January 2012

Statistical Estimation of Physiologically-based Pharmacokinetic Models: Identifiability, Variation, and Uncertainty with an Illustration of Chronic Exposure to Dioxin and Dioxin-like-compounds.

Zachary John Thompson
University of South Florida, zthompso@health.usf.edu

Statistical Estimation of Physiologically-based Pharmacokinetic Models: Identifiability, Variation, and Uncertainty with an Illustration of Chronic Exposure to Dioxin and Dioxin-like-compounds.

by

Zachary Thompson

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Epidemiology and Biostatistics
College of Public Health
University of South Florida

Major Professor: Yiliang Zhu, Ph.D
Co-Major Professor: Yangxin Huang, Ph.D
Committee member: Jeff Gift, Ph.D
Committee member: Yougui Wu, Ph.D
Committee member: Raymond Harbison, Ph.D

Date of Approval:
April 12, 2012

Keywords: PBPK, MCMC, TCDD, exposure assessment, statistical estimation

Table of Contents

List of Figures

List of Tables

List of Main Equations

Abstract

Assessment of human exposure to environmental chemicals is inherently subject to uncertainty and variability. There are data gaps concerning the inventory, source, duration, and intensity of exposure as well as knowledge gaps regarding pharmacokinetics in general. These gaps result in uncertainties in exposure assessment. The uncertainties compound further with variabilities due to population variations regarding stage of life, life style, and susceptibility, etc. Use of physiologically-based pharmacokinetic (PBPK) models promises to reduce the uncertainties and enhance extrapolation between species, between routes, from high to low dose, and from acute to chronic exposure. However, fitting PBPK models is challenging because of a large number of biochemical and physiological parameters to be estimated. Many of these model parameters are non-identifiable in that their estimates cannot be uniquely determined using statistical criteria. In practice some parameters are fixed in value and some determined through mathematical calibration or computer simulation. These estimated values are subject to substantial uncertainties. The first part of this paper illustrates the use of iteratively-reweighted-nonlinear-least-squares for fitting pharmacokinetic (PK) models, highlighting some common difficulties in obtaining statistical estimates of non-identifiable parameters and use bootstrap confidence interval to quantify uncertainties.

Statistical estimation of parameters in physiologically based pharmacokinetic (PBPK) models is a relatively new area of research. Over the past decade or so PBPK models have become important and valuable tools in risk assessment as these models are used to describe the absorption, distribution, metabolism, and excretion of xenobiotics in a biological system such as the human or rat. Because these models incorporate information on biological processes, they are well equipped to describe the kinetic behaviors of chemicals and are useful for extrapolation across dose routes, between species, from high-to-low-doses, and across exposure scenarios.

A PBPK model has been developed based on published models in the literature to describe the absorption, distribution, metabolism, and excretion of Dioxin and dioxin like compounds (DLCs) in the rat. Data from the National Toxicology Program (NTP) two year experiment TR-526 is used to illustrate model fitting and statistical estimation of the parameters. Integrating statistical methods into risk assessments is the most efficient way to characterize the variation in parameter values. In this dissertation a Markov Chain Monte Carlo (MCMC) method is used to estimate select parameters of the system and to describe the variation of the select parameters.

1 Chapter One: Introduction / Literature Review

1.1 Background

TCDD (2,3,7,8-Tetrachlorodibenzo-p-dioxin; CAS 1746-01-6) or dioxin is a chlorinated hydrocarbon, and is the prototype of the polyhalogenated aromatic hydrocarbon (PHAH) family, which includes dibenzo-p-dioxins (PCDDs), dibenzofurans (PCDFs), biphenyls, and polychlorinated biphenyls (PCB). The most toxic compound in this family is considered to be 2,3,7,8-tetrachlorodibenzo-p-dioxin or TCDD. Sometimes the term dioxins also commonly refers to chlorinated dibenzo-p-dioxins and furans to chlorinated dibenzofurans.

Dioxin is formed as an unintentional by-product of many industrial processes involving chlorine such as waste incineration, chemical and pesticide manufacturing and pulp and paper bleaching. Specifically it was a common contaminant of the herbicide 2,4,5-trichlorophenoxyacetic acid (Hamilton and Hardy, 1998). Dioxin and dioxin like chemicals (DLCs) are found everywhere around us and we are all exposed to background levels. Soil erosion and sediment transport in water and volatilization from the surfaces of soils and water bodies with subsequent atmospheric transport and deposition are thought to be the dominant mechanisms in the widespread environmental occurrence of PCBs (NRC, 2006).

Dioxin is notorious for being the primary toxic component of Agent Orange, a defoliate famous for its use in the Vietnam war (Pirkle *et al.*, 1989). It has also been found in a number of the toxic waste sties, for example, underneath Love Canal in Niagara Falls (Smith *et al.*, 1983), NY and for accidental releases in Times Beach, MO (Umbreit, 1986) and Seveso, Italy (Bertazzi *et al.*, 1998). Soil contaminated with Dioxin and DLCS caused the evacuation of Times Beach Missouri in 1983 and was the suspect in the death of local

animals and a range of human illnesses. Times beach and Love Canal are two of the three toxic waste sites that caused the creation of the Superfund law in 1980. (USDOJ, 1997)

Present day, the main sources of dioxin and dioxin like compounds releases into the environment are from combustion and incineration sources, metal smelting, refining and processing, chemical manufacturing and processing, biological and photochemical processes, and existing reservoir sources that reflect past releases. PCB mixtures were commercially produced and used in the electric power industry as dielectric insulating fluids in transformers and capacitors and used in hydraulic fluids, plastics, and paints. These toxic chemicals and chemical mixtures get into the soil (Freeman, 1986; McCrady, 1990) then erosion and runoff transport them to bodies of water subsequently fish and other marine life are exposed and the toxins bioaccumulate. This is believed to be the major mode by which PCDD/PCDFs enter the aquatic food chain (Jensen, 2000).

In 1994 the National Institute for Occupational Safety and Health (NIOSH) classified TCDD as an occupational carcinogen; thus exposure should be reduced to the lowest level possible and there is no Permissible Exposure Limit (PEL) for dioxin (Hamilton and Hardy, 1998). Also the EPA has never published a Reference Concentration (RfC) or a Reference Dose (RfD) for TCDD.

In general dioxin and dioxin-like compounds compounds have low vapor pressure, Low water solubility, low rate of biotransformation and high lipophilicity. These properties give dioxin and dioxin-like compounds (DLCs) high potential for bio-accumulation (Esposito *et al.*, 1980); they transport into and climb up our food chain. As a result, population exposure to dioxin and DLCs mostly comes from food intake, specifically through the consumption of fish, meat, and dairy products because dioxin and DLCs are fat-soluble (Schecter *et al.*, 2001).

These compounds share a common Ah receptor (AhR)-mediated mechanism of toxic and biological responses (Safe, 1986; Birnbaum, 1994). They act via an Ah receptor in both animals and humans. Exposure to dioxin and DLCs causes induction of proteins in several tissues and binding alters gene expression. The Ah receptor is present in many tissues in animals and humans most notably in the liver. Animal carcinogenesis has been correlated with dioxin's affinity for the Ah receptor (Steeland and Deddens, 2003).

The most characterized biochemical response associated with TCDD exposure is the induction of CYP1A1 (Whitlock, 1993), which involves the initial interaction of the ligand with the multimeric cytosolic AhR complex (Chen and Perdew, 1994).

Evidence indicates that TCDD acts via an intracellular protein (the aryl hydrocarbon receptor; Ah receptor), which functions as a ligand-dependent transcription factor in partnership with a second protein (known as the Ah receptor nuclear translocator; Arnt). Therefore, from a mechanistic standpoint, TCDD's adverse effects appear likely to reflect alterations in gene expression that occur at an inappropriate time and/or for an inappropriately long time. Mechanistic studies also indicate that several other proteins contribute to TCDD's gene regulatory effects and that the response to TCDD probably involves a relatively complex interplay between multiple genetic and environmental factors (Poland, 1996; Limbird and Taylor, 1998).

The Seveso accident in 1976 caused a large, populated area north of Milan, Italy, to be contaminated by 2,3,7,8- tetrachlorodibenzo-p-dioxin. In a follow up study the authors followed the exposed population for chronic effects. They report the results of the mortality follow-up extension for 1997 to 2001. The study cohort includes 278,108 subjects resident at the time of the accident or immigrating/born in the 10 years thereafter in three contaminated zones with decreasing TCDD soil levels (zone A, very high; zone B, high; zone R, low) also in a reference territory surrounding non-contaminated municipalities. Results confirmed previous findings of excesses of lymphatic and hematopoietic tissue neoplasms in the very high concentration zone (six deaths; rate ratio = 2.23, 95% confidence interval: 1.00, 4.97) and the high concentration (28 deaths; rate ratio = 1.59, 95% confidence interval: 1.09, 2.33). These zones also showed increased mortality from circulatory diseases in the years after the accident, from chronic obstructive pulmonary disease, and from diabetes mellitus among females. Thus toxic and carcinogenic risk to humans after high TCDD exposure is supported by the results of this study (Consonni *et al.*, 2008).

TCDD exposure is widely recognized to produce severe chloracne (Cook *et al.*, 1980). It may also alter liver function. Long-term exposure is linked to impairment of the immune system, endocrine system, and the reproductive system (ATSDR, 1998). Chronic exposure at low levels has resulted in several types of cancer such as Cholangiocarcinoma

and Hepatocellular adenoma in different animal species. Based on available animal data and human epidemiological data, the WHOs International Agency for Research on Cancer (IARC) classifies TCDD as "carcinogenic to human". The US Environmental Protection Agency (USEPA) also classifies TCDD as a human carcinogen (EPA, 2002).

The major acute effects from exposure of humans to high levels of $2,3,7,8-TCDD$ in air is chloracne, a severe acne-like condition that can develop within months of first exposure. (ATSDR, 1998; EPA, 1997). Animal tests in species such as dogs, monkeys, and guinea pigs have shown TCDD to have very high toxicity from oral exposure (ATSDR, 1998) .

The non cancerous Chronic Effects include Chloracne in humans while animal studies have reported hair loss, loss of body weight, and a weakened immune system from oral exposure (ATSDR, 1998). The ATSDR has calculated a chronic oral minimal risk level (MRL) of 1e-09 milligrams per kilogram body weight per day (*mg/kg/d*) based on neurological effects in monkeys. The MRL is an estimate of daily exposure to a dose of a chemical that is likely to be without appreciable risk of adverse noncancerous effects over a specified duration of exposure. The MRL is used by public health professionals as a screening tool. (ATSDR, 1998)

While there has been some evidence of reproductive and developmental effects in humans, animal studies have reported developmental effects, such as skeletal deformities, kidney defects, and weakened immune responses in the offspring of animals exposed TCDD during pregnancy (ATSDR, 1998). More over reproductive effects, including reduced production of sperm, and increases in miscarriages have been seen in animals exposed to dioxin (ATSDR, 1998).

In human studies, mostly of occupational exposure by inhalation, there is evidence of associations between Dioxin and lung cancer (Crump, 2003), soft-tissue sarcomas, and lymphomas. While animal studies have found tumors in the liver, lung, tongue, and thyroid from oral exposure (ATSDR, 1998). For Dioxin exposure, the EPA has estimated the inhalation cancer slope factor to be 1.5e-05 *mg/kg/d* and an inhalation unit risk estimate of 3.3e-05 pg/m^3 (USEPA, 1985, 1997). Also the oral cancer slope factor is estimated to be 1.5e-05 *mg/kg/d* and an oral unit risk factor of 4.5 *g/L* (USEPA, 1985, 1997).

TCDD is detectible in blood and adipose tissue of the general population (Kang *et al.*, 1991) yet the background exposure has not been associated with adverse effects. Dietary intake is generally recognized as the primary source of human exposure to CDD/CDFs (Rappe, 1992). Several studies estimated that over 90 percent of the average daily exposure to CDD/CDFs are derived from foods (Rappe, 1992; Henry *et al.*, 1992), mainly meat, dairy products, and fish.

Table 1 shows a summary of Dioxin like chemical TEQ concentrations found in the environment. Not only are toxic chemical mixtures found in air, water and food, part-per-trillion levels of CDDs/CDFs have been found in everyday materials that are contaminated with dust such as clothes dryer lint (2.4 to 6.0 *ng I-TEQDF/kg*); vacuum cleaner dust (8.3 to 12 *ng I-TEQDF/kg*); room air filters (27 to 29 *ng I-TEQDF/kg*); and house furnace filter dust (170 *ng I-TEQDF/kg*) (Berry *et al.*, 1993).

Because of the toxicity and the properties that allow the accumulation of Dioxin and DLCs up the food chain Risk assessment is necessary to know at what dose these chemicals are likely to cause harm. Dioxin and DLCs are not the only chemicals that are dangerous, the number of chemicals manufactured is in the tens of thousands and it is important to know at what dose these chemicals are likely to cause harm if in fact they are found to be toxic. Thus the focus of risk assessment is establishing the safe dose of a chemical and the probable toxic effects that may occur if that safe dose is exceeded. Risk assessment as a process can be broken down into several steps: Hazard characterization, dose-response assessment, exposure assessment, and risk characterization.

The first step in risk assessment is the hazard identification or characterization, which is a review of all biological and chemical information on a possible toxic agent that may or may not be a carcinogenic hazard. Physical and chemical properties, routes, and patterns of exposure are all examined during the hazard characterization step. Metabolic and pharmacokinetic properties along with toxicologic effects are also included. Short and long-term animal studies, and in some cases human studies are used to assess the hazard of a potential carcinogen.

The weight of evidence of carcinogenicity is basically determined from long-term animal studies and epidemiological studies on humans with the aid of other information from

short term tests, pharmacokinetic studies, and toxicity studies. The three steps to characterizing the weight of evidence for carcinogenicity in humans are the characterization of the evidence from human studies and animal studies, the combination of the characterization of the two types of studies into an indication of the overall weight of evidence and the evaluation of all supporting information.

The EPA has developed a classification of the weight of evidence yet this system is not meant to be applied rigidly. The system for the characterization of the overall weight of evidence for carcinogenicity includes five groups. Group A is designated for compounds that are considered as carcinogenic to humans. While group B are for compounds that are probably carcinogenic to humans and group C for those that are possibly carcinogenic. Finally, group D are for compounds that are not classifiable as to human carcinogenicity and group E are for those that show evidence no evidence of carcinogenicity for humans.

The EPA regards agents that fall into groups A and B are suitable for quantitative risk assessments (USEPA, 2005). Those agents in group C may be suitable for quantitative risk assessments but the decision to do so should be made on a case-by-case basis. Those agents in groups D and E have no need for quantitative risk assessments.

After the hazard characterization step is completed usually the next step is to estimate the excess cancer risk associated with a certain exposure. Of course estimates that are derived based on human epidemiologic data are preferred over those derived from animal data. The risk assessor must make a choice of mathematical model for the description of the dose-response relationship and extrapolation. Thus there is no single mathematical model or technique of estimation that is used as a gold standard for the low-dose extrapolation.

Low dose estimates obtained from animal data extrapolated to humans are confounded by uncertainty factors that are different between species. Some of these factors are life span, body size, genetic variation, pharmacokinetic effects such as metabolism and excretion. Some of the common dosing scales include *mg* per *kg* body weight per day, parts per million in the diet or water, and mg/m^2 body surface area per day. Surface area is a good scale to employ because many pharmacological effects scale according to surface area (Dedrick, 1973).

The results of a dose-response assessment should be combined with estimates of the exposures to which people are likely to be subjected in order to obtain quantitative estimate of risk. Similar to dose-response assessment there is no single technique or procedure that is appropriate for each and every case but the frequency, magnitude, and duration of the exposure are the most important factors in estimating the concentration of the carcinogen to which a subject is exposed. Often the cumulative dose received over a lifetime, expressed as an average daily exposure prorated over a short period of time is used to measure the exposure to a carcinogen. The risk assessor should always try to assess the level of uncertainty associated with the exposure assessment and provide this so that there is an understanding of the uncertainty of any final estimate.

Risk characterization is a report of the numerical estimates of risk and examination of the significance of the estimate. Numerical estimates can be presented in different ways. Possibilities include unit risk, excess lifetime risk due to a continuous constant lifetime exposure, dose corresponding to a level of risk, individual or population risks (excess individual lifetime risks or excess number of cancers per year in the exposed population). The risk characterization should include the interpretation of the estimates in order to assess the degree to which the quantitative estimates are likely to reflect the true magnitude of human risk. Risk characterization should summarize the hazard identification, the does-response assessment, exposure assessment, major assumptions scientific judgments and estimates of the uncertainties in the risk estimates.

1.2 Toxic Equivalency Factors and Quantifying Toxicity

Since we are not exposed to any one single chemical or agent but a mixture of toxins in the environment, in order to conduct risk assessment of these mixtures the toxic equivalency factor (TEF) methodology has been developed.

Chemicals that are similar in structure and show similar physio-chemical properties and cause similar adverse effects as Dioxin are considered ”dioxin-like”. The toxic potency (or TEF) of each of these dioxin like chemicals is expressed relative to TCDD and these TEFs are then used to quantify the toxicity of a mixture of chemicals relative to that of TCDD,

expressed as a toxic equivalent concentration (TEC). Not only are the structures and physiochemical properties similar in DLCSs but like Dioxin they also invoke a common set of toxic responses by a common aromatic hydrocarbon receptor (AhR) dependent mechanism in vivo.

The current TEFs are qualitative values that specify an order-of-magnitude of potency of a chemical compared with TCDD (NRC, 2006). They were established by a WHO expert scientific panel that used a large data base of REP (relative potency) estimates from in vivo and in vitro studies of biochemical and toxic effects. The TEF method takes the concentration of the individual compounds present in the mixture and these are multiplied by each specific TEF value, and the sum of all these products is expressed as the TCDD toxic equivalent quotient.

$$TEQ = \sum_i TEF_i \times C_i$$

In June 2005 a WHO-IPCS expert panel meeting was held to re-evaluate toxic equivalency factors for Dioxin like compounds. It was concluded that a combination of unweighted distributions of relative effect potencies (REP), point estimates and expert judgement. A TEF is derived for an individual PCDD, PCEF or PCB compound for producing toxic or biological effect relative to a reference compound usually 2,3,7,8-TCDD or Dioxin. The total Toxic Equivalent (TEQ) is defined as the sum of the products of the concentration of each compound multiplied by it's TEF value. The criteria for inclusion into the TEF concept a compound must show a structural relationship to the PCDDs and PCDFs, bind to the Ah receptor, elicit AhR mediated biochemical and toxic responses and finally be persistent and accumulate in the food chain (Ahlborg *et al.*, 1994).

The inherent uncertainty in the determination of REPs and sometimes high variation in the REP values for a single congener are just two reasons for periodic evaluation of validity of the concept. There are gaps in the knowledge of even the most commonly studied dioxin like chemicals such as 2,3,4,7,8-PeCDF and PCB126 (Haws *et al.*, 2006). There are multiple reasons for the variation in the REPS found in the literature for a single compound. The different values can be caused by different dosing schemes, different endpoints of interest, species, and also different methods used for calculating the REF value itself.

Thus the variations in methodologies used to derive the different REP values introduces uncertainties. The recommended method to calculate a REP is by divided the ED_{50} of the reference compound by the ED_{50} of the congener under study (Van den Berg *et al.*, 2006).

Among other considerations a study that aims to produce a REF should at the very least administer the compounds under study by the same route, to the same species, strain, gender, and age of animal (Van den Berg *et al.*, 2006). The animal subjects should be housed, fed, and maintained under consistent conditions. In vivo studies are preferred over in-vitro studies because they retain and combine the toxicokinetic and toxicodynamic properties.

The AH receptor's role in the toxicological and biochemical effects associated with dioxin and DLCs is considered to be necessary but not sufficient. There are further processes regulated by the AH receptor and AH receptor gene expression. The induction of cytochrome P450 proteins, CYP1A1 and CYP1A2, are the most studied processes associated with dioxin exposure. According to the National Center for Biotechnology Information (NCBI), these catalyze a number of reactions involved in drug metabolism and synthesis of cholesterol, steroids, and other lipids. Although there is interaction with other chemicals and other effects mediated by mechanisms that are independent of the AH receptor, they are ignored in the TEF methodology. There have been observations of changes in gene expression (Oikawa *et al.*, 2001) and toxicity in Ah receptor knock out mice. (Lin *et al.*, 2001). It is not known to what extent other mechanisms influence the biochemical effects of Dioxin and DLCs. The National Academy of Sciences (NAS) has called for future studies to consider AH receptor independent mechanisms when examining the toxicity and carcinogenicity of DLCS.

There is considerable variability in the REP values that were used to obtain the WHO TEFs and thus there is much uncertainty in the selection of TEFs values. Also the published information on how they were originally derived was not entirely clear with respect to the weighting schemes applied to the studies that provided the relative potency variables (Van den Berg *et al.*, 1998). Because of the uncertainty and variability inherent in the REPS and TEFs the assumption of consistency of DLC REP values is necessary in the TEF

methodology. This assumption requires the REP of a chemical to be equivalent for all end points and all exposure scenarios.

1.3 History of PK and PBPK models

Pharmacokinetic (PK) models date back to the late 1930s and were first developed by Torsten Teorell, a Swedish scientist whom some call the ”father of pharmacokinetics” (Obrink, 1992). These first models were developed before there were relatively painless ways to solve the sets of differential equations that explain such models. These models were expanded in the latter half of the twentieth century to accommodate dose-dependent elimination rates and flow-limited metabolism processes (Anderson, 2003). In the 1980's PK and physiologically based pharmacokinetic (PBPK) models were explored more because of the advancement in occupational toxicology and risk assessment and in the 1990's there was explosive growth in risk assessment applications of PK and PBPK models (Anderson, 2003).

Scientists have used classic pharmacokinetic models, pharmacodynamic models, and pseudophysiologic models for numerous applications but the main utility of these type of models is to describe and study the distribution of xenobiotics in a biological system such as the rat or the human. In the last few decades these models have grown into descriptive physiologically based pharmacokinetic (PBPK) models. These PBPK models are used to describe the absorption, distribution, metabolism, and excretion of xenobiotics.

Because these models incorporate information on biological processes they are better equipped to describe the kinetic behaviors of chemical and are more amenable to extrapolation across dosing routes, between species, from high-to-low-doses, and across exposure scenarios (Clewell and Andersen, 1985; Anderson, 2003).

The first pharmacokinetic models specifically for Dioxin like compounds were developed in the 1980s. These models did not include processes such as specific binding of TCDD with (non-Ah receptor) proteins, thus the estimated partition coefficients of interest were biased. The estimated Liver:Blood partition coefficient appeared to be larger than the Fat:Blood partition coefficient even though such chemicals are highly lipophilic.

Many dioxin like compounds are highly lipophilic and are stored in the body fat; however it has been shown that hepatic binding processes are the reason that dioxin is concentrating in the liver. In the late 1980s Leung *et al.*(1989a,b) explored quantitative physiologic models to study the role of the Ah-receptor in the concentration accumulation in the liver. Although this was an improvement in the biological description of the mechanism of toxicity of dioxin like compounds. The binding to the Ah receptor alone could not account for the sequestration in the liver of TCDD. Further studies eventually showed that existence of non-Ah receptor binding sites and that these non-Ah receptor binding sites are microsomal proteins with binding constants in the nano-molar range (Poland *et al.*, 1989a,b). The Ah-dioxin complex interacts with DNA to cause induction or repression of gene expression (Anderson, 2003). Modeling advances in the 1990s (Kohn *et al.*, 1993; Andersen *et al.*, 1993) expanded on these early models with the newer PBPK models for TCDD including biological detail related to DNA binding, mRNA induction and protein synthesis.

The newest US EPA cancer guidelines discuss the use of biologically based models in assessing dose response relationships. Hence one of the main uses of PBPK models is in risk assessment. Yet in order to use these models in environmental risk assessment to humans, it is necessary to extrapolate in the region of the lowest doses. This extrapolation is done by adjusting estimates of risk to be conservative with the use of uncertainty factors but by employing uncertainty factors estimates of risks for humans may end up being under-estimated. PBPK models can help address the uncertainty by describing multiple biological, biochemical, and physiological processes.

Typically parameter estimates in these models are only presented as point estimates with no margin of error or estimate of variation. Integrating statistical methods into in risk assessments would be an efficient way to characterize the variation. Uncertainty enters into risk assessment in many ways, in hazard identification the extrapolation from various species to humans is a large contributor of uncertainty. In exposure assessment lack of information on frequency, duration, and level of exposure introduce uncertainty. One important source of uncertainty is in the measure of the internal dose and specifically into an organ, the biologically effect dose. With PBPK models, the intake and uptake process of a chemical exposure is modeled as part of the absorption process. Since there are many

studies that only report point estimates of physiological parameters such as oral absorption rates and stomach emptying rates, there is no quantification of the variance.

In an oral exposure the administered dose enters the mouth (the intake process) and is termed the potential dose then the chemical makes its way through the gastrointestinal tract (the uptake process) and further metabolism creates the internal dose. The process of forecasting tissue concentrations from cumulative exposure to low level exposures can be more reliable if we can quantify the variation in physiological parameters.

This dissertation explores the difficulty of the integration of statistical methods of parameter estimation in PK and PBPK models so that estimates of the variation of key parameters can be obtained. Two of the main issues to deal with are complexity of the models and non-identifiability of parameter values. PBPK models can be very complex with multiple compartments and a large number of parameters, thus statistical estimation can be difficult because of scarce data, parameter non-identifiability, and non-convergence among others.

1.4 Specific Aims

The first step in this dissertation research is to explore parameter identification issues and incorporate statistical methods to estimate parameters in a simple two compartment pharmacokinetic model describing the transfer of a mixture of TCDD, PCB-126, and PeCDF. An oversimplfied PK model is used to illustrate and address parameter identifiability and other issues in development and estimation of the parameters because identifiability problems can only multiply and become more complex when adding equations and parameters into PK models. Fitting PK and PBPK models remains challenging due in particular to a large number of biochemical and physiological parameters. Many of these model parameters are unknown in value, are statistically non-identifiable, and have to be fixed. Mathematical calibration or computer simulation are common methods used to determine their value, and uncertainty quantification is inadequate. The first part of dissertation illustrates the use of iteratively-reweighted-nonlinear-least-squares for fitting PK models and employs bootstrap confidence intervals to quantify variation in parameter estimates.

The specific application involves a two-compartment model of rats chronically exposed to dioxin over two years and the model permits dose-dependent kinetic parameters. We show the use of simple graphs to identify region of non-identifiability and the use of the bootstrap confidence interval. Insights gained here serve to advance statistical estimation of general PBPK models and will fill some gaps in the literature by reporting on these issues.

The research then focuses on expanding this model by incorporating physiological information and leads to a second aim which is to examine assumptions employed in PBPK methodology such as the assumption of constant proportional organ weights. In toxicological experiments and pharmacological investigations the weight of an organ such as the liver is typically assumed to be a constant proportion of body weight. Such an assumption can have critical pharmacokinetic and pharmacodynamic impacts especially in the context of therapeutic intervention or environmental exposure. Yet, there is limited literature on its validity. The second part of this dissertation investigates how organ weights as a proportion of body weight change with exposure conditions during growth period. This dissertation will fill some gaps in the literature by examining organ weights, which are unlikely to be a constant proportion of body weight in a growth period or under chronic exposure condition. This non-constant relationship may have profound pharmacokinetic and pharmacodynamic implications since physiological parameters such as organ blood flow rates, surface areas, and permeability will be non-constant.

A third aim is to construct a PBPK model for chronic exposure to Dioxin (and two other DLCs, one for the mixture as well) and integrate statistical methods for parameter estimation. A markov chain monte carlo method is used to estimate key parameters and quantify their variation. Furthermore the variability and sensitivity of these key parameters to changes in other parameter values (held fixed) will be examined. Dose dependent physiological parameters for elimination are incorporated into the model.

The PBPK model will incorporate the target compartments (organs) for each chemical in a mixture of Dioxin and DLCs. In other words, there will be three ”responses” (one response in three organs at the same time) hence three regression models (implicit explicit), one for each response, namely, Y_{fat}, Y_{liver}, and Y_{blood}. First, the differential equations for all three responses simultaneously are derived. Then parameter estimation is done using

statistical methods. The issue of parameter identifiability needs to be dealt as well as issues related to assessment of model fitting, including model diagnostics and prediction.

It is clear that use of PBPK models in risk assessment requires research on model development and implementation with respect to chronic exposure of low external doses. This research fills a gap in the literature because most PBPK models are developed for acute or short term exposure. Future work includes applying these PBPK models to verify the TEF methodology. The World Health Organization (WHO) has suggested that the TEQ scheme be reevaluated every 5 years.

1.5 Tables

Table 1: *DLC* TEQ WHO Concentrations in Environmental Media and Food.

Media	n	$Mean(SD)$	References
Urban Soil, *ppt*	270	9.3 (10.2)	EPA (2000a), Pearson *et al.*(1990)
Urban Air, pg/m^3	106	0.12 (0.094)	Hoff *et al.*(1992)
Water, *ppq*	236	0.00056 (0.00079)	Meyer *et al.*(1989),Jobb *et al.*(1990)
Beef, *ppt*	63	0.18 (0.11)	Winters *et al.*(1996)
Poultry, *ppt*	78	0.068 (0.070)	Ferrario *et al.*(1997)
Marine Fish, *ppt*	158	0.26 (0.070)	Fiedler *et al.*(1997), Jensen (2000)

Table 2: Major Assumptions behind TEF approach.

Assumption	Description
Role of AHR	AHR mediates most toxicities produced by TCDD and other PCDDs, PCDFs, and coplanar PCBs that are APH agonists.
AHR Independent Mechanisms excluded	Effects mediated by other mechanisms (AHR independent) and interactions with other chemicals are ignored.
Consistancy of DLC REP values	The REP of a chemical in this group is presumed to be equivalent for all end points of concern and all exposure scenarios.
Use Of TEFs for DLC body burden	TEFs for animals exposed by dietary intake are appropriate for assessment of internal TEQ concentrations and potential toxic effects.
Additivity of DLCs	Mixtures exhibit additive toxicities based on TEFs of individual chemicals
Rodent to Human Prediction	REPs of dioxin and DLCS in rodents is predictive of REPS in humans
Natural and Synthetic Non-DLCs AHR agonists	Synthetic and non-natural non-DLC agonists do not interfere with PCDD, PCDF, and PCB dependent TEQ predictions

2 Statistical Estimation of Pharmacokinetic Models with Non-identifiable Parameters

2.1 Introduction

Evaluation of human exposure to environmental chemicals requires analysis of pollution inventory in the environment; the source, intensity, duration and routes of human contact; and pharmacokinetics upon entering human body. Knowledge gaps and data gaps inevitably generate uncertainties in exposure assessment estimates, which is compounded with inherited variability in a general population (USEPA, 1992). The term variation refers to differences in exposure due to susceptibility, life style and human activities. Uncertainty refers to changes in our exposure estimates if we change our assumptions about gaps in knowledge and data. Variations and uncertainties in exposure estimates contribute significantly to the overall uncertainty in risk assessment, hence impact regulatory policies.

The promise that physiologically-based pharmacokinetic (PBPK) models can substantially reduce uncertainties in exposure assessment has led to a growing body of research in developing and applying PBPK models (USEPA, 2006; Barton *et al.*, 2007; IPCS, 2008; Thompson *et al.*, 2008). PBPK models describe the process of absorption, distribution, metabolism, and excretion of the chemical that enters the body, and measure target tissue dose with appropriate dose metrics. A PBPK model may be used to predict tissue dose based on exposure or to back-estimate the cumulative exposure based on tissue dose. A PBPK model is a mechanism that can guide for more reliable extrapolations in estimates between species, between exposure routes, from acute to chronic exposure, or from high to low dose. Back-estimation of cumulative exposure based on a PBPK model may provide useful reference levels for environmental monitoring and regulation.

PBPK models are typically developed on the basis of experiments of acute or subchronic exposure in which a small number of animals are exposed at a few administered

dose levels. Observation periods range typically from a few hours to a few weeks to yield a small number of repeated measurements on the same subject or time-course data from animals sacrificed at different sampling points. a PBPK model is complex involving many target tissues where data are unavailable. As a result, the development and validation of PBPK models are difficult. For example, it is unclear to what extend an acute exposure model can be applied to chronic exposure.

