generalization about the population. This is the general idea behind the two-stage approach for data-fitting:

1. First, estimate model parameters for each individual in the data set.

2. Then, use individual estimates to compute statistical properties about the population.

These values may then be used to validate the model against the population, to draw conclusions about the population, or any number of other uses.

Given a collection of parameter estimates, natural estimators for the population are mean and variance. Let N be the number of subjects or pharmacokinetic profiles; let Λ be the parameter set being estimated for each subject with p parameters per subject/profile; and let $\mathbf{q}_i$ be the estimated parameters for each subject, $i = 1 \dots N$. If we consider a matrix,

$$\hat{Q} = [\mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_N],$$

where the columns are $\mathbf{q}_i$ for each subject, with entries $q_{i,j}$, then the p rows each represent a parameter estimate from each model run. This $p \times N$ matrix can then be used to compute the mean,

$$\mu_i = \frac{1}{N} \sum_{j=1}^{N} q_{i,j} \quad i = 1, \dots, p, \tag{8.5}$$

and the variance,

$$\sigma_i^2 = \frac{1}{N} \sum_{j=1}^{N} (q_{i,j} - \mu_i)^2 \quad i = 1, \dots, p, \tag{8.6}$$

with standard deviation σ_i. In estimating the variance of a population, it is generally assumed that each entry involved in the variance's computation is known with certainty. This is not the case here and each individual's parameter has some degree of bias associated with it. As a result, the two-stage method does not take into account the variability in the parameter estimation and thus inflates any variance based on these estimates [13, 20]. Mixed-effect models (Section 10.2.1) are not hindered by this variance inflation and tend to produce both mean and variance estimates that are unbiased.

8.3 Model Validation

8.3.1 Residuals and Residual Analysis

Residuals represent part of the data that a model could not reproduce. Their values and trends allow us to assess the validity of a model and improve upon it. They are the difference between the observed y-value and the predicted y-value of a model approximating the data provided:

$$e_i = y_i - f(\mathbf{x}_i; \mathbf{q}), \tag{8.7}$$

where y_i is the i^{th} response of observed values, $\mathbf{q}$ is the vector of estimated parameters, f is the regression function and $\mathbf{x}_i$ denotes the explanatory/predictor variables.

Residuals, like other error measures, are assumed to be normal and independently distributed. Further, they have zero mean and some constant variance. When the residual values of a model are plotted against the predicted values, indications of an accurate model are that the residuals are symmetrically distributed (about zero) throughout the range of fitted values and there are no clear patterns or trends. Ultimately, for a regression model to be "valid," we require randomness and unpredictability of the residuals. If these are not present, it would suggest that the model may need improvement. For example, there may be missing variables or existing terms in the model are missing some type of interaction.

Most residual analyses involve examination of graphs that allow you to identify patterns that are either poorly fit by the model or have a strong influence upon the estimated parameters. Interpretations of these graphical diagnostics are helpful to understand potential problems with the model.

Table 8.1 [20] lists plots that are often examined for systematic trends or departures from the expected values. Any trends indicate model misspecification, correlation of the residuals, or inaccurate prediction of the time trend.

8.3.2 Akaike Information Criterion

Goodness-of-fit is a measure of how well a model approximates data. In the case of multiple models, we are interested in determining which model provides the best goodness-of-fit. The Akaike Information Criterion (AIC) [4] is a model selection tool that ties together Kullback-Liebler's (1951) divergence measure to represent the amount of information

Table 8.1: Types of Graphical Analysis Plots

Vertical axis ("y" axis)	Horizontal axis ("x" axis)	What to look for
Predicted value	Residuals	Should show no systematic trend in the residuals with data appearing as a shotgun blast
Absolute or squared residuals	Predicted values	Should show no systematic trend in the residuals with data appearing as a shotgun blast
Residuals	Time	Should show random variation centered around zero and show no systematic trends
Observed values	Predicted values	Should show random variation about the line of unity
e_k, $k = 1 \ldots (n-1)$	e_{k+1}	Should show no trend in the "lag" plot

lost when approximating data and the maximum value of the log-likelihood[1] function for a model. The resulting combination is a measure of the relative goodness-of-fit of a statistical model.

In the general case, AIC is defined as

$$AIC = -2L_{max} + 2p, \tag{8.8}$$

where L_{max} is the maximized value of the log-likelihood function for the estimated model and p is the number of estimated parameters. The $2p$ here is a penalty term that effectively avoids overfitting.

In practice, we find the AIC index for each of our candidate models, then select the model with the smallest AIC value among all models. It is important to note that the AIC index does not provide any indication of whether a model is "good" or "bad;" it simply provides a means to quantitatively compare a pool of candidate models.

For analyses with least squares regression,

$$AIC = n \ln\left(\frac{\hat{\sigma}^2}{n}\right) + 2p, \tag{8.9}$$

[1]Likelihood theory - the probability that one observes a set of data given that a particular model were true

where n is the sample size (total number of data points), and

$$\hat{\sigma}^2 = \sum_{i=1}^{n} \left(y_i - f(\mathbf{x}_i; \mathbf{q})\right)^2 . \tag{8.10}$$

We modify this slightly by scaling by the maximum value of the experimental data over all time points, $\max_j(y_j)$, thus

$$\hat{\sigma}^2 = \frac{1}{\max_j(y_j)} \sum_{i=1}^{n} \left(y_i - f(\mathbf{x}_i; \mathbf{q})\right)^2 , \tag{8.11}$$

for $j \in [1, \ldots, n]$.

8.4 Existence and Uniqueness of Solutions

For any system of ordinary differential equations for which we seek a solution, we must determine if there first exists a solution and we must also determine if that solution is unique. We begin with an initial-value problem

$$\mathbf{x}' = f(t, \mathbf{x}), \tag{8.12}$$

$$\mathbf{x}(t_0) = \mathbf{x}_0. \tag{8.13}$$

Our system is autonomous, thus $f(t, \mathbf{x}) = f(\mathbf{x})$. This system will have a solution and that solution will be unique if it satisfies the Lipschitz Uniqueness Theorem [3, 9, 19].
Definition 1. $f(\mathbf{x})$ is *Lipschitz continuous* if there exists a constant L such that

$$\|f(\mathbf{x}) - f(\mathbf{y})\| \leq L\|\mathbf{x} - \mathbf{y}\|, \tag{8.14}$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

Theorem 1 (Lipschitz Uniqueness Theorem). Suppose $f(\mathbf{x})$ is *Lipschitz continuous* in some open set containing the initial data $\mathbf{x}_0$. Then a unique solution $\mathbf{x}(t)$ of Equation (8.12) exists in some interval $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$, satisfying ((8.12)) and ((8.13)).

We begin our proof with a sample model (Figure 8.1). The model has at least one representative compartment for any of the four models we have. It includes:

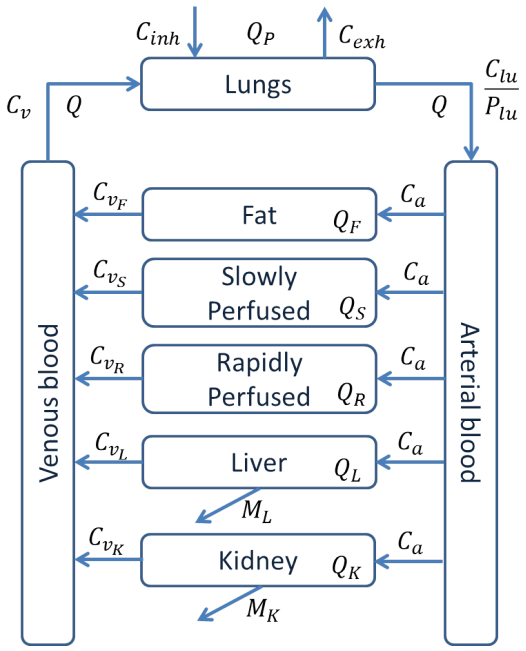


Figure 8.1: Representative PBPK system

- a compartment that has a constant input and an output (lung).
- a compartment that feeds almost all other compartments (arterial blood)
- a non-metabolizing compartment (e.g., slowly perfused, rapidly perfused, fat)
- a metabolizing compartment (e.g., liver, kidney); and
- a compartment that collects from almost all other compartments (venous blood).

We assume there is a partition coefficient, P_i, associated with each compartment. There is a blood flow rate, Q_i, for each compartment and a respiratory rate, Q_p for the lung compartment. Compartment volumes would include a constant for each equation but there is no loss of generality by ignoring them (or, rather, by assuming $vol_i = 1$). Q_i and P_i are non-zero, non-negative, and factored, where applicable, into the coefficients A_i, B_{ij}, C_{ik}, D_{il}, E_{il}, F_{im}, and G_{im} in the generalized system of equations (Equation (8.15)).

Finally, x_i is non-negative since it represents compartment concentrations.

$$vol_i \frac{dx_i}{dt} = A_i + \sum_{j=1}^{n} B_{ij} x_j - \sum_{k=1}^{n} C_{ik} x_k - \sum_{p=1}^{n_p} \sum_{l=1}^{n} \frac{D_{il} x_l}{K_{pl} + E_{il} x_l} + \sum_{p=1}^{n_p} \sum_{m=1}^{n} \frac{F_{im} x_m}{K_{pm} + G_{im} x_m} \quad (8.15)$$

To continue, let $f_i(\mathbf{x}) = vol_i \frac{dx_i}{dt}$, then

$$\begin{aligned}
&|f_i(\mathbf{x}) - f_i(\mathbf{y})| \\
&\quad = \left| \left(A_i + \sum_{j=1}^{n} B_{ij} x_j - \sum_{k=1}^{n} C_{ik} x_k - \sum_{p=1}^{n_p} \sum_{l=1}^{n} \frac{D_{il} x_l}{K_{pl} + E_{il} x_l} + \sum_{p=1}^{n_p} \sum_{m=1}^{n} \frac{F_{im} x_m}{K_{pm} + G_{im} x_m} \right) \right. \\
&\qquad \left. - \left(A_i + \sum_{j=1}^{n} B_{ij} y_j - \sum_{k=1}^{n} C_{ik} y_k - \sum_{p=1}^{n_p} \sum_{l=1}^{n} \frac{D_{il} y_l}{K_{pl} + E_{il} y_l} + \sum_{p=1}^{n_p} \sum_{m=1}^{n} \frac{F_{im} y_m}{K_{pm} + G_{im} y_m} \right) \right| \\
&\quad = \left| \sum_{j=1}^{n} B_{ij} (x_j - y_j) - \sum_{k=1}^{n} C_{ij} (x_k - y_k) - \sum_{p=1}^{n_p} \sum_{l=1}^{n} D_{il} \left(\frac{x_l}{K_{pl} + E_{il} x_l} - \frac{y_l}{K_{pl} + E_{il} y_l} \right) \right. \\
&\qquad \left. + \sum_{p=1}^{n_p} \sum_{m=1}^{n} F_{im} \left(\frac{x_m}{K_{pm} + G_{im} x_m} - \frac{y_m}{K_{pm} + G_{im} y_m} \right) \right| \\
&\quad = \left| \sum_{j=1}^{n} B_{ij} (x_j - y_j) - \sum_{k=1}^{n} C_{ij} (x_k - y_k) - \sum_{p=1}^{n_p} \sum_{l=1}^{n} D_{il} \left[\frac{K_{pl} (x_l - y_l)}{(K_{pl} + E_{il} x_l)(K_{pl} + E_{il} y_l)} \right] \right. \\
&\qquad \left. + \sum_{p=1}^{n_p} \sum_{m=1}^{n} F_{im} \left[\frac{K_{pm} (x_m - y_m)}{(K_{pm} + G_{im} x_m)(K_{pm} + G_{im} y_m)} \right] \right| \\
&\quad \leq \sum_{j=1}^{n} |B_{ij} (x_j - y_j)| + \sum_{k=1}^{n} |C_{ij} (x_k - y_k)| + \sum_{p=1}^{n_p} \sum_{l=1}^{n} \left| D_{il} \left[\frac{K_{pl} (x_l - y_l)}{K_{pl}^2} \right] \right| \\
&\qquad + \sum_{p=1}^{n_p} \sum_{m=1}^{n} \left| F_{im} \left[\frac{K_{pm} (x_m - y_m)}{K_{pm}^2} \right] \right| \\
&\quad \leq \sum_{j=1}^{n} |B_{ij}| \, |x_j - y_j| + \sum_{k=1}^{n} |C_{ij}| \, |x_k - y_k| + \sum_{p=1}^{n_p} \sum_{l=1}^{n} \left| \frac{D_{il}}{K_{pl}} \right| |x_l - y_l| \\
&\qquad + \sum_{p=1}^{n_p} \sum_{m=1}^{n} \left| \frac{F_{im}}{K_{pm}} \right| |x_m - y_m| \, .
\end{aligned} \quad (8.16)$$

By definition $\|\mathbf{x} - \mathbf{y}\| = \sum_i^n |x_i - y_i|$, thus any $|x_i - y_i| \leq \|\mathbf{x} - \mathbf{y}\|$. Continuing ((8.16)),

$$
\begin{aligned}
&|f_i(\mathbf{x}) - f_i(\mathbf{y})| \\
&\quad \leq \sum_{j=1}^{n} |B_{ij}| \, \|\mathbf{x}-\mathbf{y}\| + \sum_{k=1}^{n} |C_{ij}| \, \|\mathbf{x}-\mathbf{y}\| + \sum_{p=1}^{n_p} \sum_{l=1}^{n} \left| \frac{D_{il}}{K_{pl}} \right| \|\mathbf{x}-\mathbf{y}\| + \sum_{p=1}^{n_p} \sum_{m=1}^{n} \left| \frac{F_{im}}{K_{pm}} \right| \|\mathbf{x}-\mathbf{y}\| \\
&\quad \leq \left(\sum_{j=1}^{n} |B_{ij}| + \sum_{k=1}^{n} |C_{ij}| + \sum_{p=1}^{n_p} \sum_{l=1}^{n} \left| \frac{D_{il}}{K_{pl}} \right| + \sum_{p=1}^{n_p} \sum_{m=1}^{n} \left| \frac{F_{im}}{K_{pm}} \right| \right) \|\mathbf{x}-\mathbf{y}\| \, .
\end{aligned}
\tag{8.17}
$$

In Equation (8.17), B_{ij}, C_{ik}, D_{il}, and F_{im} are all bounded. The Michaelis-Menten constants, K_{pl} and K_{pm}, are positive, non-zero values, hence $\frac{D_{il}}{K_{pl}}$ and $\frac{F_{im}}{K_{pm}}$ are both bounded and Equation (8.17) satisfies the Lipschitz condition,

$$|f_i(\mathbf{x}) - f_i(\mathbf{y})| \leq L\|\mathbf{x}-\mathbf{y}\|,$$

with

$$L = \sum_{j=1}^{n} |B_{ij}| + \sum_{k=1}^{n} |C_{ij}| + \sum_{p=1}^{n_p} \sum_{l=1}^{n} \left| \frac{D_{il}}{K_{pl}} \right| + \sum_{p=1}^{n_p} \sum_{m=1}^{n} \left| \frac{F_{im}}{K_{pm}} \right| .$$

Thus, the system of ordinary differential equations has a solution and that solution is unique.

8.5 Sensitivity Analysis

When working with kinetic models, it is desirable to determine which parameters have the greatest effect on the model output, that is, which parameters are most influential (as well as those that have little impact). Sensitivity analysis, specifically local sensitivity analysis, is a method that allows us to determine this influence by studying the change in a model's output variable with respect to its parameters or input variables.

We assume that the model is a sufficiently reasonable description of the process we are modeling. We then compute the rate of change of the model responses with respect to the parameters. These values are the sensitivity coefficients and are required to perform parameter estimation procedures to completion. Since we wish to measure the change in model output associated with the change (perturbation) in model input (e.g., in parameter values), we determine the effects of parameters $\mathbf{q}$ on model states $\mathbf{x}$, that is, we find $\frac{\partial \mathbf{x}}{\partial \mathbf{q}}$.

Consider the initial value problem

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{f}(t, \mathbf{x}(t), \mathbf{q}), \qquad \mathbf{x}(t_0) = \mathbf{x}_0 \tag{8.18}$$

$$\mathbf{y}(t) = \mathbf{g}(t, \mathbf{x}(t), \mathbf{q}), \tag{8.19}$$

where $\mathbf{x}$ is an n-dimensional vector of state variables, $\mathbf{q}$ is the p-dimensional vector of system parameters, $\mathbf{y}$ the measurement vector, and t is the independent variable. Then

$$\frac{d}{dt}\frac{\partial \mathbf{x}}{\partial \mathbf{q}} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \mathbf{q}} + \frac{\partial \mathbf{f}}{\partial \mathbf{q}}. \tag{8.20}$$

Define the sensitivity coefficients,

$$\mathbf{S}(t) = \frac{\partial \mathbf{x}}{\partial \mathbf{q}}, \tag{8.21}$$

then (8.20) becomes

$$\frac{d}{dt}\mathbf{S}(t) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\mathbf{S}(t) + \frac{\partial \mathbf{f}}{\partial \mathbf{q}} \tag{8.22}$$

with initial condition $\mathbf{S}(t_0) = \mathbf{0}$ since the initial conditions of the model are not dependent upon the model parameters [96]. We now have a differential equation for the sensitivity coefficients where both $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$ and $\frac{\partial \mathbf{f}}{\partial \mathbf{q}}$ can be computed using automatic differentiation, a technique for systematically studying sensitivities that is completed by applying the chain rule repeatedly to numerically obtain derivatives of a function.

8.5.1 Sensitivity Rankings

Once we have computed the sensitivity coefficients, we continue our effort to determine the most (and least) influential model parameters by ranking the sensitivity coefficients. Starting with $\frac{\partial x_i}{\partial q_j}$, we multiply by q_j to non-dimensionalize the parameters. We then divide by the maximum state variable, $\max x_i$, over t_0 to t_f. Taking the L^2 norm of this expression gives the normalized sensitivity rankings, $C_{i,j}$, for the model

$$C_{i,j} = \left\| \frac{\partial x_i}{\partial q_j}\frac{q_j}{\max x_i} \right\|_2^2 = \int_{t_0}^{t_f} \left| \frac{\partial x_i}{\partial q_j}\frac{q_j}{\max x_i} \right|^2 dt \in \mathbb{R}. \tag{8.23}$$

We rank these parameter sensitivity values from largest to smallest, in order of most to least important in the model.

8.5.2 Sensitivity Coefficient Matrix

We begin by defining parametric sensitivity coefficients of the yth output of the pth parameter as [97]:

$$\frac{\partial y_r(t_n, x_n, q)}{\partial q_p}; \qquad p = 1, 2, \ldots P. \tag{8.24}$$

These coefficients indicate the magnitude of change in the response y due to perturbations in the values of the parameters and are analytically obtained as detailed in Section 8.5.

We form a sensitivity coefficient matrix, $\mathbf{S}$, from the individual coefficients:

$$S \equiv \begin{bmatrix} \left.\frac{\partial y_1}{\partial q_1}\right|_{t=t_1} & \cdots & \left.\frac{\partial y_1}{\partial q_P}\right|_{t=t_1} \\ \vdots & \ddots & \vdots \\ \left.\frac{\partial y_R}{\partial q_1}\right|_{t=t_1} & \cdots & \left.\frac{\partial y_R}{\partial q_P}\right|_{t=t_1} \\ \left.\frac{\partial y_1}{\partial q_1}\right|_{t=t_2} & \cdots & \left.\frac{\partial y_1}{\partial q_P}\right|_{t=t_2} \\ \vdots & \ddots & \vdots \\ \left.\frac{\partial y_R}{\partial q_1}\right|_{t=t_N} & \cdots & \left.\frac{\partial y_R}{\partial q_P}\right|_{t=t_N} \end{bmatrix}. \tag{8.25}$$

Scaling each element of the sensitivity matrix is required to ensure that the elements are dimensionally consistent and that some parameters do not dominate their subsequent ranking due to their large numerical values. Prior to estimability calculations, the coefficients are scaled as

$$\frac{\hat{q}_p}{\hat{y}_r|_{t_n}} \left.\frac{\partial y_r}{\partial q_p}\right|_{t_n}, \tag{8.26}$$

where $\hat{q}_p$ is either an initial guess or the best value of the pth parameter from a previous parameter estimation study, and $\hat{y}_r|_{t_n}$ is the value of the rth response variable at time t_n.

If an initial guess for a parameter is zero, an altered scaling procedure is required [51]. For our purposes, this is a scaling by $\max \hat{y}_r$, the maximum value of the rth response variable over t_0 to t_f. Henceforth, the sensitivity matrix used for identifiability and estimability analyses will contain entries in the form of Equation (8.26) scaled by $\max \hat{y}_r$.

Model parameters can be estimated if the sensitivity coefficients are not linearly dependent. If the coefficients defined in Equation (8.24) (alternatively Equation (8.26))

form columns of $\mathbf{S}$ that are linearly independent for a particular set of parameters, $\tilde{q}$, then there exists only one unique set of optimal parameter values. If, however, the sensitivity functions are linearly dependent at all values of q, then the model parameters are cannot be estimated. [97]

8.6 Identifiability

In biological systems, the model complexity along with the difficulty to make informative measurements are often at the core of parameter identifiability problems. A model is *identifiable* if and only if there is an unique input-output behavior for every parameter set. If a model has more parameters than what can be uniquely determined under ideal experimental conditions then it is *unidentifiable*.

Model predictions depend on the model parameters, some which must be estimated from experimental data. For inverse problems, estimation of parameters depends on the structure of the model as well as on the experiment performed to gather data. A model may perfectly fit the data but have unidentifiable parameters. However it still may be possible to uniquely identify a subset of parameters; not all the parameters of an unidentifiable model are unidentifiable [53]. Generally speaking, "identifiability" answers whether unique model parameters can be obtained given certain data. There are, however, four classes of identifiability: structural, theoretical, numerical, and practical [39, 53, 76, 81].

Structural identifiability is related to the structure of the underlying mathematical model and does not take into account the available experimental data. It is concerned with the algebraic properties of the structural model and whether the parameters of a model can be exactly identified from a given experiment with perfect input-output data. A structurally unidentifiable model means that the problem is *ill-posed*, whereby we can find different parameter vectors q_i and q_i' that produce the same structural predictions, i.e., $q_i \neq q_i'$ with $f(t; q_i) = f(t; q_i')$ for any $t \geq 0$. A system's structural identifiability can be determined by analyzing the modeling equations before an experiment is conducted.

Theoretical identifiability is the case where different parameter values lead to non-identical probability distributions. Different values of the model parameters generate different probability distributions of the observable variables.

Numerical identifiability is defined as the inability to obtain proper parameter estimates

from the data even if the experiment is structurally identifiable. When the data matrix (Section 8.8.1) is close to singular, we experience an inability to obtain good paramater estimates from the data. This condition is possible even when structural identifiability exists for the experiment and is associated with a deficiency in the richness of the data, referred to as ill-conditioned[2] data.

Practical identifiability is the ability to estimate a given set of parameters along with the accuracy that can be hoped (within the context of the study). It depends on the model structure and is related to the experimental conditions together with the quality and quantity of the measurements. Then, a parameter that is structurally identifiable may be practically unidentifiable given a limited amount and quality of experimental data. Additionally, a model may be theoretically identifiable but the design of the experiment may make parameter estimation difficult and imprecise. Further, if theoretical identifiability is assessed, the practical identifiability is not guaranteed since practical and theoretical parameter identifiability don't necessarily correspond to one another due to poor data quality and/or quantity that may exist and which directly affects practical identifiability.

As mentioned briefly earlier, if a model has some, but not all, unidentifiable parameters, it still may be possible to uniquely identify a subset of parameters. We discuss parameter subset selection more in Section 8.9. If it is not possible to generate additional datasets, a good mitigation scheme is to systematically eliminate the nonsensitive parameters from the estimation. This method is not completely fail-safe since parameter correlation also contributes to identifiability issues. For example, the Michaelis-Menten parameter pairs are rarely identifiable from limited data [58].

[2]Data is said to be *ill-conditioned* when the regressor variables of a model are highly correlated with each other. The condition number provides a measure of the relationship among these regressors and tells the extent to which an uncertainty in x is magnified by $f(x)$. A value of 1 indicates that the function's relative error is identical to the relative error in x. Numerically, functions with very large condition numbers are ill-conditioned and parameter estimations that use ill-conditioned data may yield unreliable parameter estimates with large standard errors [81]. Further, the literature suggests evidence of collinearity when condition numbers exceed 15 and corrective action is necessary when the condition number is around 30 or more.

8.7 Estimability

Identifiability and estimability are closely linked but are separate ideas. Identifiability is answered by the question "Are the observations sufficient to specify the parameters?" whereas estimability considers the effects of errors in measurements on the parameter estimates [45]. Specifically, it is the ability to compute parameters accurately from a given data set and experimental conditions [97] and is an often neglected first step in mathematical modeling. Both identifiability and estimability are important analysis steps that lead to estimations which include only the most important parameters to the model(s) under assessment.

To determine if the parameters can be estimated from a data set, we perform an estimability analysis from which there are three possible outcomes [38, 81]:

- The model parameters can be estimated uniquely, i.e., the model is globally identifiable;
- A finite number of alternative estimates of model parameters is possible; the model is locally identifiable (see Appendix D for Shotwell's [86] computational treatment of local identifiability); and
- An infinite number of model parameter estimates are possible and the model is unidentifiable from the data (the over-parameterized case, where a model has more parameters than what can be uniquely determined under ideal experimental conditions [46]).

In theory one tests identifiability first and then proceeds to estimate the identifiable parameters. In many cases, not all parameters can be estimated, leaving a subset of parameters that have actual influence in the response variables of the model. Performing this step first may give clues about whether the model structure is adequate, whether re-parameterization is needed, or whether the experimental design used to gather data is sufficient. For the latter, identifying estimable parameters before experimentation begins can lead to experiment designs that allow those parameters to be more precisely estimated.

8.8 Identifiability and Estimability Assessment Tools

There are a number of ways to assess estimability. One is to conduct a sensitivity analysis for a given input sequence which involves the computation of a sensitivity coefficient matrix whereby all system parameters can be estimated if the columns of the matrix are full rank [97]. Another is to plot sensitivity coefficients as functions of time [14]. One other is to compute the eigenvectors corresponding to small eigenvalues of the Fisher Information Matrix (Section 8.8.1) [46, 95]. When the normalized eigenvectors are ordered left to right corresponding to the largest-to-smallest eigenvalue, elements in the eigenvectors suggest combinations of parameters that go, in order, from easier to harder to estimate. Specifically, the largest magnitude component in the eigenvectors correspond to the most identifiable parameter.

The following sections describe additional ways in which identifiability and estimability may be assessed.

8.8.1 Fisher Information Matrix

Identifiability and estimability analyses are carried out by evaluating the Fisher Information Matrix (FIM) which is a digest of the sensitivity information and can be used to select identifiable parameter combinations. A model or a set of parameters may be identifiable but the Fisher matrix may be near singular, resulting in poor estimability. One good suggestion of this is a small overall sum of squares, indicating a good fit to the observations but with a very large standard deviation of the estimates of some parameters. In this case, some of the parameters are unidentifiable or identifiable but very poorly estimable.

We define the FIM, $\mathbf{F} = \mathbf{S}^\top\mathbf{S}$ with $\mathbf{S}$ (Equation (8.25)) evaluated at some fixed, nominal parameter value q_0. A model is locally identifiable if and only if $\det\left(\mathbf{S}^\top\mathbf{S}\right) \neq 0$, or equivalently if and only if rank of $\mathbf{S}^\top\mathbf{S}$ is equal to the number of estimated parameters, m. The rank r tells us how many linearly independent and, consequently, how many parameters can be uniquely identified. The question then becomes "which of the r parameters do we choose to create an identifiable subset?"

Additional numerical properties of the FIM give substantive information about the identifiable parameters of the model, which are selected accordingly. For example, the condition number of the FIM (i.e., the ratio of its largest and smallest eigenvalues)

indicates how numerically well-posed the estimation problem is and the rank of the FIM indicates the number of parameters that can be theoretically estimated.

8.8.2 Sensitivity Plots: Visual Interpretation of Sensitivity Functions

The *sensitivity plot* is a graphical method to assess estimability and design efficiency of a model's parameters. The sensitivity values are graphically arranged according to model parameter so that their linear dependencies are easy to see. Sensitivity functions that appear proportional indicate a correlation between the model parameters.

For simple models with few parameters and outputs, plots of sensitivity coefficients as functions of time can be used to determine estimability. (See Beck and Arnold [14] for details.) Plots of the values of the sensitivity coefficients with time can show whether there is significant correlation among the effects of different parameter on model predictions. For example, if the sum of the sensitivity coefficients for two parameters on the same response variable is zero or near zero at all times, then there is a near perfect correlation between the parameters, and these parameters cannot be estimated simultaneously [97].

8.8.3 Correlation

Yao et al. [97] used the rank of the sensitivity matrix to determine the number of estimable parameters. These results don't take in account the subsets of correlated parameters that cannot be identified. A large condition number of the FIM indicates a strong correlation between estimates and may indicate identifiability issues [53]. If parameters are correlated, they cannot be determined uniquely and independently; hence, they are not identifiable [24]. Quaiser and Mönnigmann [56, 79] proposed a ranking method of least-to-most estimable for the parameters but this method also cannot identify parameter groups in which more than two parameters are correlated together, i.e. the corresponding columns in the sensitivity matrix are linearly dependent.

There are two features that govern the estimability of a parameter [97]:

- The strong influence of that parameter on one or more of the measured responses;
- The correlation between the effects of that parameter on model predictions and the corresponding effects of all other estimable parameters.

The first feature can be determined by examining the magnitude of the column of S that corresponds to the particular parameter (See Algorithm 2). The second requires determining whether the columns corresponding to the set of estimable parameters are linearly dependent. Theoretically, the rank of S is equal to the number of estimable parameters. But when the rank is less than the number of columns, we cannot readily determine which subsets of parameters could be estimated. [97] To address this, we use the algorithm outlined in Algorithm 2 to identify the subset of estimable parameters.

8.8.4 Parameter Ranking

As discussed in Section 8.7, not all parameters of a model can be estimated but we are interested in knowing which are and to what degree. To do this, we use estimability algorithms which are able to rank the parameters from most estimable to least estimable.

Algorithms 1-2 both start with the sensitivity coefficient matrix $\mathbf{S}$ and end with a subset of estimable parameters.

However, singularity problems that might be encountered when ranking the later parameters suggest that fewer than p parameters could be simultaneously estimated using the available data, because estimating all p parameters would lead to numerical conditioning problems. A low parameter rank number indicates a parameter that should be easy to estimate because of the large amount of information in the available data whereas high rank number indicates a parameter that cannot be readily estimated due to a lack of parameter influence on the predicted responses or due to possible correlation with the effects of parameters that appear earlier in the ranked list [89].

In the absence of parameter ranking, we would need to determine a cut-off point for linear independence of the FIM. This can be done by computing the successive ratios of the magnitudes of the diagonal terms of R in Algorithm 2

$$\frac{|r_{(i,i)}|}{|r_{(i+1,i+1)}|}, \qquad i = 1 \ldots (m-1), \tag{8.27}$$

where m is the number of model parameters. [This method of a qualitative cut-off does not use R from the partitioned matrix V in Algorithm 2. Instead, the SVD steps in the algorithm have been excluded and R is calculated directly from $S^\top S\Pi = QR$ where Π is the permutation matrix.]

Both algorithms produced near identical results for parameter ranking up to the

Algorithm 1 Proposed methodology for selecting a subset of parameters for estimation.

1: Compute the sensitivity matrix S using Equation (8.25) and calculate the magnitude of each column of S.
2: Select the parameter whose column in S has the largest magnitude (the sum of squares of the elements) as the first estimable parameter. This parameter is considered to have the most effect on model predictions.
3: Mark the corresponding column as X_k ($k = 1$ for the first iteration).
4: Calculate $\hat{S}_k$, the prediction of the full sensitivity matrix, S, using the subset of columns $X_k : \hat{S}_k = X_k(X_k^T X_k)^{-1} X_k^T S$
5: Calculate the residual matrix $R_k : R_k = S - \hat{S}_k$
6: Calculate the sum of squares of the residuals in each column of R_k. The column with the largest magnitude corresponds to the next estimable parameter (among the remaining parameters) which has the largest effect on response variables and which is not correlated with the effects of parameters already selected.
7: Select the corresponding column in S, and augment the matrix X_k by including the new column. Denote the augmented matrix as X_{k+1}.
8: Advance the iteration counter by one and repeat steps 4 through 7 until the column of largest magnitude in the residual matrix is smaller than a prescribed cut-off value.

Algorithm 2 Subset selection using Singular Value Decomposition (SVD)-QR [70]

1: Given a set of parameters θ, compute the sensitivity matrix S using Equation (8.25) and use singular value decomposition to rewrite as $S = U\Sigma V^T$, where Σ is a diagonal matrix containing the singular values of S in decreasing order, and U and V are orthogonal matrices with left and right singular vectors.
2: Determine ρ, the rank of the sensitivity matrix S.
3: Partition the matrix V of eigenvectors on the form $V = [V_\rho V_{q-\rho}]$.
4: Determine a permutation matrix P by constructing a QR decomposition with column pivoting after the maximal element for V_ρ^T. (That is, determine P such that $V_\rho^T P = QR$, where Q is an orthogonal matrix and the first ρ columns of R form an upper triangular matrix with diagonal elements in decreasing order.)
5: Use P to re-order the parameter vector θ according to $\hat{\theta} = P^T\theta$.
6: Make the partition $\hat{\theta} = [\hat{\theta}_\rho \theta_{q-\rho}]$. The subset of identifiable parameters, $\hat{\theta_\rho}$, is the first ρ elements of $\hat{\theta}$.

number of estimable and identifiable parameters.

8.8.5 Cross-validation

When too many parameters are estimated using limited data, the high levels of uncertainty associated with the parameter estimates result in large variances for the model predictions. We can lower these variances by estimating on a subset of the model parameters (keeping the other parameters at their initial guesses or values found in the literature) but the model predictions and parameter estimates then become biased due to the incorrect values of the fixed parameters. We can decrease this bias by estimating more parameters. The optimal number of parameters to estimate balances the trade-off between variance and bias. A direct assessment of this trade-off is to use cross-validation which tests the predictive ability of a model by removing data from the available data set. [89]. Model parameters are then estimated and used to predict the removed data.

A low value of the objective function indicates good predictive ability of the model and the parameters. One clear benefit of cross-validation is that it provides a measure of how well the model can predict data that were not used for estimation. Another benefit is that it provides information about the sensitivity of parameter values to particular data points. [89]

8.9 Parameter Subset Selection

Low quality data or other issues may prevent some model parameters from being estimated. In this case, a subset of parameters should be selected that can be estimated reliably from the available data. The FIM can be used to decide which parameter subsets can be uniquely estimated from a given data set. We accomplish this by computing the QR decomposition of the FIM (a square matrix by definition), that is, compute orthogonal matrix Q and upper triangular matrix R with permutation matrix Π such that

$$S^\top S\Pi = QR. \tag{8.28}$$

From left to right, the columns of Π indicate the most linearly independent to least independent columns of $S^\top S$. Finally, we compute

$$\Pi^\top \cdot [1 \dots m]^\top \tag{8.29}$$

to determine the parameters corresponding to these ordered columns. While this method identifies the most-to-least important parameters in a model, it does not provide a quantitative cut-off for selecting the subset of identifiable or estimable parameters. We can, however, utilize the methods provided in Section 8.8.4 to accomplish this. Further, we may consider the rank, r, of the FIM and take the first r parameters indicated in Equation (8.29).

Recall that Algorithms 1-2 in Section 8.8.4 are additional methods that quantitatively tell which parameters are most influential on the model. When the most estimable parameters are identified, a follow-on procedure is to estimate that subset of parameters while keeping the non-identifiable parameters at their initial estimates or reparameterize the model to remove these parameters [97]. We can check the value of the objective function to compare model fits before and after applying identifiability procedures to verify that there is some improvement to the fit using our estimated subset of parameters.

By Algorithm 1, the most identifiable parameter is determined by the associated column/vector of the sensitivity matrix, S, with the highest Euclidean norm, $\max_j \|s_j\|$. The large norm with its effect on the model outcome vector makes the parameter θ_j identifiable with the available data if all other parameters are fixed. By using a Gram-Schmidt orthogonalization method, the algorithm includes identifiable parameters, one at a time, in decreasing order of Euclidean norm but also removing the effect of linear dependence of the column in the sensitivity matrix associated with the selected parameter. Overall, this estimability technique ensures that the effects of the selected parameters are not too highly correlated while also having a large impact on the measured responses.

CHAPTER 9

Model Results

In our investigation, we examine four models where our ultimate goals are to (a) determine whether the hypothesis that the inclusion of the upper respiratory tract for lung function will better describe the data provided and (b) determine any impact that adding kidney metabolism to the system(s) may have. PBPK models often include liver metabolism but do not generally consider kidney metabolism.

We use MATLAB© R2016a [59] for all programming code and time course data from Cain et al. [27] and Pleil et al. [74] to calibrate and validate the models. Initial "guesses" of Michaelis-Menten parameters to run MATLAB's `fminsearch` command were taken from Blancato et al. [18] as were relevant partition coefficients for MTBE and TBA and human physiological parameters. In the absence of given Michaelis-Menten parameters, arbitrary, but biologically reasonable, initial conditions were used. Physiological parameters not provided by the Blancato et al. paper were supplemented by Brown et al. [25].

All optimized parameter values can be found in Appendix G.

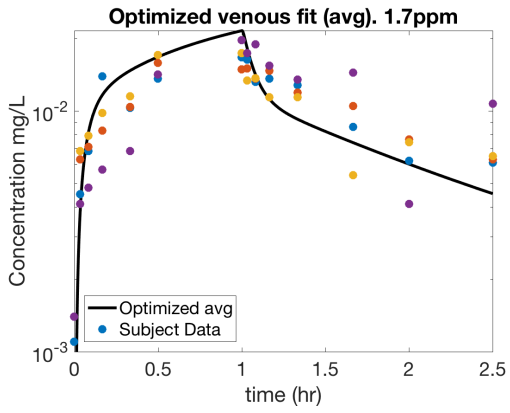
9.1 Model 1: Simple Lung

In the first model (Figure 6.1), we consider venous blood data and exhaled breath data for MTBE only, collected from subjects' constant exposure to 1.7 ppm (Cain) and 3 ppm

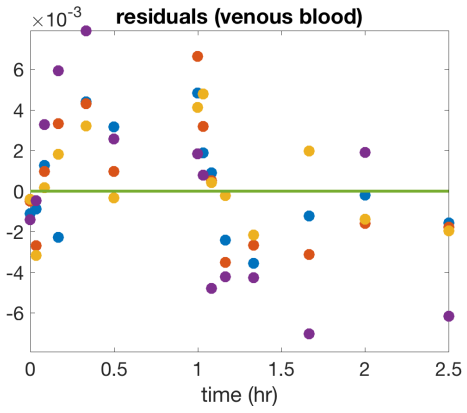
(Pleil) of MTBE, both for a duration of one hour. This first iterate builds upon concepts outlined by Blancato, et al. [18]. That is, we consider both a high- and low-affinity metabolizing pathway in the liver for MTBE (Liv_1 and Liv_2, respectfully) and we also consider a single metabolism pathway for MTBE in the kidney (Kid). This model also includes fat/adipose tissue , slowly- and rapidly-perfused tissues, arterial and venous blood compartments, and a lung compartment that assumes an inert tube model for gas exchange. Initial guesses for Michaelis-Menten parameters V_{max} and K_m for each pathway ($V_{max_{Liv1}}$, $V_{max_{Liv2}}$, $V_{max_{Kid}}$, $K_{m_{Liv1}}$, $K_{m_{Liv2}}$, and $K_{m_{Kid}}$) were also taken from the Blancato et al. paper.

Figure 9.1 shows the fit to the Cain venous blood data using the mean for the optimized parameters and the randomness of the scatter pattern of the residuals indicates a "good" fit for the given data. We see also that this model has no sensitivity to the parameters for the high-affinity liver pathway (Liv_1). The model's sensitivity to the kidney parameters are considerably low, relative to the low-affinity liver pathway (Liv_2) parameters, but its value should not be discounted.

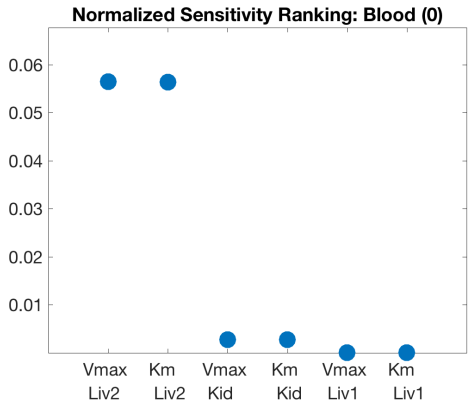
We perform a similar analysis for the Pleil data for which we have both venous blood and exhaled breath measurements. In Figure 9.2, we see the optimized fits and residuals for both data sets indicate that the model generally under-predicts breath data for the time period during inhalation of MTBE and over-predicts breath data during post-exposure to MTBE. We see generally opposite trends for the venous blood predictions for these same time periods. The sensitivity rankings show that the model is most sensitive to the Michaelis-Menten parameters for the kidney. Similar to the case for the Cain data, we see that the high-affinity liver pathway (Liv_1) is not sensitive in the model.