Because a PBPK model involves a large number of physiological and biochemical parameters, simultaneous estimation of these parameters is unlikely, and many of them are fixed at values that are uncertain and only a select few are calibrated or estimated through mathematical simulation (Wang *et al.*, 1997). The calibrated ones are nevertheless assumed to be constant over time and exposure condition. In reality, however, a physiological or biochemical parameter may change in value with exposure condition. Michalek *et al.*(2002) reported, for example, that the TCDD elimination rate increased with the initial TCDD concentration in both male and female participants of the Seveso cohort during the first three years of follow-up,but from year 3 to year 16, the rate of increase subdued in size as well as in statistical significance.

Still, the large number of parameters, complex model structure, and limited data often result in "non-identifiability" among parameters of which multiple, equally plausible estimates cannot be distinguished using statistical criteria. We call this phenomenon *hyper-parameterization* in the current context.

These issues have motivated us to explore statistical estimation of pharmacokinetic (PK) models. A case for illustration is a 2-year experiment by the National Toxicology Program (NTP) involving daily exposure of rats to a mixture of dioxin and dioxin-like compounds (DLCs). This chapter (1) illustrates statistical estimation procedures in the presence of parameter non-identifiability due to hyper-parametrization; (2) examines if kinetic parameters change in value under the chronic exposure conditions; and (3) evaluates the applicability of linear kinetics to chronic exposure. We first describe the NTP experiment providing the data, and then present a two-compartment PK model and the corresponding nonlinear regression models. We discuss statistical model fitting using iteratively reweighted non-linear least squares in conjunction with graphics to identify regions of non-

identifiability. We also discuss the bootstrap confidence interval to quantify uncertainty and variability in parameter estimates.

2.1.1 NTP's Experiment with Two-Year Chronic Exposure to TCDD

TCDD (2,3,7,8-Tetrachlorodibenzo-p-dioxin; CAS 1746-01-6) is a by-product of certain chemical synthesis processes such as the production of herbicides. Exposure to TCDD came to light in the use of the infamous "agent orange" during the Vietnam war to defoliate the rainforest (IOM, 2002). Today, the main sources of dioxins released into the environment include, but are not limited to, combustion and incineration, oil refining, chemical manufacturing and processing, and biological and photochemical processes (USEPA, 2000a), as well as existing reservoir sources from past releases.

Low water solubility, low rate of biotransformation, and high lipophilicity promote dioxin and DLCs to transport into and climb up our food chain, which accounts for a high potential for bioaccumulation (Esposito *et al.*, 1980). As a result, environmental exposure to dioxin and DLCs comes mostly from food intake, especially through the consumption of fish, meat, and dairy products because dioxin and DLCs are fat-soluble (Schecter *et al.*, 2001).

TCDD is the most toxic in the class of chlorinated hydrocarbon compounds (Van den Berg *et al.*, 1994). Short-term exposure at sufficiently high levels may result in skin lesions such as chloracne. It may also alter liver function. Long-term exposure is linked to impairment of the immune system, endocrine system, and the reproductive system (ATSDR, 1998). Chronic exposure at low levels of has resulted in increased cancer incidences such as cholangiocarcinoma and hepatocellular adenoma in different animal species. Based on available animal data and human epidemiological data, the International Agency for Research on Cancer (IARC) classifies TCDD as "carcinogenic to humans." The US Environmental Protection Agency (USEPA) also classifies TCDD as a human carcinogen (NRC, 2006).

Given the likelihood of daily exposure to dioxin and DLCs from a variety of sources and the associated potential health risk to humans, safeguard of these compounds in the

environment is of critical importance to public safety. To this end, sound risk assessment calls for better understanding of the human toxicokinetics of chronic exposure to dioxin and DLCs. In the National Research Council's evaluation of EPA's reassessment of dioxins (NRC, 2006), the committee specifically recommends the use of pharmacokinetics (PK) and physiologically-based pharmacokinetics (PBPK) in exposure assessment.

The National Toxicology Program (NTP) of the National Institute of Environmental Health Sciences conducted a series of studies to better understand the health consequences of chronic exposure to dioxin and DLCs. In one of these studies, female Harlan Sprague-Dawley rats were administered through gavage a mixture of TCDD, PeCDF, and PCB126, once daily 5-days a week, up to 104 weeks (2 years). These three compounds make up approximately half of the dioxin-like toxic activity found in human tissues (Walker *et al.*, 2005). The mixtures were prepared in a corn oil:acetone (99:1) solution at four concentration levels in the ratio of 1:2:10 parts TCDD:PeCDF:PCB126, with TCDD at levels of 3.3, 7.3, 15.2, and 33 *ng/kg* body weight, respectively. The mixture was determined according to the World Health Organization (WHO) toxic equivalent factor (TEF)(WHO 2007), so that each chemical would provide a third of the total dioxin toxic equivalent (TEQ) of the mixture. These levels were also used to maximize statistical power to test for possible interactions between the chemicals.

Eighty-one rats were randomly assigned to each dose group, in addition to a vehicle control of corn oil/acetone solution only. Eight to ten rats per dose group were sacrificed for biochemical (tissue concentration) and physiological evaluation at weeks 13, 30, and 52. The evaluation included thyroid hormone, cell proliferation, Cytochrome P450 activity, 7-ethoxyresorufin-O-deethylase and acetanilide-4-hydroxylase activities in the liver, and 7-ethoxyresorufin-O-deethylase activity in the lung. At week 104, all remaining animals were sacrificed for toxicological evaluation including carcinogenic outcomes. Among these animals, again only eight to ten were used for tissue concentrations. These animals were also weighted at baseline, weekly for 13 weeks, monthly thereafter, and at necropsy. Clinical findings were recorded on day 29, monthly thereafter, and at necropsy. See Walker *et al.*(2005) for more details on study design and clinical outcomes.

In the present analysis, we are focused on PK models for tissue concentration of TCDD in the fat and the liver. Table 2.5 summarizes the mean concentrations of TCDD in the fat along with standard deviations. Table 2.5 is a summary of concentration in the liver. TCDD concentration in each dose group clearly increases over time. Data variation shows a generally increasing dose-time trend except for fat at week 30.

2.2 Methods

2.2.1 Two-Compartment Pharmacokinetic Model

Figure 1 shows a conceptual, two-compartment model: TCDD transfers from the liver to the fat compartment at the rate K_{lf} and from the fat to the liver at the rate K_{fl}. Metabolism takes place in the liver and TCDD is eliminated at the rate K_e. Administered daily dose is assumed to enter the liver 100%. Although over-simplified, the model is adopted to illustrate statistical estimation of PK models with hyper-parametrization.

The mass balance of this model is represented by the following differential equations:

$$\begin{array}{rcl} \frac{dA_L(t)}{dt} & = & W_0 D - (K_e + K_{lf})A_L(t) + K_{fl}A_F(t), \\ \frac{dA_F(t)}{dt} & = & K_{lf}A_L(t) - K_{fl}A_F(t) \end{array} \tag{1}$$

where $A_F(t) = W_F(t)C_F(t)$ and $A_L(t) = W_L(t)C_L(t)$ are the total amount of TCDD in fat and the liver, respectively, with $C_F(t)$ and $C_L(t)$ being the respective tissue concentration, and $W_L(t)$ and $W_F(t)$ the respective tissue weight; W_0 is body weight at baseline; and D is daily administered dose. In solving the equations above, we assume that tissue weight is a constant proportion of body weight $W(t)$ over time t (ILSI, 1994). Specifically fat weight $W_F(t) = P_f W(t)$ with $P_f = 6.92\%$ and the liver weight $W_L(t) = P_l W(t)$ with $P_l = 3.63\%$ (Wang *et al.*, 1997).

The closed-form solution to equation (1) is given by

$$\begin{array}{rl} C_F(t) & = \frac{W_0}{W(t)P_F}\frac{DK_{lf}}{K_{fl}K_e}\left\{1 - \frac{\lambda_2}{\lambda_1-\lambda_2}\exp(\lambda_1 t) + \frac{\lambda_1}{\lambda_1-\lambda_2}\exp(\lambda_2 t)\right\} \\ & = \frac{W_0}{W(t)P_f}\mu_F(D,t,\mathbf{K}) \\ C_L(t) & = \frac{W_0}{W(t)P_L}\frac{D}{K_e}\left\{1 - \frac{\lambda_2(\lambda_1+K_{fl})}{K_{fl}(\lambda_1-\lambda_2)}\exp(\lambda_1 t) + \frac{\lambda_1(\lambda_2+K_{fl})}{K_{fl}(\lambda_1-\lambda_2)}\exp(\lambda_2 t)\right\} \\ & = \frac{W_0}{W(t)P_l}\mu_L(D,t,\mathbf{K}) \end{array} \tag{2}$$

where

$$\lambda_1 = -\frac{1}{2}\left\{K_{fl}+K_e+K_{lf}-\sqrt{K_{fl}^2-2K_{fl}K_e+2K_{fl}K_{lf}+K_e^2+2K_{lf}K_e+K_{lf}^2}\right\},$$

$$\lambda_2 = -\frac{1}{2}\left\{K_{fl}+K_e+K_{lf}+\sqrt{K_{fl}^2-2K_{fl}K_e+2K_{fl}K_{lf}+K_e^2+2K_{lf}K_e+K_{lf}^2}\right\},$$

and $\mathbf{K} = (K_e, K_{fl}, K_{lf})^T$

2.2.2 Statistical Models

Although the PK models are often fit through mathematical optimization and calibration, statistical estimation is advantageous in that uncertainties and variations in the kinetic parameters can be directly quantified through their interval estimates. To this end we convert the PK model (2) into a bivariate nonlinear regression model with unknown parameters $\mathbf{K} = (K_e, K_{fl}, K_{lf})^T$

$$\begin{aligned} Y_{Fijk} &= C_{Fijk}\frac{W_{ijk}(t_j)P_F}{W_{i0k}} = \mu_{Fij}(D_i, t_j, \mathbf{K}) + \boldsymbol{\varepsilon}_{\mathbf{Fijk}} \\ Y_{Lijk} &= C_{Lijk}\frac{W_{ijk}(t_j)P_L}{W_{i0k}} = \mu_{Lij}(D_i, t_j, \mathbf{K}) + \boldsymbol{\varepsilon}_{\mathbf{Lijk}} \end{aligned} \tag{3}$$

where the responses $\mathbf{Y}_{ijk} = (Y_{Lijk}, Y_{Fijk})^T$ are the per-baseline-body-weight TCDD concentration in the liver and fat, respectively, in the k^{th} rat ($k = 1, ..., n_{ij}$) of i^{th} dose group ($D_i = 3.3, 7.3, 15.2, 33$ ng/kg/day) at time t_j ($t_j = 13, 30, 52, 104$ weeks). We choose the concentrations per baseline body weight as the response variables to avoid treating weight as a time-varying covariate in the mean $\boldsymbol{\mu}_{\mathbf{ij}} = (\mu_{\mathbf{Fij}}(\mathbf{D_i}, \mathbf{t_j}, \mathbf{K}), \mu_{\mathbf{Lij}}(\mathbf{D_i}, \mathbf{t_j}, \mathbf{K}))^{\mathbf{T}}$. We also assume that the vector of errors $\boldsymbol{\varepsilon}_{\mathbf{ijk}} = (\boldsymbol{\varepsilon}_{\mathbf{Fijk}}, \boldsymbol{\varepsilon}_{\mathbf{Lijk}})^{\mathbf{T}}$ follows a bivariate normal distribution $N(0, \Sigma_{ij}^2)$, where the variance-covariance matrix

$$\Sigma_{ij} = \begin{pmatrix} \sigma_{Fij}^2 & \sigma_{FLij} \\ \sigma_{LFij} & \sigma_{Lij}^2 \end{pmatrix}$$

may vary with dose level i or time j within the same dose group. The errors are independent among rats.

Empirical evidence suggests that there is an over-dispersion in relation to the mean. We consider an variance-covariances structure with

$$\begin{aligned} \sigma_{F(ij)} &= \phi_F \mu_{Fij}, \\ \sigma_{L(ij)} &= \phi_L \mu_{Lij}, \\ \sigma_{FL(ij)} &= \rho_j \phi_F \phi_L \mu_{Fij} \mu_{Lij}, \end{aligned} \tag{4}$$

with time-specific Pearson's correlation coefficient ρ_j at time j. This variance-covariance structure captures the increasing variation in tissue concentrations with dose and/or time. The approximate variance-covariance matrices provide appropriate weights to the nonlinear least squares, and does not affect the unbiasedness of least square estimators.

It is important to note that the pharmacokinetics parameters may vary under chronic exposure conditions. Michalek *et al.*(2002) reported that the TCDD elimination rate increased significantly with the initial TCDD concentration in both male and female subjects of the Seveso cohort during the first three years of follow-up; and the rate further increased with a borderline significance in year 3-16 follow-up. To permit a dose-dependent kinetic parameter, we introduce a regression type model between the kinetic parameters and administered TCDD dose

$$\log K_m = \alpha_m + \gamma_m D_i \tag{5}$$

where $m = e, \mathrm{fl}, \mathrm{lf}$, and $D_i = 3.3, 7.3, 15.2, 33$ (*ng/kg/day*). This dose-dependent model is labeled "DD" in contrast with the dose independent (DI) model in which $\gamma_m = 0$.

2.2.3 Iteratively Re-weighted Nonlinear Least Squares

To estimate the kinetic parameters ($\mathbf{K}$), we fit the bivariate nonlinear model to the weight-scaled fat and liver concentration data simultaneously. The process relies on the iteratively re-weighted nonlinear least squares (IRNLS) to minimize the sum squares of the residuals

$$\mathrm{SSR} = \sum_{ijk} (\mathbf{y}_{ijk} - \mu_{ij})^T \Sigma_{ij}^{-1} (\mathbf{y}_{ijk} - \mu_{ij})$$

with respect to the parameters ($\mathbf{K}$). Because the variance-covariance (weight) matrices Σ_{ij} serve as weights, whose value must be updated at each iteration via updated estimation of

the means and the dispersion parameters. The IRNLS process consists of the following steps.

1. Obtain initial value $\mathbf{K}^{(0)}$ for $\mathbf{K}$. An appropriate starting value is critical to obtaining solutions to an optimization problem that involves nonlinear parameters.

2. Use previous estimate $\mathbf{K}^{(l)}$ $(l = 0, 1, 2, \ldots)$ to compute the predicted outcome $\hat{\mathbf{y}}_{ijk} = \mu_{ij}(D_i, t_j, \hat{\mathbf{K}}^{(l)})$. as well as estimate of the variance-covariance matrix $\Sigma_{ij}(\hat{\mathbf{K}}^{(l)})$. The overdispersion parameters ϕ_F and ϕ_L are obtained from a linear regression of the sample standard deviations against sample means ($\hat{\phi}_F = .199$ and $\hat{\phi}_L = .200$), and $\hat{\rho}_j$ is sample correlation coefficient between the fat and liver measurements at time j. As a proportion of the mean, the standard deviation is less sensitive to fluctuation of SSR when the mean estimate deviates from the expected value.

3. Obtain updated IRNLS estimate $\hat{\mathbf{K}}^{(l+1)}$ that minimizes

$$\text{SSR} = \sum_{ijk} (\mathbf{y}_{ijk} - \mu_{ij})^T \Sigma_{ij}^{-1} (\hat{\mathbf{K}}^{(l)}) (\mathbf{y}_{ijk} - \mu_{ij}).$$

   A number of algorithms are available for this optimization. We used the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm (Broyden, 1970; Fletcher, 1970) that is a part of the function *optim* in R (R Development Core Team, 2008). The "BFGS" algorithm is a quasi-Newton method (also known as a variable metric algorithm) capable of searching for a stationary point at which the gradient of the objective function (i.e. SSR) is 0. The stationary point is then a candidate of local optimum.

4. Repeat steps (2) and (3) until a convergence criterion is met. We employed a weak convergence criterion

$$\frac{|SSR^{(l+1)} - SSR^{(l)}|}{SSR^{(l)}} < \delta \tag{6}$$

   when the current value of SSR is within a pre-specified tolerance level $\delta > 0$ of the previous one. A negligible change in the $SSR^{(l+1)}$ from $SSR^{(l)}$ implies an overall convergence of the IRNLS algorithm to a local minimum.

2.2.4 Parameter Non-identifiability

It can occur that the objective function SSR attains its (nearly) minimum value within a contiguous neighborhood in the parameter space. Geometrically, this local neighborhood in the parameter space corresponds to an area where the surface of the objective function is "flat" - all points are equally acceptable estimates - a classic definition of parameter non-identifiability (Neath and Samaniego, 1997). In this scenario, the parameters involved cannot be uniquely determined. Non-identifiability can be structural in which the parameters are functionally dependent, i.e. one might be a function of others. Even if structurally identifiable, the parameters can still be non-identifiable when data do not provide sufficient information about the model (Hengl *et al.*, 2007). The complexity of a PBPK model dictates a large number of parameters that are functionally important, and requires highly informative data to uniquely determine the value of each and every of these parameters. In practice, however, the depth of the data is always insufficient to meet this requirement, giving rise to "hyper-parametrization".

The convergence criterion described above reflects our acceptance of non-unique estimates of the parameters. We thus develop a "clonal-expansion" algorithm in conjunction with a graphic tool to identify the contiguous region of all acceptable estimates. Specifically, expand the initial region $\mathscr{C}_\delta^*$ of acceptable estimates by including a new set $\mathbf{K}$ if the convergence criterion and the following criterion is met:

$$\frac{|SSR(\mathbf{K}) - min_{\mathscr{C}_\delta^*} SSR(\mathbf{K})|}{min_{\mathscr{C}_\delta^*} SSR(\mathbf{K})} < \delta^*. \tag{7}$$

To begin the clonal starts with the first set of estimates, then it either shifts to a better one with smaller SSR or expands to include a new member whose SSR is within the tolerance level δ^* of that of the signature member (the smallest SSR). We used $\delta^* = .005$. The contiguous region determines the scope of non-identifiable estimates from which we may pick our final estimate of the parameters, hence the final model.

2.2.5 Bootstrap Confidence Intervals

We use bootstrap confidence interval (Efron and Tibshirani, 1993) to report uncertainty and variation of the parameters. The bootstrap methods are a simulation-based resampling method that generates data to mimic the distribution underlying the original sample. Bootstrap confidence intervals are particularly useful in the present case when point estimates are not unique due to hyper-parametrization. Upon fitting the model, we obtain the residual vector, $\hat{\mathbf{e}}_{ijk} = \mathbf{y}_{ijk} - \hat{\mu}_{ij}$. For each observed concentration $\mathbf{y}_{ijk}$ we sample a residual $\hat{\mathbf{e}}^{(b)}_{ijk}$ from the same dose-time group with replacement, and add it to the predicted concentration to form a bootstrap sample data point,

$$\mathbf{y}^{(b)}_{ijk} = \hat{\mu}_{ij} + \hat{\mathbf{e}}^{(b)}_{ijk} \quad (i = 1,\ldots,I; j = 1,\ldots,J; k = 1,\ldots,n_{ij}; b = 1,2,\ldots,B). \tag{8}$$

This process yields a bootstrap copy of the original data set. The PK model is then fit to this bootstrap sample to yield a copy of bootstrap estimate $\hat{\mathbf{K}}^{(b)}$ of the kinetic parameters. We generated a total of $B = 1000$ copies for the dose independent model and $B = 1342$ for dose dependent case. (B=2000 were attempted, only copies of converging estimates for $\mathbf{K}$ were included). The estimates of $\mathbf{K}$ are obtained through α_m and γ_m via the function $\log K_m = \alpha_m + \gamma_m D$ (m=e, fl, lf).

From the bootstrap sample of estimates we then construct a bias-corrected and accelerated (BCa) percentile-intervals (Efron, 1986). The BCa confidence interval is given by

$$\theta^*_{\lfloor B\hat{\alpha}_1 \rfloor} < \quad \theta \quad < \theta^*_{\lceil B\hat{\alpha}_2 \rceil} \tag{9}$$

where the lower bound is the $\lfloor B\hat{\alpha}_1 \rfloor^{th}$ smallest bootstrap estimate determined by the floor of $B\hat{\alpha}_1$, and the upper bound is the $\lceil B\hat{\alpha}_2 \rceil^{th}$ largest bootstrap estimate determined by the ceiling of $B\hat{\alpha}_2$. The nominal levels $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are defined as the normal probability

$$\hat{\alpha}_1 = \Phi\left[\hat{z}_o + \frac{\hat{z}_o - z_{1-\alpha/2}}{1 - \hat{a}(\hat{z}_o - z_{1-\alpha/2})}\right] \text{ and } \hat{\alpha}_2 = \Phi\left[\hat{z}_o + \frac{\hat{z}_o + z_{1-\alpha/2}}{1 - \hat{a}(\hat{z}_o + z_{1-\alpha/2})}\right],$$

where $\hat{z}_o = \Phi^{-1}\left(\sum_{b=1}^{B} \frac{I(\hat{\theta}^{(b)} < \hat{\theta})}{B}\right)$ is the bias correction factor when the bootstrap estimates $\hat{\theta}^{(b)}$ deviate from $\hat{\theta}$ in median, and $\hat{a}$ is the acceleration factor which adjusts for the skewness of the bootstrap sample estimates. Let $\hat{\theta}_{(-i)}$ represent the estimate of θ when the i^{th}

observation is omitted and $\bar{\theta}$ represent the average of the $\hat{\theta}(-i)$. Thus estimated skewness is $\hat{a} = \frac{\sum_{i=1}^{B}(\bar{\theta}-\hat{\theta}(-i))^3}{6\left[(\bar{\theta}-\theta(-i))^2\right]^{3/2}}$. These adjustments make BCa confidence interval transformation invariant with better coverage than the standard percentile bootstrap confidence interval (DiCiccio and Efron, 1996).

2.3 Results

2.3.1 Model Fitting and Parameter Estimate Identifiability

To fit the PK model, we designed a large number of starting value sets along a six-dimensional grid corresponding to the six regression parameters in **K**. A non-structural "hyper-parametrization" indeed occurred among K_{fl} and K_{lf}. Using the "clonal expansion" algorithm in conjunction with the weak convergence criterion, we obtained a $n = 1173$ sets of equally acceptable estimates for the dose-independent model and $n = 1095$ sets of estimates for the dose-dependent model. These estimates are believed to be in a contiguous region corresponding to a near minimum value of the objective function SSR. The confirmation is done by projection of the six-dimensional region onto all 15 possible 2-dimensional planes (Figure 2). Although we do not know precisely the boundary of the contiguous region, we stopped expanding the grid points of starting values when convergence became increasingly rare.

Figure 2 shows clearly that for K_{fl} and K_{lf} the intercepts α_{lf} and α_{fl} are almost perfectly related (the panel on row 3 and column 5), so are the slopes γ_{lf} and γ_{fl} (the panel on row 4 and column 6). It reveals that within the 6-dimensional contiguous region of estimates corresponding to the flat maximum likelihood surface area, the estimates for α_{lf} and α_{fl} are linearly dependent, and those for γ_{lf} and γ_{fl} are also linearly dependent. This is so despite the fact that the parameters are not linearly dependent in the model itself. A limitation of the plot is that we do not know precisely the entire contiguous region of plausible estimates and cannot say where exactly the boundaries lie. Thus a warning should be noted in interpreting these values due to potential bias of not covering the entire contiguous region. More biologic and or physiological information would be required to reduce the uncertainty and resolve non-identifiability.

Also noticeable in the scatter plot, the intercept and slope of K_e (row 1 and column 2) reveals that a larger intercept is coupled with a smaller slope and vice versa, suggesting a narrow range of plausible K_e estimates even though the elimination rate itself may increase due to the value of the slope and the exposure condition. These results along with the correlation between K_{fl} and K_{lf}, indicate that the contiguous region has three edges along which the parameter estimates are not distinguishable.

It is understandable that the concentrations observed in the liver and fat compartment alone are insufficient conclude that the rate from the liver to fat (K_{lf}) is different from the rate from fat to the liver (K_{fl}) even though in theory they are structurally different. External and additional information pertaining to plausible values of these parameters are necessary in order to resolve this problem of non-identifiability. Visual inspection of the scatterplots alone cannot resolve the problem.

2.3.2 Parameter Estimation

Because we do not know precisely the entire contiguous region of estimates or the distribution within the region, it is not trivial to summarize these estimates. To a crude approximation, we use the sample mean of all estimates in the contiguous region as the central value, and then use it as the "original" estimate to construct a BCa bootstrap confidence interval to measure uncertainty and variation (Tables 3 and 4).

Table 5 shows the estimates of the regression coefficients α_m (intercept) and γ_m (slope) along with their confidence interval for the kinetic parameters in the dose-dependent models (DD). The positive slope γ_e suggests the elimination parameter K_e is increasing with exposure level. There is also some evidence that K_{fl} and K_{lf} may be varying with exposure with a positive slope γ_{fl} and negative slope γ_{lf}.

Table 6 summarizes the estimates of the kinetic parameters along with their 95% confidence intervals for each expousre level. Again, K_e is increasing with exposure, so is K_{fl}, whereas K_{lf} decreases with exposure. Note that the estimated value of K_{lf} is smaller than that of K_{fl} in both the dose-dependent and dose-independent models. TCDD bioaccumulates due to its ability to bind to Ah receptors in the liver as well as its high lipophilicity.

This simplified model does not consider binding in the liver and fat, nor uses partition coefficients to account for the lipophilicity. This is likely one reason why the estimate of K_{fl} is greater than K_{lf}.

Caution should be exercised in interpreting these summary values due to potential bias of under-covering the entire contiguous region of non-identifiability. For example, it is possible that the boundary for the slopes for K_{lf} and K_{fl} may extended beyond the range seen in the plot (Figure 2), despite our further investigation (data not shown) found that beyond this range convergence was increasingly unlikely, suggesting a practical approximation to the true boundary. The biological meaning of such statistical averages is unclear. Use a different "average" value which is statistically equally admissible may lead to a different confidence intervals as well.

Figure 3 displays the fitted model in the fat compartment under the DI model (solid lines) and the DD (dotted lines) models, respectively. Each panel in the figure represents a dose group in which a circle represents an individual animal's observed value. Data for the control group are omitted because of 0 administered dose and tissue concentration below the detection limit. Within each dose group, the tissue concentration clearly increases over time due to cumulative exposure. However, the rate of increase is subdued beyond 30 weeks, notably in the two highest dose groups, suggesting the tissue concentration reaches an equilibrium. Across the panels, the concentration in fat also increases with the administered dose level. It is important to note variation also increased as the mean concentration level increases. This heteroscedasticity supports the employment of dose-time specific weights in the iteratively re-weighted least-squares estimation.

Figure 4 displays the fitted model for TCDD concentrations in the liver. Both the model and data show increasing concentration with exposure level and time. The data also show increasing variation with the average concentration as well.

Comparison of the models with data suggests that both models impose an equilibrium concentration whereas the data (by week 104) have yet to establish such an equilibrium, especially in the two lower dose levels. Even at the highest dose level where empirical evidence does suggest an equilibrium, it does not appear until about week 52 rather than week 30. The ascending to equilibrium is markedly rapid in the DI models (solid line) compared

with DD models (dotted line). The DD models allow for a somewhat-delayed equilibrium level, especially in the lowest two dose groups. The rapid ascent to the equilibrium by the DI model resulted in overestimation of tissue concentration at the earlier times (e.g. week 13), seen clearly in the highest dose group. This disparity in the timing of reaching equilibrium can be viewed as evidence supporting for dose-dependent kinetic parameters. Moreover, the adoption of the linear kinetics may have also contributed to the inadequate model fitting. Whereas linear kinetics is often used for acute exposure studies, nonlinear kinetics as well as binding of TCDD in fat could be more appropriate (Wang *et al.*, 1997).

2.4 Conclusion

In this paper we have illustrated statistical estimation of multiple-compartment PK model using iteratively re-weighted non-linear least squares in conjunction with bootstrap confidence intervals. Our application of the models to the NTP two-year experiment of chronic exposure to TCDD mixtures showcased a common challenge in modeling biological systems - parameter non-identifiability. That is when a biological system involves a large number of parameters yet there are insufficient data or information to ascertain the values of these parameters. As a result, there will be a contiguous region within which all values are equally acceptable estimates under statistical criteria. While common statistical methods can be used to identify the region of non-identifiability, specific choice within the region will be difficult to justify by statistical criteria alone, and is subject to substantial uncertainty. We have used weak convergence criterion and graphic tools to identify the region, and use the choice of sample averages. While conventional statistical estimation remains useful, biological information will be critical to reduce the uncertainty and resolve the non-identifiability.

Because our nonlinear regression models are based on linear kinetics common in characterizing acute exposure studies, it is possible that such models may not be adequate to describe data of chronic exposure condition. Our PK models impose an equilibrium earlier than the data indicated, suggesting a nonlinear kinetics as well as exposure-altered kinetic parameters may be warranted. Although our model is an oversimplification of the underly-

ing PBPK models such as the one proposed by Wang *et al.*(1997) which involves multiple compartments and nonlinear kinetics for TCDD-AhR binding, we still found evidence that some key kinetic parameters may change over exposure conditions, including the elimination rate and transfer rates between compartments.

Although data in additional tissue compartments may help remove non-identifiability among some parameters, new compartment and new parameters also will be introduced. As a result, parameter non-identifiability will remain an issue for statistical modeling, hindering the adoption of statistical estimation. To promote statistical modeling of complex biological systems, new and more effective methods should be promoted. Methods such as Markov chain Monte Carlo (MCMC) is useful in this regard. Work is ongoing for statistical estimation of PBPK models for dioxin, following the PBPK models of Wang *et al.*(1997, 2000) and Emond (2004). Statistical modeling of complex biological systems must incorporate existing biological information to construct a model and to form a prior for the model parameters. Choice of priors in MCMC deserves attention as artificially informative prior can essentially lead to arbitrary and as uncertain as the prior. Statistical estimation is increasingly hybrid in that it will incorporate computational, mathematical, and graphic tools.

2.5 Tables

Table 3: Mean (SE) TCDD concentrations in the fat (*pg/g*)[c].

	Dose Level				
Week	Control	3.3 ng/kg	7.3 ng/kg	15.2 ng/kg	33 ng/kg
13	13[a]	305.4(12.9)	563.2(44.2)	704.7(52.6)	1399.2(103.3)
30	-[b]	326.7(12.3)	646.5(14.8)	1292.0(37.9)	2610.0(96.9)
52	-[b]	345.3(20.5)	687.6(44.4)	1098.9(75.6)	2706.3(95.0)
104	12.60 ± 0.40	553.2(33.9)	856.6(116.3)	1616.0(107.7)	2823.8(208.6)

[a] Based on one observation; [b] Level below the limit of quantitation (LOQ) of 5 *pg/g*;

[c] Number of sacrificed animals at each time point per dose group: Week 13: 10,10,10,10,10; Week 30: 10, 10, 10, 10, 10; Week 52: 8, 8, 8, 8, 8; Week 104: 8, 10, 10, 10, 10.

Table 4: Mean(SE) TCDD concentrations in the liver (pg/g)[b].

	Dose Level				
Week	Control	3.3 ng/kg	7.3 ng/kg	15.2 ng/kg	33 ng/kg
13	0.74(0.26)	613.6(33.1)	1675.5(84.3)	3625.0(166.7)	6660.0(402.9)
30	0.13[a]	727.2(47.6)	1789.9(130.9)	3383.0(200.9)	8508.0(495.0)
52	0.14[a]	779.6(32.8)	2048.8(103.8)	3836.2(181.3)	9503.8(647.3)
104	4.62(3.68)	1260.8(100.7)	2611.0(141.7)	4640.0(336.9)	10518.8(929.6)

[a] Based on one observation; [b] Number of sacrificed animals at each time point per dose group: Week 13: 10,10,10,10,10; Week 30: 10, 10, 10, 10, 10; Week 52: 8, 8, 8, 8, 8; Week 104: 8, 10, 10, 10, 10.

Table 5: Estimates (95% BCa bootstrap CIs[a]) of regression coefficients.

Coefficient	Intercept (α_m)	Slope (γ_m)
$\log(K_e)$	-2.9903 (-3.0690, -2.9678)	0.00998 (0.0082, 0.0244)
$\log(K_{fl})$	-4.1314 (-4.1657, -4.1296)	0.0059 (0.0052, 0.0095)
$\log(K_{lf})$	-4.3558 (-4.5207, -4.3502)	-0.0061 (-0.0076, -0.0038)

[a] Sample mean of all acceptable estimates was used as the original estimate for bootstrap sampling.

Table 6: Estimates (95% BCa bootstrap CIs[a]) for kinetic parameters.