(a) Optimized venous blood fit using parameter mean values

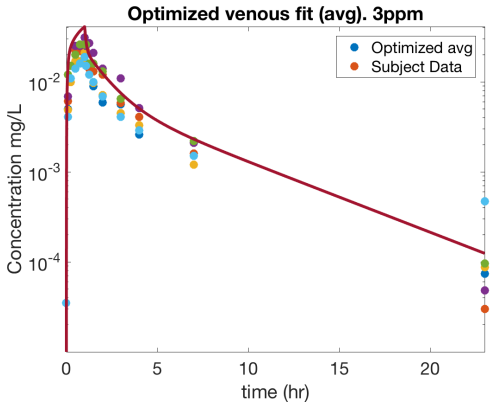


(b) Venous blood residuals

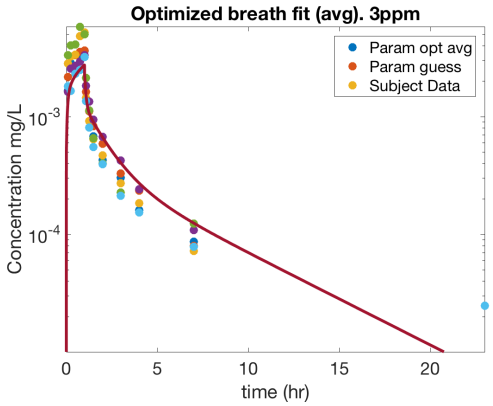


(c) Normalized sensitivity ranking; venous blood

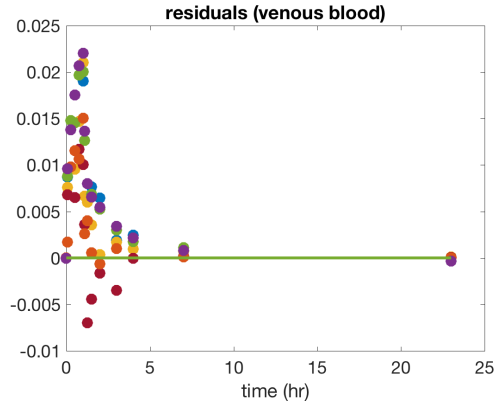
Figure 9.1: Cain (1.7 ppm), Simple lung (Model 1), MTBE-only; Venous blood fit



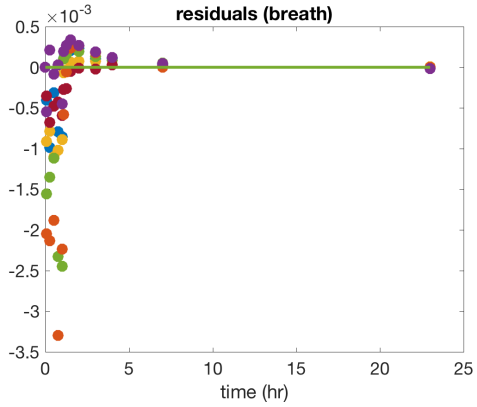
(a) Optimized venous blood fit using parameter mean values



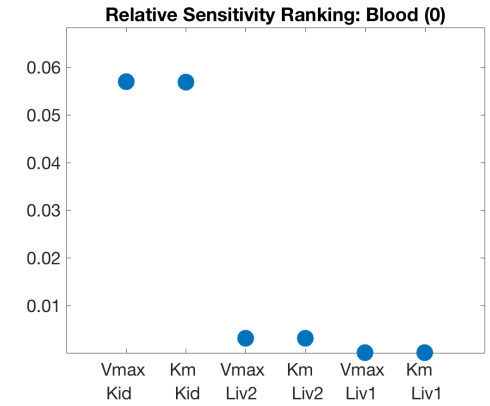
(b) Optimized exhaled breath fit using parameter mean values



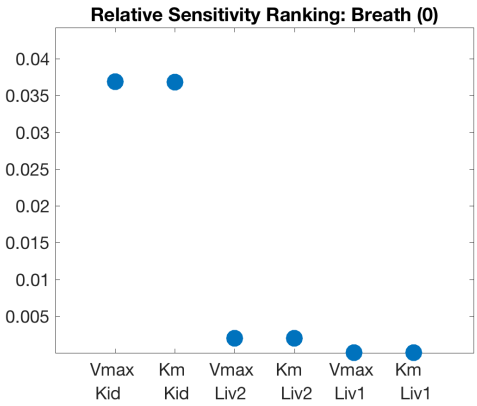
(c) Venous blood residuals



(d) Exhaled breath residuals



(e) Normalized sensitivity ranking; venous blood



(f) Normalized sensitivity ranking; exhaled breath

Figure 9.2: Pleil (3 ppm), Simple lung (Model 1), MTBE-only, Venous blood and exhaled breath fit

9.2 Model 2: Complex Lung (with Upper Respiratory Tract)

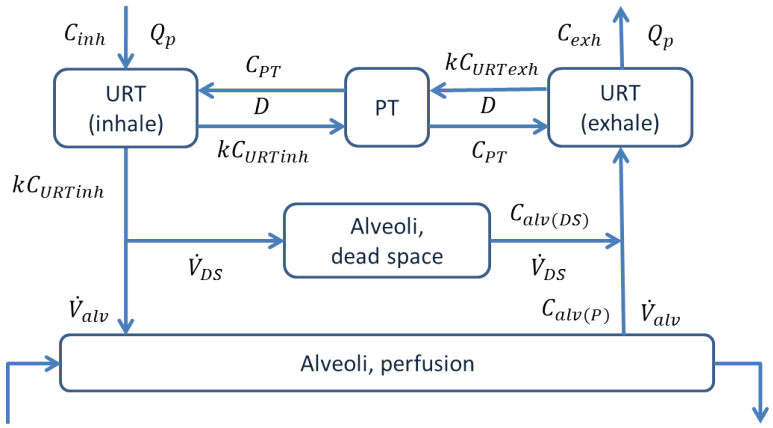


Figure 9.3: Upper Respiratory Tract system

In this model (Figure E.2), we use the upper respiratory tract (URT) over a simple lung (inert tube model). Specifically, the simple lung compartment of Figure 6.1 is replaced by the sub-compartmented lung shown in Figure 9.3. We consider the washin-washout effect where the substance gets diffused between the pulmonary tissue (PT) and the URT and that some fraction, k, of the substance leaves the URT during inhalation and exhalation. During inhalation, some amount of the substance is trapped in the alveolar dead space while the rest gets passed to the perfused alveolar space and cycles through the remaining compartments.

For the Cain data set (Figure 9.4), we have a fairly random pattern about zero for the scatter plot of the residuals (Figure 9.4b). The model generally over-predicts the data during exposure (see also Figure 9.4a) but the residual pattern does not exhibit an obvious trend following termination of active exposure ("post-exposure"). The Pleil residual patterns (Figure 9.5c and Figure 9.5d) exhibit less randomness compared to the simple lung model. Specifically, the exhaled breath data points are heavily positively

skewed for the hour during exposure and we observe a similar trend for the venous blood data.

For each data set, we again find that there is no simultaneous sensitivity of the model to both liver pathways. While the relative sensitivity ranking for the kidney pathway is near zero for the Cain data, we see this is not the case for the Pleil data where the model's sensitivity is influenced by kidney parameters and we conclude that those parameters are important for predicting MTBE concentration levels given the experimental data. This could potentially be explained by kidney metabolism being more effective at higher initial doses of MTBE (3.0 ppm vs. 1.7 ppm) but further exploration is left for future work (Section 10.2) where data for 25 ppm and 75 ppm initial doses would be examined.

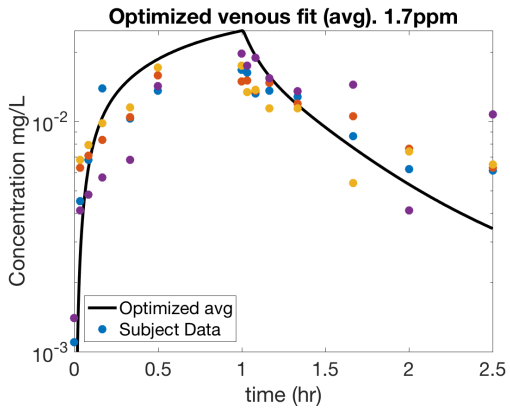
We defined $k \in (0, 1)$ as the fraction of MTBE that leaves the upper respiratory tract (URT) during inhalation and exhalation. The diffusion rate, D, is the rate at which MTBE (and TBA) is passed between the URT and pulmonary tract (PT). These values are most sensitive in the model for both sets of data, thus fluctuations in their values result in more significant changes to the output than other parameters in the model.

9.3 Model 3: Simple Lung, MTBE-to-TBA Coupled System

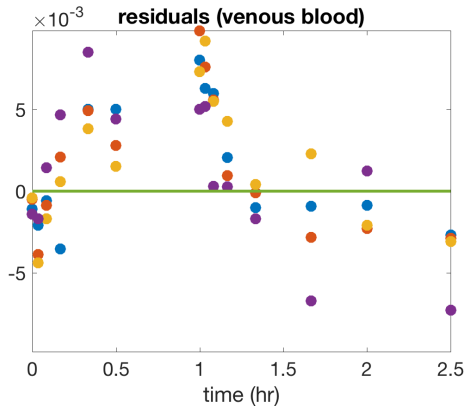
For this model, we have a coupled system representing MTBE metabolized to TBA (Figure E.4). The data include venous blood and exhaled breath measurements for both MTBE and TBA and the lung compartment is "simple" as in Model 1. Based on the relative sensitivity rankings for the Pleil data in Models 1 and 2 (Figures 9.2e, 9.2f, 9.5e, and 9.5f) indicating that the liver pathways for MTBE are not simultaneously sensitive, this model iteration employs only one MTBE liver pathway. We again use the kidney pathway for MTBE and we include a single metabolism pathway in the liver for TBA.

In Figure 9.6a, we see the model generally over-predicts the venous blood data for MTBE during exposure but more closely matches the data following termination of active exposure. The residuals here do not exhibit a random behavior. On the other hand, the model matches the TBA venous blood data quite well (Figure 9.6a) with the TBA venous blood residuals exhibiting a moderately random pattern (Figure 9.6c).

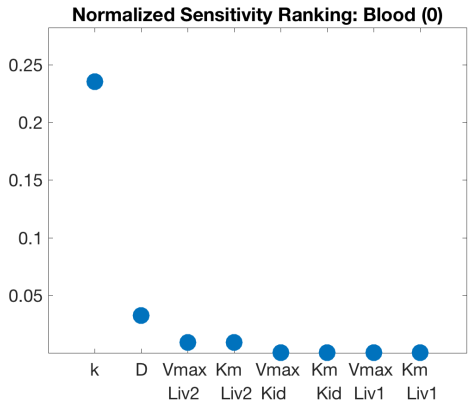
The optimized fit for the exhaled breath data is well-matched for MTBE (Figure 9.6b) but poorly fit for TBA. The residuals (Figure 9.6d) for TBA's exhaled breath data show



(a) Optimized venous blood fit using parameter mean values

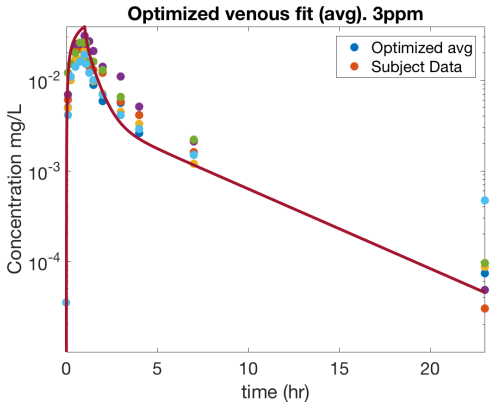


(b) Venous blood residuals

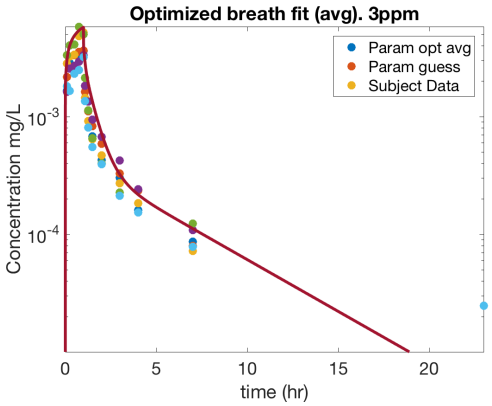


(c) Normalized sensitivity ranking; venous blood

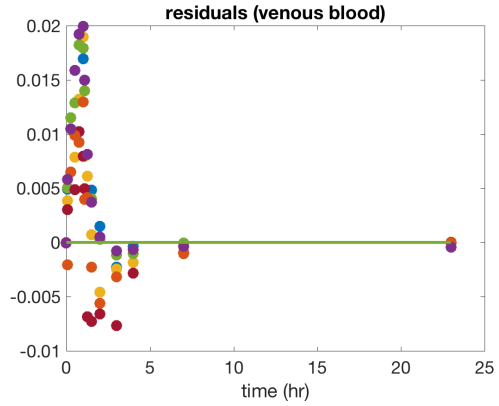
Figure 9.4: Cain (1.7 ppm), Complex lung (URT) (Model 2), MTBE-only



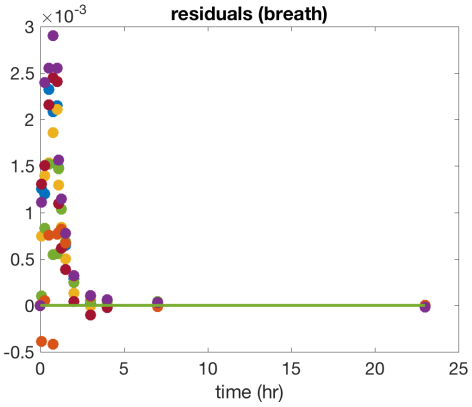
(a) Optimized venous blood fit using parameter mean values



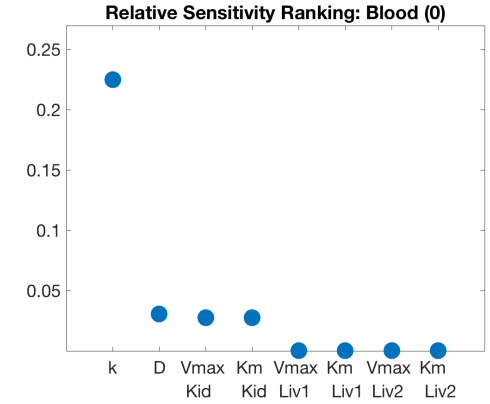
(b) Optimized exhaled breath fit using parameter mean values



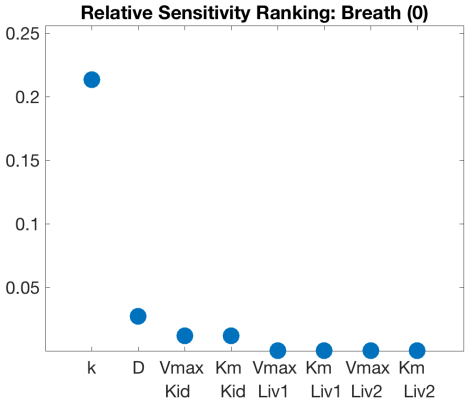
(c) Venous blood residuals



(d) Exhaled breath residuals



(e) Normalized sensitivity ranking; venous blood



(f) Normalized sensitivity ranking; exhaled breath

Figure 9.5: Pleil (3 ppm), Complex lung (URT) (Model 2), Venous blood and exhaled breath, MTBE-only

an obvious positive trend while the exhaled breath residuals are more random about zero.

9.4 Model 4: Complex Lung (with URT), MTBE-to-TBA Coupled System

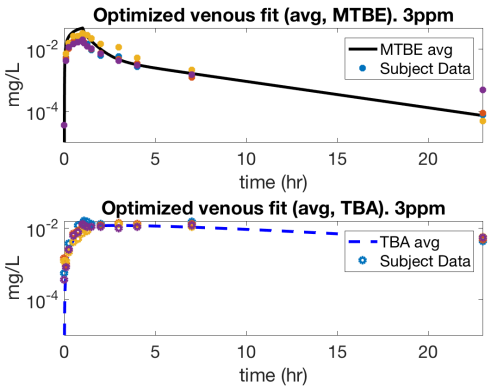
For the final model (Figure E.5), we again have a coupled system representing MTBE metabolized to TBA. The data again include venous blood and exhaled breath measurements for both MTBE and TBA but the lung compartment contains the upper respiratory tract. Following the conclusion for Model 3 that the relative sensitivity rankings for the Pleil data in Models 1 and 2 (Figures 9.2 and 9.5) indicate that the liver pathways for MTBE are not simultaneously sensitive, this model iteration also employs only one MTBE liver pathway. We again use the kidney pathway for MTBE and we include a single metabolism pathway in the liver for TBA.

The model under-predicts the venous blood data for both MTBE and TBA for all time points (Figure 9.7a). It over-predicts the exhaled breath data during exposure but more closely matches the data post-exposure (Figure 9.7b). The residuals here do not exhibit a random behavior for venous blood or exhaled breath (Figures 9.7c and 9.7d). These conditions indicate a poor fit overall for venous blood and exhaled breath data for both MTBE and TBA.

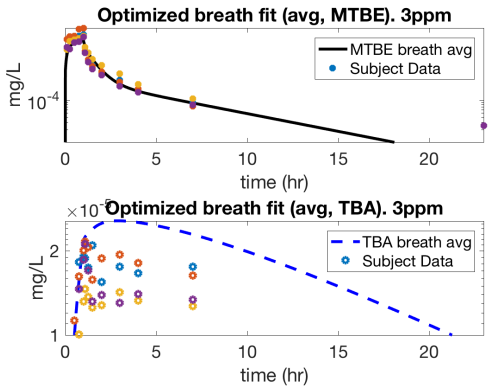
9.5 Akaike Results

Given a set of models, the Akaike Information Criterion (AIC) is used to make an assessment on which of the models is "better" in comparison to the others. Recall, the AIC does not give any indication as to whether a model is "good" or "bad," only that it is better or worse than another. A model with the smallest AIC is considered to be the "best" model of the candidate collection of models.

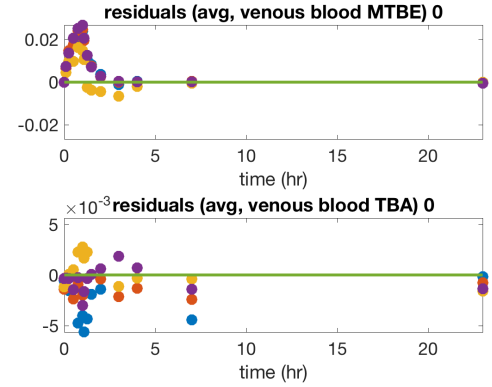
We assess each of our models using this criteria and provide the AIC values obtained for each subject in Tables (9.1)-(9.3). Recall, these values are based on the cost function (Equation (8.4)) for the optimization process but includes a penalty for models that optimize several parameters compared to those that only optimize a few parameters. With the exception of "Subject 5" from the Model 2 runs for Pleil data (Table 9.2), the "simple



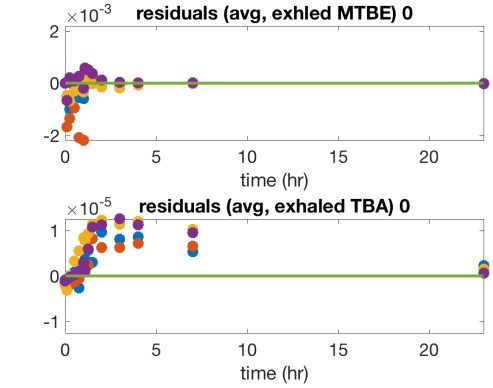
(a) Optimized venous blood fit using parameter mean values: MTBE, top. TBA, bottom.



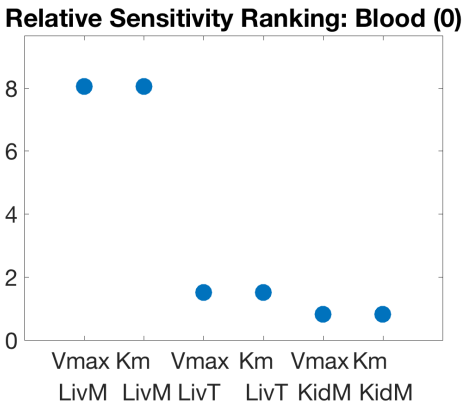
(b) Optimized exhaled breath fit using parameter mean values: MTBE, top. TBA, bottom.



(c) Venous blood residuals: MTBE, top. TBA, bottom.

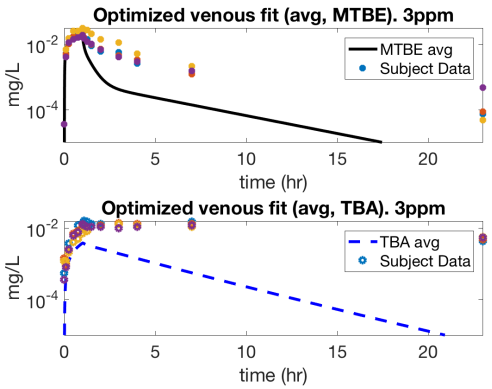


(d) Exhaled breath residuals: MTBE, top. TBA, bottom.

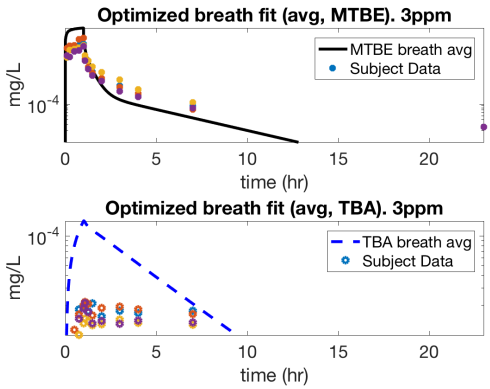


(e) Normalized sensitivity ranking; venous blood

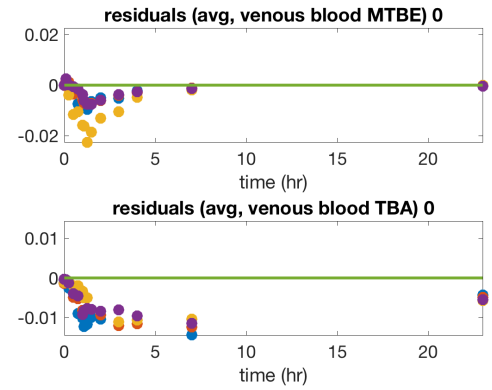
Figure 9.6: Pleil (3 ppm), Simple lung (Model 3), Venous blood and exhaled breath, MTBE-to-TBA (Coupled system)



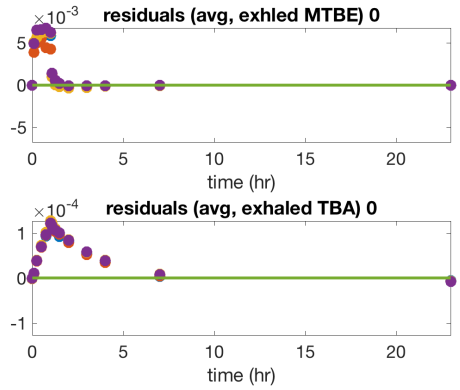
(a) Optimized venous blood fit using parameter mean values: MTBE, top. TBA, bottom.