	Dose-Independent	Dose Dependent model			
	Model	$3.3(ng/kg)$	$7.3(ng/kg)$	$15.2(ng/kg)$	$33(ng/kg)$
K_e	0.0583	0.0519	0.0540	0.0584	0.0697
CI	(0.0561, 0.0604)	(0.0477, 0.0557)	(0.0493, 0.0614)	(0.0526, 0.0745)	(0.0609, 0.1150)
K_{fl}	0.1109	0.0164	0.0168	0.0176	0.0195
CI	(0.1083, 0.1138)	(0.0158, 0.0166)	(0.0161, 0.0172)	(0.0168, 0.0186)	(0.0184, 0.0220)
K_{lf}	0.0752	0.0126	0.0123	0.0117	0.0105
CI	(0.0730, 0.0772)	(0.0106, 0.0128)	(0.0103, 0.0126)	(0.0097, 0.0122)	(0.0085, 0.0114)

2.6 Figures

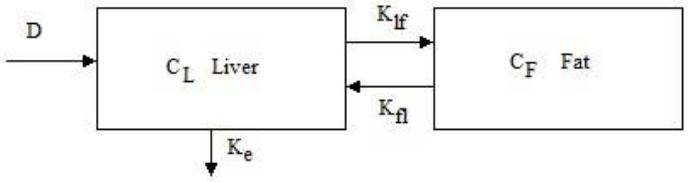

Figure 1: Schematic representation of the pharmacokinetic model.

C_L, C_F are concentrations, D is dose, and K_e, K_{fl}, and K_{lf} are kinetic parameters.

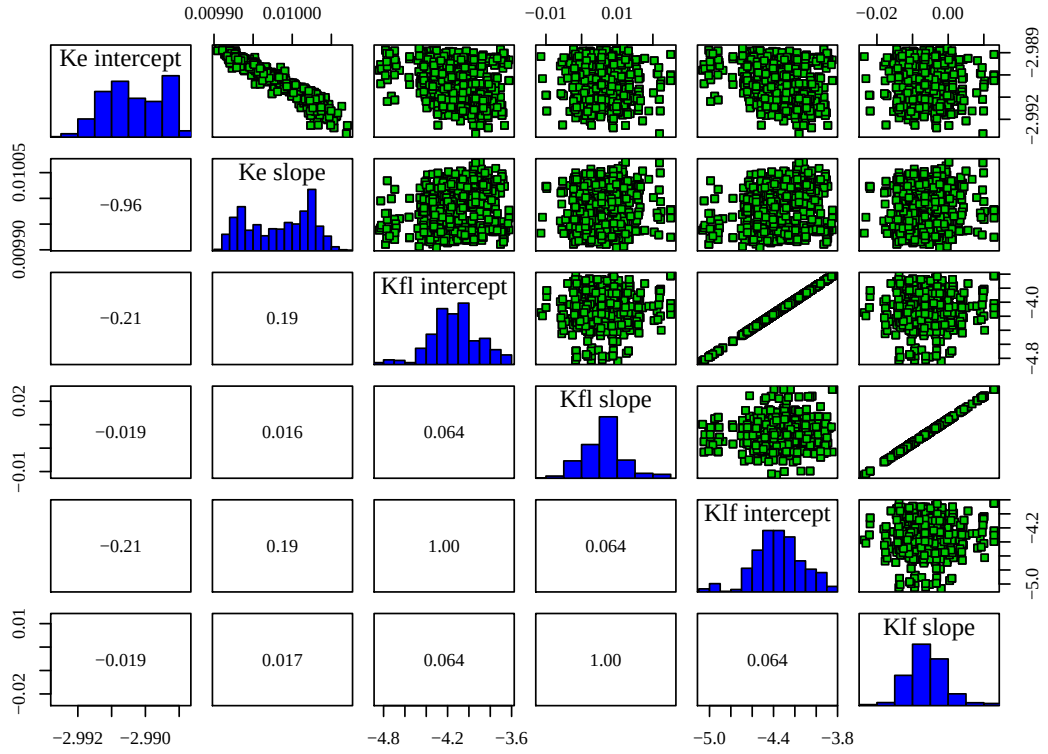

Figure 2: Scatter plots of plausible estimates of the model parameters.

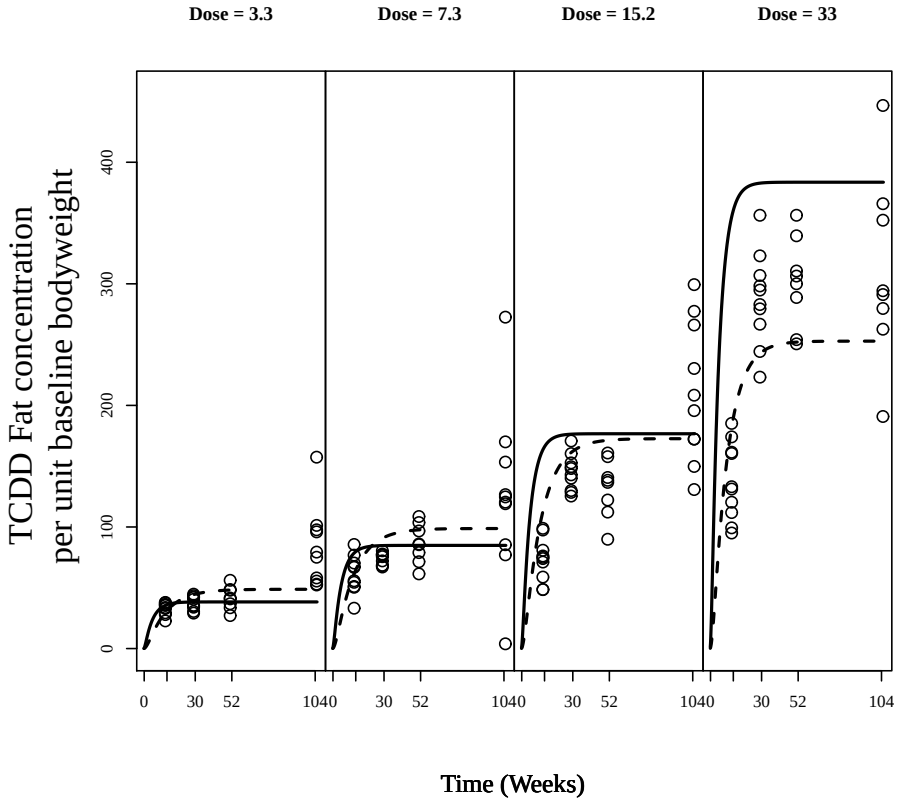

Figure 3: TCDD concentration in fat (concentration per unit baseline bw).
Constant (solid line) and dose-dependent (dotted lines) kinetic parameter models for TCDD.

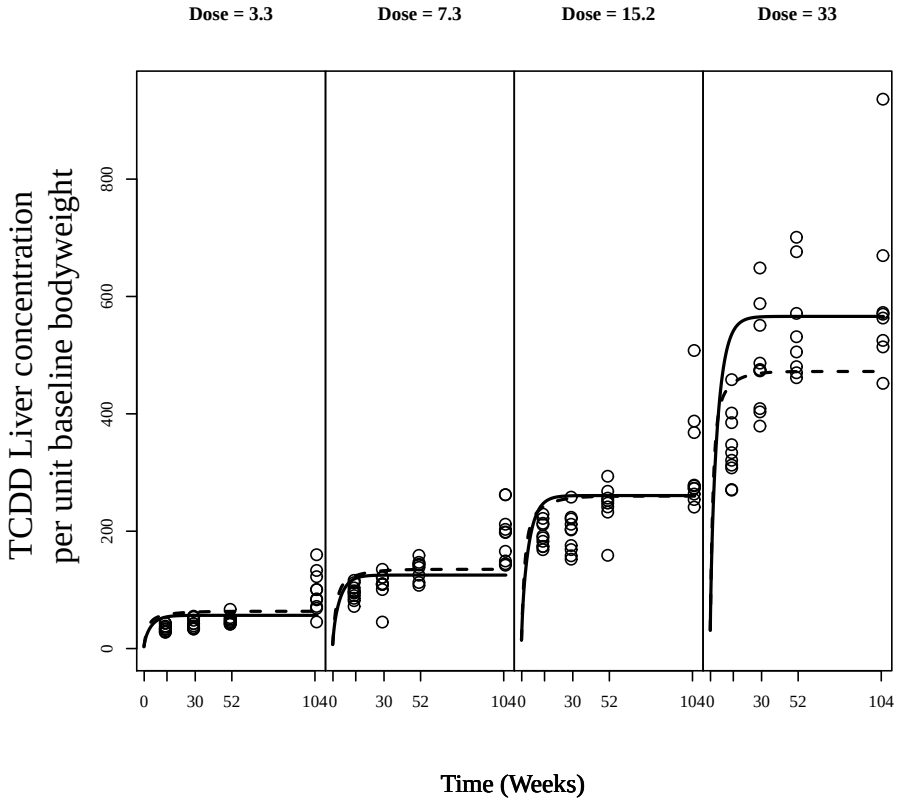

Figure 4: TCDD concentration in liver (concentration per unit baseline bw).
Constant (solid line) and dose-dependent (dotted lines) kinetic parameter models for TCDD.

3 Body and Organ Weight Growth of Female SD Rats exposed to a Mixture of DLCs

3.1 Introduction

Pharmacokinetics (toxicokinetics) and pharmacodynamics (toxicodynamics) are the main tools to study xenobiotics in a biological system of a mammal species. In the last few decades these models have grown into physiologically based pharmacokinetic (toxicokinetic) or pharmacodynamic (toxicodynamic) models by integrating realistic characteristics of a biological system (Clewell and Andersen, 1985; Anderson, 2003). Organ weight is one fundamentally important characteristic of any biological system. Because organ weight is often not directly measurable, estimation is required. One common approximation is to assume that the organ weight is constant fraction of the total body weight. For example, the International Life Science Institute and the EPA's Office of Health and Environmental Assessment suggests that the rat liver is 3.66% of the total body weight and the lungs are 0.5% of the body weight (Brown *et al.*, 1997). In most PBPK models an organ's weight is assumed to be a constant fraction of body weight, and the potential exposure or growth effects on this relation is ignored. This is the case in PBPK models of Dioxin (Wang *et al.*, 1997; Emond, 2004). These authors considered body weight growth equations, but did not consider exposure effect that might alter body weight or organ weight growth. Organ weight may hold well as a constant portion of total body weight in an acute study of short duration in which body and organ weight change is negligible. But over a longer period of active growth or significant environmental interference such as chronic exposure to toxins, the organ and body weights and their relation may not hold.

The assumption of constant organ to body weight ratio must be verified. Invalidity in the assumption can have critical impact on pharmacokinetics and pharmacodynamics especially in a sub-chronic or chronic toxicological study, including incorrect estimation of the total amount of the agent in the target tissue (Brown *et al.*, 1997). Body and organ weights and organ volumes are some of the most important physiological parameters in PBPK models, and they affect many physiological parameters, including but not limited to breathing rate, cardiac output rate, and organ perfusion rate, which in turn influence blood flow velocities, surface areas, and permeability.

The literature appears limited on the relationship between organ and body weights. Schoeffner *et al.*(1999) investigated the organ weights and fat volume of Fischer344 rats and Sprague-Dawley (SD) rats at 4, 6, 8, 10, 12, 14, and 40 weeks as the animals aged. The relative weight of the liver, spleen, kidneys, and lungs decreased between 4-week and 14-week in age. This implies that even in short term studies of acute or sub-chronic exposure, the constant organ weight fraction to body weight may not hold. In youngest rats (4-8 weeks old) the percent of body fat ranged from 5.2% to 5.8%, while in rats aged 10-14 weeks fat ranged from 6.1% to 6.5%, although the increase was not statistically significant.

This paper investigates how organ weight might change as a fraction of body weight under chronic exposure during an active growth period. We show that the growth of body weight of SD rats was affected by exposure to a dioxin mixture, and the organ weight was changing as well relative to body weight over time and across exposure level.

3.2 Methods

3.2.1 The NTP 2-Year Carcinogenicity Study of Dioxin and DLC Mixture

Body weight and organ weight data of female Harlan Sprague-Dawley rats were derived from a NTP 2-year carcinogenesis study of mixture of dioxin, PeCDF, and PCB126. TCDD ($2,3,7,8$-tetrachlorodibenzo-p-dioxin; CAS 1746-01-6) or dioxin is a chlorinated hydrocarbon, and is the prototype of the polyhalogenated aromatic hydrocarbon (PHAH) family, which includes dibenzo-p-dioxins (PCDDs), dibenzofurans (PeCDFs), biphenyls, and polychlorinated biphenyls (PCB). TCDD is the most toxic compound in this family.

The National Toxicology Program (NTP) of the National Institute of Environmental Health Sciences (NIEHS) conducted a series of studies to better understand carcinogenicity of chronic exposure to dioxin and DLCs. In one of these studies, female Harlan Sprague-Dawley rats were administered a mixture of TCDD, PeCDF, and PCB126. The mixture was administered through gavage, once daily 5 days a week, up to 104 weeks (2 years). The mixtures were prepared in a corn oil:acetone (99:1) solution at four concentration levels in the ratio of one part TCDD, two parts PeCDF, and ten parts PCB-126, with the dose of TCDD being 3.3, 7.3, 15.2, and 33 ng/kg body weight respectively. Eighty-one rats were randomly assigned to each dose group, in addition to a vehicle control of corn oil/acetone solution only. Eight to ten rats per dose group were sacrificed for biochemical (tissue concentration) and physiological evaluation, at weeks 13, 30, and 52. The animals were weighted at baseline, weekly for 13 weeks, monthly thereafter, and at necropsy. Organ weights were only collected at week 13, 30, and 52. The sampling scheme of the 2-year study is shown in Table 7. See Walker *et al.*(2005) for more details on study design and outcomes.

3.2.2 Body Weights

The overall baseline mean weight of the rats is 180.8 grams and the baseline mean weights for each dose group were 179.1, 180.0, 181.4, 181.3, and 182.1 grams in the control, 10, 22, 46, and 100 *ng TEQ/kg bw* dose groups respectively. The mean body weight by time and dose is displayed in table 2. For example the mean weight at the last time point of 104 weeks was 360.8, 380.5, 365.5, 329.4, and 287.0 grams for the control and dose groups respectively.

Figure 5 summarizes body weight at the time of sacrifice using boxplots. The boxplots show that the variation in the weights at the 104 week time point is greatest in the two lowest dose groups. Body weight growth does not seem to be slowed in the two lower dose groups compared to the control group; but the weights at the highest group were noticeably lower than the control group at all four sacrifice times. The individual rats growth curves also demonstrate this trend and are shown later. We show that exposure to dixoin and the

mixture of PCB126 and PeCDF cause a reduction in the maximum weights and growth periods of the rats at the higher exposure levels.

3.2.3 Organ Weights

Weights of the liver, lung, spleen, right kidney, thyroid, and right ovary were obtained at week 13, 30, and 52 upon sacrifice. The means and standard deviations of the observed weights are presented in Table 8. We also report the mean ratio of organ weight to body weight by dose and time in Table 9.

The mean liver weight increases as dose level increases, as well as with time. At week 13 the mean liver weight in the control group is $9.33g$ and increases to $10.72g$ in the highest dose group. At week 52 the range in mean liver weight is $10.38g$ to $12.42g$ across the dose levels.

In terms of proportions, at week 13 the control group's liver weight on average is 3.12% of total body weight, but this increased to 3.59% in the lowest dose (10 *TEQ ng/kg*) group, and to 4.12% in the highest dose (100 *TEQ ng/kg*) group. This trend suggests an increasing portion of liver weight in body weight when exposure level increases. There is a similar pattern in the ratios at weeks 30 and 52. Over time the general trend is also increasing based on these crude ratios. A formal analysis of these trends is done through the use of regression models.

Similar patterns are observed in the ratios of the lung to body weights and kidney to body weights. In the lung the mean relative weights range from 0.61% in the control group at week 13 to 0.71% in the highest dose group at week 13. At week 52 the range is 0.69% to 0.83%. While the relative kidney weight is also increasing (although not monotonically) in dose and time, from 0.26% in the control group at week 13 to 0.29% at week 52.

The change in the weight of the spleen over exposure level is opposite that of the liver, lungs, and kidney. At week 13 the mean spleen weight is $0.58g$ in the control group decreasing to $0.47g$ in the highest dose group. At week 52 the range is $0.54g$ to $0.43g$ over the dose levels. The mean spleen weight is decreasing in time. In the control group the mean

spleen weight is $0.56g$ at 13 weeks and decreases to $0.49g$ at 52 weeks. The thyroid and ovary weights also decrease in time yet changes over exposure level are not monotonic.

3.2.4 Body Weight Growth Model

To investigate and quantify the relationship between organ weights and body weight we first conduct a nonlinear regression model of body weight growth as mediated by exposure level. Body weight growth was described using a Michaelis-Menten-like kinetic function. To capture the between-subject variation in weight growth, we introduced random-effects into the coefficients which measures individual deviation in the regression coefficient from the average value. This is a statistical nonlinear mixed-effects model.

Let Y_{ijk} be the body weight of k-th rat in the i^{th} dose group at time j, where D_i ($i = 1,2,3,4,5$) corresponding to 0, 10, 22, 46, and $100ng\ TEQ\ /kg\ bw$. Figure 6 is a plot in which each panel (dose group) displays individual rats' weight growth over time. The rats' weight were measured once a week for the first 13 weeks and then once a month thereafter until sacrifice. The plot suggests an increasing trend of body weight that is suppressed somewhat by the increasing exposure level in both the growth period and peak weight. Moreover the between-subject variation is visibly increasing over time, but stays roughly the same over exposure level. These observations suggested the Michaelis-Menten kinetics function for the mean weight

$$\mu_{ij} = W_0 + \left(\frac{V_{max} \cdot t_j}{K_m + t_j}\right) \tag{10}$$

where V_{max} and K_m are functions of dose to permit the exposure effect on maximum growth as well as on rate of growth:

$$V_{max} = exp(V_a + V_b \cdot D_i) \tag{11}$$

$$K_m = exp(K_a + K_b \cdot D_i) \tag{12}$$

Note that the coefficient K_m is the time when half-maximal growth is achieved. The complete statistical model for body weight growth is therefore

$$W_{ijk} = \mu\left(D_i, t_j, W_0, V_{max}, K_m\right) + \varepsilon_{ijk} \tag{13}$$

where the error term $\varepsilon_{ijk} \sim N\left(0, \sigma^2\mathbf{I}\right)$ To capture the increasing variation in body weight as the animal grew older, we proposed $\sigma = exp^{(\beta \cdot t)}$. For simplicity the errors from the same subject were considered independent, hence the variance covariance matrix involves the identity matrix $\mathbf{I}$. To further allow for between-subject variation, random effects are introduced so that W_0, and K_a, K_b can be replaced with

$$W_0 + \gamma_{1k}$$

$$K_a + \gamma_{2k}$$

and

$$K_b + \gamma_{3k}$$

respectively.

3.2.5 Organ and Relative Organ Weight Models

We modeled the relationship between organ weight and body weight through regression models with exposure level and time as mediation effects. Two approaches were considered. The first is a linear model in which organ weight is thought of as a function of body weight with adjustment by animals age and exposure level.

Thus we initially fit a model of the form.

$$Y_{ijk} = \beta(D_i, t_j)W_{ijk} = \left(\beta_0 + \beta_1 \cdot D_i + \beta_2 \cdot t_j\right) W_{ijk} + \varepsilon_{ijk} \tag{14}$$

where Y_{ijk} represents the weight of the organ of interest for the k^{th} rat in the i^{th} dose group and week t_j ($i = 1 \ldots 5$ and $j = 1, 2, 3$). The model can be extended to include nonlinear term (e.g. quadratic term) or cross product of dose and time along the classic linear regression framework. The function $\beta(D_i, t_j)$ is the mean relative weight of the organ of interest and it is this quantity that will be derived and reported.

The second approach is a variation of the first model in which we directly model the ratio of an organ weight to body weight.

$$\log \frac{Y_{ijk}}{W_{ijk}} = \beta_0 + \beta_1 \cdot D_i + \beta_2 \cdot t_j + \varepsilon_{ijk} \quad (15)$$

This model can also be expanded to include nonlinear term or cross-product of dose and time. This second approach is attractive in its ability to deal with hetereocadesticity - variance is not constant over dose or time. In both models, we can consider

$$\varepsilon_{ijk} \sim N(0, \sigma_{ij}^2 \mathbf{I})$$

where we allow σ_{ij} to be distinct between dose levels and time points, including the more formal function $\sigma_{ij} = exp(\lambda_1 \cdot D_i + \lambda_2 \cdot t_j)$.

3.3 Results

3.3.1 Body Weight

Regression estimates and standard deviations of the random effects of the nonlinear mixed-effects model for body weight growth are presented in Table 10. The regression coefficients for the slopes (V_b and K_b) of the dose terms in the maximum weight (V_{max}) and half-life of growth (K_m) equations are both highly significant (p-value $<$ 1.00e-04) and negative. The values of -9.81e-03 and -0.02 for V_b and K_b respectively confirm the reduced maximum weight and shortened growth spurt of the rats. The standard deviations of the random effects of the slope and intercept of the K_m term are small, 4.89e-03 and 4.07e-05 for K_a and K_b respectively, indicating small between subject variability.

Using the regression model we can calculate the value of V_{max} and K_m for each dose group and these are presented in Table 11. The model suggests that body weight growth was reduced by approximately 28% in the highest dose group compared to the control group. The growth spurt reached it's half maximum value earlier in the highest dose group, about 46% earlier then the control group.

In this case, ignoring the dependence of body weight on exposure would cause misleading estimates of other related parameters. Especially in the context of a PK or PBPK

model. Examining subject growth in body weight should be done on a case by case basis in applications of model building and risk assessment.

3.3.2 Organ Weight as a fraction of body weight

The regression coefficients and associated summary statistics for the organ weight models are displayed in Table 12. The only organ model to have a significant quadratic dose effect was the liver model and it is denoted as β_3. Since both of the dose terms are highly significant in the liver model, we have evidence that exposure modifies the weight of the liver relative to the body weight. The dose coefficient is 2.28e-04 (p-value = 2.63e-13) and the quadratic coefficient is -1.33e-06 (p-value = 1.73e-06), thus the relative liver weights are increasing as dose increases. Given that the sign of the dose coefficient is positive and the sign of the quadratic term is negative, the change in relative liver weight is increasing yet the rate of growth slowed as dose increased. Liver weight as a fraction of body weight also changes over time. Given the observations and these results we conclude liver weight is not a constant proportion of body weight over the first two years of life for SD rats. The predicted values of the relative weights from all the organ models are found in Table 13.

In the lung and spleen models there is a dose effect. In the lung model the estimated coefficient for dose is 9.57e-06 (p-value = 4.28e-05), which implies an increase of 9.57e-06 in the relative lung weight for every unit increase in the dose level. The coefficient for dose in the spleen model is -1.08e-06 (p-value = 2.95e-02), this means we observe a decreasing effect as dose increases, opposite to that of the liver and lung as time is held constant.

The same models were fit to the kidney, thyroid, and ovary weight data. In the kidney model the dose coefficient is borderline significant (9.55e-07, p-value = 7.72e-02) but in the thyroid and ovary models dose is not statistically significant. Although the dose terms are not significant, the regression coefficients in the thyroid and ovary models are negative, implying a decreasing effect on the size of the organ as dose increases.

In all the organ models, time is highly significant. The relative weight of the liver, lung, and kidney as a fraction of body weight increase with the animal's age; however, in the spleen, thyroid, and ovary the relative weight is decreased over time. In the lung for

example, the mean relative weight is increasing by 1.43e-05 per week and the spleen is decreasing by -7.62e-06 per week.

3.3.3 Organ-Body Weight Ratios

The observed liver weight to body weight ratios were plotted by dose and time point and is displayed in Figure 7. The ratio is directly modeled on a log-scale. The regression coefficients and associated summary statistics are displayed in Table 14. The predicted values from all of the organ-body ratio models are found in Table 15. The only organ model to have a significant quadratic dose term is the liver model and it is denoted as β_3.

Since both dose terms are highly significant in the liver model, we have evidence that exposure modifies the ratio of liver to body weight. The dose coefficient is 6.11e-03 (p-value = 2.30e-13) and the quadratic dose coefficient is -3.59e-05 (p-value = 1.13e-06), thus the relative liver weights are increasing as dose increases but the rate of change slowed as dose increased (as indicated by the negative value of the coefficient for the quadratic term).

Given the signs of the dose coefficients, the change in the log of the ratio is increasing yet concave with a downwards curvature as dose increases and time is held constant. By exponentiating the intercept β_0 we estimate that the liver is approximately 3.06% of the body weight at baseline.

The observed relative liver weight ranged from 3.12% in the control group to 4.12% in the highest dose group at week 13 and at the end of one year the mean ranged from 3.21% for the control group to 4.32% for the highest dose group. The mean predicted values are not far from the observed mean values. The quadratic liver model predicted 3.13% in the control at week 13 and 4.28% for the highest dose group at the end of one year. A difference of only .04% in the highest dose group at week 52.

Not only did time and dose modify the growth trajectory for the relative liver weights but an increasing effect was seen in the lung and a decreasing effect in the spleen. Growth over time was statistically significant in all organs. Both the time and dose terms are highly significant in the relative weight model for the lung and spleen. For the lung, the estimate of the dose effect is 1.42e-03 (p-value = 7.72e-06). For the spleen, the estimate of the dose

effect is -6.95e-04 (p-value = 8.50e-03), suggesting a decreasing relative weight as dose increases.

The same analysis was conducted on the relative weights of the kidney, thyroid, and ovary. The dose effects are not statistically significant at the 5% level in the these organs. There is very little change in the relative right kidney weights over the one year time span. At week 13 in the control group the mean relative weight was .26% while in the highest dose group the mean was .29%. At the end of one year, the control group's mean was .30%, while the highest dose group's mean was .31%.

Although not statistically significant, the dose effects on the relative weight of the thyroid and ovary are negative; suggesting a decreasing trend in the relative weight, this is opposite from the effect observed in the liver.

The two different models (equations 14 and 15) give similar predictions of relative organ weights. For the liver, the models predict the week 13 control group means to be 3.10% and 3.13%, in the highest dose group at week 13 the predictions are 4.05% and 4.03%. At the end of one year the model predictions are almost equal to each other, in the control group the models predict 3.34% and 3.33% and in the highest dose group the predictions are 4.29% and 4.28%. Given the similar predictions the model in equation 15 is favored because of its ability to deal with hetereocadesticity, as it models the log of the organ to body weight ratio.

3.4 Discussion and Conclusions

Using the assumption of organ weight being a constant proportion of body weight can cause "pharmacokinetic bias" when predicting amounts or concentrations. Since PBPK models attempt to describe or predict the way the body absorbs, distributes, metabolizes, and excretes chemicals or drugs; any measurement of concentration or amount of substance in an organ will be biased when using a constant proportion as the organ weight, especially in chronic exposure scenarios.

Sprague-Dawley rat strains are commonly used to evaluate adverse effects as a result of environmental exposure to toxic chemicals and PBPK models are commonly based on

animal experiments. This paper illustrates that organ weights may change disproportional to body weight over the growth period or under chronic exposure to DLCs. Thus in PBPK models physiological parameters such as organ volumes and organ weight may be affected by exposure conditions and should not be assumed to be a constant proportion of body weight. Organ volumes and adipose tissue content can substantially influence the disposition of drugs and chemicals in physiologically based pharmacokinetic models (Sotaniemi, 1992).

There is strong evidence that liver and lung weights, as a proportion of body weight, increase with the exposure level and growth age of the animals. The proportion of other organs such as the spleen to body weight also changed statistically significantly. Since organ weights are unlikely to be a constant proportion of body weight in the first year of growth or under chronic exposure condition, this non-constant relationship may have profound pharmacokinetic and pharmacodynamic implications. This is because physiological parameters such as organ blood flow rates and surface areas will be non-constant. The use of accurate physiological values in PK/PBPK models is essential for meaningful predictions of the pharmacokinetic behavior. Also the increased use of monte carlo analysis in conjunction with PBPK modeling requires that appropriate ranges of the physiological values in question be determined in order to develop valid probability distributions (lipton *et al.*, 1995) for parameters. Our results in this paper not only confirm the likelihood of organ weight change in disproportion to body weight, but also illustrate the range of relative weights when experiment subjects are exposed to mixtures of DLCs.

Allometric equations within PBPK models are often used to describe the change in cardiac output. For example in Emonds (2004) PBPK model for developmental exposures to TCDD in the rat, an equation of the form $Q_c(ml/h) = QCC \cdot 60 \cdot ((BodyWeight)/1000)^{(0.75)}$ is employed for predicting the cardiac output, where QCC $= 311.4\ ml/min/kg$. If body weight is assumed to grow independently of exposure, then in cases where exposure or therapeutic intervention modifies or inhibits the growth trajectory, cardiac output predictions will be biased. We have also shown that assuming body weight growth is *not* modified by exposure can lead to over estimating the body weight of animals in high dose groups. This can have a direct effect on cardiac output and blood flow rates. Cardiac output and blood

flow rates are important physiologic parameters in any PBPK model, they help determine the total amount and rate at which blood is moved in the organism.

Every substance has a different density and will have a different volume to weight conversion, in PBPK models the density of each organ is usually considered to be 1 (Brown *et al.*, 1997), thus no conversion from mass to volume is used and $Density = \frac{Mass}{Volume}$. Furthermore, this implies that the weight is assumed to be the volume. Consider the formula for amount of a chemical (i.e., in the Liver) with respect to concentration and volume, we have $A_L = \frac{C_L}{V_L}$. If we let $\hat{V}_L$ be an estimate of liver volume when assuming liver weight is a constant proportion of body weight and let $V_L(t,d)$ be an estimate of liver volume based on a model using body weight that includes exposure and time as modifiers of the relationship, then in the context of the observed results of the NTP study (the relative weight of the liver increases with exposure), using $\hat{V}_L$ would cause over-estimation of A_L for a given C_L since $\hat{V}_L < V_L(t,d)$.

This paper has demonstrated that the assumption of organ weights being a constant proportion of body weight is not valid and can possibly bias estimates of concentrations or amounts of chemicals when using PBPK models. We have also demonstrated that growth in SD rats is slowed by exposure to the dioxin and dioxin like compounds and that exposure effects must be incorporated into the growth equations of PBPK models in hopes of eliminating bias and improving accuracy.

3.5 Tables

Table 7: Sampling scheme of the 2-year NTP study of TCDD mixture.

	Administered Dose Level ($ng\ TEQ\ /\ kg$)				
Week	Control	10	22	46	100
14	10	10	10	10	10
30	10	10	10	10	10
52	8	8	8	8	8
104	8	10	10	10	8

Table 8: Observed mean (SD) of the organ and body weights (g) by time and dose.

		Administered Dose Level ($ng\ TEQ\ /\ kg$)				
Organ	Week	Control	10	22	46	100
Whole Body	0	179.11(9.98)	179.96(7.40)	181.39(9.94)	181.33(7.88)	182.12(6.41)
	13	298.77(23.61)	280.44(20.59)	281.99(12.98)	277.42(17.93)	259.83(19.21)
	30	290.47(18.96)	297.00(14.71)	299.50(21.06)	293.79(25.23)	294.65(17.86)
	52	323.94(13.08)	307.86(24.70)	316.35(25.13)	320.05(30.45)	287.46(14.26)
	104	360.75(54.91)	380.47(68.73)	365.51(66.36)	329.43(58.92)	287.01(47.05)
Liver						
	13	9.33(0.83)	10.07(0.87)	9.78(1.06)	10.11(1.01)	10.72(1.16)
	30	8.94(1.19)	10.69(0.86)	10.65(0.71)	11.56(1.63)	11.99(1.34)
	52	10.38(0.79)	11.27(1.10)	11.30(1.16)	14.09(2.00)	12.42(1.36)
Lung						
	13	1.83(0.27)	1.86(0.31)	1.85(0.23)	1.93(0.35)	1.85(0.16)
	30	1.91(0.33)	1.89(0.28)	2.07(0.22)	2.12(0.33)	2.10(0.24)
	52	2.25(0.46)	2.12(0.23)	2.16(0.26)	2.35(0.31)	2.38(0.26)
Kidney						
	13	0.77(0.06)	0.82(0.06)	0.79(0.03)	0.77(0.04)	0.76(0.06)
	30	0.82(0.06)	0.91(0.08)	0.84(0.06)	0.86(0.08)	0.84(0.05)
	52	0.96(0.08)	0.92(0.08)	0.91(0.08)	1.01(0.07)	0.88(0.07)

Table 8 continued. Observed mean (SD) of the organ and body weights (g) by time and dose.

		Administered Dose Level (*ng TEQ / kg*)				
Organ	Week	Control	10	22	46	100
Spleen	13	0.58(0.04)	0.56(0.08)	0.56(0.07)	0.52(0.04)	0.47(0.06)
	30	0.51(0.06)	0.50(0.06)	0.51(0.04)	0.52(0.07)	0.48(0.07)
	52	0.54(0.03)	0.49(0.07)	0.50(0.05)	0.55(0.16)	0.43(0.05)
Thyroid						
	13	2.65e-02(2.68e-03)	2.47e-02(2.91e-03)	3.13e-02(6.41e-03)	2.78e-02(7.93e-03)	2.29e-02(4.61e-03)
	30	2.34e-02(4.25e-03)	2.39e-02(6.40e-03)	2.69e-02(3.57e-03)	2.41e-02(3.35e-03)	2.61e-02(3.67e-03)
	52	2.27e-02(9.18e-03)	1.85e-02(2.62e-03)	2.01e-02(3.09e-03)	1.76e-02(2.92e-03)	1.61e-02(2.10e-03)
Ovary						
	13	6.40e-02(1.08e-02)	6.00e-02(1.89e-02)	5.60e-02(1.35e-02)	5.60e-02(1.35e-02)	5.30e-02(1.64e-02)
	30	5.00e-02(8.16e-03)	5.60e-02(1.43e-02)	5.40e-02(1.35e-02)	5.90e-02(1.59e-02)	5.50e-02(8.50e-03)
	52	5.37e-02(1.19e-02)	5.12e-02(6.41e-03)	4.88e-02(6.41e-03)	6.38e-02(1.85e-02)	4.50e-02(9.26e-03)

Table 9: Observed mean (SD) of the relative organ weight[a] by time and dose (Expressed as percentage).

		Administered Dose Level (*ng TEQ / kg*)				
Organ	Week	Control	10	22	46	100
Liver	13	3.12(0.12)	3.59(0.21)	3.47(0.32)	3.64(0.24)	4.12(0.24)
	30	3.07(0.26)	3.60(0.15)	3.56(0.17)	3.92(0.32)	4.06(0.29)
	52	3.21(0.22)	3.66(0.18)	3.58(0.32)	4.40(0.45)	4.32(0.42)
Lung						
	13	0.61(0.06)	0.66(0.09)	0.66(0.07)	0.70(0.12)	0.71(0.05)
	30	0.66(0.09)	0.64(0.10)	0.69(0.08)	0.72(0.10)	0.72(0.08)
	52	0.69(0.13)	0.69(0.09)	0.70(0.16)	0.74(0.10)	0.83(0.11)
Kidney						
	13	0.26(0.02)	0.29(0.01)	0.28(0.01)	0.28(0.02)	0.29(0.01)
	30	0.28(0.02)	0.31(0.03)	0.28(0.02)	0.29(0.02)	0.29(0.01)
	52	0.30(0.02)	0.30(0.03)	0.29(0.02)	0.32(0.05)	0.31(0.02)

Table 9 continued. Observed mean (SD) of the relative organ weight[a] by time and dose (Expressed as percentage).

		Administered Dose Level (*ng TEQ / kg*)				
Organ	Week	Control	10	22	46	100
Spleen	13	0.19(0.01)	0.20(0.02)	0.20(0.02)	0.19(0.01)	0.18(0.01)
	30	0.18(0.02)	0.17(0.01)	0.17(0.01)	0.18(0.02)	0.16(0.02)
	52	0.17(0.01)	0.16(0.02)	0.16(0.01)	0.17(0.05)	0.15(0.02)
Thyroid						
	13	8.80e-03(1.03e-03)	8.80e-03(1.03e-03)	1.12e-02(2.30e-03)	1.02e-02(3.22e-03)	9.00e-03(1.49e-03)
	30	7.90e-03(1.52e-03)	8.00e-03(2.31e-03)	8.90e-03(1.37e-03)	8.20e-03(1.4e-03)	9.00e-03(1.63e-03)
	52	6.88e-03(2.70e-03)	6.25e-03(1.04e-03)	6.25e-03(1.04e-03)	5.62e-03(9.2e-04)	5.50e-03(7.60e-04)
Ovary						
	13	2.10e-02(3.16e-03)	2.00e-02(8.16e-03)	2.00e-02(6.67e-03)	2.10e-02(5.68e-03)	2.10e-02(5.68e-03)
	30	1.90e-02(3.16e-03)	2.00e-02(4.71e-03)	2.10e-02(3.16e-03)	2.00e-02(4.71e-03)	1.90e-02(3.16e-03)
	52	1.75e-02(4.63e-03)	1.87e-02(3.54e-03)	1.63e-02(5.18e-03)	2.00e-02(7.56e-03)	1.50e-02(5.35e-03)

Table 10: Regression coefficients and statistics for Nonlinear mixed effects body weight model.

Term	Estimate	Std. Error	t value	Pr($>\|t\|$)	SD of random effect estimates
Wint	185.22	1.36	136.11	<1.00e-04	9.89e-02
V_a	-0.20	0.01	-30.38	<1.00e-04	-
V_b	-9.81e-03	1.00e-03	-21.25	<1.00e-04	-
K_a	1.86	0.07	24.92	<1.00e-04	4.89e-03
K_b	-0.02	2.8e-03	-6.65	<1.00e-04	4.07e-05

$\sigma_j = \exp(\beta \cdot t_j)$ with the estimate $\hat{\beta} = 0.012$.

Table 11: Estimates of V_{max} and K_m by dose level.

Administered Dose level *ng TEQ / kg bw*	V_{max}	K_m
Control	0.82	6.33
10	0.79	5.95
22	0.76	5.52
46	0.70	4.76
100	0.59	3.41

Table 12: Regression statistics of organ weight models.

Organ	Term	Estimate	Std. Error	t value	Pr($>\|t\|$)
Liver (quadratic model)					
	β_0	3.02e-02	6.91e-04		
	$\beta_1(Dose)$	2.28e-04	2.81e-05	8.11	2.63e-13***
	$\beta_2(Week)$	6.23e-05	1.64e-05	3.79	2.26e-04***
	$\beta_3\left(Dose^2\right)$	-1.33e-06	2.66e-07	-5.00	1.73e-06***
Lung					
	β_0	6.12e-03	1.93e-04		
	$\beta_1(Dose)$	9.57e-06	2.26e-06	4.23	4.28e-05
	$\beta_2(Week)$	1.43e-05	5.02e-06	2.85	5.10e-03
Kidney					
	β_0	2.69e-03	4.56e-05		
	$\beta_1(Dose)$	9.55e-07	5.36e-07	1.78	7.72e-02
	$\beta_2(Week)$	5.26e-06	1.19e-06	4.42	1.96e-05
Spleen					
	β_0	2.03e-03	4.19e-05		
	$\beta_1(Dose)$	-1.08e-06	4.93e-07	-2.20	2.95e-02
	$\beta_2(Week)$	-7.62e-06	1.09e-06	-6.97	1.23e-10
Thyroid					
	β_0	1.09e-04	3.56e-06		
	$\beta_1(Dose)$	-1.82e-08	4.19e-08	-0.43	6.67e-01
	$\beta_2(Week)$	-8.78e-07	9.29e-08	-9.45	1.23e-16
Ovary					
	β_0	2.18e-04	8.32e-06		
	$\beta_1(Dose)$	1.32e-08	9.77e-08	0.14	8.93e-01
	$\beta_2(Week)$	-9.74e-07	2.17e-07	-4.49	1.49e-05

$$Y_{ijk} = \beta(D_i, t_j)W_{ijk}$$

Table 13: Means (SE) of estimated relative weight based on organ weight regression models

		Administered Dose Level (*ng TEQ / kg*)				
Organ	Week	Control	10	22	46	100
Liver (quadratic model)						
	13	3.10e-02(5.52e-04)	3.33e-02(4.37e-04)	3.54e-02(4.47e-04)	3.87e-02(5.70e-04)	4.05e-02(6.74e-04)
	30	3.21e-02(4.64e-04)	3.43e-02(3.15e-04)	3.64e-02(3.25e-04)	3.97e-02(4.78e-04)	4.16e-02(6.03e-04)
	52	3.34e-02(5.77e-04)	3.57e-02(4.61e-04)	3.78e-02(4.65e-04)	4.11e-02(5.80e-04)	4.29e-02(6.91e-04)
Lung						
	13	6.30e-03(1.44e-04)	6.40e-03(1.33e-04)	6.52e-03(1.25e-04)	6.75e-03(1.25e-04)	7.26e-03(1.93e-04)
	30	6.55e-03(1.10e-04)	6.64e-03(9.58e-05)	6.76e-03(8.35e-05)	6.99e-03(8.37e-05)	7.50e-03(1.69e-04)
	52	6.86e-03(1.51e-04)	6.96e-03(1.40e-04)	7.07e-03(1.32e-04)	7.30e-03(1.32e-04)	7.82e-03(1.98e-04)
Kidney						
	13	2.76e-03(3.41e-05)	2.77e-03(3.16e-05)	2.78e-03(2.95e-05)	2.81e-03(2.96e-05)	2.86e-03(4.57e-05)
	30	2.85e-03(2.61e-05)	2.86e-03(2.27e-05)	2.87e-03(1.98e-05)	2.90e-03(1.98e-05)	2.95e-03(4.00e-05)
	52	2.97e-03(3.57e-05)	2.98e-03(3.32e-05)	2.99e-03(3.13e-05)	3.01e-03(3.13e-05)	3.06e-03(4.68e-05)

Table 13 continued. Means (SE) of estimated relative weight based on organ weight regression models.

		Administered Dose Level (*ng TEQ / kg*)				
Organ	Week	Control	10	22	46	100
Spleen						
	13	1.93e-03(3.13e-05)	1.92e-03(2.90e-05)	1.90e-03(2.72e-05)	1.88e-03(2.72e-05)	1.82e-03(4.20e-05)
	30	1.80e-03(2.40e-05)	1.79e-03(2.09e-05)	1.77e-03(1.82e-05)	1.75e-03(1.82e-05)	1.69e-03(3.68e-05)
	52	1.63e-03(3.28e-05)	1.62e-03(3.06e-05)	1.61e-03(2.88e-05)	1.58e-03(2.88e-05)	1.52e-03(4.30e-05)
Thyroid						
	13	9.72e-05(2.66e-06)	9.70e-05(2.47e-06)	9.68e-05(2.31e-06)	9.64e-05(2.31e-06)	9.54e-05(3.57e-06)
	30	8.23e-05(2.04e-06)	8.21e-05(1.77e-06)	8.19e-05(1.55e-06)	8.14e-05(1.55e-06)	8.05e-05(3.13e-06)
	52	6.30e-05(2.79e-06)	6.28e-05(2.60e-06)	6.26e-05(2.45e-06)	6.21e-05(2.45e-06)	6.11e-05(3.65e-06)
Ovary						
	13	2.05e-04(6.21e-06)	2.05e-04(5.75e-06)	2.06e-04(5.39e-06)	2.06e-04(5.40e-06)	2.07e-04(8.34e-06)
	30	1.89e-04(4.76e-06)	1.89e-04(4.14e-06)	1.89e-04(3.61e-06)	1.89e-04(3.61e-06)	1.90e-04(7.30e-06)
	52	1.67e-04(6.50e-06)	1.67e-04(6.06e-06)	1.68e-04(5.71e-06)	1.68e-04(5.71e-06)	1.69e-04(8.52e-06)

Table 14: Regression statistics for organ weight relative to body weight.

Organ	Term	Estimate	Exp(Est)	Std. Error	t value	Pr($>\|t\|$)
Liver (quadratic model)						
	β_0	-3.49	3.06e-02	1.82e-02		
	$\beta_1(Dose)$	6.11e-03	1.01	7.52e-04	8.13	2.30e-13
	$\beta_2(Week)$	1.59e-03	1.00	4.37e-04	3.63	4.00e-04
	$\beta_3\left(Dose^2\right)$	-3.59e-05	1.00	7.04e-06	-5.10	1.13e-06
Lung						
	β_0	-5.10	6.09e-03	2.61e-02		
	$\beta_1(Dose)$	1.42e-03	1.00	3.05e-04	4.65	7.72e-06
	$\beta_2(Week)$	2.18e-03	1.00	7.00e-04	3.11	2.26e-03
Kidney						
	β_0	-5.91	2.70e-03	1.51e-02		
	$\beta_1(Dose)$	3.45e-04	1.00	1.76e-04	1.96	5.17e-02
	$\beta_2(Week)$	1.78e-03	1.00	4.03e-04	4.42	2.01e-05
Spleen						
	β_0	-6.19	2.05e-03	2.23e-02		
	$\beta_1(Dose)$	-6.95e-04	1.00	2.60e-04	-2.67	8.50e-03
	$\beta_2(Week)$	-4.55e-03	1.00	5.98e-04	-7.60	4.10e-12
Thyroid						
	β_0	-9.09	1.13e-04	4.12e-02		
	$\beta_1(Dose)$	-2.75e-04	1.00	4.80e-04	-0.57	5.68e-01
	$\beta_2(Week)$	-1.15e-02	1.00	1.10e-03	-10.50	3.69e-19
Ovary						
	β_0	-8.45	2.13e-04	4.22e-02		
	$\beta_1(Dose)$	-5.83e-05	1.00	4.92e-04	-0.12	9.06e-01
	$\beta_2(Week)$	-4.97e-03	1.00	1.13e-03	-4.39	2.21e-05

$$\log \frac{Y_{ijk}}{W_{ijk}} = \beta_0 + \beta_1 \cdot D_i + \beta_2 \cdot t_j$$

Table 15: Estimated means (SE) of relative organ weights.

		Administered Dose Level (*ng TEQ / kg*)				
Organ	Week	Control	10	22	46	100
Liver (quadratic model)						
	13	3.13e-02(1.46e-02)	3.32e-02(1.13e-02)	3.52e-02(1.14e-02)	3.84e-02(1.48e-02)	4.03e-02(1.69e-02)
	30	3.21e-02(1.25e-02)	3.42e-02(8.40e-03)	3.61e-02(8.60e-03)	3.95e-02(1.27e-02)	4.14e-02(1.51e-02)
	52	3.33e-02(1.57e-02)	3.54e-02(1.27e-02)	3.74e-02(1.28e-02)	4.09e-02(1.59e-02)	4.28e-02(1.79e-02)
Lung						
	13	6.27e-03(1.95e-02)	6.36e-03(1.80e-02)	6.47e-03(1.67e-02)	6.69e-03(1.65e-02)	7.22e-03(2.55e-02)
	30	6.50e-03(1.54e-02)	6.60e-03(1.34e-02)	6.71e-03(1.16e-02)	6.94e-03(1.13e-02)	7.50e-03(2.24e-02)
	52	6.82e-03(2.16e-02)	6.92e-03(2.03e-02)	7.04e-03(1.92e-02)	7.28e-03(1.90e-02)	7.86e-03(2.71e-02)
Kidney						
	13	2.77e-03(1.12e-02)	2.78e-03(1.04e-02)	2.79e-03(9.65e-03)	2.81e-03(9.53e-03)	2.87e-03(1.47e-02)
	30	2.85e-03(8.85e-03)	2.86e-03(7.71e-03)	2.88e-03(6.70e-03)	2.90e-03(6.53e-03)	2.95e-03(1.29e-02)
	52	2.97e-03(1.25e-02)	2.98e-03(1.17e-02)	2.99e-03(1.11e-02)	3.01e-03(1.09e-02)	3.07e-03(1.56e-02)

Table 15 continued. Estimated means (SE) of relative organ weights.

		Administered Dose Level (*ng TEQ / kg*)				
Organ	Week	Control	10	22	46	100
Spleen						
	13	1.93e-03(1.67e-02)	1.92e-03(1.54e-02)	1.90e-03(1.43e-02)	1.87e-03(1.41e-02)	1.80e-03(2.18e-02)
	30	1.79e-03(1.31e-02)	1.78e-03(1.14e-02)	1.76e-03(9.94e-03)	1.73e-03(9.68e-03)	1.67e-03(1.92e-02)
	52	1.62e-03(1.85e-02)	1.61e-03(1.73e-02)	1.59e-03(1.64e-02)	1.57e-03(1.62e-02)	1.51e-03(2.32e-02)
Thyroid						
	13	9.70e-05(3.08e-02)	9.68e-05(2.84e-02)	9.65e-05(2.64e-02)	9.59e-05(2.61e-02)	9.45e-05(4.01e-02)
	30	7.98e-05(2.42e-02)	7.96e-05(2.11e-02)	7.93e-05(1.83e-02)	7.88e-05(1.78e-02)	7.77e-05(3.54e-02)
	52	6.20e-05(3.41e-02)	6.18e-05(3.20e-02)	6.16e-05(3.02e-02)	6.12e-05(2.99e-02)	6.03e-05(4.27e-02)
Ovary						
	13	2.00e-04(3.15e-02)	2.00e-04(2.91e-02)	2.00e-04(2.71e-02)	1.99e-04(2.67e-02)	1.99e-04(4.11e-02)
	30	1.84e-04(2.48e-02)	1.84e-04(2.16e-02)	1.84e-04(1.88e-02)	1.83e-04(1.83e-02)	1.83e-04(3.62e-02)
	52	1.65e-04(3.50e-02)	1.65e-04(3.28e-02)	1.65e-04(3.10e-02)	1.64e-04(3.07e-02)	1.64e-04(4.38e-02)

3.6 Figures

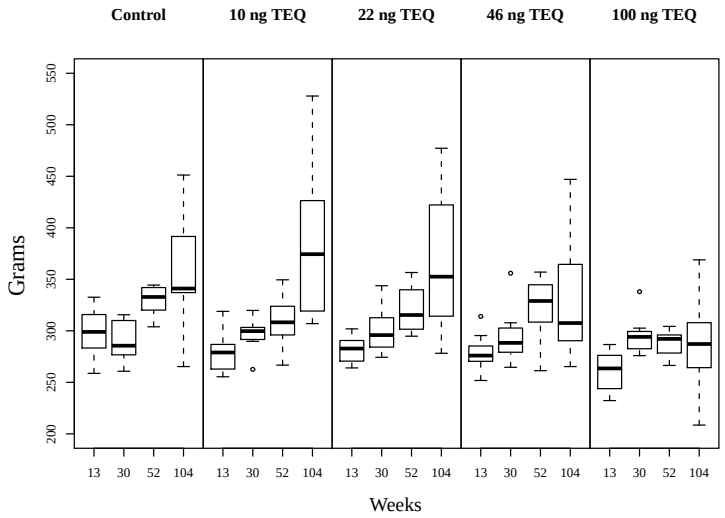

Figure 5: Boxplots of weight at time of sacrifice (by dose and time).

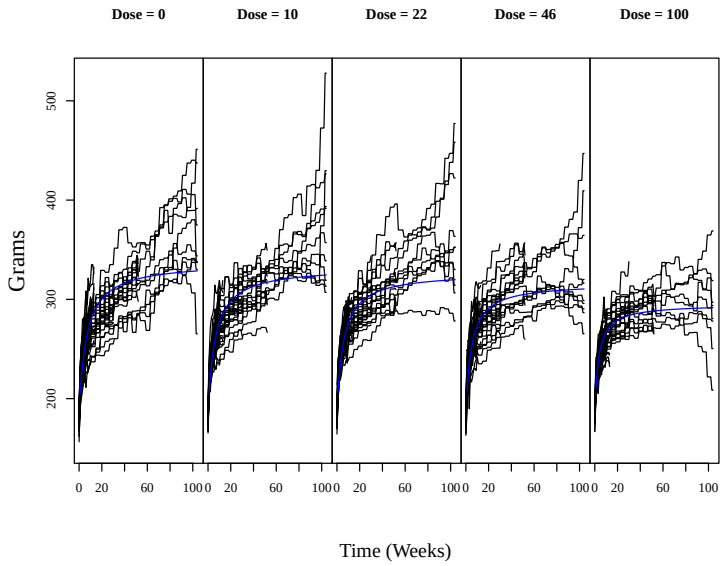

Figure 6: Weights over time.

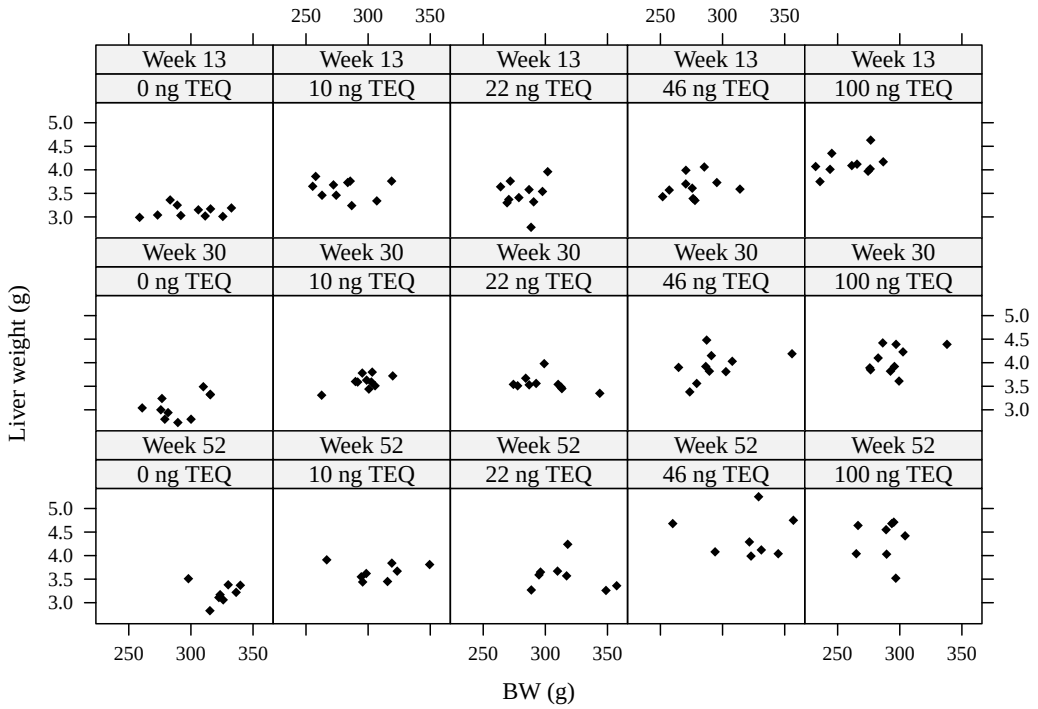

Figure 7: Scatter plot of relative liver weights (by dose and time).

4 Statistical Estimation of PBPK Model Parameters.

4.1 Introduction

Statistical estimation of parameters in physiologically based pharmacokinetic (PBPK) models is a relatively new area of research. Over the past decade or so PBPK models have become important and valuable tools in risk assessment as these models are used to describe the absorption, distribution, metabolism, and excretion of xenobiotics in a biological system such as the human or rat. Because these models incorporate information on biological processes, they are well equipped to describe the kinetic behaviors of chemicals and are useful for extrapolation across dose routes, between species, from high-to-low-doses, and across exposure scenarios (Clewell and Andersen, 1985; Anderson, 2003).

A PBPK model has been developed based on published models in the literature to describe the absorption, distribution, metabolism, and excretion of Dioxin and dioxin like compounds (DLCs) in the rat. Data from the National Toxicology Program (NTP) two year experiment TR-526 are used to illustrate model fitting and statistical estimation of the parameters. Typically, parameter estimates in these models are only presented as point estimates with no margin of error or estimate of variation. Integrating statistical methods into risk assessments is the most efficient way to characterize the variation. In this chapter a Bayesian approach via a Markov Chain Monte Carlo (MCMC) algorithm applied to a PBPK model described by a system of ordinary differential equations is used to estimate select parameters and to describe the variation of the select parameters of the system. A future aim will be to use this model and data from NTP's study to evaluate the toxic equivalency methodology with PBPK models.

Figures 8 and 9 describe the conceptual structure of the PBPK model. The model contains 4 main compartments: fat, liver, blood, and the rest of the body. The target organs for this research are the fat and liver. There are also equations for the lumen and stomach

which are used to introduce the xenobiotics into the system and the parameters in these two equations will not be estimated as there are a number of previous studies that provide information such as a stomach emptying rate and absorption rate.

4.2 Model Equations Specified by a System of ODEs

4.2.1 The Data - Doses and Amounts

The NTP of the National Institute of Environmental Health Sciences (NIEHS) conducted a series of studies to better understand the health consequences of chronic exposure to dioxin and DLCs. In one of these studies, female Harlan Sprague-Dawley rats were administered through gavage a mixture of TCDD (CAS NO. $1746-01-6$), $2,3,4,7,8$-penta chlorodibenzofuran (PeCDF) (CAS NO. $57117-31-4$), and $3,3N,4,4N,5$-pentachloro biphenyl (PCB 126) (CAS NO. $57465-28-8$), once daily 5 days a week, for up to 104 weeks (2 years). The mixtures were prepared in a corn oil:acetone (99:1) solution at four concentration levels in the ratio of one part TCDD, two parts PeCDF, and ten parts PCB-126. The dose of TCDD was 3.3, 7.3, 15.2, and 33 *ng/kg* body weight, respectively. These doses were established using the World Health Organization (WHO) toxic equivalency factor (TEF) values. The mixture study was designed so each chemical would provide a third of the total dioxin toxic equivalent (TEQ) of the mixture. These levels were also used to maximize statistical power to test for possible interactions between the chemicals. These three chemical compounds make up approximately half of the dioxin-like toxic activity found in human tissue (Walker *et al.*, 2005).

This study was designed with a chronic exposure scenario. The dosing schedule for the rats in the NTP Tr526 study was 5 times a week (Monday through Friday) for up to 104 weeks. The actual amount of chemical feed to each rat is determined by the rats body weight. There were 5 dose groups in the study including a control group, which was removed for parameter estimation. Body weights were measured for the first 13 weeks then about once every 3-4 weeks after that.

The body weights were used to calculate a dose per day and for each rat there is one dose vector of input, each element of which is a daily dose. Initial attempts to input individual

dosing vectors made computing time to predict the expected concentrations for each rat in the study far too long to be practical. To speed up simulation, the average dose by dose level and time group is used and the predicted value will be repeated as many times as necessary to match the number of observed values at each time point. There were 8 to 10 observations at each design point. To further speed up computing time the dosing data is input into the model on a weekly scale as opposed to daily. The sum of the five doses per week was used as input and this further reduced computing time to a manageable level.

Dose units are nanograms per kilogram of body weight per day (*ng/kg bw/day*) and the measured concentrations are in picograms per gram of tissue (*pg/g*) while organ and body weights are measured in grams (*g*). Multiplying the weight of the organ at the time of sacrifice (observation time) by the concentration of the chemical yields the amount. The amounts are then divided by 1000 to get the mass in nanograms, which is consistent with the units of the doses.

Figures 14, 15, and 16 show the observed amounts in the fat, liver, and blood compartments. In these figures the control groups are not included as the amounts in the control groups are negligible and they will not be considered when fitting the model as the expected outcomes are zero. The dose is increasing from right to left over the panels as well as time increasing within each panel, observed time points are 13, 30, 52, and 104 weeks. The obvious pattern is that amounts are increasing as dose level increases. Also, within each dose level, amounts in the tissues increase with time, and the variability of the amounts increase with time as well.

After an oral dose TCDD and DLCs are delivered to the blood via the lymphatic system as well as from the GI tract. The differential equations that are commonly used in PBPK models to describe the absorption of DLCS (Wang *et al.*, 1997; Emond, 2004, 2006) through the stomach and lumen are below. Fortunately, a closed form solution to this subsystem of two differential equations exists and was used in the numerical solution. The solution is found in equation 16.

$$\frac{dA_{Lum}}{dt}(ng/week) = K_{st}A_{St} - K_{abs}A_{Lum}$$

$$\frac{dA_{St}}{dt}(ng/week) = -K_{st}A_{St} + Dose$$

$$\text{at } t = 0, A_{Lum} = A_{St} = 0.$$

Where K_{abs} and K_{st} are the absorption rate and the stomach emptying rate respectively. Solving the two differential equations for the lumen and stomach we can use the equation below for the amount in the lumen.

$$A_{Lum} = \frac{Dose}{K_{abs}} - \left(\frac{Dose}{K_{abs}-K_{st}}\right) e^{-K_{st}\cdot t} + \left(\frac{Dose}{K_{abs}-K_{st}} - \frac{Dose}{K_{abs}}\right) e^{-K_{abs}\cdot t} \quad (16)$$

where A_{St} and A_{Lum} are the amounts of TCDD in the stomach and lumen.

4.2.2 Body Weight over time and Cardiac Output

Body weight over time is modeled with a type of Michaelis Menten kinetics equation. There is considerable variation in the weights especially as time increases; therefore a nonlinear mixed effect model for the weights over time with a variance function was used to predict the mean weight. Details are found in the previous chapter and the estimated regression coefficients are found in table 10.

$$BW_t = BW_{t0} + BW_{t0} \cdot \frac{V_{max}\cdot t}{K_m + t} \quad (17)$$

Body weight is a very important parameter in PBPK modeling. Cardiac output and organ volumes are often calculated using the body weight, so the consequences of a poor fit to the weight data are possibly very far reaching though-out the model. Since wasting occurs in rats exposed to Dioxin and DLCS (Tuomisto, 1999), V_{max} and K_m are functions of dose, where $V_{max} = V_a + V_b \times dose$ and $K_m = K_a + K_b \times dose$.

Once the weight is predicted then the cardiac output can be predicted using an allometric type of equation. The cardiac output is used in turn to calculate the tissue blood flows as fractions of the total cardiac output and tissue permeabilities. Again the weight of the rat plays a very important role in a PBPK model. Using a sub-system or sub-model for the rat weights including dose as a factor is an improvement over Wang's 1997 model and Emond's 2004 and 2006 models. Only time and not dose level is considered as a covariate in these previously published models.

$$Q_c(ml/week) = QCCAR \cdot 10080 \cdot ((BW_t)/1000)^{(0.75)} \quad (18)$$

where QCCAR = 311.4 *ml/min/kg*

4.2.3 Liver Compartment

The liver compartment follows the membrane influenced condition and two differential equations are used in describing the compartment. Equation 19 describes the change in the amount of chemical in the liver tissue. Equation 20 describes the change in the Liver blood amount. At time zero the amount is considered to be zero. The input to the liver tissue is the liver blood and the outputs are the elimination of the free liver concentration and output of free concentration back to the blood compartment.

In the liver tissue blood sub-compartment the inputs are the amounts from the blood, liver tissue blood, and the amount in the lumen while the outputs are the amounts in the liver blood and the free liver concentration.

Liver tissue (intra-cellular sub-compartment):

$$\frac{dA_L}{dt}(ng/week) = M_L(C_{LB} - C_{Lf}) - K_e A_{Lf} \tag{19}$$

$$C_L = \frac{A_L}{W_L}; \text{ at t = 0, } A_L = 0$$

Liver Tissue Blood (Extra-cellular sub-compartment):

$$\frac{dA_{LB}}{dt}(ng/week) = Q_L(C_B - C_{LB}) - M_L(C_{LB} - C_{Lf}) + (1-a)K_{abs}A_{Lum} \tag{20}$$

$$C_{LiB} = \frac{A_{LB}}{W_{LB}}; \text{ at t = 0, } A_{LB} = 0$$

Where C_B, C_{LB}, and C_{Lf} are the blood, liver blood, and free liver concentrations. M_L is the liver permeability and Q_L is the liver tissue blood flow rate. The empirical parameter a is the fraction of the dose entering the system circulation via the lymphatic flow and K_{abs} is the absorption rate constant of TCDD from lumen to the system circulation. K_e is the dose dependent elimination from the liver. A_{Lum}, A_L, A_{LB} are amounts of TCDD in the luman, liver, and liver blood respectively.