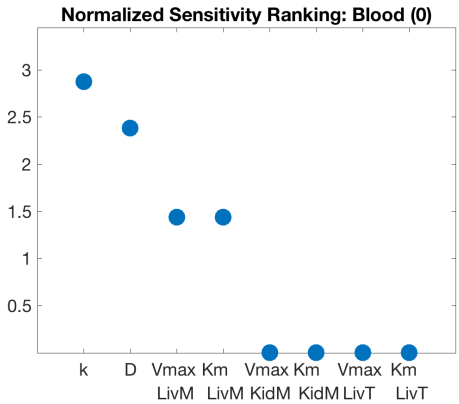
(b) Optimized exhaled breath fit using parameter mean values: MTBE, top. TBA, bottom.



(c) Venous blood residuals: MTBE, top. TBA, bottom.



(d) Exhaled breath residuals: MTBE, top. TBA, bottom.



(e) Normalized sensitivity ranking; venous blood

Figure 9.7: Pleil (3 ppm), Complex lung (Model 4), Venous blood and exhaled breath, MTBE-to-TBA (Coupled system)

Table 9.1: AIC Indices: Cain data; MTBE-only models

AIC (Cain, MTBE-only)	Subj 1	Subj 2	Subj 3	Subj 4
Simple lung (Model 1)	-96.0783	-89.3240	-100.1247	-85.6269
Complex lung (Model 2)	-79.9659	-75.2802	-81.5990	-77.7229

Table 9.2: AIC Indices: Pleil data; MTBE-only models

AIC (Pleil, MTBE-only)	Subj 1	Subj 2	Subj 3	Subj 4	Subj 5	Subj 6
Simple lung (Model 1)	60.8029	58.0720	67.4135	42.0550	63.1668	77.5742
Complex lung (Model 2)	74.5630	68.1170	67.8261	73.7999	59.6890	92.1001

lung" for all other cases—both Cain and Pleil—is shown to be the better model since the AIC values are for Model 1 are generally less than the AIC values for Model 2.

9.6 Validation and Residual Analysis

We discussed the patterns and trends of our models' residuals versus time in Sections 9.1-9.4. These residuals include only the subject data used to optimize the model parameters. In this section, we extend the residual analysis to include subject data that was not used to obtain the optimized parameters. Doing this is a method of model validation as outlined in Section 8.3. Note that this validation and extended analysis is only done for the Pleil data set since all subjects in the Cain data set were used to calibrate the model.

Of the methods provided, we turn our attention to one plot type in particular: observed versus predicted data values. The closer the predicted values are to the observed values,

Table 9.3: AIC Indices: Pleil data; MTBE-to-TBA models

AIC (Pleil, MTBE-to-TBA)	Subj 1	Subj 2	Subj 3	Subj 4
Simple lung (Model 3)	-61.9364	-57.8143	-71.3504	-50.5250
Complex lung (Model 4)	-48.7236	-53.7672	-46.0784	-49.4288

the more the points align around the unity reference line in an observed data vs. predicted data plot. This gives an indication of an accurate model since there is a strong correlation between the model's predictions and the actual data. We can, in fact, compute the correlation coefficients for each subject to get a quantitative perspective on how well each model's observed data vs. predicted data plot exhibits a linear behavior.

We see in Figure 9.11 that the MTBE and TBA exhaled breath observed and predicted values for Model 4 (complex lung, coupled system) are far less correlated than those for Model 3 (Figure 9.10; simple lung, coupled system). The figures indicate a consistent over-prediction by Model 4 for both MTBE and TBA exhaled breath. Additionally, the figures indicate a consistent under-prediction by Model 4 for MTBE and TBA venous blood. Based on Figures 9.8-9.10, models 1-3 provide predictions that generally fit well to the data.

Referring to the computed correlation coefficients in Tables 9.4-9.7, there is no significant quantitative difference across all models for the fits to MTBE venous blood exhaled breath data. Contrast this with the correlation coefficients for TBA and we see a marked difference in the fits between the simple and complex lung models for both TBA venous blood and exhaled breath data. This suggests that a simple lung model for the coupled MTBE-to-TBA system gives better overall fits to the observed data, and specifically to the TBA data, than the complex lung model for the coupled MTBE-to-TBA system.

Where the correlation coefficients indicate how linear the relationship between observed data and predicted data is, well-matched data points should show random variation about the line of unity ($m = 1$). Given a set of observed-vs.-predicted data points, we need only compute the slope of the line of best fit through those points and compare it to $m = 1$. This is accomplished by a two-stage approach, first using MATLAB's `polyfit` command on each set of data points, then finding the average slope (and intercept) across sets. These values are given in Table 9.8 and plots of the lines of best fit using the average slope are superimposed along with the reference unity line in each of Figures 9.8-9.11.

Finally, we compute the relative error for each averaged slope value. These errors are provided in Table 9.9. Comparing the error values between Model 1 and Model 2, we see that Model 1 has the smallest error for both venous blood and exhaled breath fits, indicating the best-fit line through the set of observed data values vs. predicted data values for Model 1 is closer to unity than the best-fit line for Model 2. Comparing the relative error values between Model 3 and Model 4, we see that Model 4 gives the best-fit

Table 9.4: Model Correlation Values: Pleil MTBE Venous Blood Data

Venous blood, MTBE	Correlation coeff. range	mean	std
Simple lung (Model 1)	$0.8672 - 0.9942$	0.9523	0.0380
Complex lung (Model 2)	$0.8701 - 0.9847$	0.9464	0.0375
Simple lung, coupled (Model 3)	$0.8805 - 0.9850$	0.9511	0.0345
Complex lung, coupled (Model 4)	$0.7104 - 0.9469$	0.8518	0.0731

Table 9.5: Model Correlation Values: Pleil MTBE Exhaled Breath Data

Exhaled breath, MTBE	Correlation coeff. range	mean	std
Simple lung (Model 1)	$0.9545 - 0.9961$	0.9751	0.0140
Complex lung (Model 2)	$0.9730 - 0.9949$	0.9853	0.0088
Simple lung, coupled (Model 3)	$0.9290 - 0.9858$	0.9573	0.0183
Complex lung, coupled (Model 4)	$0.9461 - 0.9803$	0.9645	0.0132

line for MTBE venous blood data, but Model 3 is the overall better model, with best-fit lines closer to unity, since the relative error provided by this model is smaller for MTBE exhaled breath data and TBA data for both venous blood and exhaled breath.

9.7 Estimability and Sensitivity Plots

Algorithms 1-2 along with the ratio test in Equation (8.27) provide an abundance of quantitative information about the identifiability and estimability of the parameters in each of the four models. Where one algorithm was used to determine the most identifiable or estimable parameters, another was used to validate the results. Recall from Section

Table 9.6: Model Correlation Values: Pleil TBA Venous Blood Data

Venous blood, TBA	Correlation coeff. range	mean	std
Simple lung, coupled (Model 3)	$0.9276 - 0.9835$	0.9619	0.0175
Complex lung, coupled (Model 4)	$0.3876 - 0.8364$	0.6151	0.1341

Table 9.7: Model Correlation Values: Pleil TBA Exhaled Breath Data

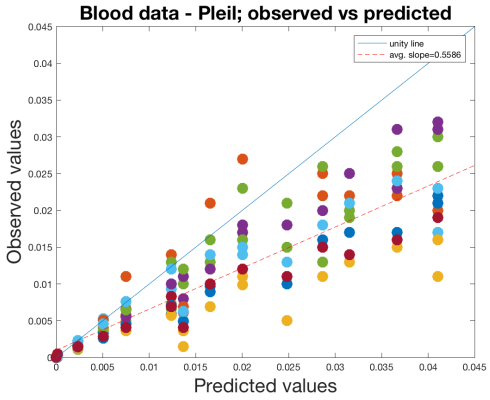
Exhaled breath, TBA	Correlation coeff. range	mean	std
Simple lung, coupled (Model 3)	$0.8675 - 0.9798$	0.9210	0.0504
Complex lung, coupled (Model 4)	$0.6789 - 0.8247$	0.7674	0.0543

Table 9.8: Model Estimated Slope Values

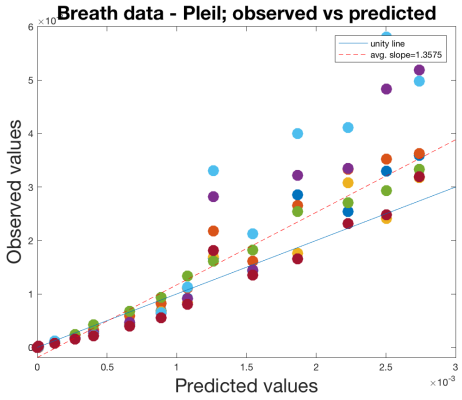
	MTBE		TBA	
	Blood	Breath	Blood	Breath
Simple lung (Model 1)	0.5586	1.3575	-	-
Complex lung (Model 2)	0.5456	0.6078	-	-
Simple lung, coupled (Model 3)	0.4726	1.1853	1.103	0.5799
Complex lung, coupled (Model 4)	1.1574	0.3134	2.2881	0.0917

Table 9.9: Model Estimated Slope Values: Relative Error

	MTBE		TBA	
	Blood	Breath	Blood	Breath
Simple lung (Model 1)	0.4414	0.3575	-	-
Complex lung (Model 2)	0.4544	0.3922	-	-
Simple lung, coupled (Model 3)	0.5274	0.1853	0.103	0.4201
Complex lung, coupled (Model 4)	0.1574	0.6866	1.2881	0.9083

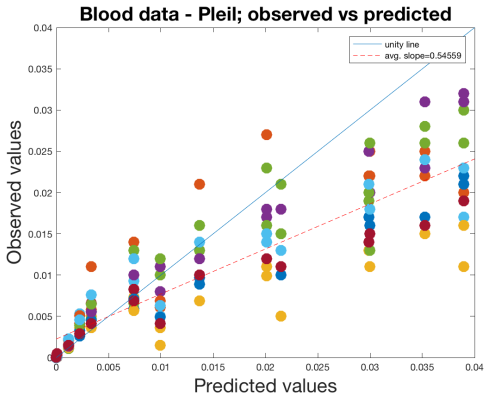


(a) Venous blood data: observed vs. predicted

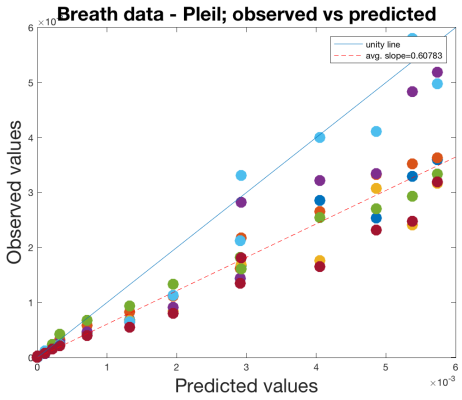


(b) Exhaled breath data: observed vs. predicted

Figure 9.8: Pleil (3 ppm), Simple lung (Model 1); (Validation) Data: observed vs. predicted

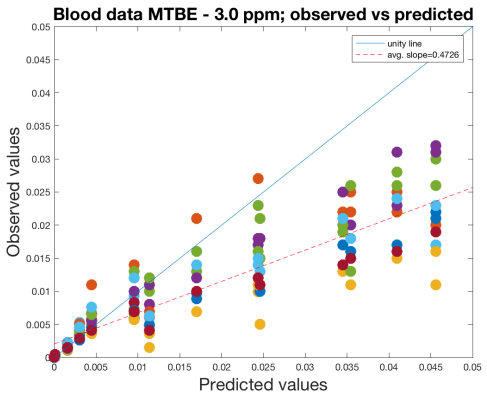


(a) Venous blood data: observed vs. predicted

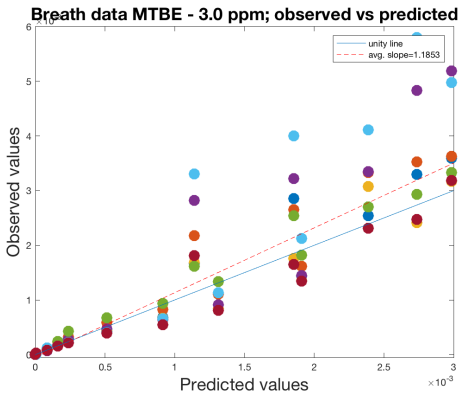


(b) Exhaled breath data: observed vs. predicted

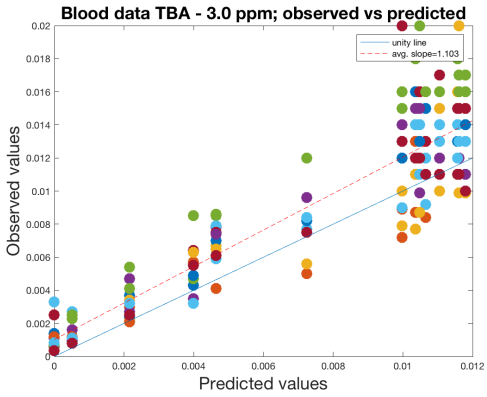
Figure 9.9: Pleil (3 ppm), Complex lung (Model 2); (Validation) Data: observed vs. predicted



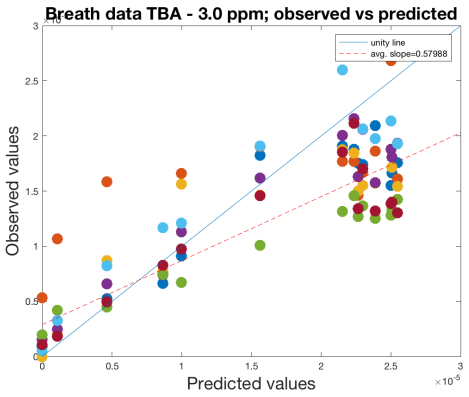
(a) MTBE Venous blood data: observed vs. predicted



(b) MTBE Exhaled breath data: observed vs. predicted

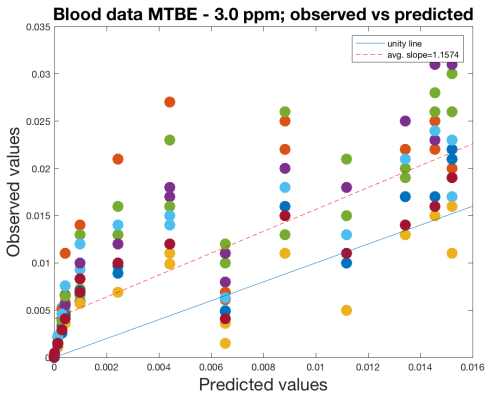


(c) TBA Venous blood data: observed vs. predicted

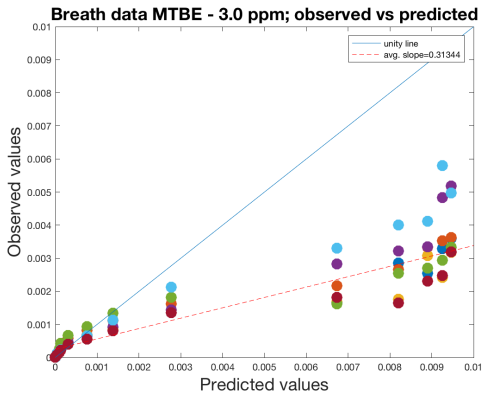


(d) TBA Exhaled breath data: observed vs. predicted

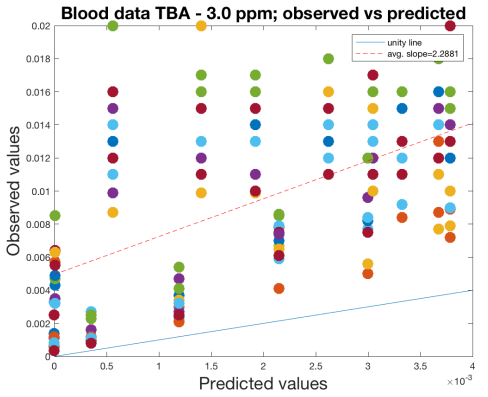
Figure 9.10: Pleil (3 ppm), Simple lung MTBE-to-TBA (Model 3); (Validation) Data: observed vs. predicted



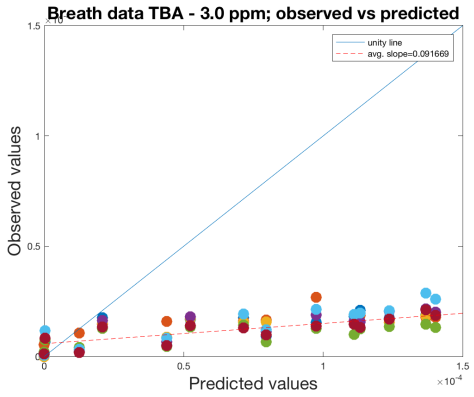
(a) MTBE venous blood data: observed vs. predicted



(b) MTBE exhaled breath data: observed vs. predicted



(c) TBA venous blood data: observed vs. predicted



(d) TBA exhaled breath data: observed vs. predicted

Figure 9.11: Pleil (3 ppm), Complex lung MTBE-to-TBA (Model 4); (Validation) Data: observed vs. predicted

8.5.2 that the rank of the sensitivity matrix, S (Equation (8.25)), is a determination of the number of identifiable parameters in a model while the rank of the Fisher Information Matrix (FIM), $S^\top S$, gives the number of estimable parameters. We examine each of the quantities to make qualitative assessments of the estimated parameters in the models.

In Model 1, we estimate six parameters: $V_{max_{Liv1}}$, $V_{max_{Liv2}}$, $V_{max_{Kid}}$, $K_{m_{Liv1}}$, $K_{m_{Liv2}}$, and $K_{m_{Kid}}$.

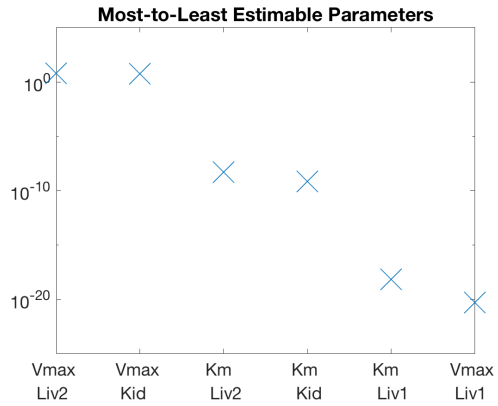
- For the Cain 1.7 ppm data, Figure 9.12a is a visual representation of the estimability order with the largest difference in the ratios occurring between the second and third ordered parameters, thus, without any knowledge of parameter ranking, we would make a qualitative cut-off after the second parameter. More effort, particularly trial and error of parameter combinations, is required to determine which two parameters to estimate. However, the subset selection algorithms indicate that, prior to estimating any parameters, five of the parameters are identifiable and four are estimable, ordered from most to least estimable: $V_{max_{Liv2}}$, $V_{max_{Kid}}$, $K_{m_{Liv2}}$, and $K_{m_{Kid}}$.

- For the Pleil 3.0 ppm data, the largest difference in the ratios also occurs after the second parameter (Figure 9.12b). The subset selection algorithms indicate that all six parameters are identifiable, of which, five are estimable, ordered from most to least estimable: $V_{max_{Kid}}$, $V_{max_{Liv2}}$, $K_{m_{Liv1}}$, $K_{m_{Kid}}$, $V_{max_{Liv1}}$.

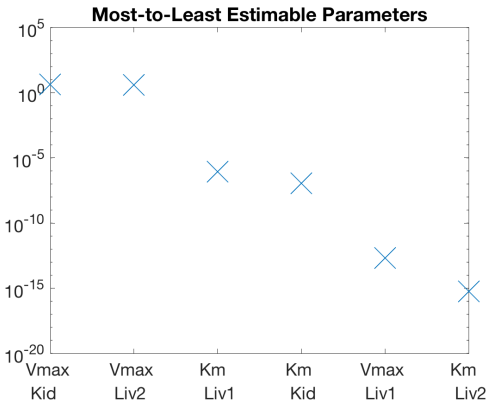
In Model 2, we estimate eight parameters: $V_{max_{Liv1}}$, $V_{max_{Liv2}}$, $V_{max_{Kid}}$, $K_{m_{Liv1}}$, $K_{m_{Liv2}}$, $K_{m_{Kid}}$, k, and D.

- For the Cain 1.7 ppm data, Figure 9.12c shows the largest difference in the ratios occurring between the fourth and fifth ordered parameters, thus, in the absence of parameter ranking knowledge, we would make a qualitative cut-off and estimate on four parameters. The subset selection algorithms generally agree with this assessment, indicating five of the parameters are identifiable and four are estimable. They are, ordered from most to least estimable: $V_{max_{Liv2}}$, $V_{max_{Kid}}$, $K_{m_{Liv2}}$, and $K_{m_{Kid}}$.