The free liver concentration is described by a non linear function solved by using the bisection method. The two right most terms on the right side of the equation are the amount of TCDD bound to the Ah receptor and the amount of TCDD bound to the CYP1a2 protein, respectively.

$$C_{Lf}(ng/gliver) = \frac{1}{P_L}\left[C_L - \frac{C_{AhLi}C_{Lf}}{K_{Ahd}+C_{Lf}} - \frac{C_{CYP1A2}C_{Lf}}{K_{CYP1A2d}+C_{Lf}}\right] \quad (21)$$

where C_{AhLi} is the Ah receptor level, C_{Lf} is the free liver concentration, K_{Ahd} is the TCDD-AH dissociation constant, C_{CYP1A2} is the concentration of CYP1A2 in the liver and $K_{CYP1A2d}$ is the TCDD-CYP1A2 dissociation constant. Calculation of free concentration in liver is done by numerically solving the non linear equation. More details are found in the numerical methods section. The total concentration is the sum of the specific bound and free concentrations, ie. Total concentration = Specific bound + Free concentration. The non linear components of the equation are the specific bound terms.

4.2.4 Adipose Tissue Compartment

The adipose tissue follows the membrane limited condition. Membrane-limitations are likely for large molecules, which is the case when the resistance of convection is negligible compared to the membrane transfer resistance. Two differential equations describe the fat compartment. The input into the tissue is from the fat blood and the output is the concentration in the fat divided by a partition coefficient. Elimination of TCDD from the adipose tissue does not occur.

Tissue amount (intra-cellular sub-compartment):

$$\frac{dA_F}{dt}(ng/week) = M_F(C_{FB} - \frac{C_F}{P_F}) \quad (22)$$

$$C_F = \frac{A_F}{W_F}; \text{ at t = 0}, A_F = 0$$

Fat tissue blood (extra-cellular sub-compartment):

$$\frac{dA_{FB}}{dt}(ng/week) = Q_F(C_B - C_{FB}) - M_F(C_{FB} - \frac{C_F}{P_F}) \quad (23)$$

$$C_{FB} = \frac{A_{FB}}{W_{FB}}; \text{ at t = 0}, A_{FB} = 0$$

Where C_B is the blood concentration and C_{FB} is the TCDD concentration in the adipose tissue blood. M_F is the adipose tissue permeability and Q_f is the adipose tissue blood flow rate. A_{FB} and A_F are the amounts in the fat tissue blood and fat respectively. The input into the fat tissue blood is from the blood and the amount in the fat tissue. The output is the amount in the fat blood multiplied by a permeability parameter.

4.2.5 Blood Compartment

The rate of change of the amount of chemicals in the blood is described by one differential equation. The inputs are from the lumen, fat, liver and rest of the body compartments. There is a urinary clearance term K_u which is one parameter that will be considered as linearly dose dependent, and an intercept and slope will be both be estimated.

$$\begin{aligned} \frac{d(A_B)}{dt}(ng/week) &= Q_L(C_{LB}-C_B)+Q_F(C_{FB}-C_B)+Q_R(C_{RB}-C_B) \\ &\quad +aK_{abs}A_{Lum}-K_uA_B \\ &\quad \text{at } C_B=\tfrac{A_B}{W_B}\text{; at t = 0, } A_B=0 \end{aligned} \tag{24}$$

where K_{abs} is the oral absorption rate and A_{Lum} is the amount in the lumen. K_u is the dose dependent elimination from the blood also known as urinary clearance. Q_F, Q_L, and Q_R are the blood flow rates for the adipose tissue, liver, and rest of the body. $C_{LB}, C_B, C_{FB}, C_{RB}$ are the TCDD concentrations in the liver tissue blood, blood, adipose tissue blood, and the rest of the body respectively.

4.2.6 System of ODEs

All of the equations above are solved simultaneously with the details found in the numerical methods section. The predicted amounts from the ODE system, specifically the amounts in the liver, fat, and blood are the means of the statistical model specified later. Physiological parameter descriptions, abbreviations, and values can be found in tables 16 and 17.

There are sub-equations in the liver compartment that describes the binding and induction of enzymes. These are found in appendix B. Acetanilide-4-hydroxylase ativity (Ah) activity was measured in the liver (*nmol/min/mg*) and is used in describing the process of binding with the Cytochrome P450 1A2 (CYP1A2) protein. Also ethoxyresorufin deethylase (EROD) activity was measured in the liver (*pmol/min/mg*). EROD is considered a biomarker for CYP1A1 activity and the equations describe the change in amount in the liver compartment just as Wang *et al.*(1997) and Anderson (1997). Since the amount of CYP1A2 varies with the TCDD, the specific binding to CYP1A2 also varies. This results in a non-linear binding. The change of CYP1A2 with time is described by the model for a

stimulation process proposed by Dayneka *et al.*(1993). The process of TCDD specifically bound is simplified using the equilibrium binding equation.

4.3 Methods

One characteristic of the objective is to develop a model of appropriate measure of tissue amount and examine how statistical methodology can be integrated into the modeling process. There is a need for clear strategies to incorporate statistical methodology into risk assessment and PBPK modeling and to address the uncertainty and variability in parameter estimation.

The R system and its programming language is a powerful tool used in the building, application, and analysis of dynamic models. Frequently used in combination with object orientated programming languages such as $C++$ or Fortran, R is becoming being used for model simulations in different fields such as epidemiology and microbiology.

Denoting the system of ordinary differential equations (ODE) similar to the documentation for the *simecol* package developed for the R language (Petzoldt and Rinke, 2007) we have:

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{f}(t, \mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) \\ \mathbf{y}(t) &= \mathbf{g}(t, \mathbf{x}(t), \mathbf{u}(t), \mathbf{p}) \end{aligned} \tag{25}$$

where $\mathbf{x}(t)$ is the state of the system, and $\dot{x}$ is its first derivative, t is time, $\mathbf{u}(t)$ is the input function. $\mathbf{p}$ is a vector of constant parameters. The functions $\mathbf{f}$ and $\mathbf{g}$ are the state transition function and the observation function respectively. In this setting $\mathbf{x}(t)$ is the output from the LSODA function (the ODE solver) and $\mathbf{g}$ would be the objective function such as the sum of squares.

Since there is no easy way to obtain an explicit solution to the function $\mathbf{f}$ when the dynamic system becomes large or complicated, a numeric solution is obtained by using a algorithm. Dynamic systems and continuous models require interpolation of input data especially when using ODE solvers. There are functions in the base package of R that conduct interpolation and *"approxfun"* is utilized. The function *"approxfun"* returns a

function performing (linear or constant) interpolation of the given data points. For a given set of x values, this function will return the corresponding interpolated values. The inputs are time and dose thus when the ODE solver picks up a time other then the exact dosing time there is an interpolated value that is used as input for the integration.

When using numerical integration to solve a dynamic system of ODEs over a certain time period, the method must jump from time point to the next successive time point in a discrete manner. One needs to specify starting values of the concentrations or amounts ie. the initial conditions at time t_0. Then at each point in time the value of the derivatives are estimated. These are the rates of change of the amounts and the numerical integration uses the previous values of the amounts plus the rates of change to estimate the value of the amounts at the next time point of interest. Figure 10 is a flowchart that describes the numerical integration. A compact expression of the update for each time point can be written as follows.

$$A^{t+\Delta t} = f\left(A^t, \frac{dA^t}{dt}, \Delta t\right) \tag{26}$$

where A is an amount and Δ is a very small change. Since a closed form solution of most PBPK models is intractable, a numerical solution to the system of differential equations must be utilized. When parameter estimation is attempted a numerical solution to the PBPK model will be calculated at every iteration, thus the numerical solution must be as fast as possible to avoid excessive computer usage and waiting time.

The R function *lsoda* provides an interface to the Fortran ODE solver of the same name and the system of ODEs is written in R. The solver is only one component of the simulation step. The *simecol* package has been developed to give users an interactive environment to implement, modify, and simulate dynamic models without writing long and complex programs. The package uses an object oriented approach, which provides a flexible way to implement simulation models of different types.

4.3.1 Calculation of free concentration in liver

In PBPK models the system of nonlinear ODEs can be very complex, involving numerous equations with a large number of parameters and the analytical solutions are not

available. This can happen because many biological phenomenons are described with non linear terms such as binding. When modelers encounter these types of situations a numerical method must be used.

Since there is at least one non-linear function in **f** another numerical algorithm must be used to solve it. The calculation of the free concentration in the liver is defined as Free concentration = Total concentration - Specific bound.

$$C_{Lif}(ng/gliver) = \frac{1}{P_L}\left[C_L - \frac{C_{AhLi}C_{Lf}}{K_{Ahd}+C_{Lf}} - \frac{C_{CYP1A2}C_{Lf}}{K_{CYP1A2d}+C_{Lf}}\right] \tag{27}$$

The root finding algorithm known as the Bisection method is used to solve the equation for the free liver concentration. We can denote it as $C_{Lf}(t|C_L(t),p)$ where C_L is the current total liver concentration and p are the kinetic parameters involved in the binding process.

1. The numerical solution can run up to the point where C_{Lf} is needed
2. Take the current liver concentration $C_L(t)$ and the specified parameters to calculate the free concentration. Use the bisection method, root will be in $(0, C_L)$.
3. Thus $C_{LiFree}(t|C_L(t),p) = C_{Lf}(t)$.
4. Use $C_{Lf}(t)$ in the binding equations and continue the numerical solution.

4.3.2 Statistical model

The statistical model can be written as 3 separate equations, one for each major compartment, liver, fat, and the blood. Each is a function of dozens of parameters, dose, and time. Each also has an error term distributed normally.

$$\begin{aligned} Y_L &= A_L(\theta,d,t)+\varepsilon_L \qquad \varepsilon_L \sim N\left(0,\sigma_L^2\right) \\ Y_F &= A_F(\theta,d,t)+\varepsilon_F \qquad \varepsilon_F \sim N\left(0,\sigma_F^2\right) \\ Y_B &= A_B(\theta,d,t)+\varepsilon_B \qquad \varepsilon_B \sim N\left(0,\sigma_B^2\right) \end{aligned} \tag{28}$$

The responses $\mathbf{Y}_{ijk} = (Y_{Lijk}, Y_{Fijk}, Y_{Bijk})^T$ are the amounts of TCDD in the liver, fat, and blood compartments, respectively, in the k^{th} rat ($k = 1, ..., n_{ij}$) of i^{th} dose group ($d_i = 3.3, 7.3, 15.2, 33$ *ng/kg/day*) at time t_j ($t_j = 13, 30, 52, 104$ weeks). Denote the mean by $\mu_{\mathbf{ij}} = (\mu_{\mathbf{Fij}}(\mathbf{d}, \mathbf{t}, \theta), \mu_{\mathbf{Lij}}(\mathbf{d}, \mathbf{t}, \theta), \mu_{\mathbf{Bij}}(\mathbf{d}, \mathbf{t}, \theta))^{\mathbf{T}}$. The error terms are a combination of measurement errors and model error prediction.

There are dozens of physiological and biochemical parameters in the model but six have been chosen to be estimated. The remaining parameters will be fixed at values found in previously published papers (Anderson, 1997; Emond, 2004, 2006; Safe, 1988; Wang 1997). A dose dependent elimination rate from the liver K_e is defined as a linear function of dose ($K_e = \alpha_e + \beta_e \times Dose$) thus both the slope and intercept will be estimated. Similarly, urinary clearance K_u is also linearly dose dependent with intercept α_u and slope β_u. The fat and liver partition coefficients will also be estimated.

4.3.3 Bayesian approach for parameter estimation via MCMC analysis

When the number of parameters to estimate in a model increases, non-identifiability problems become more and more complex; and trying to use common optimization techniques with the method of least squares is not efficient. Also the method of least squares only produces point estimates while using MCMC methods can generate estimates of the means and variances of the probability distributions of physiological parameters of interest.

MCMC methods can be used to sample values for parameters from complex probabilistic models. In a high dimensional space sampling is an important step for simulating, estimating parameters, and optimization. Prior distributions are specified for each parameter of interest and the hyperparameters of the prior distributions can be specified from published studies and literature.

In the current PBPK model there are over three dozen parameters. We will only choose a subset of them to illustrate the algorithm. We will consider estimating six key parameters grouped in two tiers. In the first tier we have elimination from the liver K_e (for all doses simultaneously as a function of dose $K_e = \alpha_e + \beta_e \times Dose$ and urinary clearance K_u, also linearly dose dependent with intercept α_u and slope β_u. These parameters help determine

the amount in the tissue and tissue blood sub-compartments of their respective compartments. They also indirectly affect the concentration in other compartments and between the tissue blood compartment and the tissue compartment.

The second tier of parameters consists of two partitions coefficients. P_F is the partition coefficient for the fat compartment and for the liver compartment, we denote it by P_L. These parameters help determine the amounts in the tissue blood and tissue sub-compartments of their respective compartments.

Most, if not all PBPK models reported in the literature for exposure to Dioxin and dioxin like compounds are for acute exposure scenarios and only a few have been used to describe a sub-chronic exposure (defined as up to a few months). This PBPK model is chronic in nature as it attempts to describe the absorption, distribution, elimination and excretion over a two year exposure period, the parameter values in the literature can only be used as initial guesses as to the true parameter values for this chronic exposure model. Since these existing parameter values are not well defined for the current problem and/or unmeasured, there is a need to estimate these values. However, there are many problems or challenges to this including non-identifiable parameters, the estimation of the variation, uncertainty in the parameters, and the effect of these uncertainties on predicted outcomes is unknown.

This PBPK model like other complex dynamic systems includes non-identifiable parameters which cannot be uniquely determined with a high degree of precision (Vajda *et al.*, 1989). Such non-identifiability is usually self inflicted by related parameters, meaning the effect of modifying the value of one parameter can be reversed or partially undone by modifying some other parameter or parameters. This type of over-parametrization is commonplace for complex models in PBPK and ecological modeling; it is nearly unavoidable (Mieleitner and Reichert, 2006). In order for the parameter estimation algorithm to converge, with estimates of decent precision, the parameter set must be identifiable.

Bayesian probabilistic methodology has been reported as a method to quantitatively address both variability and uncertainty in PBPK modeling (Jonsson and Johanson, 2003). However, this has not been done yet in chronic studies of exposure to Dioxin and DLCS with PBPK models.

Figure 11 is a flow chart that describes the process of the estimation of the model parameters. First, the numerical solution of the PBPK model is calculated, then the simulated values are compared with the observed values and a value of the objective function is realized. MCMC sampling is done to pick up a set of parameter values and then the process starts over after accepting or rejecting the new set of parameter values.
The MCMC algorithm is summarized as follows.

1. Initial chain $t = 0$ with initial values of θ^0. Where $\theta^0 = (\alpha_e^0, \beta_e^0, \alpha_u^0, \beta_u^0, P_f^0, P_{Li}^0)$

2. Pick up a proposal set of parameters θ^*. This new set θ^* is accepted as θ^{t+1} if

$$u < \frac{P(\theta^*)Q(\theta^t;\theta^*)}{P(\theta^t)Q(\theta^*;\theta^t)} \tag{29}$$

   where u is a value drawn from $U[0,1]$ and Q is a multivariate normal proposal distribution. P is the posterior distribution which is proportional to a function of the sum of squared residuals between the observed values and the model predictions with the current value of θ and the sum of squares of the differences between the current value of θ and the prior distribution means.

$$P\left(\theta^*|y,\sigma^2\right) \propto \exp\left(-0.5\left(\frac{SS(\theta^*)}{\sigma^2} + SS_{piror}(\theta^*)\right)\right)$$

   with

$$SS(\theta^*) = \sum_{ijk}(\mathbf{y}_{ijk} - \mu_{ij})^T \Sigma_{ij}^{-1}(\mathbf{y}_{ijk} - \mu_{ij})$$

   and

$$SS_{piror}(\theta^*) = \sum_i \left(\frac{\theta^* - \theta^0}{|\theta^0|/10}\right)$$

   If this new set of parameter values θ^* is not accepted stay at the same place and $\theta^{t+1} = \theta^t$.

3. Then return to the numerical solution to the PBPK model using as parameter values θ^{t+1}.

The MCMC sampler was implemented using the flexible modeling environment (FME) package written in the R statistical software language (Soetaert and Petzoldt, 2010). The

MCMC scheme for drawing samples from the posterior distributions of parameters to be estimated in the models was obtained by iterating the algorithm above. After collecting the final samples we drew statistical inference for the parameters θ. After 38,000 burn in iterations every 10th sample was retained out of the next 50,000. The number of iterations was stopped at 88,00 total due to the traceplots indicating convergence in all six parameters.

All prior distributions for the parameters we estimated are normal with the standard deviation being set at 10% of the mean. Details are shown in table 18. Furthermore, all the prior distributions for the inverse of error variances are gamma distributions denoted as $P\left(\sigma_i^{-2}|y,\theta\right) \sim \Gamma\left(\frac{n}{2}, \frac{SS(\theta)}{2}\right)$ where i is the index of parameters.

To conduct a sensitivity analysis additional MCMC scenarios were run to estimate parameters when fixing selected parameters at different values, such as body fat composition, bioavailability, and cardiac output. The results are in the next section. The above algorithm is used but one difference is that the standard deviations of the prior distribution of each parameter is now 50% of the mean for the intercepts and partition coefficients and 25% of the mean for the slopes of the dose dependent elimination terms K_e and K_u.

4.4 Results

MCMC analysis allows for advantages in PBPK modeling, including but not limited to the use of experimental data and results from previous published efforts as starting values or prior distributions, the capability to update multiple variables simultaneously; and the ability to separately consider parameter variability (i.e., population heterogeneity) and uncertainty (i.e., lack of information, which can be addressed by collection of additional data).

4.4.1 Model fit and estimated parameters

Figures 12 and 13 show the traceplots and the histograms of the MCMC analysis. When the horizontal line on each plot of the trace plots is stabile we consider convergence reached. This line is the cumulative average of the parameter estimate. The histograms show the distribution of each parameter and all look normally distributed with the exception of P_L,

the liver partition coefficient. This looks to have a bimodal distribution with tails similar to a normal distribution. This is a clear example of parameter non identifiability, there are two modes in this distribution with the approx values of 24.8 and 25.5. Usually in the literature partition coefficients are reported as whole numbers and the mean of this distribution is 25.25 so we can use a value of 25 when reporting this parameter.

Table 18 displays the means and standard deviations of the prior distributions of the MCMC parameters and table 19 displays the estimated posterior means and associated 95% CIs. The mean of the intercept and slope of the elimination term in the liver are 2.536 and -0.00106 respectively. After transformation $exp(intercept + dose \times slope)$, this gives values of 12.42, 12.16, 11.66, and 10.61 for each dose group respectively. The negative slope indicates we have a decreasing elimination rate from the liver. The range of the slope of the liver elimination is -1.42e-03 to -7.60e-04 and the 95% CI for the slope is (-1.27e-03, -8.70e-04). Since zero is not included we can conclude the slope is statistically different from zero and the elimination from the liver slows as the dose increases. This is consistent with Emonds (2004) PBPK model where the half life is increasing with increasing exposure.

The estimated values of the intercept and slope of the urinary clearance term are 0.5189 and 0.0001. These give values of 1.683, 1.686, 1.693, and 1.708 for the four dose groups respectively. These are very close to the point estimate reported by Emond (2004) for the use in a PBPK model, Emond reports a value of .01 *1/hr* which in terms of weeks is 1.68 *1/week*. The positive slope indicates we have an increasing urinary clearance. The range of the slope of the urinary clearance in the MCMC estimates is 6.19e-05 to 1.38e-04 and the 95% CI for the slope is (8.08e-05, 1.19e-04). Although close to zero the lower bound is positive thus we can conclude the slope is statistically different from zero and increasing.

The mean value of the partition coefficient for the fat compartment is 200.8 with a 95% CI of (187.8, 213.9). This is higher than used in Wang's model (1997) where the value is 100. However, published partition coefficient values from different studies reveal that they can range from 80 to 260 (Murphy *et al.*, 1995). Leung *et al.*(1990a) used a value of 350. So this value of 200.5 is biologically plausible. The mean partition coefficient for the liver is 25.25 with a 95% CI of (24.2, 26.0). This is also higher than reported values in the

literature. Wang *et al.*(1997) and Emond *et al.*(2004) use a value of 6, but again this values can range in the literature from 3.3 (Murphy *et al.*, 1995) to 20 (Leung *et al.*, 1990a).

4.4.2 Sensitivity in Parameters

A number of MCMC simulations scenarios were run to estimate parameters when fixing select parameters such as body fat composition, bioavailability, and cardiac output. The parameter estimates are found in table 20. Twelve different scenarios were considered. Combinations of bioavailability at 50, 60 and 75% with percentage of body fat being either 0.069% or 0.169% (the latter is in line with a human being), and cardiac output being fixed at either 311 or 342 *ml/min/kg*.

As the value of the bioavailability increases there is an increasing trend in the intercept of the liver elimination (α_e). When cardiac output and fat percentage are held at 311 *ml/min/kg* and 16.9%, the value of α_e is 2.29, 2.49, and 2.53 for bioavailability values of 50, 66 and 75 respectively. Thus a 33% reduction in the bioavailability results in a 9.5% change in the intercept. When the cardiac output and percentage of body fat is held at those same levels the slope of the dose dependent parameter β_e is -0.0053, -0.0034, and -0.0040 for bioavailability values of 50, 66 and 75 respectively.

There is a decreasing trend in the intercept of the urinary clearance when bioavailability increases and cardiac output and fat percentage are held at 342 $ml/min/kg$ and 6.9% respectively. The value of α_u is 0.57, 0.55, and 0.49 for bioavailability values of 50, 66 and 75 respectively. When the cardiac output and percentage of body fat is held at those same levels the slope of the dose dependent parameter β_u stays constant to four decimal places.

The values of the partition coefficients for the liver (P_l) in the different scenarios is constant and the value of P_f is consistent but different when the fat percentage is 6.9% compared with 16.9%. When fat percentage is 6.9%, P_f varies between 194 and 196; but when fat percentage is 16.9%, P_f varies between 78 and 80. This difference in the P_f is expected as both of these parameters are in the equations of the fat compartment and both directly influence the amount of chemicals in the compartment.

A few notes must be made about the limitations of the simulations and the parameter estimates presented. First, the length of the runs for each MCMC is limited, due to excessive computation time a maximum of $140,000$ was obtained. It is difficult to declare convergence of the parameter estimates with relatively short runs when there is still large variability in the trace plots. The exception are the trace plots for the partition coefficients, these show a stabile cumulative mean. This may be attributed to relatively wide prior distributions as specified by standard deviations being 50% for the intercepts of the dose dependent parameters and the partition coefficients, while the standard deviations of the slopes of the dose dependent parameters is 25%. The acceptance rate range from between 22.3% and 29.8% for the twelve different scenarios.

4.4.3 Uncertainty in Fixed Parameters

PBPK models often contain a large number of parameters. In order to estimate the parameters in which we do not have values for, other parameters need to be fixed. This introduces uncertainty into the model which is unavoidable. Bioavailability is a parameter for which multiple values have appeared in the literature. In Sprague-Dawley rats given just one single oral dose of 1.0 *ug [14C]-2,3,7,8-TCDD/kg bw* in a acetone:corn oil mixture, the fraction absorbed ranged from 66% to 93% (Rose *et al.*, 1976). In a study under a repeated oral dosing of rats at 1.0 *ug/kg/day* for 5 days per week for seven weeks, the gastrointestinal absorption of TCDD was observed to be approximately that observed for a single oral exposure (Rose *et al.*, 1976). Yet oral exposure of Sprague-Dawley rats to a larger dose of TCDD in acetone:corn oil (50 *ug/kg*) resulted in an average absorption of 70% of the administered dose (Piper *et al.*, 1973) also Diliberto *et al.*(1996) reported an 88% absorption of TCDD following oral exposure.

The estimated values in Wang's 1997 study showed that the estimated fraction of TCDD absorption was 90, 85, 80, 80, 70, and 70% following a single exposure to a dose from 0.01, 0.1, 0.3, 1.0, 10.0, to 30 *ug TCDD/kg bw*. On the nanogram scale, this ranges from 10 to 30,000 *ng TCDD/kg bw*. These results indicate that bioavailability may be dose dependent and decreasing as the administered dose increases. The doses in the NTP study data used

in this paper are on the nanogram scale and at the very low end, ranging from only 3.3 to 33 *ng TCDD/kg bw/day*. Given all these different published values of bioavailability uncertainty is introduced when fixing this parameter as a constant. Using this information the bioavailability for this model was set at 88% to do the parameter estimation for which the results are presented. Future analysis will include a sensitivity analysis on the bioavailability parameter and possibly estimating a dose dependent bioavailability.

The cardiac output parameter provides more uncertainty in the model. In studies with acute or sub-chronic exposure to dioxin like compounds the cardiac output may not need to vary, yet cardiac output cannot be expected to be a single fixed value for all subjects as there is evidence of variation in cardiac output due to age and gender in rats and humans. For example, average resting cardiac output is 5.6 *L/min* for a human male and 4.9 *L/min* for a female (Guyton, 2006). Delp *et al.*(1998) studied the effects of age and body composition on cardiac output in Fischer-344 rats and concluded that cardiac output changes with age. They grouped the rats by age into three categories: juvenile (2-mo-old), adult (6-mo-old), and aged (24-mo-old) rats. Cardiac output was found to be lower in juvenile rats (51 ± 4 *ml/min*) than in adult (106 ± 5 *ml/min*) or aged (119 ± 10 *ml/min*) groups. Since the NTP study is a chronic study conducted over two years starting with rats with an average age of 8 weeks a constant cardiac may not describe the biological process well.

In fact each fixed parameter adds a layer of uncertainty which cannot be avoided. Parameters such as permeabilities (as fraction of tissue blood flow) and partition coefficients for the fat, liver, and the rest of body have numerous values published yet the true values are unknown.

4.5 Discussion and Conclusion

PBPK models describe the absorption, distribution, metabolism, and excretion of xenobiotics in the body. The key factors in the absorption of Dioxin and dioxin like chemicals are the gastric non absorption rate (ie. stomach emptying rate), the oral absorption rate, and the bioavailability. The main factors of the distribution and metabolism are the cardiac output, tissue blow flows and tissue volumes. Tissue permeability and partition coefficients

also help determine how chemicals are distributed in tissue and tissue blood compartments. Finally, elimination depends on the diffusion in and out of the fat tissue, metabolism, sequestration in the liver, and elimination from the liver and urinary clearance.

Figure 14 displays the model fit in the Liver compartment. The model, with the parameter estimates found in table 19 fits the data in the liver compartment best in the two highest dose groups. The four panels from right to left display the four dose groups in increasing order. The model overestimates the early time point observations in the two lowest dose groups. Similar to the liver compartment, the model does not fit the data obtained from the fat compartment perfectly. It under estimates the amount in the lowest dose group at the last time point, 104 weeks. From figure 15 we see that the predicted curve for the highest dose group (33 *TCDD ng/kg bw*) over estimates the amount at the early time points. The model fits the second highest dose group (7.3 *ng/kg bw*) best upon visual inspection. Finally in the blood compartment the predicted values of the model describe the observed data well with the exception of the two earliest time points in the highest dose group.

There are two possible explanations for the overestimation of the amounts at the early time points. The parameter values or the structure of the equations. In most of the compartments the equations do not have a nonlinear kinetic component. In the fat and blood compartment there are only linear kinetics describing the rate in change of the amounts. Adding non linear components to the model may improve the overall fit, also using different target tissue proportions of body weight by dose and time point groups would improve model fit but both of these would increase the number of parameters necessary to estimate.

One limitation to this model is the lack of data at time points before 13 weeks and lack of information between one year and two year measurement of tissue concentrations.

The PBPK model presented in this paper is one of the first PBPK models for chronic exposure to Dioxin for a period of two years time. Previously models (Wang *et al.*, 1997; Emond, 2004) used data from studies which observation time ranged from only a few hours up to 30 weeks and are considered sub-chronic at best.

Although other published models have shown a dose-dependent elimination of TCDD (Andersen, 1993; Kohn, 2001) this is one of only a few PBPK models to present evidence using a PBPK model that kinetic parameters are dose dependent with a chronic exposure

scenario, which would be the closest to the actual type and duration of environmental exposure of such chemicals to humans.

Emond *et al.*(2006) modified their 2004 model to include an inducible elimination term in the liver based on induction of CYP1A2. Still there is a lack of understanding of the enzymes metabolizing TCDD and the role of other processes and hepatotoxicity in the pharmacokinetics of TCDD. A dose-dependent elimination of dioxins can influence exposure assessments in epidemiologic studies assessing the potential adverse health effects of dioxins and mixtures of dioxin like chemicals.

In the past epidemiologic studies that have addresses the relationship between TCDD exposure and adverse health effects using a first-order elimination rate from current measured body concentrations to back-calculate TCDD body burdens at the time of initial exposure (Crump *et al.*, 2003; Steenland *et al.*, 2001). Emond *et al.*(2005) suggest that using a model with dose-dependent elimination can result in nonlinear relationships between measured and predicted body burdens. Applying PBPK models that include dose dependent or inducible elimination rates to the epidemiologic data may result in quantitatively different relationships between exposure and adverse health effects observed in these studies.

Future research will include sensitivity analysis with the current model and testing of the current TEF methodology by applying this model to the other chemicals in the NTP study. One purpose of building this model is to test the Toxic equivalency factor method of quantifying the toxicity of a mixture of chemicals. The NTP study from which the data used in this paper was designed to generate results that could allow the testing of the TEQs. This model will be fit to the other two dioxin like chemicals in the study, PeCDF, and PCB 126 and also the TEQ of all three chemicals.

Applying the models separately to dioxin, furan, and PCB, as well as to the mixture of the three based on TEQ we can assess if the TEQ methodology holds true. For example, we predict the internal dose of the mixture (TEQ) by the PBPK model for the mixture, and call this the predicted internal dose of the mixture. Also separately we predict the internal dose of dioxin, furan, and PCB using the PBPK models for each of these compounds, and then obtain the mixture of the internal doses of the three compounds and denote this mixture as the predicted internal doses. We can then check the consistency or agreement between

the predicted internal dose of the mixture with the mixture of the predicted internal doses. While the true TEQ has to be tested against health outcome using PBPD models, this will be a first step in verification using PBPK models.

4.6 Tables

Table 16: Physiological Parameters for the PBPK model for DLCs in Rat.

Parameter	Abbreviation	Value	Reference
QACCR	*QACCR*	311.4 *ml/min/kg*	Emond (2004)
Tissue blow flows (frac of cardiac output)			
Total	Q_c	*ml/week*	Calculated
Liver	Q_L	.1830	Emond (2004)
Fat	Q_F	.0693	Emond (2004)
Rest of Body	Q_R	.7487	Emond (2004)
Tissue weights (fraction of Bw)			
Blood	WBo	0.0760	Wang (1997)
Liver	Wlo	0.0362	Emond (2004)
Fat	WFo	0.0690	Emond (2004)
Rest of Body	WRo	0.7290	Emond (2004)
Tissue Blood weights (fraction of organ)			
Liver	$WliBo$	0.2660	Emond (2004)
Fat	$WFBo$	0.0503	Emond (2004)
Rest of Body	$WRBo$	0.0300	Emond (2004)
Permeability (fraction of tissue blood flow)			
Fat	M_F	0.091	Emond (2004)
Liver	M_{Li}	0.350	Emond (2004)
Rest of Body	M_R	0.0298	Emond (2004)
Partition coefficient			
Fat	P_F	100	Wang (1997)
Liver	P_L	6	Wang (1997)
Rest of body	P_R	1.5	Wang (1997)
Gastric non absorption rate	K_S	0.36 *1/hr*	Roth *et al.*(1993)
Oral Absorption rate	K_{Abs}	0.20 *1/hr*	Wang (1997)
Elimination Liver	K_e	Estimating	
Urinary Clearance	K_u	0.01 *ml/hr*	Emond (2004)

Table 17: PBPK model parameters for binding and enzyme activity.

Parameter or Variable	Abbreviations	Value	Reference
CYP1a2 associated ACOH activity			
TCDD-AH Dissociation	K_{Ahd}	0.1 *nM*	Gasiewicz (1984)
TCDD-CYP1A2 Dissociation			
	$K_{CYP1A2d}$	0.1 *nM*	Gasiewicz (1984)
Ah receptor level	C_{AhLi}	.35 *nM*	Safe (1988)
ACOH Basal induction rate	K_{0ACOH}	160 *nM/hr*	Webber (1993)
Degradation rate	K_{2ACOH}	0.1 *liter/hr*	Wang (1997)
		16.8 *liter/wk*	
Maximum Induction Fold	In_{ACOH}	0.4 *nmol/g/hr*	Anderson (1997)
TCDD-Ah-DNA	IC_{ACOH}	130 *nm*	Tritscher (1992)
Hill coefficient	h	.6	Wang (1997)
CYP1a1 associated EROD activity			
Basal induction rate	K_{0EROD}	160 *nM/hr*	Webber (1993)
		1100 *pM/hr*	adjusted
Degradation rate	K_{2EROD}	0.1 liter/hr	Wang (1997)
		16.8 *liter/wk*	
Maximum Induction Fold	In_{EROD}	0.365 *nmol/g/hr*	Anderson (1997)
		365 *pmol/g/hr*	
TCDD-Ah-DNA	IC_{EROD}	130 *nm*	Anderson (1997)
		1300 *pm*	Anderson (1997)

Units converted to weeks where necessary in the model.

Table 18: Specified prior distributions and starting values.

Parameter	Abbreviation	Prior distribution	Hyper-parameters
K_e Intercept	α_e	$N(\alpha_e^0, (\lvert\alpha_e^0\rvert/10)^2)$	2.550
K_e Slope	β_e	$N(\beta_e^0, (\lvert\beta_e^0\rvert/10)^2)$	0.001
K_u Intercept	α_u	$N(\alpha_u^0, (\lvert\alpha_u^0\rvert/10)^2)$	0.518
K_u Slope	β_u	$N(\beta_u^0, (\lvert\beta_u^0\rvert/10)^2)$	0.001
Fat partition coefficient			
	P_f	$N(P_f^0, (\lvert P_f^0\rvert/10)^2)$	190
Liver partition coefficient			
	P_l	$N(P_l^0, (\lvert P_l^0\rvert/10)^2)$	26

Table 19: Estimated posterior means and 95% equal-tail credible intervals.

Parameter	Abbreviation	Mean	95% CI
K_e Intercept	α_e	2.536	(2.494, 2.578)
K_e Slope	β_e	-1.061e-03	(-1.266e-03, -8.74e-04)
K_u Intercept	α_u	0.519	(0.5092, 0.5284)
K_u Slope	β_u	1.0014e-04	(8.082e-05, 1.19e-04)
Fat partition coefficient			
	P_f	200.8	(187.8, 213.9)
Liver partition coefficient			
	P_l	25.2	(24.2 , 25.9)

Table 20: Estimated posterior means and 95% equal-tail credible intervals for sensitivity analysis.

bio	cardiac output	Fat %	α_e	β_e	α_u	β_u	P_f	P_l
75	342	0.169	2.35 (2.27, 2.47)	-3.34e-03 (-4.761e-03, -2.209e-03)	0.39 (0.27, 0.50)	1e-04 (5.00e-05, 1.50e-04)	79 (72.7, 86.1)	25 (24.3, 26.6)
		0.069	2.47 (2.36, 2.56)	-3.93e-03 (-4.62e-03, -3.258e-03)	0.49 (0.38, 0.58)	1e-04 (5.10e-05, 1.48e-04)	193 (177.7, 209.6)	25 (24.2, 26.4)
	311	0.169	2.53 (2.41, 2.65)	-4.03e-03 (-5.01e-03, -3.19e-03)	0.34 (0.21, 0.51)	1e-04 (4.90e-05, 1.48e-04)	79 (73, 85.8)	25 (24.2, 26.4)
		0.069	2.56 (2.45, 2.66)	-5.15e-03 (-5.75e-03, -4.55e-03)	0.43 (0.35, 0.5)	1e-04 (5.10e-05, 1.5e-04)	191 (175.3, 207.7)	25 (24.2, 26.4)
60	342	0.169	2.40 (2.3, 2.49)	-4.36e-03 (-5.34e-03, -3.75e-03)	0.36 (0.29, 0.48)	1e-04 (5.10e-05, 1.49e-04)	80 (73.6, 86.2)	25 (24.1, 26.1)
		0.069	2.53 (2.48, 2.61)	-5.34e-03 (-5.89e-03, -4.93e-03)	0.55 (0.47, 0.59)	1e-04 (5.00e-05, 1.46e-04)	194 (178.3, 209.1)	25 (24.1, 26.1)
	311	0.169	2.49 (2.4, 2.58)	-3.39e-03 (-4.05e-03, -2.62e-03)	0.68 (0.55, 0.85)	1e-04 (5.30e-05, 1.48e-04)	79 (72.4, 86.1)	26 (24.4, 26.7)
		0.069	2.34 (2.21, 2.47)	-4.69e-03 (-6.08e-03, -3.75e-03)	0.28 (0.19, 0.39)	1e-04 (4.9e-05, 1.46e-04)	193 (177.8, 208.4)	25 (24.2, 26.2)
50	342	0.169	2.35 (2.25, 2.43)	-6.26e-03 (-7.08e-03, -5.53e-03)	0.39 (0.34, 0.44)	1e-04 (5.10e-05, 1.5e-04)	78 (72.4, 85.1)	25 (24.1, 26.1)
		0.069	2.32 (2.16, 2.46)	-5.06e-03 (-5.89e-03, -4.33e-03)	0.57 (0.5, 0.65)	1e-04 (4.90e-05, 1.48e-04)	195 (179.9, 209.4)	25 (24, 26)
	311	0.169	2.29 (2.25, 2.33)	-5.33e-03 (-6.223e-03, -4.63e-03)	0.48 (0.4, 0.55)	1e-04 (5.20e-05, 1.47e-04)	79 (72.9, 85.7)	25 (24.1, 26.1)
		0.069	2.34 (2.26, 2.44)	-5.52e-03 (-6.45e-03, -4.86e-03)	0.50 (0.39, 0.62)	1e-04 (5.10e-05, 1.52e-04)	194 (179.4, 208.5)	25 (24.1, 26.1)

4.7 Figures

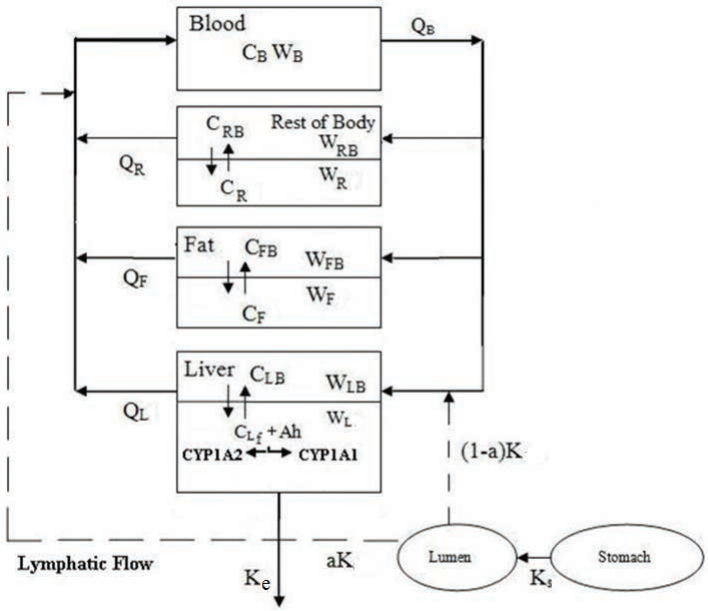

Figure 8: Conceptual PBPK model for mixture of Dioxin and DLCs in SD rat.

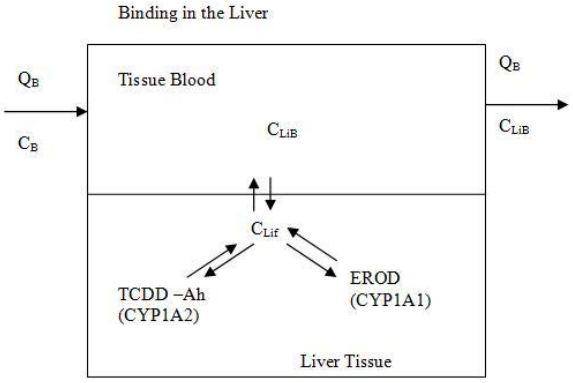

Figure 9: Specific binding processes in the liver.

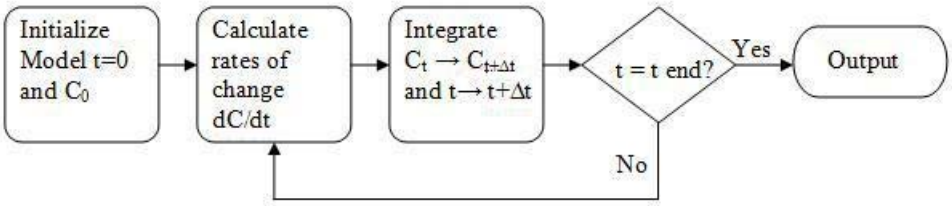

Figure 10: Flow chart for numerical integration

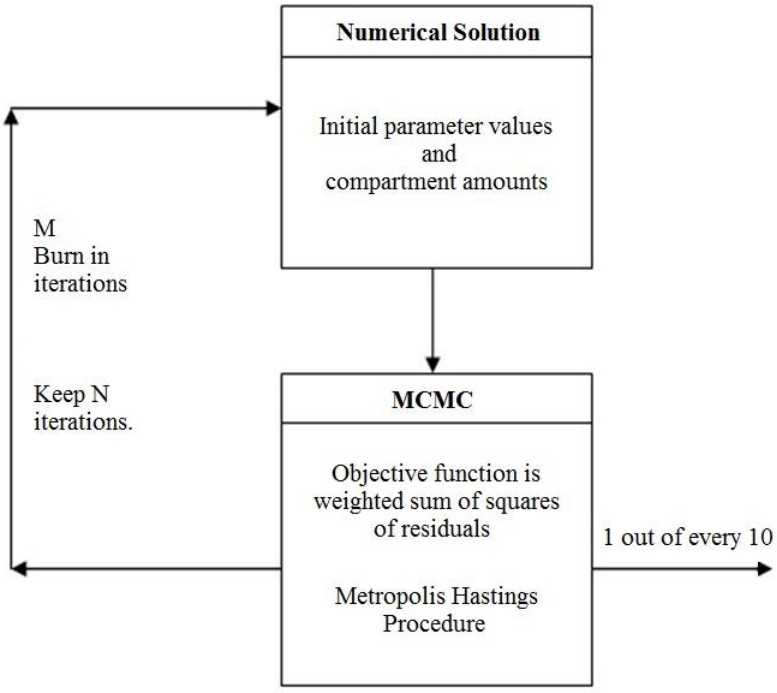

Figure 11: Flow chart computer algorithm.

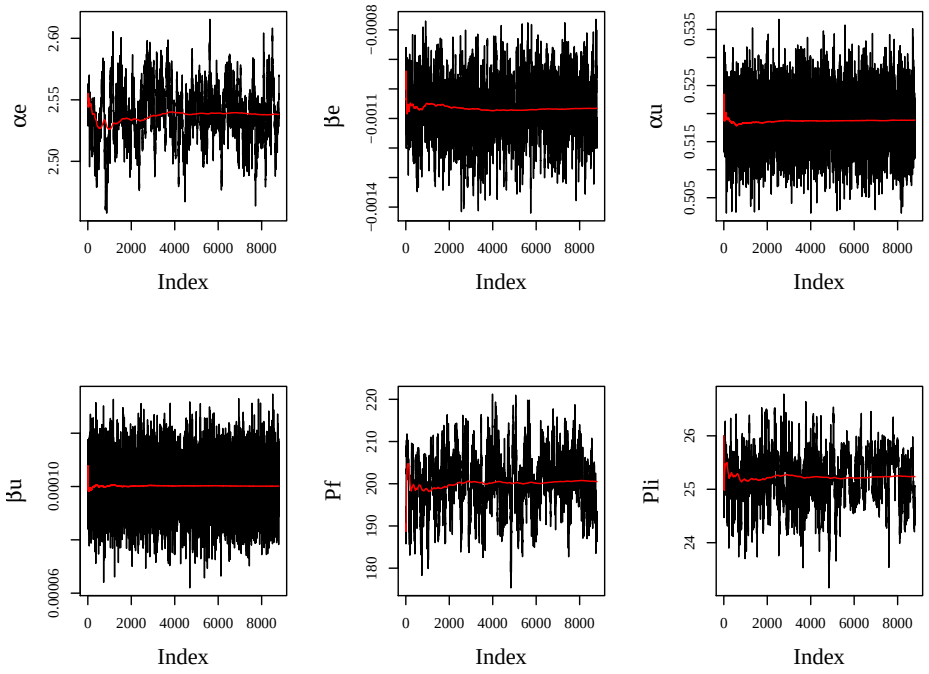

Figure 12: Traceplots of MCMC parameter estimates.

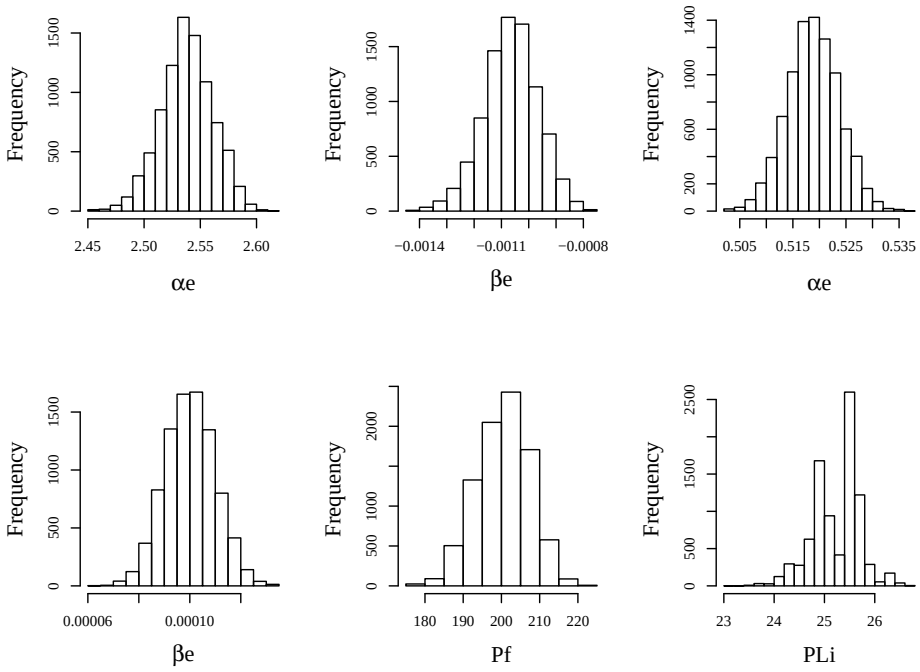

Figure 13: Histograms of MCMC parameter estimates.

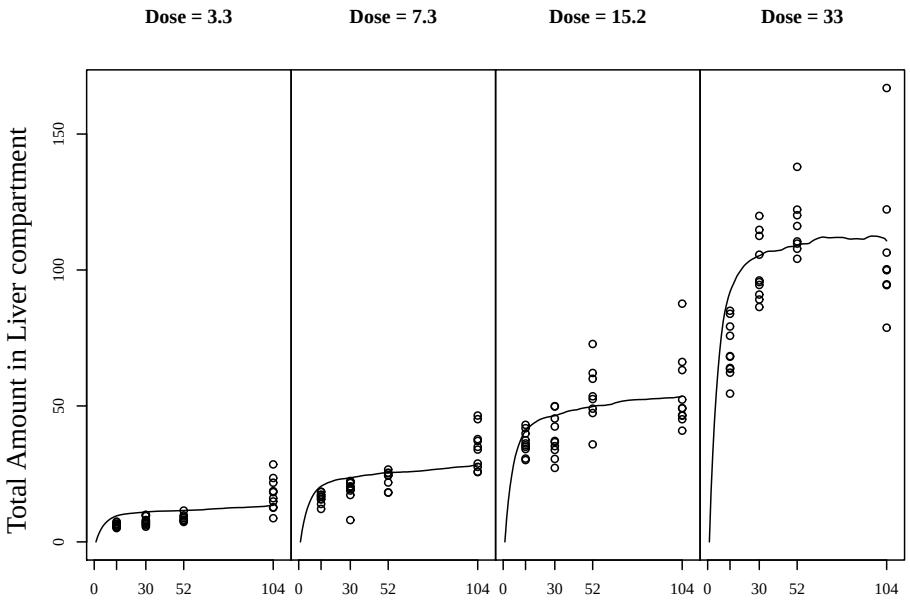

Figure 14: Model fit - Liver compartment.

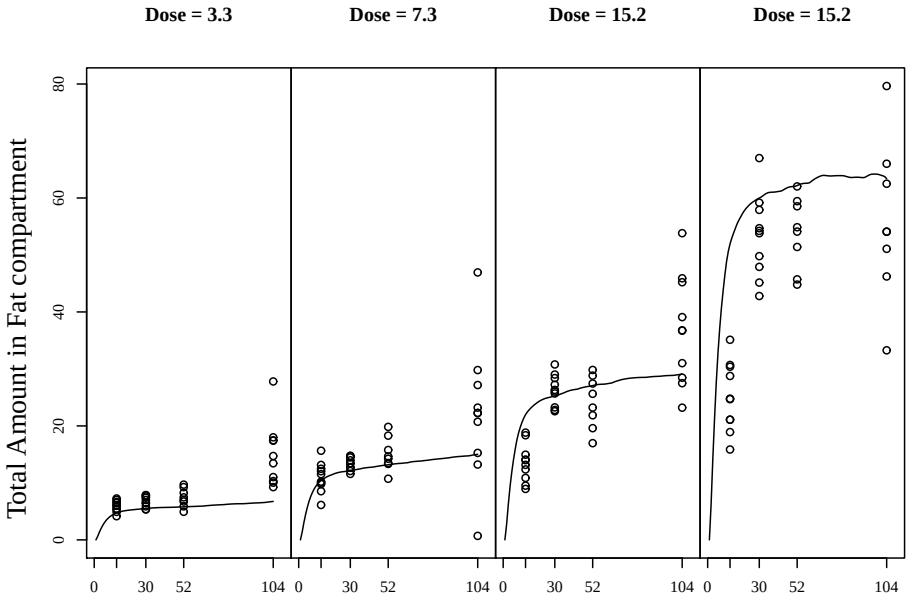

Figure 15: Model fit- Fat compartment.

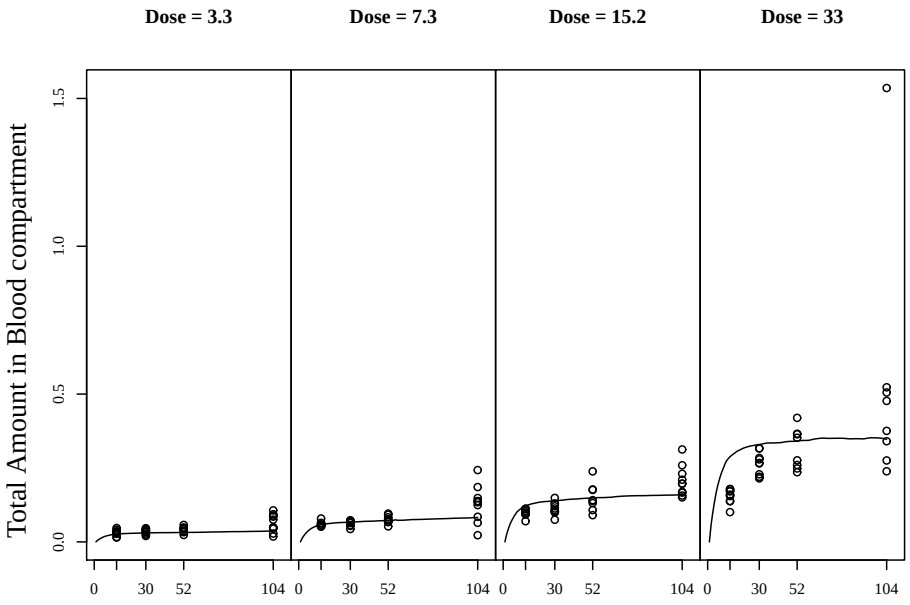

Figure 16: Model fit - Blood compartment.

5 Overall Discussion and Conclusions

In human health risk assessment, exposure assessment, especially in the context of Dioxin and DLCs, is inherently variable and uncertain as there are little or no human data to estimate the internal dose of a toxicant in a target tissue. Use of physiologically based pharmacokinetic (PBPK) models to estimate the internal dose (Anderson, 1995; Clark, 2004) has been proposed to reduce uncertainties, especially regarding interspecies, route to route, high to low dose, and acute to chronic exposure extrapolation. PBPK models quantify the internal dose in target tissues by using appropriate dose metrics and considering the process of absorption, distribution, metabolism, and excretion of chemicals that enters the body. There has been growing efforts in developing, validating, and applying PBPK models in risk assessment. PBPK models have been used to describe and study the distribution of chemicals in the body for decades but they have only been applied to toxicology and risk assessment since the 1980s when models were developed for polychlorinated biphenyls and other persistent lipophilic compounds (Anderson, 1995).

Traditionally, PBPK models are based on animal studies using a small number of subjects under acute or subchronic exposure. Because PBPK models typically involve a large number of physiological and biochemical parameters, information from published literature is often utilized to fix selected parameters as if they are known. Statistical estimation of unknown parameters is often not performed, instead mathematical simulation is relied upon. Moreover, kinetic parameters are typically assumed to be constants, free of exposure level. The first part of the dissertation research focused on a simple pharmacokinetic model to describe internal tissue (fat and liver) concentrations of dioxin (2,3,7,8-tetrachlorodibenzo-p-dioxin), a human carcinogen, using single-point time course data from a 2-year chronic exposure study. Statistical estimation was conducted via nonlinear re-weighted least squares method and the fitted models are used to examine whether the kinetic parameters are dose-dependent. The statistical modeling process is an attempt to gain insights in statistical

estimation of PBPK models with multiple tissue compartments and their application in risk assessment.

We fit the PK model using statistical approaches such as iterative re-weighted least squares and bootstrap confidence intervals to quantify the variation and uncertainty in the parameter estimates. It is evident that even in the simple case of a two compartment PK model that there are no unique solutions to the mathematical equations that describe the model. Employing weighted least squares is important as variability of the tissue concentrations increases with dose and time. Extrapolating to large PBPK models with multiple compartments and dozens of parameters will only increase the complexity of the problem.

This research fills a gap in the literature in more than one way. It explores and highlights the uncertainty in the mathematical solution of differential equations that govern PK and PBPK models, while showing that physiological parameters can vary by exposure level and should not be treated as constants during long term studies. These are two avenues of model development that need to be traveled to improve risk assessment and the modeling process that risk assessment relies upon.

In 2004 the National Research Council (NRC) reviewed the EPA's Exposure and Human Health reassessment of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) and related compounds. Recommendations were made for improvement in the quantitative approaches used to characterize risk. These recommendations included improving the way uncertainty and variability are handled and communication of the methods used to do so.

PBPK models for exposure assessment can be used to enhance quantification and reduction of uncertainties concerning inter-species, high-to-low dose, route-to-route extrapolation (Wang *et al.*, 2000; Anderson 2003). Within the context of pharmacokinetic modeling, uncertainties can be categorized as model uncertainty or parameter uncertainty. The choice of model for pharmacokinetic analysis is termed as model uncertainty as it is structural in nature. Model uncertainty is examined through analysis with alternative model assumptions or choices to give a range of outputs (Krupnick *et al.*, 2006).

In the present context, the decision to exclude a pathway or a tissue compartment in the PK model, or whether to treat some of the kinetic parameters as dose-dependent or constant are examples of model uncertainty. Our choice of a two-compartment PK model

for the TCDD tissue concentrations obviously is an over-simplification, hence it is a source of significant uncertainty and the analysis of dose-dependent kinetic parameters versus constant parameters illustrates a case of structural uncertainty.

Parameter uncertainty arises from errors or difficulties in either measuring data or applying data from the measured source. Measurement errors and population variations are among sources of parameter uncertainty. Among parameter uncertainty, some elements such as systematic bias are unknown and cannot be quantified; inherent uncertainty such as population variation (variation) can be quantified, but is irreducible. The overall parameter uncertainty is conveniently and efficiently quantified with bootstrap confidence intervals.

In this paper we have illustrated statistical estimation of multiple tissue compartment PK models, through which, we also have quantified parameter uncertainty using confidence intervals. In particular, we have investigated parameter uncertainty associated with a structural relationship - whether kinetic parameters are dose-dependent.

Nonlinear PK models make conventional methods for estimating standard errors and confidence intervals difficult and less reliable. In contrast, the bootstrap bias corrected intervals utilize the sampling distribution of the bootstrap samples, are less dependent on conventional assumptions such as symmetry, and thus are robust. The applicability of bootstrap methods in deriving confidence intervals for PBPK models is particularly meaningful as conventional approaches might be extremely difficult when PBPK models are defined by a set of differential equations without a closed-form solution.

Although the purpose of the early research is to demonstrate statistical approach to PK modeling and parameter uncertainty, the ultimate objective was to expand this approach to PBPK modeling. This required the incorporation of all relevant tissues, a large number of physiological, biochemical, and kinetic parameters. In such settings the PBPK models are typically hyper-parameterized in the sense that available data are likely insufficient for statistically estimating all parameters. Current practice is to fix selected parameters, and estimate or calibrate the remaining parameters based on mathematical simulations.

A significant part of any PBPK model is to model the size or weight of the animal, because typically, if not always, organs volumes and weights and other physiological parameters such as cardiac output are based directly on the weight of the animal.

Since there is uncertainty and variation in all parameter values there is variation in physiological parameters such as organ weight or volume and we must model the rat growth. The female Harlan Sprague-Dawley rat weight data from the 2-year carcinogenesis study conducted by the NTP/NIEHS was used to model the organ and body weight growth. Body weight growth was modeled using a Michaelis-Menten-like kinetic function in conjunction with random-effects for between-subject variation and was incorporated into the PBPK model.

The results of the weight growth modeling indicate that the estimated slopes of the V_{max} and K_m parameters are dose dependent and they decrease with increasing dose level. This is more evidence that kinetic parameters are not constant over exposure level. Also, this suggests that without incorporating exposure effects into the weight function of a PBPK model can result in over estimation of the weight, which in turn would cause other parameters dependent on the body weight to be overestimated as well. Examining the growth of an experimental subject over time should be done on a case by case basis. In this experiment exposure to the mixture of TCDD, PCB126, and PeCDF, reduced body weight growth and shortened the growth period. According to the estimates of V_{max} and K_m for the control group and the highest dose group, body weight growth was reduced by approximately 27%. Also the growth period was shortened by approximately 46%. These estimates are calculated from the percent change from the control group.

The organ weights were also modeled as a function of body weights while adjusting for animals' age and exposure level. Similar models were developed for other organs including lung, spleen, and ovary. There is strong evidence that liver weight as a proportion of body weight increased with the exposure level and growth age of the animals and thus are not constant with respect to body weight. The proportion of other organs (ie. lung, spleen) to body weight also changed statistically significantly.

The final section of the dissertation describes the development and parameter estimation of a PBPK model. An algorithm for estimating the parameters of a dynamic system described by a system of differential equations using a markov chain monte carlo method was presented, illustrated using data from the NTP study.

The PBPK model presented is one of the first for chronic exposure to Dioxin for a period of two years time. Previously models (Wang *et al.*, 1997; Emond 2004) used data from studies which observation time ranged from only a few hours up to 30 weeks and are considered sub-chronic at best.

Although other published models have shown a dose-dependent elimination of TCDD (Andersen, 1993; Kohn, 2001), the model presented is one of only a few PBPK models to present evidence that kinetic parameters are dose dependent in a chronic exposure scenario, which would be the closest to the actual type and duration of environmental exposure of such chemicals to humans.

Emond *et al.*(2006) modified their 2004 model to include an inducible elimination term in the liver based on induction of CYP1A2. Still there is a lack of understanding of the enzymes metabolizing TCDD and the role of other processes and hepatotoxicity in the pharmacokinetics of TCDD. A dose-dependent elimination of dioxins can influence exposure assessments in epidemiologic studies assessing the potential adverse health effects of dioxins and mixtures of dioxin like chemicals.

Future research will include sensitivity analysis with the current model and testing of the current TEF methodology by applying this model to the other chemicals in the NTP study. The NTP TR526 study was designed to generate results that could allow the testing of the TEQs. This model will be fit to the other two dioxin like chemicals in the study, $2,3,4,7,8$-pentachlorodibenzofuran (PeCDF) and $3,3N,4,4N,5$-pentachlorobiphenyl (PCB 126) and also the TEQ of all three chemicals.

Applying the models separately to dioxin, furan, and PCB, as well as to the mixture of the three based on TEQ we can assess if the TEQ methodology holds true. For example, we predict the internal dose of the mixture (TEQ) by the PBPK model for the mixture, and call this the predicted internal dose of the mixture. Also separately we will predict the internal dose of TCDD, PeCDF, and PCB using the PBPK models for each of these compounds and then obtain the mixture of the internal doses of the three compounds and denote this mixture as the predicted internal doses. We can then check the consistency or agreement between the predicted internal dose of the mixture with the mixture of the predicted internal

doses. While the true TEQ has to be tested against health outcome using PBPD models, this will be a first step in verification using PBPK models.

One implication of this dissertation within the context of existing methodology is that it has shown that uncertainty in the solutions to the systems of equations that make up the PBPK model should be dealt with using a range or confidence interval generated by a resampling technique as the bootstrap method or a probability distribution through markov chains. Another implication of this dissertation is that tissue compartment volumes in PBPK models should not be considered as always proportional to body weight.

Implications and plans for future research based on this dissertation include further refinement of the PBPK model, examining the difference in predictions when using different dosing schemes, for example an individual dosing vectors for input vs average dose vector by dose and time; and examining the results when simulation is on different time scales, such as days vs weeks. Computing time is a large factor/obstacle.

Long term pharmacokinetic studies designed with more frequent measures of tissue conentrations/amounts, although expensive, would be ideal for constructing PBPK models in which to test the toxic equivalence methodology. This research utilized data from only 4 time points months apart. Tissue concentrations were measured at 13, 30, 52 and 104 weeks. Since the PBPK model built in this dissertation was applied on a weekly scale, weekly observations would have provided more information for parameter estimation.

The take home message from this research should include the fact that there is a large amount of uncertainty in the current state of science of PBPK modeling and risk assessment employing PK and PBPK models. No one single model will describe the exact distribution, absorption, metabolism and elimination of TCDD and DLCS, our current level of knowledge of the molecular processes involved includes much too uncertainty.

Also from a mathematical and statistical point of view, the non identifiability of the actual parameter values used in PBPK models adds variability to the uncertainty. Only recently, over just the last decade or so, scientists have begun to realize that kinetics are not constant. This research reinforces that fact by showing elimination in a PBPK model fit to tissue amounts in a chronic exposure study change as dose (and concentration) increases.

The statistical estimation of parameters in a PBPK model built for risk assessment or for testing assumptions of methods used in risk assessment, such as the toxic equivalence method, must incorporate a multidisciplinary approach instead of a strict biological or mathematical approach. The approach must include expert knowledge of toxicology, biology, mathematics, statistics, and last but not least computer science to implement the algorithms in a convenient and efficient manner.

List of References

Agency for Toxic Substances and Disease Registry (ATSDR). (1998). Toxicological Profile for Chlorinated Dibenzo-p-Dioxins. Public Health Service, U.S. Department of Health and Human Services, Atlanta, GA.
*http://www.atsdr.cdc.gov/tfacts*104.*pdf* [Accessed August 17 2010].

Ahlborg UG, Becking GC, Brinbaum LS, Brouwer A, Derks HJGM, Feeley M, Golor G, Hanberg A, Larsen JC, Liem AKD, Safe SH, Schlatter C, Waern F, Younes M, Yrjanheikki E. (1994). Toxic Equivalancy Factors for Dioxin-Like PCBs. *Chemosphere*. **28***(6)*:1049-1067.

Andersen ME, Birnbaum LS, Barton HA, Eklund CR. (1997). Regional Hepatic CYP1A1 and CYP1A2 Induction with 2,3,7,8-Tetrachlorodibenzo-p-dioxin Evaluated with a Multicompartment Geometric Model of Hepatic Zonation. *Toxicology and Applied Pharmacology*. **144**:145155.

Andersen ME, Clewell HJ III, Gargas ML, Smith FA, Reitz RH. (1987). Physiologically based pharmacokinetics and the risk assessment process for methylene chloride. *Toxicology and Applied Pharmacology*. **87**:185-205.

Andersen ME, Mills JJ, Gargas ML, Kedderis L, Birnbaum LS, Neubert D, Greenlee WF. (1993). Modeling receptor-mediated processes with dioxin: implications for pharmacokinetics and risk assessment. *J. Risk Anal*. **13**:25-36.

Andersen ME. (1995). Development of physiologically based pharmacokinetic and physiologically based pharmacodymamic models for applications in toxicology and risk assessment. *Toxicology Letters*. **79**:35-44.