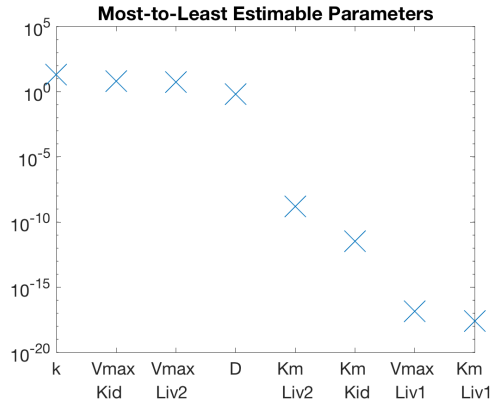
- For the Pleil 3.0 ppm data (Figure 9.12d), the largest difference in the ratios also occurs after the fourth parameter. The subset selection algorithms indicate that all eight parameters are identifiable, of which, six are estimable, ordered from most to least estimable: k, $V_{max_{Kid}}$, $V_{max_{Liv1}}$, D, $K_{m_{Kid}}$, and $K_{m_{Liv1}}$.



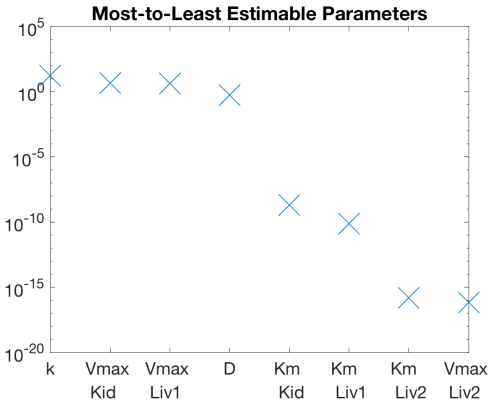
(a) Cain 1.7 ppm; Simple lung; MTBE-only (Model 1)



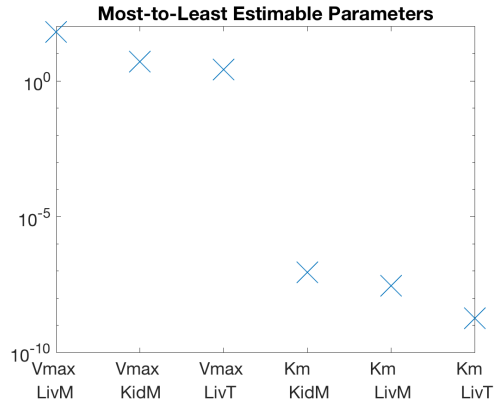
(b) Pleil 3.0 ppm; Simple lung; MTBE-only (Model 1)



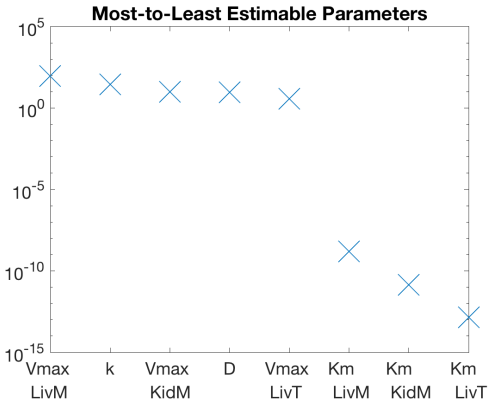
(c) Cain 1.7 ppm; Complex lung; MTBE-only (Model 2)



(d) Pleil 3.0 ppm; Complex lung; MTBE-only (Model 2)



(e) Pleil 3.0 ppm; Simple lung; MTBE-to-TBA coupled system (Model 3)



(f) Pleil 3.0 ppm; Complex lung; MTBE-to-TBA coupled system (Model 4)

Figure 9.12: Most-to-least estimable parameters (diagonal elements of R from QR factorization labeled with estimability results of Algorithm 2)

In Model 3, we estimate six parameters: $V_{max_{LivM}}$, $V_{max_{KidM}}$, $V_{max_{LivT}}$, $K_{m_{LivM}}$, $K_{m_{KidM}}$, and $K_{m_{LivT}}$, where 'M' and 'T' indicate the pathways that process MTBE and TBA, respectively. For the Pleil 3.0 ppm data, the largest difference in the ratios occurs after the third parameter (Figure 9.12e). But the subset selection algorithms indicate that all six parameters are identifiable, of which, all six are estimable.

Lastly, in Model 4, we estimate again eight parameters: $V_{max_{LivM}}$, $V_{max_{KidM}}$, $V_{max_{LivT}}$, $K_{m_{LivM}}$, $K_{m_{KidM}}$, $K_{m_{LivT}}$, k, and D. For the Pleil 3.0 ppm data, the largest difference in the ratios occurs after the fifth parameter (Figure 9.12f), indicating five parameters are estimable. The subset selection algorithms indicate that all eight parameters are identifiable, of which, seven are estimable, ordered from most to least estimable: $V_{max_{LivM}}$, k, $V_{max_{KidM}}$, D, $V_{max_{LivT}}$, $K_{m_{LivM}}$, and $K_{m_{KidM}}$.

For each subject, across all models, we completed an identifiability and estimability check. In each case, the check included assessing the identifiability and estimability prior to parameter estimation. That involved conducting the assessments using the "initial guess" values. Once the most estimable parameters were identified, we ran the models again to obtain the optimized parameters using only the most estimable parameters. We defer further analysis of those results and computation of confidence intervals and prediction intervals for future work.

CHAPTER 10

Conclusions and Future Work

10.1 Summary of Results and Conclusions

"Pharmacokinetics is the quantitative study of factors that control the time course for absorption, distribution, metabolism, and elimination of chemicals within the body [33]." Physiologically-based pharmacokinetic (PBPK) modeling adds knowledge about physiological processes along with physiochemical properties of the substance under study in order to predict or simulate the time-course behavior in venous blood and tissues of the substance that has been introduced into the body.

PBPK modeling involves the use of a number of substance-dependent and physiology-dependent parameters: (a) species-specific physiologic parameters (e.g., body weight, blood flow rate), (b) physiochemical properties of compounds of interest (tissue partition coefficients), and (c) relevant kinetics parameters (Michaelis-Menten V_{max} and K_m values for metabolism).

Many models consider each organ in the body as one or more compartments with the following assumptions [68]:

- Only the essential tissues will be used;
- Each organ is a well-mixed compartment;

- The substance is uniformly distributed in the plasma and tissues;
- The body represented by a series of tissues;
- The substance moves in and out of tissues and is satisfied with a system of ordinary differential equations (ODEs) and conservation equations; and
- Groups of tissues have similar blood flow and substance affinity.

Generally, major organs of the body are represented in a PBPK model but some models may use a "lumping" approach where different organs of the body are represented as a single kinetic system. This may be useful where an organ is of particular interest (e.g., the liver for metabolic processes) but kinetics in the remainder of the body is less important and can be represented as one or two lumped compartments such as rapidly or slowly perfused tissues. [91]

In Figure 10.1, each box corresponds to a well-mixed compartment and each arrow represents an input or output to the compartment. The transport to/from each compartment is determined by the blood flow, which for all models discussed, follows flow-limited perfusion.

Methyl Tertiary-Butyl Ether and Tert Butyl Alcohol

Methyl tertiary-butyl ether is a fuel oxgenate that saw widespread use as a result of the Clean Air Act Amendments of 1990. Its intended use was to make gasoline burn better and decrease carbon monoxide emissions but individuals exposed to MTBE at gasoline pumps, from occupational exposure, or other means reported having headaches, nausea, and dizziness. Additonally, MTBE was contaminating groundwater and reservoirs due to leaks in the underground containers used to store it. The adverse health effects and the potential large-scale hazard due to groundwater contamination led to a number of carcinogenicity studies to assess the risks and long-term effects of MTBE on animals [15, 16, 17, 26, 28, 84]. The overall results of those studies were that MTBE is possibly a human carcinogen [6, 44, 67, 90, 93].

Tert-butyl alcohol is a known metabolite of MTBE and has been shown to induce tumors in both rats and mice [65]. According to EPA cancer guidelines [93], evidence of carcinogenicity across species, sex, strain, site, or exposure route tends toward a descriptor of "likely to be carcinogenic to humans." Thus MTBE and TBA pose a potential carcinogenic risk to humans.

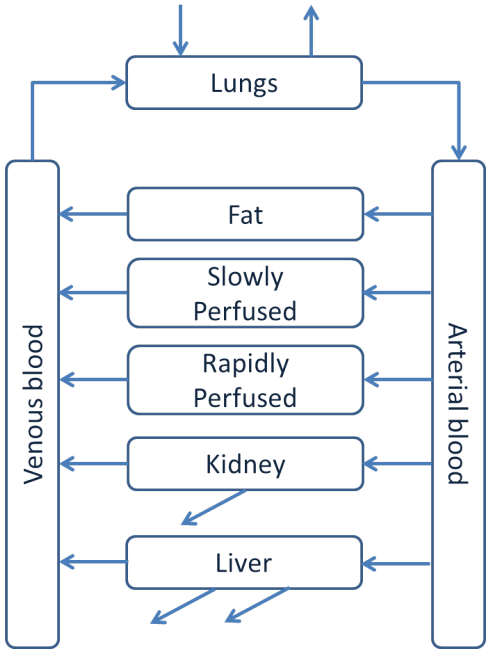


Figure 10.1: Eight-compartment PBPK model

Research Aims

We address three primary concerns with this work.

First, is the lung's upper respiratory tract (URT) required to fit data for volatile gases? There is general suggestion that, when the washin-washout effect (absorption in the respiratory tract of a substance during both inhalation and exhalation) is present, the URT is required to fit the data. During the washin-washin process, the full exposure amount of a substance does not reach the blood stream, thus less substance is circulated throughout the body. The hypothesis is that a substance that has a blood:air partition coefficient greater than 1000 ($P_{b:a} > 100$) should be modeled with the inclusion of the URT [47] versus a simple inert tube model.

Second, does including kidney metabolism help explain renal carcinogenicity endpoints? Chemically-induced $\alpha_{2\mu}$-globulin nephropathy is one mechanism for the development of kidney tumors in rodents. However, there is no evidence to indicate that MTBE caused $\alpha_{2\mu}$-globulin accumulation in rodents and it was determined that the increased tumor incidences was not related to the protein accumulation [35]. In the absence of this evidence and since there is no production of an analogous $\alpha_{2\mu}$-globulin protein in humans, we

consider the possibility that the kidney tumors (seen primarily in male rats) were induced by some other non-species-specific mechanism and we explore a PBPK model that includes kidney metabolism.

Last, can we improve upon model design with results and interpretations from sensitivity analysis techniques? A number of factors drive the initial design of a PBPK model, most notably tissues of interest and biokinetics processes. After a model is constructed, one should conduct analysis techniques on its results. One particularly relevant technique is sensitivity analysis which ultimately leads to a parameter ranking process to assess which model parameters and/or state variables, if any, can be eliminated from the model, thus resulting in a simpler model.

Model Simulation and Data Fitting

We developed a series of PBPK models for MTBE and TBA:

- ***Simple lung model for MTBE (Model 1, Figure E.1)***. Venous and arterial blood compartments are always included in a PBPK model but are generally reduced to an algebraic expression for the model. Because we have experimental data for MTBE (and TBA) concentration in venous blood, we leave this compartment as a differential equation expression. The lungs are the port of entry for MTBE and the tissue of interest in our hypothesis. Addtionally, we have experimental data for the concentration of MTBE (and TBA) in exhaled breath. Since MTBE is a liphophilic substance, we include an adipose tissue compartment. The liver is generally the primary organ for metabolism and we include a low- and high-affinity pathway to model this metabolism. Finally, the kidney is an organ of interest in our hypothesis and we include a single metabolism pathway.

- ***Complex lung model for MTBE (Model 2, Figure E.2)***. This model moves from an inert tube for the lung to a lung that includes the URT and a diffusion process between the URT and the pulmonary tissue to account for the washin-washin effect. It also accounts for the loss of substance that gets trapped in alveolar dead space. It further assumes that there is some fraction of the substance that remains in the URT during inhalation and exhalation. All other tissues are included for the reasons given for the simple lung model.

- ***Simple lung model for MTBE and its metabolite, TBA (Model 3, Figure***

E.4). This model takes the simple lung system provided for MTBE and duplicates it for TBA. Based on sensitivity analysis results for the MTBE-only models above (details in the following section), we reduce the liver pathways for MTBE in the liver to a single pathway. Since TBA is a metabolite of MTBE, the system is a coupled system, linked together by the liver pathways processing MTBE outputting TBA into the venous blood compartment of the second system. We assume a single metabolism pathway for TBA in the liver and no metabolism pathway for TBA in the kidney.

- ***Complex lung model for MTBE and its metabolite, TBA (Model 4, Figure E.5)***. This final model replicates the MTBE-only complex lung model for TBA. Again, based on sensitivity analysis results which showed no simultaneous sensitivity for two metabolism pathways, we have a single metabolism pathway for MTBE in the liver and continue with a single metabolism pathway for TBA in the liver and no pathway for TBA in the kidney.

All parameters for our models were taken from the literature with the exception of the Michaelis-Menten paramaters which were estimated using MATLAB's `fminsearch` command. "Initial guesses" required by `fminsearch` were either taken from the literature or assigned an arbitrary, but biologically plausible, value.

Data values for concentration of MTBE and TBA in venous blood and exhaled breath from the Cain et al. [27] and Pleil et al. [74] papers on inhalation exposure to MTBE were used to calibrate and validate the models.

Conclusions

We performed a sensitivity analysis for each subject (results not provided) and for the group of subjects with fitted parameters. We were particularly interested to see that the kidney parameters were shown to be important to the model in one of four subjects for the Cain (1.7 ppm) venous blood data set and five of six subjects for the Pleil data set (3 ppm) with venous blood and exhaled breath data. The results indicate that the inclusion of kidney metabolism in the model contributes meaningfully to further exploring and explaining tumor development in the kidneys and affirm that kidney metabolism is potentially a means for inducing these tumors.

Another significant result from the sensitivity analysis is that, in no individual subject,

did the two-pathway liver models show simultaneous sensitivity to both pathways. That is, we saw only a high sensitivity ranking for one pathway—either high affinity or low affinity—at one time. These results drove model design for the coupled systems, reducing kinetics from two pathways to one, thus allowing a simpler model and confirming that sensitivity analysis can—and does—lead to improved model design. The results also indicated that additional parameters could be reduced from the model(s) due to their low sensitivity analysis rankings. Further considerations for those model reductions are left for future work.

Finally, we address whether the inclusion of the upper respiratory tract (URT) is required to fit data for volatile gases. The Akaike Information Criterion helps answer this question. We saw in Section 9.5 and Tables (9.1)-(9.3) that the "simple lung" model for all but one subject was the preferred model since the AIC values for each of those were less than those for the "complex lung" models. This was expected for the models of the Cain data since we only considered MTBE which has a blood:air partition coefficient of 17.7. This is far lower than the suggested blood:air partition coefficient threshold for gases with $P_{b:a} > 100$ to require the inclusion of the URT.

For the Pleil data, we have a coupled system that includes MTBE ($P_{b:a} = 17.7$) and TBA ($P_{b:a} = 462$). The AIC values indicate that the URT is not needed to describe the MTBE or TBA data. In this case, the available data is best explained by an inert tube model thus not proving the hypothesis that gases with $P_{b:a} > 100$ require the upper respiratory tract.

10.2 Future Work

10.2.1 Nonlinear Mixed Effect Models (NLME)

We established in Section 8.2 that the two-stage approach to parameter estimation may fail to take into account the variability in the estimated parameter $\hat{\theta}$, particularly if few samples are collected during the absorption phase of a concentration-time profile. Further the variance and covariance for the parameter estimates are often exaggerated because the variance of the average parameters depend on both intra- and interindividual variations. [20, 21, 71]

We consider two types of models: fixed effects and random effects. Fixed-effect

parameters represent the population typical value, or central tendency, of the model parameter for the population. There is one true effect and the assumption is that all differences in observed effects are due to sampling error. The fixed effects describe the relationship between explanatory variables (e.g., age, body weight) and pharmacokinetic outcomes (e.g., plasma or tissue concentration of a substance). The values for random-effect parameters represent the variance of a distribution of some element of the model. The random-effect parameters themselves help to quantify the magnitude of unexplained variability in parameters or error in model predictions.

For population models, there are often levels of random effects. The first level addresses the difference in the model parameters between subjects (intersubject) or the differences between occasions (interoccasion) while second-level effects quantify the differences between observed and predicted values of the dependent variable. [71]

The functions, $f(x, \theta)$, under consideration in nonlinear mixed effects (NLME) models are nonlinear in the model parameters θ. Rather than model data from each individual in a set, population pharmacokinetics studies the PK of the entire set simultaneously. Nonlinear mixed effects models account for the intersubject, interoccasion and other elements of variability and inherently consider both fixed and random effects analysis. As a result, unlike two-stage estimation models, NLME models often produce unbiased mean and variance estimates and are better at detecting and characterizing nonlinearities in the physiological system [20]. Other advantages of NLME models are that rich data is not required (e.g., the number of data points may be small for each individual), the data may have irregular sampling times, and information from multiple individuals is used simultaneously.

All parameter estimation techniques thus far used the two-stage approach. In order to reduce the amount of variability in the parameter estimates, future work would include employing NLME models.

10.2.2 Parameter Estimation and Subset Selection

In Section 9.7, we discussed assessing the identifiability and estimability of our model parameters prior to estimation. Based on the estimability results of Algorithms 1-2 in Section 8.8.4, we estimate model parameters using only the most estimable parameters. Additional future work in this area would include analysis of the results (e.g., Is there an improvement in fitting the data?) and computation of the associated confidence intervals

and prediction intervals before and after parameter subset selection.

10.2.3 More Human Data: Vainiotalo, 25 ppm and 75 ppm

In addition to the Cain and Pleil data used to calibrate our models, we have available venous blood data at initial MTBE concentrations of 25 ppm and 75 ppm.

The subjects for this study were four males, aged 21-24. Volunteers were exposed on separate occasions to 25 ppm and 75 ppm of substance in air for 4 hours. Venous blood samples were collected: one just before exposure; four during exposure; and 6-7 over 20 hours post-exposure (at t= 0, 35, 100, 160, 225, 255, 270, 290, 360, 620, and 1525 minutes for 25 ppm and at t= 0, 80, 150, 205, 270, 290, 305, 330, 355, 573, and 1360 minutes for 75 ppm). The exposures were carried out in a dynamically controlled chamber measuring 15m^3. Exposure sessions were separated by a time interval of at least one week. Measurements were performed by gas chromatography using salting out technique.

Future work for these data sets would include parameter estimation via both two-stage and NLME techniques.

REFERENCES

[1] Mtbe. *ICIS Chemical Business*, 292(4):37, Aug 2017. Copyright - Copyright Reed Business Information UK Aug 11-Aug 17, 2017; Last updated - 2017-08-22.

[2] Agency for Toxic Substances and Disease Registry. Methyl-tert-butyl ether, 1996. `http://www.atsdr.cdc.gov/ToxProfiles/tp91-c1-b.pdf`.

[3] Shair Ahmad and Antonio Ambrosetti. *A textbook on ordinary differential equations*, volume 88. Springer, 2015.

[4] Hirotugu Akaike. Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory*, pages 276–281. Academiai Kiado, 1973.

[5] Richard A Albanese, HT Banks, Marina V Evans, and Laura K Potter. Physiologically based pharmacokinetic models for the transport of trichloroethylene in adipose tissue. *Bulletin of mathematical biology*, 64(1):97–131, 2002.

[6] American Cancer Society. *MTBE and Cancer Risk*, 2014. `https://www.cancer.org/cancer/cancer-causes/mtbe.html`.

[7] Joseph C Anderson, Albert L Babb, and Michael P Hlastala. Modeling soluble gas exchange in the airways and alveoli. *Annals of biomedical engineering*, 31(11):1402–1422, 2003.

[8] Soren Anderson and Andrew Elzinga. *A Ban on One is a Boon for the Other: Strict Gasoline Content Rules and Implicit Ethanol Blending Mandates*. Energy Institute at Haas, 2012 (accessed February 3, 2014). `https://ei.haas.berkeley.edu/research/papers/WP235.pdf`.

[9] Panos J Antsaklis and Anthony N Michel. *A linear systems primer*, volume 1. Birkhäuser Boston, 2007.

[10] RC Aster, B Borchers, and CH Thurber. Parameter estimation and inverse problems: Elsevier academic. 2005.

[11] KP Baetcke, GC Hard, IS Rodgers, RE McGaughy, and LM Tahan. Alpha 2u-globulin: Association with chemically induced renal toxicity and neoplasia in the male rat. In *US Environmental Protection Agency, Risk Assessment Forum, Washington, DC*, volume 20460, 1991. EPA/630/P-03/001B.

[12] HT Banks, Yanyuan Ma, and Laura K Potter. A simulation-based comparison between parametric and nonparametric estimation methods in pbpk models. *Journal of Inverse and Ill-posed Problems jiip*, 13(1):1–26, 2005.

[13] Stuart L Beal and Lewis B Sheiner. Methodology of population pharmacokinetics. *Drug Fate and Metabolism: Methods and Techniques*, 5:135–183, 1985.

[14] James Vere Beck and Kenneth J Arnold. *Parameter estimation in engineering and science*. James Beck, 1977.

[15] F Belpoggi, M Soffritti, and C Maltoni. Pathological characterization of testicular tumours and lymphomasleukemias, and of their precursors observed in sprague-dawley rats exposed to methyltertiary butyl ether (mtbe). *Eur. J. Oncol*, 3(3):201–206, 1998.

[16] Fiorella Belpoggi, Morando Soffritti, and Cesare Maltoni. Methyl-tertiary-butyl ether (mtbe), a gasoline additive, causes testicular and lympho haematopoietic cancers in rats. *Toxicology and Industrial Health*, 11(2):119–149, 1995.