Anderson ME. (2003). Toxicokinetic modeling and its applications in chemical risk assessment. *Toxicology Letters*. **138**:9-27.

Barton HA, Chiu WA, Setzer RW, Andersen ME, Bailer AJ, Bois FY, DeWoskin RS, Hays S, Johanson G, Jones N, Loizou G, MacPhail RC, Portier CJ, Spendiff M, and Tan Y. (2007). Characterizing Uncertainty and Variability in Physiologically Based Pharmacokinetic Models: State of the Science and Needs for Research and Implementation. *Toxicological Sciences* **99***(2)*: 395-402.

Berry RM, Lutke CE, Voss RH. (1993). Ubiquitous nature of dioxins: a comparison of the dioxins content of common everyday materials with that of pulps and papers. *Environmental Science and Technology*. **27***(6)*:1164-1168.

Bertazzi PA, Bernucci I, Brambilla G, Consonni D, and Pesatori AC. (1998). The Seveso studies on early and long-term effects of dioxin exposure: a review. *Environ Health Perspect*. **106***(Suppl 2)*:625633.

Birnbaum L, DeVito MJ. (1995). Use of toxic equivalency factors for risk assessment for dioxins and related compounds. *Toxicology* **105**:391-401.

Birnbaum L. (1994). The Mechanism of Dioxin Toxicity. Relationship to Risk Assessment. *Environmental Health Perspectives Vol. 102, Supplement 9: Toxicological Evaluation of Chemical Interactions*. **102**:157-167.
Stable URL: http://www.jstor.org/stable/3838284

Bischoff KB. (1975). Some fundamental considerations of the applications of pharmacokinetics to cancer chemotherapy. *Cancer Chemother. Rep.*. **59**:777-793.

Brown RP, Delp MD, Lindstedt SL, Rhomberg LR, Beliles RP. (1997). Physiological Parameter Values for Physiologically Based Pharmacokinetic Models *Toxicol Ind Health*. **13**: 407-484.

Broyden CG. (1970). The Convergence of a Class of Double-rank Minimization Algorithms. *Journal of the Institute of Mathematics and Its Applications*. **6**:76-90.

Chen HS, Perdew GH. (1994). Subunit composition of the heteromeric cytosolic aryl hydrocarbon receptor complex. *J Biol Chem*. **269***(44)*:27554-27558.

Clark LH, Woodrow Setzer R, and Barton HA. (2004). Framework for Evaluation of Physiologically-Based Pharmacokinetic Models for Use in Safety or Risk Assessment. *Risk Analysis*. **24**:1697-1717.

Cleverly DH, Monetti M, Pilliphs L, Cramer P, Heit M, McCarthy S, O'Rourke K, Stanley J, and Winters D. (1996). A time trends study of the occurrences and levels of CDDs, CDFs, and Dioxin like PCBs in sediment cores for 11 geographically distributed lakes in the United States. *Organohalogen Compunds*. **28**:77-82.

Cleverly DH, Winters D, Ferrario J, Schaum J, Schweer G, Buchert J, Greene C, Dupuy A, Byrne C. The National Dioxin Air Monitoring Network (NDAMN): Results of the First Year of Atmospheric Measurements of CDDs, CDFs, and Dioxin-Like PCBs in Rural and Agricultural Areas of the United States: June 1998 - June 1999.

Clewell HJ III, Andersen ME. (1985). Risk assessment extrapolations and physiological modeling. *Toxicol. Ind. Health* **1**:111-131.

Consonni D, Pesatori AC, Zocchetti C, Sindaco R, D'Oro LC, Rubagotti M, Bertazzi PA. (2008). Mortality in a population exposed to dioxin after the Seveso, Italy, accident in 1976: 25 years of follow-up. *Am J Epidemiol*. **167***(7)*:847-58.

Cook RR, Townsend JC, Ott MG, and Silverstein LG. (1980). Mortality experience of employees exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) *J. Occup Med*. **22**:530-532.

Crump KS, Canady R, Kogevinas M. (2003). Meta-analysis of dioxin cancer dose response for three occupational cohorts. *Environ. Health Perspect.* **111**:681-68.

Dayneka NL, Garg V, Jusko WJ. (1993). Comparison of four basic models of indirect pharmacodynamic response. *J. Pharmacokinet. Biopharm.* **21**: 457-478.

Dedrick RL. (1973). Animal scale-up. *Journal of Pharmacokinetics and Pharmacodynamics*. **1***(5)*:435-461.

Delp MD, Evans MV, Duan C. (1998). Effects of aging on cardiac output, regional blood flow, and body composition in Fischer-344 rats. *J. Appl. Physiol.* **85**: 1813-1822.

DiCiccio TJ, Efron B. (1996). Bootstrap Confidence Intervals. *Statistical Science.* **11**:189-229.

Diliberto JJ, Jackson JA, Birnbaum LS. (1996). Comparison of 2,3,7,8-Tetrachloro dibenzo-p-dioxin (TCDD) Disposition Following Pulmonary, Oral, Dermal, and Parenteral Exposures to Rats. *Toxicology and Applied Pharmacology.* **138**: 158-168.

Efron B, Tibshiran R. (1986). Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy. *Statistical Science.* **1**:54-75.

Efron B, Tibshiran R. (1993). *An Introduction to the Bootstrap*. London: Chapman and Hall.

Emond C, Birnbaum LS, DeVito M. (2004). Physiologically based pharmacokinetic model for developmental exposures to TCDD in the rat. *Toxicol Sci.* **80**:115-133.

Emond C, Birnbaum LS, DeVito M. (2006). Use of a Physiologically Based Pharmacokinetic Model for Rats to Study the Influence of Body Fat Mass and Induction of CYP1A2 on the Pharmacokinetics of TCDD. *Environmental Health Perspectives.* **114***(9)*: 1394-1400

EPA (U.S. Environmental Protection Agency). (2000a). Characterization of Dioxins, Furans, and PCBs in Soil Samples Collected from the Denver Fort Range Area. Region VIII, U.S. Environmental Protection Agency Denver Co.

Esposito MP, Tiernan TO, Drydent FE. (1980). Dioxins. Prepared for the U.S. Environmental Protection Agency, Cincinnati, OH. EPA-600/2-80-197.

Ferrario J, Bryne C, Lorber M, Saunders P, Leese W, Dupuy A, Winters D, Cleverly D, Schaum J, Pinsky P, Deyrup C, Ellis R, and Walcott J. (1997). A statistical survey of dioxin like compounds in the United States poultry fat. *Organohalogen Compounds*. **32**:245-251.

Fiedler H, Cooper KR, Bergek S, Hjelt M, Rappe C. (1997). Polychlorinated dibenzo-p-dioxins and polychlorinated dibenzofurans (PCDD/PCDF) in food samples collected in southern Mississippi, USA. *Chemosphere*. **34**:1411-1419.

Finley BL, Conner KL, Scott PK. (2003). *Journal of Toxicology and Environmental Health, Part A*. **66**:533-550.

Fletcher R. (1970). A New Approach to Variable Metric Algorithms. *The Computer Journal*. **13**:317-322.

Freeman RA, Schroy JM. (1986). Modeling the transport of 2,3,7,8-TCDD and other low volatility chemicals in soils. *Environmental Progress*. **5(1)**:28-33

Guyton AC. Edward J. (2006). Textbook Of Medical Physiology (11th ed.). Philadelphia: Elsevier Inc. ISBN 0-7216-0240-1.

Hamilton, Hardy H. (1998). *Industrial Toxicology* 5th edition. Raymond D Harbison Eds. St. Louis: Mosby.

Haws LC, Su SH, Harris M, DeVito MJ, Walker NJ, Farland WH, Finley BF, Brinbaum LS. (2006). Development of a Refined Database of Mammalian Relative Potency Estimates for Dioxin-like Compounds. *Toxicological Sciences*. **89***(1)*: 4-30.

Hengl S, Kreutz C, Timmer J and Maiwald T (2007). Data-based identifiability analysis of non-linear dynamical models *Bioinformatics* **23***19*:2612-2618.

IARC (International Agency for Research on Cancer). (1997). Monographs on the Evaluation of Carcinogenic Risks to Humans Volume 69 Polychlorinated Dibenzo-para-Dioxins and Polychlorinated Dibenzofurans Summary of Data Reported and Evaluation.

Institute of Medicine (IOM) (2002). *Veterans and Agent Orange: update 2002*: Committee to review he health effects in Vietnam Veterans of Exposure to Herbicides. The National Academies Press, Washington DC.

International Programme on Chemical Safety (IPCS) (2008). Principles of characterizing and applying physiologically-based pharmacokinetic and toxicokinetic models in risk assessment. World Health Organization.

Jensen E, Canady R, Bolger PM. (2000). Exposure assessment for dioxins and furans in seafood and dairy products in the United States, 1998-99. *Organohalogen Compounds*. **47**:318-321.

Jobb, B, M. Uza, R. Hunsinger, K. Roberts, H. Tosine, R. Clemnt an dB. Bobbie, G. LeBel, D. Williams and B. Lau. (1990). A Survey of drinking water supplies in the Province of Ontario for Dioxins and Furans. *Chemosphere*. **20***(10-12)*:1553-1558.

Jonsson F, Johanson G. (2003). The Bayesian population approach to physiological toxicokinetictoxicodynamic modelsan example using the MCSim software. *Toxicol. Lett.* **138 (8)**: 143-150.

Kang HK, Watanabe KK, Breen J, Remmers J, Conomos MG, Stanley J and Flicker M. (1991). Dioxins and dibenzofurans in adipose tissue of US Vietnam veterans and controls. *Am J Public Health*. **81**:344-249.

King FG, Dedrick RL, Collins JM, Matthews HB, Birnbaum LS. (1983). Physiological model for the pharmacokinetics of 2,3,7,8-tetrachlorodibenzofuran in several species. *Toxicol. Appl. Pharmacol*. **67**:390-400.

Kohn MC, Lucier GW, Clark GC, Sewall C, Tritscher A, Portier CJ. (1993). A mechanistic model of the effects of dioxin on gene expression in the rat liver. *Toxicol. Appl. Pharmacol*. **120**:138-154.

Kohn MC, Walker NJ, Kim AH, Portier CJ. (2001). Physiological modeling of a proposed mechanism of enzyme induction by TCDD. *Toxicology.* **162**: 193-208.

Leung HW, Paustenbach DJ, Murray FJ, and Andersen ME. (1990). A physiological pharmacokinetic description of the tissue distribution and enzyme-inducing properties of 2,3,7,8-tetrachlorodibenzo-p dioxin in the rat. *Toxicol. Appl. Pharmacol.* **103**:399-410.

Leung HW, Poland AP, Paustenbach DJ, Andersen ME. (1989b). Dose-dependent pharmacokinetics of [125I]2-iodo-3,7,8-trichlorodibenzo-p-dioxin in mice: analysis with a physiological modeling approach. *Toxicol. Appl. Pharmacol*. **103**:399-410.

Limbird LE, Taylor P. (1998). Endocrine disruptors signal the need for receptor models and mechanisms to inform policy. *Cell*. **93(2)**:157-63.

Lin TM, Ko K, Moore RW, Buchanan DL, Cooke PS, and Peterson RE. (2001). Role of the aryl hydrocarbon receptor in development of control and 2,3,7,8-tetrachloro-dibenzo-p-dioxin-exposed male mice. *J Toxicol. Environ. Health Part A*. **64***(4)*:327-342.

Lipton J, Shaw WD, Holmes J, and Patterson A. (1995). Selecting input distributions for use in Monte Carlo simulations. *Regul. Toxicol. Pharmacol.* **21**:192-198.

Marino DJ, Clewell HJ, Gentry PR, Covington TR, Hack CE, David RM, Morgott DA. (2006). Revised assessment of cancer risk to dichloromethane: Part I Bayesian PBPK and doseresponse modeling in mice. *Regulatory Toxicology and Pharmacology*. **45**:44-54.

McCrady JK, McFarlane C, Gander LK. (1990). The transport and fate of 2,3,7,8-TCDD in soybean and corn. *Chemosphere*. **21***(3)*:359-376.

Mes J, Newsome WH, Conacher HBS. (1991). Levels of specific polychlorinated biphenyl congeners in fatty foods from five Canadian cities between 1986 and 1988. *Food Additives and Contaminants*. **8***(3)*:351-361.

Meyer C, O'Keefe P, Hilker D, Rafferty L, Wilson L, Conner S, Aldous K, Markussen K, and Slade K. (1989). A survey of twenty community water systems in New York State for PCDDs and PCDFs. *Chemosphere*. **19***(1-6)*:21-26.

Michalek JE, Pirkle JL, Needham LL, Patterson DG Jr., Caudill SP, Tripathi RC, Mocarelli P. (2002). Pharmacokinetics of 2,3,7,8-tetrachlorodibenzo-p-dioxin in Seveso adults and veterans of operation Ranch Hand *Journal of Exposure Analysis and Environmental Epidemiology.* **12**:44-53.

Mieleitner J, Reichert P. (2006). Analysis of the Transferability of a Biogeochemical Lake Model to Lakes of Different Trophic State. *Ecological Modeling*. **194**:49-61.

Murphy JE, Janszen DB, and Gargas L. (1995). An in vitro method for determination of tissue partition coefficients of non-volatile chemicals such as 2,3,7,8-tetrachlorodibenzo-p-dioxin and estradiol. *J. Appl. Toxicol.* **15**: 147-152.

National Research Council (NRC) Committee on EPA's Exposure and Human Health Reassessment of TCDD and Related Compounds. (2006). Health Risks from Dioxin and Related Compounds: Evaluation of the EPA Reassessment. Washington, DC: The National Academies Press.

National Research Council (NRC). (1983). Risk Assessment in the Federal Government: Managing the Process. Washington, DC: The National Academies Press.

National Research Council. (1983). Risk Assessment in the Federal Government: Managing the Process. National Academy Press, Washington, DC.

Neath AA, Samaniego FJ (1997). On the Efficacy of Bayesian Inference for Nonidentifiable Models. The American Statistician, **51**:225-232

NTP. NTP Technical report on the toxicology and carcinogenesis studies of a mixture of 2,3,7,8 TETRACHLORODIBENZO-p-DIOXIN (TCDD) (CAS NO. 1746-01-6),

2,3,4,7,8 PENTACHLORODIBENZOFURAN (PeCDF) (CAS NO. 57117-31-4), AND 3,3N,4,4N,5 PENTACHLOROBIPHENYL (PCB 126) (CAS NO. 57465-28-8) in Female Harlan Sprague-Dawley Rats. (2006). NIH Publication No. 06-4462. Accessible at $http://ntp.niehs.nih.gov/files/526_{W}eb_{F}inal.pdf$

Obrink KJ, Nigel M. (1992). Torsten Teorell (1905-1992). *The Physiologist* **35**:306.

Oikawa K, Ohbayashi T, Mimura J, Iwata R, Kameta A, Evine K, Iwaya K, Fujii-Kuriyama, Kuroda M, Mukai K. (2001). Dioxin suppresses the checkpoint protein, MAD2, by an aryl hydrocarbon receptor-independent pathway. *Cancer Res.* **61***(15)*:5707-5709.

Pearson RG, McLaughlin DL, and McIvenn WD. (1990). Concentrations of PCDD and PCDF in Ontario soils from the vicinity of refuse and sewage sludge incinerators and remote rural and urban locations. *Chemosphere*. **20***(10-12)*:1543-1548.

Petzoldt T, Rinke K. (2007). Simecol: An Object-Oriented Framework for Ecological Modeling in R. *Journal of Statistical Software*, **22***(9)*:1-31.

Piper WN, Rose JQ, Gehring PJ. (1973). Excretion and tissue distribution of 2,3,7,8-tetrachlorodibenzo-p-dioxin in the rat. *Environ. Health Perspectives* **5**: 241-244.

Pirkle JL, Wolfe WH, Patterson DG, Needham LL, Michalek JE, Miner JC, Peterson MR, Phillips DL. (1989). Estimates of the half-life of 2,3,7,8-tetrachlorodibenzo-p-dioxin in Vietnam veterans of operation ranch hand. *Journal of Toxicology and Environmental Health*. **27***(2)*:165-171.

Poland A, Teitelbaum P, Glover E, Kende A. (1989b). Stimulation of in vivo hepatic uptake and in vitro hepatic binding of ['251]-2-iodo-3,7,8-trichlorodibenzo-p-dioxin by the administration of agonists for the Ah receptor. *Mol. Pharmacol*. **36**:121-127.

Poland A, Teitelbaum P, Glover E. (1989a). The [125I]-2iodo-3,7,8-trichlorodibenzo-p -dioxin binding species in mouse liver induced by agonists for the Ah receptor: characterization and identification. *Mol. Pharmacol*. **36**:113-120.

Poland A. (1996). Meeting report: receptor-acting xenobiotics and their risk assessment. *Drug Metab Dispos*. **24***(12)*:1385-1388.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0,

URL: http://www.R-project.org.

Rappe C. (1992). Sources of PCDDs and PCDFs. Introduction. Reactions, levels, patterns, profiles and trends. *Chemosphere*. **25***(1-2)*:41-44.

Rose JQ, Ramsey JC, Wentzler TH, Hummel RA, Gehring PJ. (1976). The fate of 2,3,7,8-tetrachlorodibenzo-p-dioxin following single and repeated oral doses to the rat. *Toxicology and Applied Pharmacology.* **36**: 209-226.

Safe S, Bandiera S, Sawyer T, Robertson L, Safe L, Parkinson A, Thomas PE, Ryan DE, Reik LM, Levin W. (1985). PCBs: Structure-function relationships and mechanism of action. *Environ. Health Perspect.* **60**:47-56.

Safe SH. (1986). Comparative toxicology and mechanism of action of polychlorinated dibenzo-p-dioxins and dibenzofurans. *Annu. Rev. Pharmacol. Toxicol.* **26**:371-399.

Schecter A, Cramer P, Boggess K. (2001). Intake of Dioxins and related compounds from food in the U.S. population. *J. Toxicol. Environ. Health* **63**:1-18.

Schoeffner DJ, Warren DA, Muralidhara S, Brucker JV, Simmions JE. (1999). Organ Weights and Fat Volume in Rats as a Function of Strain and Age. *Journal of Toxicology and Environmental Health, Part A*. **56***(7)*:449-462.

Smith RM, O'Keefe PW, Aldous KM, Hilker DR, O'Brien JE. (1983). 2,3,7,8 Tetrachloro dibenzo-p-dioxin in sediment samples from Love Canal storm sewers and creeks. *Environ. Sci. Technol*. **17***(1)*:610.

Soetaert K, Petzoldt T. (2010). Inverse Modeling, Sensitivity and Monte Carlo Analysis in R Using Package FME. *Journal of Statistical Software.* **33***(3)*:1-28.

Sotaniemi EA, Pirttiaho HI, Stenbeck FG. (1992). Drug interaction involving changes in liver size and fibrogenesis. *Int. J. Clin. Pharmacol. Ther. Toxicology*. **11**:523-525.

Steenland K, Deddens J, Piacitelli L. (2001). Risk Assessment for 2,3,7,8-tetrachloro dibenzo-p-dioxin based on an epidemiologic study. *Am. J. Epidemiol.* **154**:451-458.

Steenland K, Deddens J. (2003). Dioxin: exposure-response analyses and risk assessment. *Ind Health*. **41***(3)*:175-180.

Thompson CM, Sonawane B, Barton HA, DeWoskin RS, Lipscomb JC, Schlosser P, Weih-

sueh A. Chiu WA, and Krishnan K. (2008). Approaches for applications of physiologically based pharmacokinetic models in risk assessment. *Journal of Toxicology and Environmental Health*, Part B. **11**:519-547

Toyoshiba H, Walker N, Bailer AJ, and Portier C. (2004). Evaluation of toxic equivalency factors for induction of cytochromes P450 CYP1a1 and CYP1a2 enzyme activity by dioxin like compounds. *Toxicology and Applied Pharmacology*. **194**:156-168.

Tuomisto J, Pohjanvirta R, Unkila M, Tuomisto J. (1999). TCDD Induced Anorexia and Wasting Syndrome in Rats: Effects of Diet-Induced Obesity and Nutrition. *Pharmacology Biochemistry and Behavior*. **64***(4)*:735-742.

Tuomisto J, Pohjanvirta R. (1994). Short-Term Toxicity of 2,3,7,8 Tetrachloro dibenzo-p-dioxin in Laboratory Animals: Effects, Mechanisms, and Animal Models. *Pharmacological reviews*. **46**:483-549.

U.S. Department of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry (ATSDR). (1998). Toxicological Profile for Chlorinated Dibenzo-p-dioxins. Agency for Toxic Substances and Disease Registry (ATSDR). Url $http://www.atsdr.cdc.gov/toxprofiles/tp.asp?id=366\&tid=63$

U.S. Environmental Protection Agency (USEPA). (1992). Guidelines for Exposure Assessment. Federal Register. **57***(104)*:22888-22938.

U.S. Environmental Protection Agency (USEPA). (2000). Exposure and Human Health Reassessment of 2,3,7,8-Tetrachlorodiobenzo-p-dioxin (TCDD) and Related Compounds. Part I: Estimating Exposure to dioxin-like compounds. Volume 2: Sources of dioxin-like compounds in the United States. EPA/600/P-00/001 Bb. National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency, Washington, DC.

U.S. Environmental Protection Agency (USEPA). (2006). Use of Physiologically Based Pharmacokinetic Models to Quantify the Impact of Human Age and Interindividual Differences in Physiology and Biochemistry Pertinent to Risk. EPA/600/R-06/014A

Umbreit TH, Hesse EJ, Gallo MA. (1986). Comparative toxicity of TCDD contaminated soil from Times Beach, Missouri, and Newark, New Jersey. *Chemosphere*. **15***(9-*

12):2121-2124

USDOJ (1997). Press release *www.justice.gov/opa/pr/*1997/*July*97/281*enr.htm* [accessed August 12, 2010].

USEPA (U.S. Environmental Protection Agency). (1985). Health Assessment Document for Polychlorinated Dibenzo-p-Dioxin. Environmental Criteria and Assessment Office, Office of Health and Environmental Assessment, Office of Research and Development, Cincinnati, OH. EPA $600/8-84-014F$.

USEPA (U.S. Environmental Protection Agency). (1997). Health Effects Assessment Summary Tables. FY 1997 Update. Environmental Criteria and Assessment Office, Office of Health and Environmental Assessment, Office of Research and Development, Cincinnati, OH.

USEPA (U.S. Environmental Protection Agency). (2005). Guidelines for Carcinogen Risk Assessment. U.S. Environmental Protection Agency, Washington, DC, EPA $630/P-03/001F$.

USEPA (U.S. Environmental Protection Agency) Exposure Factors Handbook 2011 Edition (Final). (2011). U.S. Environmental Protection Agency, Washington, DC, $EPA/600/R-09/052F$.

Vajda S, Rabitz H, Walter E, Lecourtier Y. (1989). Qualitative and Quantitative Identifability Analysis of Non-Linear Chemical Kinetic Models. *Chemical Engineering Communication*. **83**:191-219.

Van den Berg M, Birnbaum LS, Bosveld AT, Brunstrm B, Cook P, Feeley M, Giesy JP, Hanberg A, Hasegawa R, Kennedy SW, Kubiak T, Larsen JC, van Leeuwen FX, Liem AK, Nolt C, Peterson RE, Poellinger L, Safe S, Schrenk D, Tillitt D, Tysklind M, Younes M, Waern F, Zacharewski T. (1998). Toxic equivalency factors for PCBs, PCDDs, PCDFs for humans and wildlife. *Environ Health Perspective.* **106***(12)*:775-92.

Van den Berg M, Birnbaum LS, Michael Denison M, De Vito M, Farland W, Feeley M, Fiedler H, Hakansson H, Hanberg A, Haws L, Rose M, Safe S, Schrenk D, Tohyama C, Tritscher A, Tuomisto J, Tysklind M, Walker N, and Peterson RE. (2006). The 2005

World Health Organization Reevaluation of Human and Mammalian Toxic Equivalency Factors for Dioxins and Dioxin-Like Compounds. *Toxicological Sciences*. **93***(2)*:223-241.

Van den Berg M, de Jongh J, Poiger H, Olson JR. (1994). The toxicokinetics and metabolism of polychlorinated dibenzo-pdioxins (PCDDs) and dibenzofurans (PCDFs) and their relevance for toxicity. *Crit. Rev. Toxicol.* **24**:1-74

Walker NJ, Crockett PW, Nyska A, Brix AE, Jokinen MP, Sells DM, Hailey JR, Easterling M, Haseman JK, Ming Yin M, Wyde ME, Bucher JR, and Portier CJ. (2005). Dose-Additive Carcinogenicity of a Defined Mixture of Dioxin-like Compounds. *Environmental Health Perspectives*. **113**, 43-48.

Wang X, Santostefano MJ, DeVito MJ, Birnbaum LS. (2000). Extrapolation of a PBPK model for Dioxins across Dosage Regimen, Gender, and Species. *Toxicological Letters*. **56**: 49-60.

Wang X, Santostefano MJ, Evans MV, Richardson VM, Diliberto JJ, Birnbaum LS. (1997). Determination of parameters responsible for pharmacokinetic behavior of TCDD in female spraguedawley rats. *Toxicology and Applied Pharmacology*. **147**:151-168

Whitlock JP Jr. (1993). Mechanistic aspects of dioxin action. *Chem Res Toxicol*. **6***(6)*:754-763.

Winters D, Cleverly D, Lorber M, Meier K, Dupuy A, Byrne C, Deyrup C, Ellis R, Ferrario J, Lesse W, Schuam J, and Walcott J. (1996b). Coplaner polychlorinated biphenyls (PCBs) in national sample of beef in the United Stats: Preliminary Results. *Organohalogen Compounds*. **23**:350-354.
http://*www.epa.gov*/*ncea*/*pdfs*/*pcbbeef.pdf* [accessed August 10, 2010]

Winters D, Cleverly D, Meier K, Dupuy A, Byrne C, Deyrup C, Ellis R, Ferrario J, Harless R, Lesse W, Lorber M, McDaniel D, Schuam J, and Walcott J. (1996a). A statistical survey of dioxin like compounds in United States Beef: A progress report. *Chemosphere*. **32(3)**:469-478.

World Health Organization Fact sheet N225, May (2010).
http://*www.who.int*/*mediacentre*/*factsheets*/*fs*225/*en*/ [accessed March 26, 2012]

Appendix A:
Additional PK model equations

Appendix A

Denote the amount of a chemical in the liver and the fat by $A_L(t) = W_L(t)C_L(t)$ and $A_F(t) = W_F(t)C_F(t)$, where W is the organ weight C is the organ-specific concentration, we the following set of ordinary differential equations (ODE) to describe the mass balance:

$$\begin{aligned}\frac{dA_L(t)}{dt} &= W_oD - (K_e + K_{lf})A_L(t) + K_{fl}A_F(t),\\ \frac{dA_F(t)}{dt} &= K_{lf}A_L(t) - K_{fl}A_F(t),\end{aligned}$$

with baseline body weight W_0, administered dose D, and kinetic parameters K. The general solution to the ODE can be written as

$$A(t) = A_0 + b_1\exp(\lambda_1 t)\eta_1 + b_2\exp(\lambda_2 t)\eta_2$$

where $A(t) = (A_L(t), A_F(t))^T$, $A_0 = (a_{0L}, a_{0F})^T$, and b_1 and b_2 are coefficients solved from the system with the initial value $A(0) = 0$. Further, λ_1 and λ_2 are the eigenvalues, and $\eta_1 = (\eta_{11}, \eta_{21})^T$ and $\eta_2 = (\eta_{21}, \eta_{22})^T$ are the corresponding eigenvectors of the matrix of the ODE coefficients.

The eigenvectors are

$$\begin{aligned}\eta_{11} &= \frac{-Kfl+K_e+K_{lf}-\sqrt{K_{fl}^2-2K_{fl}K_e+2K_{lf}K_{fl}+K_e^2+2K_eK_{lf}+K_{lf}^2}}{-2K_{lf}}\\ \eta_{12} &= \frac{-Kfl+K_e+K_{lf}+\sqrt{K_{fl}^2-2K_{fl}K_e+2K_{lf}K_{fl}+K_e^2+2K_eK_{lf}+K_{lf}^2}}{-2K_{lf}}\\ \eta_{21} &= \eta_{22} = 1\end{aligned}$$

and the eigenvalues are given by

$$\begin{aligned}\lambda_1 &= -.5K_{fl} - .5K_e - .5K_{lf}\\ &\quad +.5\sqrt{K_{fl}^2 - 2K_{fl}K_e + 2K_{fl}K_{lf} + K_e^2 + 2K_{lf}K_e + K_{lf}^2}\\ \lambda_2 &= -.5K_{fl} - .5K_e - .5K_{lf}\\ &\quad -.5\sqrt{K_{fl}^2 - 2K_{fl}K_e + 2K_{fl}K_{lf} + K_e^2 + 2K_{lf}K_e + K_{lf}^2}\end{aligned}$$

Note $\lambda_1 - \lambda_2 = \sqrt{K_{fl}^2 - 2K_{fl}K_e + 2K_{fl}K_{lf} + K_e^2 + 2K_{lf}K_e + K_{lf}^2}$

To solve for the steady state amounts a_{0L} and a_{0F}, the steady-state equations

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -(K_e + K_{lf}) & K_{fl} \\ K_{lf} & -K_{fl} \end{pmatrix} \begin{pmatrix} a_{0L} \\ a_{0F} \end{pmatrix} + \begin{pmatrix} W_0 D \\ 0 \end{pmatrix}$$

give the solution $a_{0L} = \frac{W_0 D}{K_e}$, and $a_{0F} = \frac{K_{lf} W_0 D}{K_{fl} K_e}$ The solution to b_1 and b_2 is from the initial condition $\mathbf{A}(0) = 0$.

$$\begin{aligned} 0 &= \frac{W_0 D}{K_e} + \left(-b_2 - \frac{K_{lf} W_0 D}{K_{fl} K_e} \right) \eta_{11} - b_2 \eta_{12} \\ 0 &= \frac{W_0 D}{K_e} - \frac{K_{lf} W_0 D}{K_{fl} K_e} \eta_{11} + b_2 (\eta_{12} - \eta_{11}) \end{aligned}$$

Specifically,

$$b_1 = \frac{W_0 D (K_{fl} - K_{lf} \eta_{12})}{K_e K_{fl} (\eta_{12} - \eta_{11})} \text{ and } b_2 = \frac{W_0 D (K_{lf} \eta_{11} - K_{fl})}{K_e K_{fl} (\eta_{12} - \eta_{11})}$$

Thus the solution to ODE is

$$\begin{aligned} A_F(t) &= \frac{K_{lf} W_0 D}{K_{fl} K_e} + \frac{W_0 D (K_{fl} - K_{lf} \eta_{12})}{K_e K_{fl} (\eta_{12} - \eta_{11})} \exp(\lambda_1 t) + \frac{W_0 D (K_{lf} \eta_{11} - K_{fl})}{K_e K_{fl} (\eta_{12} - \eta_{11})} \exp(\lambda_2 t) \\ A_L(t) &= \frac{W_0 D}{K_e} + \frac{W_0 D (K_{fl} - K_{lf} \eta_{12})}{K_e K_{fl} (\eta_{12} - \eta_{11})} \exp(\lambda_1 t) \eta_{11} + \frac{W_0 D (K_{lf} \eta_{11} - K_{fl})}{K_e K_{fl} (\eta_{12} - \eta_{11})} \exp(\lambda_2 t) \eta_{12} \end{aligned}$$

Utilizing the relationships

$$\eta_{12} - \eta_{11} = \frac{-(\lambda_1 - \lambda_2)}{K_{lf}}, \; K_{fl} - K_{lf} \eta_{11} = -\lambda_1, \; K_{fl} - K_{lf} \eta_{12} = -\lambda_2$$

as well as

$$\frac{(K_{fl} - K_{lf} \eta_{11})}{(\eta_{12} - \eta_{11})} = \frac{-\lambda_1 K_{lf}}{\lambda_1 - \lambda_2}, \; \frac{(K_{fl} - K_{lf} \eta_{12})}{(\eta_{12} - \eta_{11})} = \frac{-\lambda_2 K_{lf}}{\lambda_1 - \lambda_2}, \text{ and } A(t) = W(t) C(t)$$

we have the simplified version given in (3).