[17] MG Bird, HD Burleigh-Flayer, JS Chun, JF Douglas, JJ Kneiss, and LS Andrews. Oncogenicity studies of inhaled methyl tertiary-butyl ether (mtbe) in cd-1 mice and f-344 rats. *Journal of Applied Toxicology*, 17(S1), 1997.

[18] Jerry N Blancato, Marina V Evans, Fred W Power, and Jane C Caldwell. Development and use of pbpk modeling and the impact of metabolism on variability in dose metrics for the risk assessment of methyl tertiary butyl ether (mtbe). *J. Environ. Prot. Sci*, 1:29–51, 2007.

[19] P Blanchard, RL Devaney, and GR Hall. Differential equations (2nd edn.) brooks/cole (2002). *MARK McCARTNEY and SHARON GIBSON School of Computing and Mathematics, University of Ulster, Shore Road, Newtownabbey, Northern Ireland.*

[20] Peter L Bonate and Jean-Louis Steimer. *Pharmacokinetic-pharmacodynamic modeling and simulation*. Springer, 2006.

[21] Michael Borenstein, Larry V Hedges, Julian Higgins, and Hannah R Rothstein. A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, 1(2):97–111, 2010.

[22] SJ Borghoff, JE Murphy, and MA Medinsky. Development of a physiologically based pharmacokinetic model for methyl tertiary-butyl ether and tertiary-butanol in male fischer-344 rats. *Toxicological Sciences*, 30(2):264–275, 1996.

[23] SJ Borghoff and TM Williams. Species-specific tumor responses following exposure to methyl tert-butyl ether. *CIIT ACTIVITIES*, 20(2), 2000.

[24] D Brockmann, K-H Rosenwinkel, and E Morgenroth. Practical identifiability of biokinetic parameters of a model describing two-step nitrification in biofilms. *Biotechnology and bioengineering*, 101(3):497–514, 2008.

[25] Ronald P Brown, Michael D Delp, Stan L Lindstedt, Lorenz R Rhomberg, and Robert P Beliles. Physiological parameter values for physiologically based pharmacokinetic models. *Toxicology and industrial health*.

[26] HD Burleigh-Flayer, JS Chun, and WJ Kintigh. Methyl tertiary butyl ether: vapor inhalation oncogenicity study in cd-1 mice. *Bushy Run Research Center, Export, PA, BRRC report 91N0013A, October*, 15, 1992.

[27] William S Cain, Brian P Leaderer, Gary L Ginsberg, Larry S Andrews, J Enrique Cometto-Muniz, Janneane F Gent, Marion Buck, Larry G Berglund, Vahid Mohsenin, Edward Monahan, et al. Acute exposure to low-level methyl tertiary-butyl ether (mtbe): Human reactions and pharmacokinetic response. *Inhalation toxicology*, 8(1):21–48, 1996.

[28] JS Chun, HD Burleigh Flayer, and WJ Kintigh. MTBE: vapor inhalation oncogenicity study in Fisher 344 rats, 1992.

[29] Bobbie Clark. Mtbe. *ICIS Chemical Business*, 291(21):38, May 2017. Copyright - Copyright Reed Business Information UK May 26-Jun 1, 2017; Last updated - 2017-06-07.

[30] Samuel M Cohen, ME Meek, James E Klaunig, Dorothy E Patton, and Penelope A Fenner-Crisp. The human relevance of information on carcinogenic modes of action: overview. *Critical reviews in toxicology*, 33(6):581–589, 2003.

[31] California Energy Commission et al. Supply and cost of alternatives to mtbe in gasoline. *Technical Appendix. Ethanol Blending Properties for Task*, 3, 1998.

[32] JJ DiStefano and Elliot M Landaw. Multiexponential, multicompartmental, and noncompartmental modeling. i. methodological limitations and physiological interpretations. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 246(5):R651–R664, 1984.

[33] JJ DiStefano and Elliot M Landaw. Title. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 246(5):R651–R664, 1984.

[34] Energy Information Administration (EIA DOE). MTBE, Oxygenates, and Motor Gasoline, 2000. `http://www.eia.gov/forecasts/steo/special/pdf/mtbe.pdf`.

[35] Adriana M Doi, Georgette Hill, John Seely, James R Hailey, Grace Kissling, and John R Bucher. α2u-globulin nephropathy and renal tumors in national toxicology program studies. *Toxicologic pathology*, 35(4):533–540, 2007.

[36] Vicky Ellis. Mtbe. *ICIS Chemical Business*, 291(21):7, May 2017. Copyright - Copyright Reed Business Information UK May 26-Jun 1, 2017; Last updated - 2017-06-07.

[37] International Agency for Research on Cancer et al. Some chemicals that cause tumours of the kidney or urinary bladder in rodents and some other substances. In *Some chemicals that cause tumours of the kidney or urinary bladder in rodents and some other substances*. International Agency for Research on Cancer, 1999.

[38] Keith Godfrey. Compartmental models and their application. In *Compartmental models and their application*. Academic Press, 1983.

[39] Jérémie Guedj, Rodolphe Thiébaut, and Daniel Commenges. Practical identifiability of hiv dynamics models. *Bulletin of mathematical biology*, 69(8):2493–2513, 2007.

[40] Halim Hamid and Mohammed Ashraf Ali. *Handbook of MTBE and other gasoline oxygenates*. CRC Press, 2004.

[41] Jun-Yan Hong, Chung S Yang, Maojung Lee, Yong-Yu Wang, Wei-qun Huang, Yizheng Tan, Christopher J Patten, and Flordeliza Y Bondoc. Role of cytochromes p450 in the metabolism of methyl tert-butyl ether in human livers. *Archives of toxicology*, 71(4):266–269, 1997.

[42] Duncan E Hutcheon. Disposition, metabolism, and toxicity of methyl tertiary butyl ether, an oxygenate for reformulated gasoline. *Journal of Toxicology and Environmental Health Part A*, 47(5):453–464, 1996.

[43] Independent Chemical Information Service (ICIS). *TIMELINE: A very short history of MTBE in the US*, 2006 (accessed February 3, 2014). `https://www.icis.com/resources/news/2006/07/05/1070674/timeline-a-very-short-history-of-mtbe-in-the-us/`.

[44] International Agency for Research on Cancer. *IARC MONOGRAPHS ON THE EVALUATION OF CARCINOGENIC RISKS TO HUMANS*, 2006. `http://monographs.iarc.fr/ENG/Preamble/index.php`.

[45] John A Jacquez and Timothy Perry. Parameter estimation: local identifiability of parameters. *American Journal of Physiology-Endocrinology and Metabolism*, 258(4):E727–E736, 1990.

[46] Barath Ram Jayasankar, Amos Ben-Zvi, and Biao Huang. Identifiability and estimability study for a dynamic solid oxide fuel cell model. *Computers & Chemical Engineering*, 33(2):484–492, 2009.

[47] Gunnar Johanson. Modelling of respiratory exchange of polar solvents. *Annals of occupational hygiene*, 35(3):323–339, 1991.

[48] Gunnar Johanson, Annsofi Nihlén, and Agneta Löf. Toxicokinetics and acute effects of mtbe and etbe in male volunteers. *Toxicology letters*, 82:713–718, 1995.

[49] Kenneth A Johnson and Roger S Goody. The original michaelis constant: translation of the 1913 michaelis–menten paper. *Biochemistry*, 50(39):8264–8269, 2011.

[50] Julian King, Karl Unterkofler, Gerald Teschl, Susanne Teschl, Pawel Mochalski, Helin Koç, Hartmann Hinterhuber, and Anton Amann. A modeling-based evaluation of isothermal rebreathing for breath gas analyses of highly soluble volatile organic compounds. *Journal of breath research*, 6(1):016005, 2012.

[51] Bo Kou, Kim B McAuley, CC Hsu, David W Bacon, and K Zhen Yao. Mathematical model and parameter estimation for gas-phase ethylene homopolymerization with supported metallocene catalyst. *Industrial & engineering chemistry research*, 44(8):2428–2442, 2005.

[52] EM Landaw and JJD DiStefano. Multiexponential, multicompartmental, and noncompartmental modeling. ii. data analysis and statistical considerations. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 246(5):R665–R677, 1984.

[53] Marc Lavielle and Leon Aarons. What do we mean by identifiability in mixed effects models? *Journal of pharmacokinetics and pharmacodynamics*, 43(1):111–122, 2016.

[54] Chia-Wei Lee, Sandra N Mohr, and Clifford P Weisel. Toxicokinetics of human exposure to methyl tertiary-butyl ether (mtbe) following short-term controlled exposures. *Journal of exposure analysis and environmental epidemiology*, 11(2):67–78, 2000.

[55] Dongmei Li, Chuntao Yuan, Yi Gong, Yufeng Huang, and Xiaodong Han. The effects of methyl tert-butyl ether (mtbe) on the male rat reproductive system. *Food and chemical toxicology*, 46(7):2402–2408, 2008.

[56] Pu Li and Quoc Dong Vu. Identification of parameter correlations for parameter estimation in dynamic biological models. *BMC systems biology*, 7(1):1, 2013.

[57] Panos Macheras and Athanassios Iliadis. *Modeling in biopharmaceutics, pharmacokinetics, and pharmacodynamics*, volume 30 of *Interdisciplinary Applied Mathematics*. Springer, 2016.

[58] Johan Mailier, A Delmotte, M Cloutier, M Jolicoeur, and A Vande Wouwer. Parametric sensitivity analysis and reduction of a detailed nutritional model of plant cell cultures. *Biotechnology and bioengineering*, 108(5):1108–1118, 2011.

[59] The Mathworks, Inc., Natick, Massachusetts. *MATLAB version 9.0.0.341360 (R2016a)*, 2016.

[60] R McKee, M Molyneux, and BJ Simpson. *The health hazards and exposures associated with gasoline containing MTBE*. CONCAWE, 1997.

[61] Leonor Michaelis and Maud Menten. Die kinetik der invertinwirkung. *Biochem Z*, 49:333–369, 1913.

[62] Gen Nakamura and Roland Potthast. Inverse modeling. *December 2015*, 2015.

[63] National Center for Biotechnology Information. tert-Butanol, 2017. `https://pubchem.ncbi.nlm.nih.gov/compound/6386`.

[64] National Center for Biotechnology Information. tert-Butyl methyl ether, 2017. `https://pubchem.ncbi.nlm.nih.gov/compound/15413`.

[65] National Toxicology Program. *Toxicology and carcinogenesis studies of t-butyl alcohol (CAS NO. 75-65-0) in f344/n rats and b6c3f1 mice (drinking water studies)*, 1995. `https://ntp.niehs.nih.gov/ntp/htdocs/lt_rpts/tr436.pdf`.

[66] National Toxicology Program. *Scientific Review of Methyl-t-Butyl Ether (MtBE)*, 1999. `https://ntp.niehs.nih.gov/go/19669`.

[67] National Toxicology Program. *The criteria for listing an agent, substance, mixture, or exposure circumstance in the Report on Carcinogens (RoC)*, 2011. `https://ntp.niehs.nih.gov/pubhealth/roc/criteria/index.html`.

[68] Ivan Nestorov. Whole body pharmacokinetic models. *Clinical pharmacokinetics*, 42(10):883–908, 2003.

[69] College of Chemisty and Biochemisty. *Derivation of the Michaelis-Menten Equation*. University of Oklahoma, 2003 (accessed February 3, 2014). `http://www.ou.edu/OpenEducation/ou-resources/biochemical-methods/lab-11/michaelis-menten-derivation.pdf`.

[70] Mette S Olufsen and Johnny T Ottesen. A practical approach to parameter estimation applied to model predicting heart rate regulation. *Journal of mathematical biology*, 67(1):39–68, 2013.

[71] Joel S Owen and Jill Fiedler-Kelly. *Introduction to population pharmacokinetic/pharmacodynamic analysis with nonlinear mixed effects models*. John Wiley & Sons, 2014.

[72] Sheila Annie Peters. *Physiologically-based pharmacokinetic (PBPK) modeling and simulations: principles, methods, and applications in the pharmaceutical industry*. John Wiley & Sons, 2012.

[73] JR Peterson. Alkylate is key for cleaner burning gasoline. *Preprints of Papers, American Chemical Society, Division of Fuel Chemistry*, 41(CONF-960807–), 1996.

[74] JD Pleil, D Kim, JD Prah, and SM Rappaport. Exposure reconstruction for reducing uncertainty in risk assessment: example using mtbe biomarkers and a simple pharmacokinetic model. *Biomarkers*, 12(4):331–348, 2007.

[75] Torka S Poet and Susan J Borghoff. In vitrouptake of methyltert-butyl ether in male rat kidney: Use of a two-compartment model to describe protein interactions. *Toxicology and applied pharmacology*, 145(2):340–348, 1997.

[76] Kese Pontes Freitas Alberton, André Luís Alberton, Jimena Andrea Di Maggio, Vanina Gisela Estrada, María Soledad Díaz, and Argimiro Resende Secchi. Simultaneous parameters identifiability and estimation of an e. coli metabolic network model. *BioMed research international*, 2015, 2015.

[77] James Prah, David Ashley, Benjamin Blount, Martin Case, Teresa Leavens, Joachim Pleil, and Frederick Cardinali. Dermal, oral, and inhalation pharmacokinetics of methyl tertiary butyl ether (mtbe) in human volunteers. *Toxicological Sciences*, 77(2):195–205, 2004.

[78] Judith S Prescott-Mathews, Douglas C Wolf, Brian A Wong, and Susan J Borghoff. Methyltert-butyl ether causes α2u-globulin nephropathy and enhanced renal cell proliferation in male fischer-344 rats. *Toxicology and applied pharmacology*, 143(2):301–314, 1997.

[79] Tom Quaiser and Martin Mönnigmann. Systematic identifiability testing for unambiguous mechanistic modeling–application to jak-stat, map kinase, and nf-κ b signaling pathway models. *BMC systems biology*, 3(1):50, 2009.

[80] Micaela Reddy, RS Yang, Melvin E Andersen, and Harvey J Clewell III. *Physiologically based pharmacokinetic modeling: science and applications*. John Wiley & Sons, 2005.

[81] T Agami Reddy. *Applied data analysis and modeling for energy engineers and scientists*. Springer Science & Business Media, 2011.

[82] Elisabeth Reese and Renate D Kimbrough. Acute toxicity of gasoline and some additives. *Environmental health perspectives*, 101(Suppl 6):115, 1993.

[83] Stephen M Roberts, Robert C James, and Phillip L Williams. *Principles of toxicology: environmental and industrial applications*. John Wiley & Sons, 2014.

[84] M Robinson, RH Bruner, and GR Olson. Fourteen-and ninety-day oral toxicity studies of methyl tertiary-butyl ether in sprague-dawley rats. *Journal of the American College of Toxicology*, 9(5):525–540, 1990.

[85] Joy Bellis Sakai. *Practical pharmacology for the pharmacy technician*. Lippincott Williams & Wilkins, 2009.

[86] Matthew S Shotwell and Richard A Gray. Estimability analysis and optimal design in dynamic multi-scale models of cardiac electrophysiology. *Journal of Agricultural, Biological, and Environmental Statistics*, 21(2):261–276, 2016.

[87] American Cancer Society. MTBE, 2014. `http://www.cancer.org/cancer/cancercauses/othercarcinogens/pollution/mtbe`.

[88] Albert Tarantola. *Inverse problem theory and methods for model parameter estimation*. SIAM, 2005.

[89] Duncan E Thompson, Kim B McAuley, and P James McLellan. Parameter estimation in a simplified mwd model for hdpe produced by a ziegler-natta catalyst. *Macromolecular Reaction Engineering*, 3(4):160–177, 2009.

[90] Patty Toccalino. *Human-Health Effects of MTBE: A Literature Summary*. U.S. Geological Survey, 2013. `http://sd.water.usgs.gov/nawqa/vocns/mtbe_hh_summary.html`.

[91] Richard N Upton, David JR Foster, and Ahmad Y Abuhelwa. An introduction to physiologically-based pharmacokinetic models. *Pediatric Anesthesia*, 26(11):1036–1046, 2016.

[92] U.S. Energy Information Administration. *Status and Impact of State MTBE Ban*, 2003 (accessed February 3, 2014). `http://www.eia.gov/oiaf/servicerpt/mtbeban/`.

[93] US Environmental Protection Agency. *Guidelines for Carcinogen Risk Assessment*, March 2005. EPA/630/P-03/001B.

[94] Sinikka Vainiotalo, Vesa Riihimäki, Kaija Pekari, Eija Teräväinen, and Antero Aitio. Toxicokinetics of methyl tert-butyl ether (mtbe) and tert-amyl methyl ether (tame) in humans, and implications to their biological monitoring. *Journal of occupational and environmental hygiene*, 4(10):739–750, 2007.

[95] S Vajda, H Rabitz, E Walter, and Y Lecourtier. Qualitative and quantitative identifiability analysis of nonlinear chemical kinetic models. *Chemical Engineering Communications*, 83(1):191–219, 1989.

[96] P Valko and Sandor Vajda. An extended ode solver for sensitivity calculations. *Computers & Chemistry*, 8(4):255–271, 1984.

[97] K Zhen Yao, Benjamin M Shaw, Bo Kou, Kim B McAuley, and DW Bacon. Modeling ethylene/butene copolymerization with multi-site catalysts: parameter estimability and experimental design. *Polymer Reaction Engineering*, 11(3):563–588, 2003.

APPENDICES

APPENDIX A

Table of Abbreviations and Symbols

Table A.1: Tissue/Organ/Compartment Abbreviations

Tissue/Organ/Compartment Abbreviations			
K, Kid	Kidney	S, SPT	Slowly perfused tissue
F	Fat	R, RPT	Rapidly perfused tissue
L, Liv	Liver	URTinh	Upper respiratory tract (inhalation)
Lu	Lung	URTexh	Upper respiratory tract (exhalation)
PT	Pulmonary Tissue	alv(DS)	Alveoli, dead space
v	Venous	alv(P)	Alveoli, perfused
a	Arterial		

Table A.2: Equation Abbreviations

Equation Abbreviations	
A_i^x	Amount of chemical x in compartment i (mg)
C_{exh}^x	Concentration of chemical x in exhaled air (mg/L)
C_i^x	Concentration of chemical x in tissue i (mg/L)
C_{inh}^x	Inhaled concentration of chemical x during the exposure and zero after the exposure ends (mg/L)
$C_{v_i}^x$	Concentration of chemical x in venous blood from tissue i (mg/L)
k	The fraction of chemical leaving the URT ($0 < k < 1$)
$K_{max_i}^x$	(Michaelis-Menten) Concentration of chemical x in tissue i (mg/L) at $\frac{1}{2}V_{max_i}^x$
$Metab_i^x$	Metabolism rate of chemical x in tissue i (mg/hr)
$P_{b:a}^x$	Blood:air partition coefficient for chemical x
P_i^x	Tissue:blood partition coefficient for chemical x in tissue i
Q	Cardiac blood output (L/hr)
Q_i	Blood flow rate to tissue i (L/hr)
Q_p	Respiration rate (L/hr)
V_i	Volume of tissue i (L)
$\dot{V}_j$	Alveolar blood flow rates $j \in \{alv, DS\}$ (L/hr)
$V_{max_i}^x$	(Michaelis-Menten) Maximum velocity of enzyme reaction of chemical x in tissue i (mg/hr/kg)
f	Breathing frequency (breaths/hr)
V_T	Tidal volume (L)
V_D	Dead space volume (L)

APPENDIX B

MTBE and TBA Properties

Table B.1: MTBE Properties

MTBE Property	Value
Other names	Methyl *tertiary*-butyl ether; Methyl *tert*-butyl ether; Methyl *t*-butyl ether; *tert*-Butyl methyl ether; MTBE;
CAS Number	1634-04-4
Chemical formula	$C_5H_{12}O$
Molecular Weight	88.15 g/mol
Color	clear, colorless
Physical state	liquid

Table B.2: TBA Properties

TBA Property	Value
Other names	*tert*-Butyl alcohol; *t*-Butyl alcohol; tert-Butanol; Trimethyl carbinol; 2-Methyl-2-propanol; TBA
CAS Number	75-65-0
Chemical formula	$C_4H_{10}O$
Molecular Weight	74.12 g/mol
Color	clear, colorless
Physical state	liquid

APPENDIX C

Michaelis-Menten Equation Derivation

The Michaelis-Menten kinetics equation begins with the following chemical reaction [69]:

$$S + E \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} P + E \tag{C.1}$$

where S is substrate, E is enzyme, P is product, and ES is enzyme-substrate complex. k_1 is the forward reaction rate where the substrate and enzyme form the ES complex. k_{-1} is the backward reaction rate where the complex dissociates back to the substrate and enzyme. Lastly, k_2 is the rate for the creation of the product.

The following assumptions apply:

- The early stages of the reaction so little product is formed that the reverse reaction can be ignored
- The concentration of the substrate is much greater than that of total enzyme ($[S] \gg E_t]$) so it can be treated as a constant.
- For enzyme kinetics, the rate of change of enzyme-substrate is zero, that is, there is a steady-state condition, $\frac{d[ES]}{dt} = 0$. Alternatively stated, the rate of formation of $[ES]$ equals the rate of breakdown of $[ES]$,

$$k_1[E][S] = k_{-1}[ES] + k_2[ES] = (k_{-1} + k_2)[ES]. \tag{C.2}$$

We start with the rate of product concentration,

$$\frac{dP}{dt} = v = k_2[ES]. \tag{C.3}$$

$[ES]$ is a generally unmeasurable so we need an alternative expression for it. Let E_t be the concentration of the total enzyme, $[E_t] = [ES] + [E]$ where $[ES]$ is bound enzyme and $[E]$ is unbound enzyme. Then the fraction of bound enzyme $[ES]$ to total enzyme $[E_t]$ is

$$\frac{[ES]}{[E_t]} = \frac{[ES]}{[ES] + [E]}. \tag{C.4}$$

Rewriting Equation (C.4), we have

$$[ES] = \frac{[E_t]}{1 + \frac{[E]}{[ES]}}, \tag{C.5}$$

then substituting for the ratio $\frac{[E]}{[ES]}$ using Equation (C.2), we have

$$[ES] = \frac{[E_t]}{1 + \frac{k_{-1}+k_2}{k_1[S]}}. \tag{C.6}$$

Define the Michaelis constant, a measure of the substrate concentration required for effective catalysis to occur,

$$K_m = \frac{k_{-1} + k_2}{k_1}, \tag{C.7}$$

then, simplifying Equation (C.6), we have

$$[ES] = \frac{[E_t][S]}{K_m + [S]}. \tag{C.8}$$

Substituting this into the product concentration rate, Equation (C.3),

$$v = k_2[ES] = \frac{k_2[E_t][S]}{K_m + [S]}, \tag{C.9}$$

and considering the case when $[S] \gg K_m$, we have $v \approx k_2[E_t] = k_{cat}[E_t]$ where k_2 behaves as the catalytic constant, k_{cat}, called the "turnover number" (the number of substrate molecules turned over per enzyme molecule per second) and gives a direct measure of the

catalytic production of product under optimum conditions (saturated enzyme). When $[S]$ is "saturating," the enzyme is functioning as fast as it can (all enzyme molecules are tied up with $[S]$, or $[ES] = [E_t]$)and we can define $k_2[E_t] = k_{cat}[E_t] = V_{max}$ where V_{max} is the maximum production rate (maximum velocity) that can be obtained. Then Equation (C.9) can be rewritten into the more familiar Michaelis-Menten equation,

$$v = k_2[ES] = \frac{V_{max}[S]}{K_m + [S]}. \tag{C.10}$$

Now, consider when $v = \frac{1}{2}V_{max}$. Under this condition and simplifying Equation (C.10), we find $K_m = [S]$. In addition to the formal definition of K_m as in Equation (C.7), K_m is also the substrate concentration that gives half-maximal velocity of the enzyme-catalyzed reaction. Concluding, the Michaelis-Menten equation given by Equation (C.10) describes the kinetic behavior of an enzyme that acts according to the simple model (Equation (C.1)).

APPENDIX D

Local Identifiability

According to Shotwell et al. [86], a model is *locally identifiable* at θ' if $y(\cdot, \theta) = y(\cdot, \theta)$ implies $\theta = \theta'$ for all possible θ holds for θ in a neighborhood of θ', which suggests that the problem can be analyzed by linearization of $y(x, \theta)$ about θ'. The first-order Taylor series approximation of $y(x, \theta)$ in a neighborhood of θ' is

$$y(x, \theta) \approx y(x, \theta') + J(x, \theta')(\theta - \theta'),$$

where the elements of the Jacobian $J(x, \theta')$ are

$$\mathbf{J}(x, \theta') = \left[\frac{\partial y_i(x, \theta)}{\partial \theta_j}\right]_{\theta=\theta'}, \tag{D.1}$$

for $i = 1, \ldots, m$ and $j = 1, \ldots, p$. Thus, locally identifiability is verified when, for some x, there are no nontrivial solutions (i.e., $\theta \neq \theta'$) to the linear system,

$$\mathbf{J}(x, \theta')(\theta - \theta') = \mathbf{0}.$$

By definition, this requires linear independence in the columns of $\mathbf{J}(x, \theta')$. Local identifiability can then be verified by searching for two or more experiments that jointly

ensure no nontrivial solutions to

$$\mathbf{J}(x, \theta')(\theta - \theta') = \mathbf{0}, \tag{D.2}$$

where $\mathbf{J}(x, \theta')$ is an $nm \times p$ matrix consisting of n blocks of $m \times p$ Jacobians and x are the n associated experimental conditions. Local identifiability about θ is then determined by checking for column-wise linear independence in $\mathbf{J}(x, \theta')$, the sensitivity matrix, or equivalently that $\mathbf{F}(x, \theta') = \mathbf{J}(x, \theta')^T \mathbf{J}(x, \theta')$, the FIM, is nonsingular.

APPENDIX E

Models

E.1 MTBE-only system, SIMPLE lung. Cain 1.7ppm, Pleil 3ppm

Lung:

$$V_{lu}\frac{dC_{lu}}{dt} = C_v Q + C_{inh}Q_p - C_{exh}Q_p - \frac{C_{lu}}{P_{lu}}Q \tag{E.1}$$

Liver:

$$V_L\frac{dC_L}{dt} = Q_L(C_a - C_{v_L}) - \frac{V_{max_{L_1}}C_{v_L}}{K_{m_{L_1}} + C_{v_L}} - \frac{V_{max_{L_2}}C_{v_L}}{K_{m_{L_2}} + C_{v_L}} \tag{E.2}$$

Kidney:

$$V_K\frac{dC_K}{dt} = Q_K(C_a - C_{v_K}) - \frac{V_{max_K}C_{v_K}}{K_{m_K} + C_{v_K}} \tag{E.3}$$

Fat, SPT, RPT:

$$V_i\frac{dC_i}{dt} = Q_i(C_a - C_{v_i}) \tag{E.4}$$

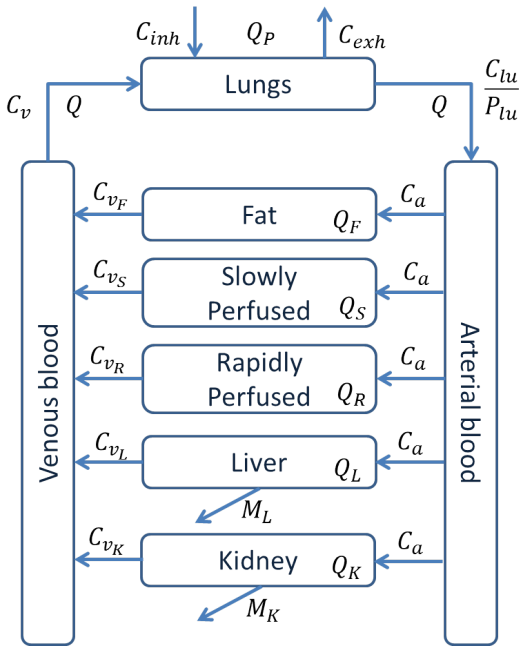


Figure E.1: MTBE-only system with simple lung

Venous:

$$V_v \frac{dC_v}{dt} = Q_L C_{v_L} + Q_K C_{v_K} + Q_F C_{v_F} + Q_S C_{v_S} + Q_R C_{v_R} - Q C_v \tag{E.5}$$

Arterial:

$$V_a \frac{dC_a}{dt} = Q \frac{C_{lu}}{P_{lu}} - (Q_L C_a + Q_K C_a + Q_F C_a + Q_S C_a + Q_R C_a) \tag{E.6}$$

E.1.1 Assumptions

A1. There is no steady-state condition assumed for any compartment. Compartment concentrations are determined solely from their differential equations.

A2. Metabolism/clearance takes place in both the kidneys and liver.

E.1.2 Additional Equations

E1. The outgoing concentration for each tissue is given by the concentration of the chemical in the tissue divided by the tissue:blood partition coefficient for that

chemical:

$$C_{v_i} = \frac{C_i}{P_i} \tag{E.7}$$

In each of (E.2)-(E.5), we can replace C_{v_i} using (E.7).

E2. Total cardiac flow:

$$Q = Q_L + Q_K + Q_F + Q_S + Q_R \tag{E.8}$$

which reduces ((E.6)) to

$$V_a \frac{dC_a}{dt} = Q\left(\frac{C_{lu}}{P_{lu}} - C_a\right) \tag{E.9}$$

E3. Total body volume:

$$V = V_L + V_K + V_F + V_S + V_R + V_{lu} + V_v + V_a \tag{E.10}$$

E4. Exhaled concentration with $P_{lu:air}$ computed from $P_{lung:blood}$ and $P_{blood:air}$.

$$C_{exh} = \frac{C_{lu}}{P_{lu:air}} \tag{E.11}$$

E.2 MTBE-only system, COMPLEX lung. Cain 1.7ppm, Pleil 3ppm, Vainiotalo 25ppm and 75ppm

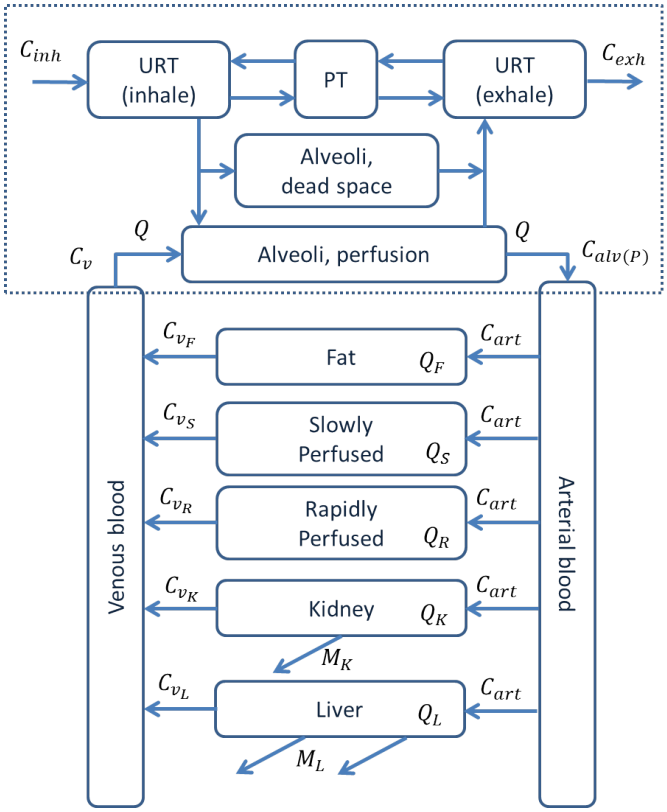


Figure E.2: MTBE-only system with complex lung

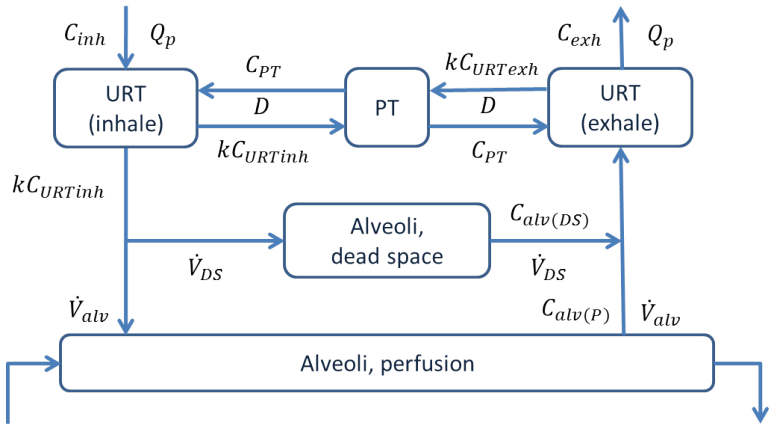


Figure E.3: Upper Respiratory Tract (URT) system

E.2.1 Equations

Upper Respiratory Tract (URT), inhalation:

$$V_{URT}\frac{dC_{URTinh}}{dt} = C_{inh}Q_p + (C_{PT} - kC_{URTinh})\,D - kC_{URTinh}\dot{V}_{alv} - kC_{URTinh}\dot{V}_{DS} \tag{E.12}$$

Pulmonary Tissue (PT):

$$V_{PT}\frac{dC_{PT}}{dt} = (kC_{URTinh} + kC_{URTexh} - 2C_{PT})\,D \tag{E.13}$$

Upper Respiratory Tract (URT), exhalation:

$$V_{URT}\frac{dC_{URTexh}}{dt} = (C_{PT} - kC_{URTexh})\,D - C_{exh}Q_p + C_{alv(P)}\dot{V}_{alv} + C_{alv(DS)}\dot{V}_{DS} \tag{E.14}$$

Alveoli with Perfusion:

$$V_{alv(P)}\frac{dC_{alv(P)}}{dt} = C_vQ + \hat{C}_{URTinh}\dot{V}_{alv} - C_{art}Q - C_{alv(P)}\dot{V}_{alv} \tag{E.15}$$

Alveoli without Perfusion (Alveolar deadspace):

$$V_{alv(DS)}\frac{dC_{alv(DS)}}{dt} = kC_{URTinh}\dot{V}_{DS} - C_{alv(DS)}\dot{V}_{DS} \tag{E.16}$$

Liver:

$$V_L \frac{dC_L}{dt} = Q_L(C_{art} - C_{v_L}) - \frac{V_{max_{L_1}} C_L}{K_{m_{L_1}} + C_L} - \frac{V_{max_{L_2}} C_L}{K_{m_{L_2}} + C_L} \tag{E.17}$$

Kidney:

$$V_K \frac{dC_K}{dt} = Q_K(C_{art} - C_{v_K}) - \frac{V_{max_K} C_K}{K_{m_K} + C_K} \tag{E.18}$$

Fat, SPT, RPT:

$$V_i \frac{dC_i}{dt} = Q_i(C_{art} - C_{v_i}) \tag{E.19}$$

Venous:

$$V_v \frac{dC_v}{dt} = Q_L C_{v_L} + Q_K C_{v_K} + Q_F C_{v_F} + Q_S C_{v_S} + Q_R C_{v_R} - Q C_v \tag{E.20}$$

Arterial:

$$V_a \frac{dC_{art}}{dt} = Q C_{alv(P)} - (Q_L C_{art} + Q_K C_{art} + Q_F C_{art} + Q_S C_{art} + Q_R C_{art}) \tag{E.21}$$

E.2.2 Variable Relationships

$$\dot{V}_{alv} = f(V_T - V_D) \tag{E.22}$$

$$\dot{V}_{DS} = f(V_D) \tag{E.23}$$

$$\Lambda = \{F, R, S, K, L\} \tag{E.24}$$

E.2.3 Assumptions

A1. Metabolism/clearance takes place in both the kidneys and liver.

A2. Total cardiac flow: $Q = Q_L + Q_K + Q_F + Q_S + Q_R$ which reduces ((E.21)) to

$$V_a \frac{dC_{art}}{dt} = Q\left(C_{alv(P)} - C_{art}\right) \tag{E.25}$$

A3. Total body volume:

$$V = V_L + V_K + V_F + V_S + V_R + V_{URT} + V_{PT} + V_{alv(P)} + V_{alv(DS)} + V_v + V_a \tag{E.26}$$

A4. The exiting concentration for each tissue is given by the concentration of the chemical in tissue i divided by the tissue:blood partition coefficient for that chemical:

$$C_{v_i} = \frac{C_i}{P_i}, i \in \Lambda \tag{E.27}$$

In each of (E.17)-(E.20), we can replace C_{v_i} using (E.27).

E.2.4 Further considerations

- EPA TCE paper also has clearance at the Pulmonary Tissue. No clearance at that compartment is shown here.
- On inhalation and exhalation, some fraction of the chemical remains in the URT. Since $Amount = Concentration * Volume$ and volume is constant, we define the concentrations leaving the URT

$$\hat{C}_{URTinh} = kC_{URTinh} \tag{E.28}$$

and

$$C_{exh} = \hat{C}_{URTexh} = kC_{URTexh} \tag{E.29}$$

where k is the fraction of chemical leaving the URT ($0 < k < 1$). (Compare this use of k to that of partition coefficients.)

E.3 MTBE-to-TBA coupled system, SIMPLE lung. Pleil 3ppm

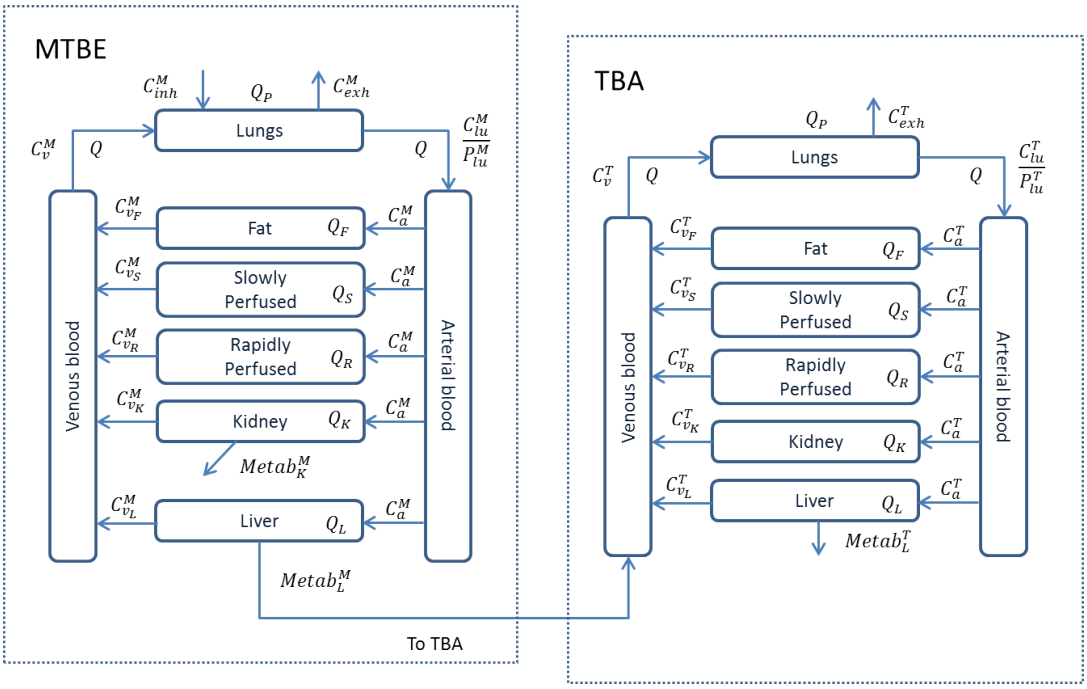


Figure E.4: MTBE-to-TBA coupled system with simple lung

E.3.1 Equations

Lung:

$$V_{lu}\frac{dC_{lu}^{M}}{dt} = C_{v}^{M}Q + C_{inh}^{M}Q_{p} - C_{exh}^{M}Q_{p} - \frac{C_{lu}^{M}}{P_{lu}^{M}}Q \tag{E.30}$$

$$V_{lu}\frac{dC_{lu}^{T}}{dt} = C_{v}^{T}Q - C_{exh}^{T}Q_{p} - \frac{C_{lu}^{T}}{P_{lu}^{T}}Q \tag{E.31}$$

Liver:

$$V_L \frac{dC_L^M}{dt} = Q_L(C_a^M - C_{v_L}^M) - Metab_L^M \tag{E.32}$$

$$V_L \frac{dC_L^T}{dt} = Q_L(C_a^T - C_{v_L}^T) - Metab_L^T \tag{E.33}$$

Kidney:

$$V_K \frac{dC_K^M}{dt} = Q_K(C_a^M - C_{v_K}^M) \tag{E.34}$$

$$V_K \frac{dC_K^T}{dt} = Q_K(C_a^T - C_{v_K}^T) \tag{E.35}$$

Fat, SPT, RPT:

$$V_i \frac{dC_i^M}{dt} = Q_i(C_a^M - C_{v_i}^M) \tag{E.36}$$

$$V_i \frac{dC_i^T}{dt} = Q_i(C_a^T - C_{v_i}^T) \tag{E.37}$$

Venous:

$$V_v \frac{dC_v^M}{dt} = Q_L C_{v_L}^M + Q_K C_{v_K}^M + Q_F C_{v_F}^M + Q_S C_{v_S}^M + Q_R C_{v_R}^M - Q C_v^M \tag{E.38}$$

$$V_v \frac{dC_v^T}{dt} = Metab_L^M + Q_L C_{v_L}^T + Q_K C_{v_K}^T + Q_F C_{v_F}^T + Q_S C_{v_S}^T + Q_R C_{v_R}^T - Q C_v^T \tag{E.39}$$

Arterial:

$$V_a \frac{dC_a^M}{dt} = Q \frac{C_{lu}^M}{P_{lu}^M} - (Q_L C_a^M + Q_K C_a^M + Q_F C_a^M + Q_S C_a^M + Q_R C_a^M) \tag{E.40}$$

$$V_a \frac{dC_a^T}{dt} = Q \frac{C_{lu}^T}{P_{lu}^T} - (Q_L C_a^T + Q_K C_a^T + Q_F C_a^T + Q_S C_a^T + Q_R C_a^T) \tag{E.41}$$

E.3.2 Assumptions

A1. There is no steady-state condition assumed for any compartment. Compartment concentrations are determined solely from their differential equations.

A2. Metabolism/clearance takes place in both the liver and the kidneys.

A3. The lung does not have the upper respiratory tract (URT) included.

E.3.3 Additional Equations

E1. The outgoing concentration for each tissue is given by the concentration of the chemical in the tissue divided by the tissue:blood partition coefficient for that chemical:

$$C_{v_i}^x = \frac{C_i^x}{P_i^x} \tag{E.42}$$

In each of (E.32)-(E.39), we can replace $C_{v_i}^x$ using (E.42).

E2. Total cardiac flow:

$$Q = Q_L + Q_K + Q_F + Q_S + Q_R \tag{E.43}$$

reduces (E.40) to

$$V_a \frac{dC_a^M}{dt} = Q\left(\frac{C_{lu}^M}{P_{lu}^M} - C_a^M\right) \tag{E.44}$$

and reduces (E.41) to

$$V_a \frac{dC_a^T}{dt} = Q\left(\frac{C_{lu}^T}{P_{lu}^T} - C_a^T\right) \tag{E.45}$$

E3. Total body volume:

$$V = V_L + V_K + V_F + V_S + V_R + V_{lu} + V_v + V_a \tag{E.46}$$

E4. Exhaled concentration with $P_{lu:air}^x$ computed from $P_{lung:blood}^x$ and $P_{blood:air}^x$.

$$C_{exh}^M = \frac{C_{lu}^M}{P_{lu:air}^x} \tag{E.47}$$

$$C_{exh}^T = \frac{C_{lu}^T}{P_{lu:air}^x} \tag{E.48}$$

E5. Liver metabolism (Michaelis-Menten kinetics)

$$Metab_L^M = \frac{V_{max_L}^M C_{v_L}^M}{K_{m_L}^M + C_{v_L}^M} \tag{E.49}$$

$$Metab_L^T = \frac{V_{max_L}^T C_{v_L}^T}{K_{m_L}^T + C_{v_L}^T} \tag{E.50}$$

E.4 MTBE-to-TBA coupled system, COMPLEX lung. Pleil 3ppm

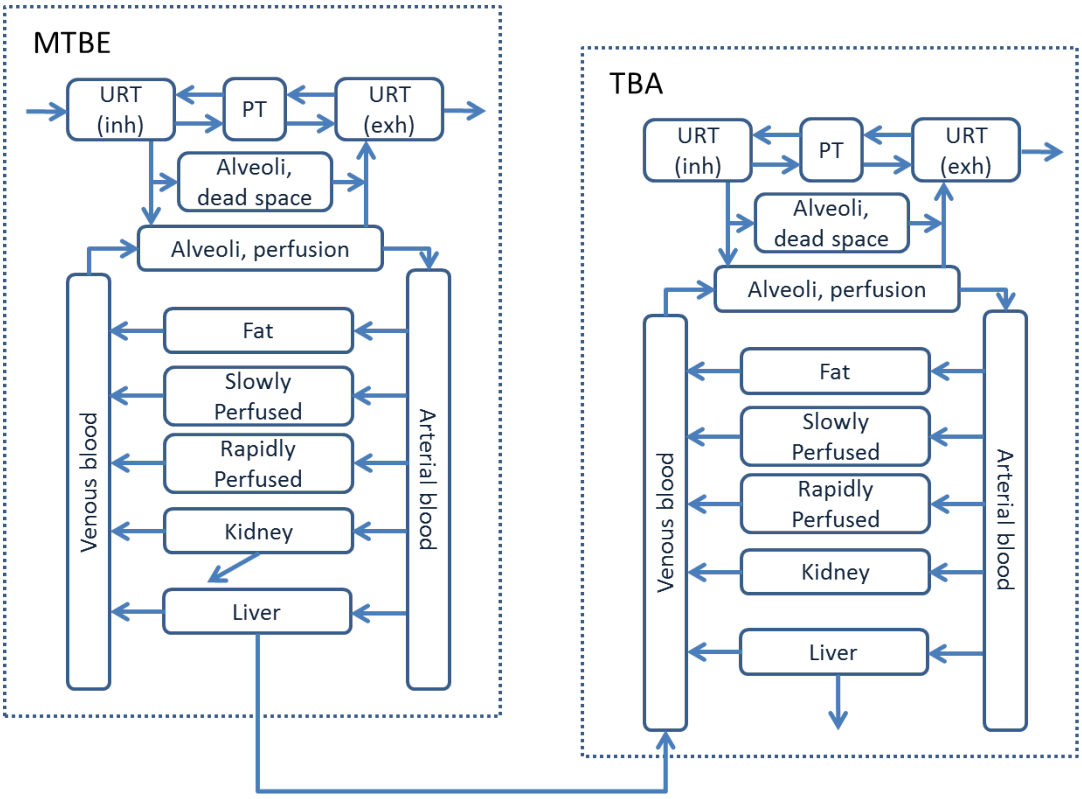


Figure E.5: MTBE-to-TBA coupled system with URT (complex lung)

E.4.1 Equations

Refer to Figures (E.3) and (E.4) for general route labeling.

Upper Respiratory Tract (URT), inhalation:

$$V_{URT}\frac{dC^M_{URT_{inh}}}{dt} = C^M_{inh}Q_p + C^M_{PT}D - kC^M_{URT_{inh}}\dot{V}_{alv} - kC^M_{URT_{inh}}\dot{V}_{DS} - kC^M_{URT_{inh}}D \quad \text{(E.51)}$$

$$V_{URT}\frac{dC^T_{URT_{inh}}}{dt} = C^T_{PT}D - kC^T_{URT_{inh}}\dot{V}_{alv} - kC^T_{URT_{inh}}\dot{V}_{DS} - kC^T_{URT_{inh}}D \quad \text{(E.52)}$$

Pulmonary Tissue (PT):

$$V_{PT}\frac{dC^M_{PT}}{dt} = kC^M_{URT_{inh}}D + kC^M_{URT_{exh}}D - 2C^M_{PT}D \tag{E.53}$$

$$V_{PT}\frac{dC^T_{PT}}{dt} = kC^T_{URT_{inh}}D + kC^T_{URT_{exh}}D - 2C^T_{PT}D \tag{E.54}$$

Upper Respiratory Tract (URT), exhalation:

$$V_{URT}\frac{dC^M_{URT_{exh}}}{dt} = C^M_{PT}D + C^M_{alv(P)}\dot{V}_{alv} + C^M_{alv(DS)}\dot{V}_{DS} - kC^M_{URT_{exh}}D - C^M_{exh}Q_p \tag{E.55}$$

$$V_{URT}\frac{dC^T_{URT_{exh}}}{dt} = C^T_{PT}D + C^T_{alv(P)}\dot{V}_{alv} + C^T_{alv(DS)}\dot{V}_{DS} - kC^T_{URT_{exh}}D - C^T_{exh}Q_p \tag{E.56}$$

Alveoli without Perfusion (Alveolar deadspace):

$$V_{alv(DS)}\frac{dC^M_{alv(DS)}}{dt} = kC^M_{URT_{inh}}\dot{V}_{DS} - C^M_{alv(DS)}\dot{V}_{DS} \tag{E.57}$$

$$V_{alv(DS)}\frac{dC^T_{alv(DS)}}{dt} = kC^T_{URT_{inh}}\dot{V}_{DS} - C^T_{alv(DS)}\dot{V}_{DS} \tag{E.58}$$

Alveoli with Perfusion:

$$V_{alv(P)}\frac{dC^M_{alv(P)}}{dt} = kC^M_{URT_{inh}}\dot{V}_{alv} + C^M_v Q - C^M_{alv(P)}Q - C^M_{alv(P)}\dot{V}_{alv} \tag{E.59}$$

$$V_{alv(P)}\frac{dC^T_{alv(P)}}{dt} = kC^T_{URT_{inh}}\dot{V}_{alv} + C^T_v Q - C^T_{alv(P)}Q - C^T_{alv(P)}\dot{V}_{alv} \tag{E.60}$$

Liver:

$$V_L\frac{dC^M_L}{dt} = Q_L(C^M_a - C^M_{v_L}) - Metab^M_L \tag{E.61}$$

$$V_L\frac{dC^T_L}{dt} = Q_L(C^T_a - C^T_{v_L}) - Metab^T_L \tag{E.62}$$

Kidney:

$$V_K\frac{dC^M_K}{dt} = Q_K(C^M_a - C^M_{v_K}) - Metab^M_K \tag{E.63}$$

$$V_K\frac{dC^T_K}{dt} = Q_K(C^T_a - C^T_{v_K}) \tag{E.64}$$

Fat, SPT, RPT:

$$V_i \frac{dC_i^M}{dt} = Q_i(C_a^M - C_{v_i}^M) \tag{E.65}$$

$$V_i \frac{dC_i^T}{dt} = Q_i(C_a^T - C_{v_i}^T) \tag{E.66}$$

Venous:

$$V_v \frac{dC_v^M}{dt} = Q_L C_{v_L}^M + Q_K C_{v_K}^M + Q_F C_{v_F}^M + Q_S C_{v_S}^M + Q_R C_{v_R}^M - Q C_v^M \tag{E.67}$$

$$V_v \frac{dC_v^T}{dt} = Metab_L^M + Q_L C_{v_L}^T + Q_K C_{v_K}^T + Q_F C_{v_F}^T + Q_S C_{v_S}^T + Q_R C_{v_R}^T - Q C_v^T \tag{E.68}$$

Arterial:

$$V_a \frac{dC_a^M}{dt} = C_{alv(P)}^M Q - (Q_L C_a^M + Q_K C_a^M + Q_F C_a^M + Q_S C_a^M + Q_R C_a^M) \tag{E.69}$$

$$V_a \frac{dC_a^T}{dt} = C_{alv(P)}^T Q - (Q_L C_a^T + Q_K C_a^T + Q_F C_a^T + Q_S C_a^T + Q_R C_a^T) \tag{E.70}$$

E.4.2 Assumptions

A1. There is no steady-state condition assumed for any compartment. Compartment concentrations are determined solely from their differential equations.

A2. Metabolism/clearance takes place in both the liver and the kidneys.

E.4.3 Additional Equations

E1. The outgoing concentration for each tissue is given by the concentration of the chemical in the tissue divided by the tissue:blood partition coefficient for that chemical:

$$C_{v_i}^x = \frac{C_i^x}{P_i^x} \tag{E.71}$$

In each of (E.61)-(E.68), we can replace $C_{v_i}^x$ using (E.71).

E2. Total cardiac flow:

$$Q = Q_L + Q_K + Q_F + Q_S + Q_R \tag{E.72}$$

reduces (E.69) to

$$V_a \frac{dC_a^M}{dt} = Q\left(C_{alv(P)}^M - C_a^M\right) \tag{E.73}$$

and reduces (E.70) to

$$V_a \frac{dC_a^T}{dt} = Q\left(C_{alv(P)}^T - C_a^T\right) \tag{E.74}$$

E3. Total body volume:

$$V = V_L + V_K + V_F + V_S + V_R + V_{URT} + V_{PT} + V_{alv(P)} + V_{alv(DS)} + V_v + V_a \tag{E.75}$$

E4. Blood flow rates for alv(P) and alv(DS)

$$\dot{V}_{alv} = f(V_T - V_D) \tag{E.76}$$

where f is breathing frequency, V_T is tidal volume, and V_D is dead space volume.

$$\dot{V}_{DS} = f(V_D) \tag{E.77}$$

E5. Exhaled concentration: k is the fraction of concentration leaving the compartment, $0 < k < 1$

$$C_{exh}^M = kC_{URT_{exh}}^M \tag{E.78}$$

$$C_{exh}^T = kC_{URT_{exh}}^T \tag{E.79}$$

E6. Liver metabolism (Michaelis-Menten kinetics), $i \in \{K, L\}$

$$Metab_i^M = \frac{V_{max_i}^M C_{v_i}^M}{K_{m_i}^M + C_{v_i}^M} \tag{E.80}$$

$$Metab_L^T = \frac{V_{max_L}^T C_{v_L}^T}{K_{m_L}^T + C_{v_L}^T} \tag{E.81}$$

APPENDIX F

Data Values

Table F.1: Cain 1.7 ppm. MTBE venous blood measurements

MTBE ($\mu g/L$)				
Time (min)	S01	S02	S03	S04
0	1.1	0.5	0.4	1.4
2	4.5	6.3	6.8	4.1
5	6.8	7.1	7.9	4.8
10	13.9	8.3	9.8	5.7
20	10.3	10.4	11.5	6.8
30	13.6	15.8	17.1	14.2
60	16.7	14.9	17.4	19.7
62	16.3	15	13.4	17.4
65	13.2	13.6	13.7	18.9
70	13.6	14.7	11.4	15.4
80	12.8	11.9	11.4	13.5
100	8.6	10.5	5.40	14.4
120	6.2	7.6	7.4	4.1
150	6.1	6.3	6.5	10.7

Table F.2: Pleil 3 ppm. MTBE venous blood measurements

MTBE ($\mu g/L$)														
Time (hr)	S08	S09	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20	S21
-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0.035
-0.92	5	6.1	1.5	8	10	5	-	4.9	6.9	3.6	11	12	6.3	4.1
-0.75	15	15	5	18	21	11	-	10	15	13	18	15	13	11
-0.5	17	22	13	-	19	17	20	17	25	14	25	20	21	14
-0.25	22	22	15	23	28	15	-	17	25	15	31	26	24	16
0	22	20	11	32	30	17	-	21	31	16	31	26	23	19
0.08	16	22	13	-	13	15	15	16	25	11	20	26	18	15
0.25	14	14	9.9	18	23	14	-	12	27	11	17	16	15	12
0.5	8.9	13	10	-	13	14	-	9.7	21	6.9	12	16	14	10
1	5.9	12	7.3	10	6.6	9.3	8.3	7.1	14	5.7	10	13	12	6.9
2	5.6	5.8	5	-	4.7	7.6	6.6	4.5	11	3.6	5.5	6.5	7.6	4.1
3	2.6	4.1	3.6	3.9	2.9	5.3	-	3.3	5.1	2.9	3.9	3.9	4.6	2.9
6	1.2	1.6	1.2	-	1.1	2.3	1.6	1.2	2.1	1.2	1.4	2.2	2.2	1.5
22	0.074	0.03	0.051	-	0.047	0.09	0.045	0.087	0.048	0.24	0.17	0.096	0.1	0.47

Table F.3: Pleil 3 ppm. TBA venous blood measurements

TBA ($\mu g/L$)														
Time (hr)	S08	S09	S10	S11	S12	S13	S14	S15	S16	S17	S18	S19	S20	S21
-1	0.54	0.56	1.1	1.1	0.7	3.3	2.5	1.4	1.2	0.56	0.61	0.7	0.82	0.36
-0.92	1.2	1.1	1.1	1.6	2.3	2.7	-	0.94	1.2	0.97	0.98	2.5	1.1	0.8
-0.75	3.7	3	2.2	4.7	5.4	4.1	-	2.4	2.1	3.4	2.7	4.1	3.2	2.5
-0.5	6.1	8.5	6.7	7.8	8.5	7.9	7.5	7	4.1	6.5	7.4	8.6	5.9	6.1
-0.25	12	8.4	9.6	9.6	12	7.7	-	8.2	5	5.6	12	12	8.4	7.5
0	14	8.9	10	15	15	9	20	12	7.2	7.9	14	16	13	13
0.08	16	13	11	15	14	12	14	12	8.7	7.7	14	18	14	12
0.25	15	12	11	12	16	9.2	13	11	8.4	11	12	16	12	11
0.5	13	13	15	14	14	13	17	11	11	10	12	16	14	11
1	13	14	16	14	15	14	15	12	11	11	12	18	12	11
2	14	14	16	15	17	16	15	14	13	9.9	11	16	13	10
3	13	15	20	16	17	-	15	13	12	9.9	12	16	13	11
6	15	15	16	15	14	14	16	13	11	8.7	9.9	20	11	12
22	4.3	6.4	5.6	5.5	4.7	6.3	6.4	4.9	5.7	6.3	3.5	8.5	3.2	5.5

Table F.4: Pleil 3 ppm. MTBE and TBA exhaled breath measurements

MTBE breath measurements (ppbv)							
Time (hr)	S08	S09	S11	S15	S16	S19	S21
-1	0.08	0.15	0.4	0.14	0.18	0.12	0.45
-0.92	462.73	602.73	468.1	782.26	447.59	917.48	502.31
-0.75	791.15	736.05	488.67	892.65	705.58	1109.83	458.68
-0.5	704.34	923.04	853.64	928.41	750.14	1140.06	641.57
-0.25	914.11	977.39	669.31	1340.8	813.86	1609.18	687.09
0	996.94	1006.77	879.65	1438.35	924.57	1380.32	884.53
0.08	397.88	448.02	374.95	400.58	504.09	589.54	373.98
0.25	253.78	307.76	234.15	253.31	370.31	313.62	223.44
0.5	188.77	228.83	184.23	178.03	260.57	181.16	152.42
1	118.63	162.75	123.57	129.88	186.55	111.04	109.23
2	83.45	91.31	74.37	75.31	117.78	62.87	59.05
3	44.57	64.94	51.07	50.72	66.68	-	42.56
6	24.01	22.03	22.25	20.02	30.05	34.2	21.6
22	1.75	0.53	0.81	1.85	1.21	1.4	6.86
TBA breath measurements (ppbv)							
Time (hr)	S08	S09	S11	S15	S16	S19	S21
-1	0.3	1.76	0	0.51	0.65	0.17	0.35
-0.92	0.61	3.52	0.65	0.81	1.38	1.08	0.6
-0.75	1.73	5.22	2.86	2.17	1.48	2.71	1.63
-0.5	3.01	5.48	5.15	3.72	2.22	3.99	3.21
-0.25	6.01	6.28	4.82	5.34	3.32	6.29	4.81
0	6.29	5.83	6.18	6.61	4.34	8.57	6.12
0.08	6.2	5.84	6.09	7.11	4.81	9.45	6.97
0.25	5.74	5.51	5.12	6.8	4.51	6.81	5.62
0.5	6.9	6.14	5.19	5.2	4.13	6.52	4.35
1	5.11	8.85	4.56	6.2	4.22	7.05	4.58
2	5.79	5.31	5.09	6.38	4.7	6.37	4.3
3	5.49	-	5.65	5.96	4.38	-	4.62
6	5.78	4.82	4.95	5.38	4.18	-	4.42
22	2.18	2.49	2.55	2.44	2.45	3.85	2.73

APPENDIX G

Optimized Parameters

Table G.1: Cain, Simple lung (Model 1), MTBE-only Optimized Parameters

mg/kg/hr	Guess	Mean	std	mg/L	Guess	Mean	std
$V_{max,liv_{Hi}}$	0.4	0.2475	0.4624	$K_{m,liv_{Hi}}$	0.2	225.3374	439.8150
$V_{max,liv_{Lo}}$	44	74.9627	69.9288	$K_{m,kid_{Lo}}$	110	57.5274	48.7685
$V_{max,kid}$	10	20.0462	39.3053	$K_{m,kid}$	10	110.8519	90.8109

Table G.2: Pleil, Simple lung (Model 1), MTBE-only Optimized Parameters

mg/kg/hr	Guess	Mean	std	mg/L	Guess	Mean	std
$V_{max,liv_{Hi}}$	0.4	0.0350	0.07388	$K_{m,liv_{Hi}}$	0.2	2.3759	3.2250
$V_{max,liv_{Lo}}$	44	55.50001	50.6082	$K_{m,kid_{Lo}}$	110	420.6519	457.4823
$V_{max,kid}$	10	26.0400	40.1210	$K_{m,kid}$	10	24.1377	28.0630

Table G.3: Cain, Complex lung (Model 2), MTBE-only Optimized Parameters

mg/kg/hr	Guess	Mean	std	mg/L	Guess	Mean	std
$V_{max,liv_{Hi}}$	0.4	17.3538	4.5014	$K_{m,liv_{Hi}}$	0.2	1869.3264	67.5791
$V_{max,liv_{Lo}}$	44	23.9530	27.4814	$K_{m,kid_{Lo}}$	110	76.5274	80.7503
$V_{max,kid}$	10	24.5989	49.0205	$K_{m,kid}$	10	1648.0958	497.1423
k	0.62	0.9839	0.03187	D	0.802	997.9073	4.0154

Table G.4: Pleil, Complex lung (Model 2), MTBE-only Optimized Parameters

mg/kg/hr	Guess	Mean	std	mg/L	Guess	Mean	std
$V_{max,liv_{Hi}}$	0.4	24.1705	16.2476	$K_{m,liv_{Hi}}$	0.2	780.8265	214.8892
$V_{max,liv_{Lo}}$	44	2.4324	3.8565	$K_{m,kid_{Lo}}$	110	396.7448	143.8971
$V_{max,kid}$	10	115.3580	13.3927	$K_{m,kid}$	10	140.2041	58.8659
k	0.62	0.9999	1.20837E-06	D	0.802	999.9970	0.0059

Table G.5: Pleil, Simple lung (Model 3), MTBE-to-TBA Optimized Parameters

mg/kg/hr	Guess	Mean	std	mg/L	Guess	Mean	std
$V_{max,liv_{MTBE}}$	44	47.9650	13.4404	$K_{m,liv_{MTBE}}$	110	140.1759	48.5721
$V_{max,kid_{MTBE}}$	10	35.0993	14.4730	$K_{m,kid_{MTBE}}$	10	31.2901	17.5617
$V_{max,liv_{TBA}}$	0.4	2.7775	2.7222	$K_{m,liv_{TBA}}$	0.2	36.8895	41.2321

Table G.6: Pleil, Complex lung (Model 4), MTBE-to-TBA Optimized Parameters

mg/kg/hr	Guess	Mean	std	mg/L	Guess	Mean	std
$V_{max,liv_{MTBE}}$	44	62.6879	66.6498	$K_{m,liv_{MTBE}}$	110	139.9255	167.1362
$V_{max,kid_{MTBE}}$	10	0.7079	1.1626	$K_{m,kid_{MTBE}}$	10	536.7700	756.4230
$V_{max,liv_{TBA}}$	0.4	0.0742	0.1044	$K_{m,liv_{TBA}}$	0.2	1285.1121	362.0945
k	0.62	0.9999	3.09E-05	D	0.802	49.9998	0.0001