Appendix B:
Additional PBPK model equations

Appendix B

In the NTP Tr526 experiment, information was collected on individual animal tissue chemistry at weeks 13, 30, 52. Acetanilide-4-hydroxylase Activity (A4H, Ah) activity was measured in the Liver (*nmol/min/mg*) and is used in describing the process of binding with the CYP1A2 protein. Also Ethoxyresorufin deethylase (EROD) activity was measured in the liver (*pmol/min/mg*). EROD is considered a biomarker for CYP1A1 activity.

For the Ah (CYP1A2) binding in the liver the change in concentration as Wang 97 and Anderson 97 did. Since the amount of CYP1A2 varies with the TCDD, the specific binding to CYP1A2 also varies. This results in a non-linear binding. The change of CYP1A2 with time is described by the model for a stimulation process proposed by Dayneka *et al.*(1993). The process of TCDD specifically bound is simplified using the equilibrium binding equation. Amount of TCDD occupied by Ah receptor. C_{AhLi} is not observed in TR-526.

$$C_{AH-TCDD}(ng/gliver) = \frac{C_{AhLi}C_{Lf}}{K_{Ahd}+C_{Lf}}$$

Amount of acetanilide-4-hydroxylase (ACOH):

$$\frac{dA_{ACOH}}{dt}(nmol/week/gliver) = (S_{ACOH}(t)K_{0ACOH} - K_{2ACOH}C_{ACOH}))W_L$$

$$\text{at t = 0, } C_{ACOH}(t) = C_{ACOHBS}$$

Stimulation function

$$S_{ACOH}(t) = 1 + In_{ACOH}\left(\frac{C^h_{AH-TCDD}}{C^h_{AH-TCDD}+IC^h_{ACOH}}\right)$$

Amount of TCDD occupied by CYP1A2:

$$C_{CYP-TCDD}(ng/gliver) = \frac{f(C_{ACOH})C_{Lf}}{K_{CYP1A2d}+C_{Lf}}$$

A link between acetanilide-4-hydroxylase (ACOH) concentration, C_{ACOH}, and CYP1A2, is assumed to be linear.

$$C_{CYP1A2} = f(C_{ACOH}) = K_{ACOHcyp} C_{ACOH}$$

Two bind relationships are described by the two "equations" below. TCDD binds with the Ah receptor, which in turn binds to DNA sites. Thus there are two different disassociation constants K_{Ahd2} and K_d.

$$\begin{array}{rcl} & \overset{K_{Ahd2}}{} & \\ AhR + TCDD & \longleftrightarrow & [AhR - TCDD] \\ & \overset{K_d}{} & \\ [AhR - TCDD]^n + DNA^n & \longleftrightarrow & [AhR - TCDD - DNA]^n \end{array}$$

K_{mx} is the maximum CY1A2 synthesis rate and K_0 is the basal synthesis rate as a fraction of K_{Mx}; h is the coefficient in the Hill model $(h > 0)$. Denote A_{A2} as the total CYP1A2 amount in the intra-cellular domain, and $A_{A2} = W_{Li} P_{Li} C_{A2}$

C_{A2BS} is the Basal level of C_{A2}. While K_0 is the zero order basal rate for CYP1A2 synthesis and K_2 is the first order constant rate for CYP1A2 degradation. Note that when $TCDD = 0$ then $C_{Lf} = 0$ and $C_{Ah-TCDD} = 0$, thus C_{A2} will not vary, hence $\frac{dA_{A2}}{dt} = 0$.

We need a value for K_{mx} so we can use an estimate of a basal synthesis rate to get K_0 as a fraction for K_{mx}. We can use $Ah_{li} = .35$ *nmol* as reported in Safe (1988) and $K_{Dah} = 7.5$ *nm* for CyP1A2 from Anderson *et al.*(1997).

Biding to CYP1A1 (EROD activty) is described using the same form of equilibrium binding equation as the CYP1a2 binding process. All the parameter notations are the same with an Er prefix in the subscript.

Amount of EROD:

$$\frac{dA_{EROD}}{dt}(pmol/week/gliver) = (S_{EROD}(t) K_{0EROD} - K_{2EROD} C_{EROD})) W_L$$

at t = 0, $C_{EROD}(t) = C_{ERODBS}$ also $K_{0EROD} = K_{2EROD} C_{ERODBS}$

Stimulation function for EROD:

$$S_{EROD}(t) = 1 + In_{EROD} \left(\frac{C^h_{AH-TCDD}}{C^h_{AH-TCDD} + IC^h_{EROD}} \right)$$

K_{Ermx} is maximum CY1A1 synthesis rate and K_{Er0} is the basal synthesis rate as a fraction of K_{Mx}. Denote A_{Er} as the total CYP1A1 amount in the intra-cellular domain, and $A_{Er} = W_L P_L C_{Er}$. C_{ErBS} is the Basal level of C_{Er}. While K_{Er0} is the zero order basal rate for CYP1A1 synthesis and K_{Er2} is the first order constant rate for CYP1A1 degradation. Note that when $TCDD = 0$ then $C_{Lf} = 0$ and $C_{Er-TCDD} = 0$, thus C_{Er} will not vary, hence $\frac{dAEr}{dt} = 0$.

The data from the control group in the TR526 data will provide baseline and max synthesis values for the binding process. K_{Ermx} so we can use an estimate of a basal synthesis rate to get K_{Er0} as a fraction for K_{Ermx}. We can write the amount of Ah receptor occupied by TCDD as:

$$C_{Er-TCDD} \quad = \quad (Ah_{Li} C_{Lf})/(K_{Erd2} + C_{Lf})$$

With $Ah_{li} = .35$ *nmol* as reported in Safe (1988).

The last set of equations describes the rate of change in the rest of the body.

Rest of Body tissues (Intra-cellular sub-compartment):

$$\frac{dA_R}{dt}(ng/week) \quad = \quad M_R(C_{RB} - \frac{C_R}{P_R})$$

$$C_R = \frac{A_R}{W_R}; \text{ at t = 0, } A_R = 0$$

Rest of body tissue blood (Extra-cellular sub-compartment):

$$\frac{dA_{RB}}{dt}(ng/week) \quad = \quad Q_R(C_B - C_{RB}) - M_R(C_{RB} - \frac{C_R}{P_R})$$

$$C_{RB} = \frac{A_{RB}}{W_{RB}}; \text{ at t = 0, } A_{RB} = 0$$

Appendix C:

R code for all models and parameter estimation.

```
#CLOSED FORM SOLUTION FOR THE 2 COMP PK MODEL
#These are used in the functions for the fat and liver amounts.##
eta11fun<-function(K){K[2]/(.5*(-K[2]+K[1]+K[3]+sqrt((K[2]^2)-
2*K[2]*K[1]+2*K[3]*K[2]+(K[1]^2)+2*K[1]*K[3]+(K[3]^2))))}
eta12fun<-function(K){K[2]/(.5*(-K[2]+K[1]+K[3]-sqrt((K[2]^2)-
2*K[2]*K[1]+2*K[3]*K[2]+(K[1]^2)+2*K[1]*K[3]+(K[3]^2))))}
lambda1fun<-function(K){(.5*(-K[2]-K[1]-K[3]+sqrt((K[2]^2)-
2*K[2]*K[1]+2*K[3]*K[2]+(K[1]^2)+2*K[1]*K[3]+(K[3]^2))))}
lambda2fun<-function(K){(.5*(-K[2]-K[1]-K[3]-sqrt((K[2]^2)-
2*K[2]*K[1]+2*K[3]*K[2]+(K[1]^2)+2*K[1]*K[3]+(K[3]^2))))}
# fat model ##########################################################
FATmodelY2 <- function(theta,tdymat,dosedep=FALSE,
fixedfl=TRUE,fixedlf=TRUE){
#fat concentration must be the 3 rd column of the tdymat
            #time is col 1 and dose is col2.
 if(!dosedep)  K1<-exp(theta[1]+0*tdymat[,2]) else
 K1<-exp(theta[1]+theta[2]*tdymat[,2])
if(fixedfl)     K2<-exp(theta[3]+0*tdymat[,2]) else
K2<-exp(theta[3]+theta[4]*tdymat[,2])
if(fixedlf)     K3<-exp(theta[5]+0*tdymat[,2]) else
K3<-exp(theta[5]+theta[6]*tdymat[,2])
eta12<-eta12fun(c(K1,K2,K3))
 eta11<-eta11fun(c(K1,K2,K3))
lambda1<-lambda1fun(c(K1,K2,K3))
lambda2<-lambda2fun(c(K1,K2,K3))
tempvec <- (tdymat[,2]/(K1*K2))*(K3+((K2-K3*eta12)/(eta12-eta11))
             *exp(lambda1*tdymat[,1])
    -((K2-K3*eta11)/(eta12-eta11))*exp(lambda2*tdymat[,1])) #fat
        return(tempvec)} #end of fat function
 ## Liver model   ###############################################
LivermodelY2  <- function(theta,tdymat,dosedep=FALSE,
fixedfl=TRUE,fixedlf=TRUE){
  # LIVER concentration must be the 4th column of the tdymat!!!
  #time is col 1 and dose is col2.
 if(!dosedep)  K1<-exp(theta[1]+0*tdymat[,2]) else
 K1<-exp(theta[1]+theta[2]*tdymat[,2])
if(fixedfl)     K2<-exp(theta[3]+0*tdymat[,2]) else
K2<-exp(theta[3]+theta[4]*tdymat[,2])
if(fixedlf)     K3<-exp(theta[5]+0*tdymat[,2]) else
K3<-exp(theta[5]+theta[6]*tdymat[,2])
eta12<-eta12fun(c(K1,K2,K3))
 eta11<-eta11fun(c(K1,K2,K3))
lambda1<-lambda1fun(c(K1,K2,K3))
lambda2<-lambda2fun(c(K1,K2,K3))
tempvec2<-(tdymat[,2]/K1)*(1+(((K2-K3*eta12)/((K2)*(eta12-eta11)))
           *eta11*exp(lambda1*tdymat[,1]))
 -(((K2-K3*eta11)/((K2)*(eta12-eta11)))*eta12*exp(lambda2*tdymat[,1])))
         return(tempvec2)
}#end of liver function
 ## ###################################################################
 ## function to convert theta to K ###################################
theta2k <- function(theta, d, dosedep=FALSE){ #old version of function
   #input: theta = parameter estimates on log scale, d = doses.
  if(!dosedep){
    if(length(theta)!=3) stop("Theta is not of length 3.")
  K1<-exp(theta[1])
  K2<-exp(theta[2])
```

```
  K3<-exp(theta[3])
  return(c(K1, K2, K3))
  }else{    if(length(theta)!=6) stop("Theta is not of length 6.")
  K1<-exp(theta[1]+theta[2]*d)
  K2<-exp(theta[3]+theta[4]*d)
  K3<-exp(theta[5]+theta[6]*d)
  return(as.matrix(data.frame(Ke=K1, Kfl=K2, Klf=K3)))
       }  } #end o'function
theta2k2 <- function(theta, d, dosedep=FALSE,fixedfl=TRUE,
fixedlf=TRUE)#input:theta=parameter estimates log scale,d=doses.
    # is Kfl fixed = fixedfl, Is Klf fixed = fixedlf
if(!dosedep) K1<-exp(theta[1]+0*d) else K1<-exp(theta[1]+theta[2]*d)
if(fixedfl)  K2<-exp(theta[3]+0*d) else  K2<-exp(theta[3]+theta[4]*d)
if(fixedlf)  K3<-exp(theta[5]+0*d) else  K3<-exp(theta[5]+theta[6]*d)
    return(as.matrix(data.frame(Ke=K1, Kfl=K2, Klf=K3)))
  }#end o'function
 ## ######################################################################
 ## Calculate the weight matrices from RESIDUALS.######################
newweightmat <- function(theta,tdymat,dosedep=T,fixedfl=F,fixedlf=F){
# input: theta = parameter estimates. tdymat = DATA. and options
if(!dosedep) pke<-1 else pke <-2   #is Ke dose dependent?
if(fixedfl) pkfl<-1  else pkfl<-2  #is Kfl fixed?
if(fixedlf) pklf<-1 else pklf<-2   #is Klf fixed?
 p<-pke+pkfl+pklf ;   #total # of parameters.
fittedy <- FATmodelY2(theta,tdymat, dosedep=dosedep,fixedfl=fixedfl,
fixedlf=fixedlf) #pred fat
fittedy2<-LivermodelY2(theta,tdymat, dosedep=dosedep,fixedfl=fixedfl,
fixedlf=fixedlf)#pred liver
res.f <- tdymat[,3]-fittedy    #fat residuals
res.l<- tdymat[,4]-fittedy2    #liver residuals
resids<-data.frame(res.f,res.l)  #dataframe of residuals
nt<-sort(unique(tdymat[,1]))    #unique times
nd<-sort(unique(tdymat[,2]))    #unique doses
lnd<-length(nd)                 #number of unique doses
lnt<-length(nt)                 #number of unique times
pred<- data.frame(fittedy ,fittedy2)
#save weight matrices in an array.
#will use as input to weighted objective function and unlist and convert
# to matrix when needed
saveweights<-array(data = list(matrix(c(1,0,0,1),nrow=2),
                               matrix(c(1,0,0,1),nrow=2),
  matrix(c(1,0,0,1),nrow=2),matrix(c(1,0,0,1),nrow=2),
  matrix(c(1,0,0,1),nrow=2),matrix(c(1,0,0,1),nrow=2),
  matrix(c(1,0,0,1),nrow=2),matrix(c(1,0,0,1),nrow=2),
  matrix(c(1,0,0,1),nrow=2),matrix(c(1,0,0,1),nrow=2),
  matrix(c(1,0,0,1),nrow=2),matrix(c(1,0,0,1),nrow=2),
  matrix(c(1,0,0,1),nrow=2),matrix(c(1,0,0,1),nrow=2)),dim = c(4,4))
allcors<-rep(c(0.80973,0.94639,0.96306,.90011),4)
 ocormat<-  matrix(allcors,nrow=4,ncol=4,byrow=F)
 for(i in 1:lnt) {  #times loop
   for( j in 1:lnd){  #doses loop
   temp<-pred[tdymat[,1]==nt[i]&tdymat[,2]==nd[j],];
  corFL<-ocormat[i,j];#cor(temp[,1],temp[,2]);
    SDf<-.19901*mean(temp[,1]);
   varf<-SDf*SDf;
```

```
 SDl<-.19980*mean(temp[,2]);
 varl<-SDl*SDl;
  covFL<-corFL*.19901*.19980*mean(temp[,1])*mean(temp[,2]);
 etestmat<-matrix(c(varf,covFL,covFL,varl),nrow=2);  #put in matrix
 #var & cov of residuals in temp data frame
saveweights[i,j]<- list(etestmat)
        }#end dose loop
    } #end time loop
     return(saveweights)
     }#End o'function to calculate weight matrices.
## OBJECTIVE FUNCTIONS ##############################################
   #USED FIRST TO GET INITIAL ESITMATES FROM STARTING VALUES
   # for 8 groups.. observed times 1,2,3 together by the 4 dose groups
   vs time 4 by the 4 dose groups.
mult.model.obj.newweightmat <- function(theta,tdymat,dosedep=T,
fixedfl=F,fixedlf=F){
#input: theta = parameter estimates, tdymat = DATA... options...
 if(!dosedep) pke<-1 else pke <-2    #is Ke dose dependent?
 if(fixedfl) pkfl<-1  else pkfl<-2   #is Kfl fixed?
 if(fixedlf) pklf<-1 else pklf<-2    #is Klf fixed?
 p<-pke+pkfl+pklf ;  #total number of parameters
 fittedy <- FATmodelY2(theta,tdymat, dosedep=dosedep,fixedfl=fixedfl,
 fixedlf=fixedlf) #pred fat
 fittedy2<-LivermodelY2(theta,tdymat, dosedep=dosedep,fixedfl=fixedfl,
 fixedlf=fixedlf)#pred liver
 res.f <- tdymat[,3]-fittedy      #fat residuals
 res.l<- tdymat[,4]-fittedy2      #liver residuals
 resids<-data.frame(res.f,res.l)    #dataframe of residuals
 nt<-sort(unique(tdymat[,1]))      #unique times
 nd<-sort(unique(tdymat[,2]))      #unique doses
 lnd<-length(nd)                   #number of unique doses
 lnt<-length(nt)                   #number of unique times
 pred<- data.frame(fittedy ,fittedy2)
#save weight matrices in an array.
 #will use as input to weighted objective function and unlist and
 # convert to matrix when needed
saveweights<-array(data = list(matrix(c(1,0,0,1),nrow=2),
                      matrix(c(1,0,0,1),nrow=2),
  matrix(c(1,0,0,1),nrow=2),matrix(c(1,0,0,1),nrow=2),
  matrix(c(1,0,0,1),nrow=2),matrix(c(1,0,0,1),nrow=2),
  matrix(c(1,0,0,1),nrow=2),matrix(c(1,0,0,1),nrow=2),
  matrix(c(1,0,0,1),nrow=2),matrix(c(1,0,0,1),nrow=2),
  matrix(c(1,0,0,1),nrow=2),matrix(c(1,0,0,1),nrow=2),
  matrix(c(1,0,0,1),nrow=2),matrix(c(1,0,0,1),nrow=2)),dim = c(4,4))
allcors<-rep(c(0.80973,0.94639,0.96306,.90011),4)
ocormat<- matrix(allcors,nrow=4,ncol=4,byrow=F)
savemat<- matrix(0,nrow=lnt,ncol=lnd)
for(i in 1:lnt) {  #times loop
   for( j in 1:lnd){  #doses loop
  temp<-pred[tdymat[,1]==nt[i]&tdymat[,2]==nd[j],];
  corFL<-ocormat[i,j];#cor(temp[,1],temp[,2]);
  SDf<-.19901*mean(temp[,1]);
   varf<-SDf*SDf;
```

```
   SDl<-.19980*mean(temp[,2]);
   varl<-SDl*SDl;
   covFL<-corFL*.19901*.19980*mean(temp[,1])*mean(temp[,2]);
   vcmat<-matrix(c(varf,covFL,covFL,varl),nrow=2);
   invvcmat<- ginv(vcmat)#solve(vcmat)
   #use residuals with var-cov of observed data...
    residstemp<-resids[tdymat[,1]==nt[i]&tdymat[,2]==nd[j],];
   testsum<- sum(apply(residstemp,1,function(x){
 test<-as.numeric(t(x)%*%invvcmat%*%as.matrix(x))}));
 savemat[i,j]<-testsum
     }#end dose loop
 } #end time loop
    return(sum(savemat) )
    } #End o'function.
 ## reweighted OBJECTIVE FUNCTION        used in IRLS
w.mult.model.obj.newweightmat <- function(theta,tdymat,w,dosedep=T,
                                          fixedfl=F,fixedlf=F){
 #input: theta = parameter estimates, tdymat = DATA... options..
  fittedy <- FATmodelY2(theta,tdymat, dosedep=dosedep,
   fixedfl=fixedfl,fixedlf=fixedlf) #pred fat
  fittedy2<-LivermodelY2(theta,tdymat, dosedep=dosedep,
   fixedfl=fixedfl,fixedlf=fixedlf) #pred liver
  res.f <- tdymat[,3]-fittedy        #fat residuals
  res.l<- tdymat[,4]-fittedy2        #liver residuals
  resids<-data.frame(res.f,res.l)      #dataframe of residuals
  nt<-sort(unique(tdymat[,1]))        #unique times
  nd<-sort(unique(tdymat[,2]))        #unique doses
  lnd<-length(nd)                     #number of unique doses
  lnt<-length(nt)                     #number of unique times
  savemat<-matrix(0,nrow=lnt,ncol=lnd)
 for(i in 1:lnt) {  #times loop
    for( j in 1:lnd){  #doses loop
 temp<-resids[tdymat[,1]==nt[i]&tdymat[,2]==nd[j],];
    vcmat<-matrix(unlist(w[i,j]),nrow=2);  #get Weight matrix for i,j
    and unlist and put into matrix
    invvcmat<- ginv(vcmat)#solve(vcmat)   # get the sum of the X^{T} *
    #W^{-1} * X for all rows in data group i,j
        testsum<- sum(apply(temp,1,function(x){
        test<-as.numeric(t(x)%*%invvcmat%*%as.matrix(x))}));
        savemat[i,j]<-testsum  #save results by i and j.
          }#end times loop
            }#end doses loop
   return(sum(savemat) )   # SUM ALL together
    } #End o'reweighted objective function!
 ## function for the Max Element Difference of the weight
 #matrices over each iteration.
MaxElementDiff<-function(x,y){#input x old & y new weight matrices
   r<-dim(x)[1];  nc<-dim(x)[2]
   med<-rep(0,r*nc)
   p<-1
  for(i in 1:r){
for( j in 1:nc){med[p]<-max(abs(matrix(unlist(x[i,j]),nrow=2)
-matrix(unlist(y[i,j]),nrow=2))); p<-p+1;
   } #end row loop }    #end col loop.
          return(med)     }
```

```
#simple SSR############################
theSSR<- function(theta,tdymat,dosedep=T,fixedfl=F,fixedlf=F){
#input theta = parameter estimates, tdymat = DATA.
  if(!dosedep) pke<-1 else pke <-2  #is Ke dose dependent?
  if(fixedfl) pkfl<-1  else pkfl<-2 #is Kfl fixed?
  if(fixedlf) pklf<-1 else pklf<-2  #is Klf fixed?
   p<-pke+pkfl+pklf ;  #total number of parameters.
  fittedy <- FATmodelY2(theta,tdymat, dosedep=dosedep,fixedfl=fixedfl,
     fixedlf=fixedlf)    #pred Fat
fittedy2<-LivermodelY2(theta,tdymat, dosedep=dosedep,fixedfl=fixedfl,
     fixedlf=fixedlf)  #pred Liver
  res.f <- tdymat[,3]-fittedy     #fat residuals
  res.l<- tdymat[,4]-fittedy2     #liver residuals
  resids<-data.frame(res.f,res.l)   #dataframe of residuals
    return(sum(res.f*res.f+res.l*res.l,na.rm=T));
        }
#################################################################
#put it all together now.#put it all together now.
HybridIRLSnewweightmat<-function(start.theta,tdymat,
 #start.theta = starting param values, tdymat=DATA
dosedep=dosedep,fixedfl=fixedfl,fixedlf=fixedlf,  #options
xtol,reltol,pg.tol,fact.r,itermax=50,
#reltol for odiff or vdiff, pg and fact for optim.
criterion=c("vdiff","odiff"), method=c("nlminb", "optim"), trace=6){
 # odiff = objective function difference, vdiff = max differenece in
 # weight matrices between iter.
  if(method=="nlminb"){
 unw.fit<-nlminb(start=start.theta,obj=mult.model.obj.newweightmat,
 tdymat=tdymat,
 dosedep=dosedep,fixedfl=fixedfl,fixedlf=fixedlf,
 #lower = start.theta-abs(start.theta)*1.1,
 upper =start.theta+abs(start.theta)*1.1,
 control=list(iter.max=itermax,x.tol=xtol,rel.tol=reltol))
   }
  else{    print("optim #1");
  LB <- c(max(-6,as.numeric(start.theta[1])-3),-.025,
  max(-6,as.numeric(start.theta[3])-3),
  -.025,max(-6,as.numeric(start.theta[5])-3),-.025);
  UB <- c(min(6,as.numeric(start.theta[1])+3),.025,
  min(6,as.numeric(start.theta[3])+3),
  .025,min(6,as.numeric(start.theta[5])+3),.025);
 unw.fit<-optim(par=start.theta,fn=mult.model.obj.newweightmat,
 tdymat=tdymat,
 dosedep=dosedep,fixedfl=fixedfl,fixedlf=fixedlf,
#lower = start.theta-abs(start.theta)*1.1,
#upper=start.theta+abs(start.theta)*1.1,
  lower = LB ,
  upper = UB,
 method="L-BFGS-B",
 control=list(maxit=2,factr=fact.r,pgtol=pg.tol,trace=trace),hessian=F)
 print("NOT REWEIGHTED RESULTS:");
 print(unw.fit);
 unw.fit.SSR<-theSSR(unw.fit$par,tdymat,dosedep=dosedep,
 fixedfl=fixedfl,fixedlf=fixedlf);  }
        okgo<-0
```

```
#if((unw.fit$convergence%in%c(0,1)&unw.fit$value<1500)){if(okgo==0){
    print("DOING IRLS");
 p <- length(start.theta)
 neww <-newvar.by.dt <-try(newweightmat(unw.fit$par,tdymat,dosedep=dosedep,
 fixedfl=fixedfl,fixedlf=fixedlf))
 i <- 0
 odiff <- 1000    #set large to enter while loop
 vdiff <- 100000    #set large to enter while loop.
 newtheta <- unw.fit$par # CONVERGED STARTING VALUES!!!
 newobjective<-mult.model.obj.newweightmat(newtheta,tdymat,dosedep=dosedep,
   fixedfl=fixedfl,fixedlf=fixedlf)
while(eval(as.name(criterion))>reltol & i<=itermax){#While loop for IRLS
 i <- i+1
 cat("IRLS ITERATION: i =", i, "\n")
 if(method=="nlminb"){
      w.fit <- nlminb(start=newtheta,obj=w.mult.model.obj.newweightmat,
      tdymat=tdymat,w=neww,
                      dosedep=dosedep,fixedfl=fixedfl,fixedlf=fixedlf,
                      lower = start.theta-abs(start.theta)*1.1,
      upper = start.theta-abs(start.theta)*1.1,
      control=list(iter.max=500,x.tol=xtol,rel.tol=reltol))
      oldobjective <- newobjective
      newobjective <- w.fit$objective
                   }
   else{
     LB <- c(max(-6,as.numeric(newtheta[1])-3),-.025,
     max(-6,as.numeric(newtheta[3])-3),
     -.025,max(-6,as.numeric(newtheta[5])-3),-.025);
     UB <- c(min(6,as.numeric(newtheta[1])+3),.025,
     min(6,as.numeric(newtheta[3])+3),.025,
     min(6,as.numeric(newtheta[5])+3),.025);
w.fit <- try(optim(par=newtheta,fn=w.mult.model.obj.newweightmat,
     tdymat=tdymat,w=neww,
 dosedep=dosedep,fixedfl=fixedfl,fixedlf=fixedlf,
 lower = LB, upper = UB,
 method="L-BFGS-B",
 control=list(maxit=itermax,factr=fact.r,pgtol=pg.tol,
 trace=trace),hessian = F))
 if(inherits(w.fit,"try-error")||inherits(w.fit,
            "Error in solve.default(vcmat)")){
   outlist<-list(NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
   unw.fit$par,unw.fit$convergence, unw.fit$message,
   unw.fit$value,  unw.fit.SSR) #
   names(outlist)<-c("startingvalues",
   "parameters",   "K",
   "objective",   "Iterations",
   "Message",   "call",   "convergence",
   "weights",   "SSR",   "RealStartVals",
   "conv",   "mes",
"unw.obj",            "unwSSR")
      return(outlist)       }
else {
 oldobjective <- newobjective #save old objective function values
newobjective <- w.fit$value#save new objective values   }
 oldtheta <- newtheta  #save old parameter estimates
 newtheta <- w.fit$par #save new parameter estimates
 oldw <- neww          #replace weights
```

```
 #GET THE NEW WEIGHTS!
 neww <-newvar.by.dt<-try(newweightmat(newtheta,tdymat,dosedep=dosedep,
 fixedfl=fixedfl,fixedlf=fixedlf))
 #objective function difference.
    odiff <- abs(newobjective-oldobjective)/oldobjective
    oldvar.by.dt <- newvar.by.dt
   #difference of var-covar matrices.
   vdiff <- max( MaxElementDiff(newvar.by.dt,oldvar.by.dt))
   print(vdiff)
 } } # end while loop
print("SSR");
SSR<-theSSR(newtheta,tdymat,dosedep=dosedep,fixedfl=fixedfl,
fixedlf=fixedlf);
print(SSR);
 K <- if(dosedep) theta2k2(w.fit$par,unique(tdymat[,2]),dosedep=dosedep,
 fixedfl=fixedfl,fixedlf=fixedlf) else exp(w.fit$par);
outlist<-list(start.theta,
   w.fit$par,
  K, switch(method, nlminb=w.fit$objective, optim=w.fit$value),
  i+1,  w.fit$message,
  match.call(),    w.fit$convergence,
    newvar.by.dt,    SSR,     unw.fit$par,
    unw.fit$convergence,    unw.fit$message,
    unw.fit$value,      unw.fit.SSR) #
 names(outlist)<-c("startingvalues",
        "parameters",        "K",
        "objective",        "Iterations",
     "Message",      "call",
     "convergence",     "weights",
     "SSR",       "RealStartVals",
     "conv",     "mes",
     "unw.obj",     "unwSSR")
     }
     else { print("SORRY CHARLIE BETTER TASTING TUNA NEXT TIME.");
   outlist<-list(NA,NA,NA,NA,NA,NA,NA,NA,NA,NA,
   unw.fit$par,
   unw.fit$convergence,
   unw.fit$message,
   unw.fit$value,
   unw.fit.SSR) #
   names(outlist)<-c("startingvalues",
   "parameters",
   "K",   "objective",
   "Iterations",   "Message",
   "call",   "convergence",
   "weights",   "SSR",
   "RealStartVals",   "conv",
   "mes",    "unw.obj",   "unwSSR")
        }
      return(outlist)
}#end o'function.
#########################################################################
# BOOTSTRAPPING!!! for PBPK using IRLS
#########################################################################
DDbootstrapPBPK <- function(theta,tdymat,dosedep=FALSE,RANDOMsv=FALSE,
B=1000, start.theta,
```

```
 xtol, reltol,pg.tol,fact.r,criterion=c("vdiff","odiff"),
 method=c("nlminb", "optim"),trace=0,conf=0.95){
    ptm <- proc.time()
    datetimestart<-date()
    fittedy <- FATmodelY2(theta,tdymat, dosedep=dosedep)
    fittedyL<- LivermodelY2(theta,tdymat, dosedep=dosedep)
    res <- tdymat[,3]-fittedy
    resL <- tdymat[,4]-fittedyL
         tn <- table(tdymat[,1]);    # time group sample size
    print(table(tdymat[,1]));
    tp <- as.numeric(names(tn));
 xy<-table(tdymat[,2],tdymat[,1])#time and dose group sample size
     print(xy);
      print(tp);
    dn <- table(tdymat[,2])      # group sample size
    d <- as.numeric(names(dn))
    nd <- length(d)
    bootres <- res
    bootresL <- resL
    boottheta <- matrix(0, nrow=B, ncol=length(theta))
    bootresid <- matrix(0, nrow=B, ncol=length(res))
    IRLSiter <- rep(0,B)
    probs <- c((1-conf)/2, 0.5+conf/2)
    startvalues<-start.theta
    for(b in 1:B){
 cat("b =", b, "\n")
 for(i in 1:nd){
  for(j in 1:4){ #time loop j
bootres[tdymat[,2]==d[i]&tdymat[,1]==tp[j]]<-sample(res[(tdymat[,2]==d[i]&
tdymat[,1]==tp[j])], size=xy[i,j], replace=TRUE)
bootresL[tdymat[,2]==d[i]&tdymat[,1]==tp[j]]<-
                              sample(resL[(tdymat[,2]==d[i]&
tdymat[,1]==tp[j])], size=xy[i,j], replace=TRUE)
   } }        #end of for loop with i
 bootresid[b,] <- bootres
 boottdymat <- tdymat
      boottdymat[,3] <- fittedy+bootres
      boottdymat[,4] <- fittedyL+bootresL
    if(RANDOMsv==FALSE){startvalues<-start.theta}
      else {
              if(b==1){startvalues<-start.theta}
        else {
  if (is.na(boottheta[b-1,1])==TRUE) {startvalues<-start.theta}
 else {startvalues<-boottheta[b-1,]    } }       }
bootfit <-try(HybridIRLSnewweightmat(start.theta =theta,tdymat=boottdymat,
dosedep = T,fixedfl=F,fixedlf=F,
    xtol = xtol, reltol = reltol ,fact.r=fact.r,pg.tol=pg.tol,itermax=35,
    criterion = criterion, method = method, trace = 2)  )
if( inherits(bootfit, "try-error")){boottheta[b,]<-rep(NA,length(theta));
IRLSiter[b]<-rep(NA,1) }
 else {boottheta[b,]<-bootfit$par;IRLSiter[b]<-bootfit$Iterations}
                }
        mean2<-function(x){mean(x,na.rm=T)}
       var2<-function(x){var(x,na.rm=T)}
    bias <- apply(boottheta, 2, mean2) - theta
    se <- sqrt(apply(boottheta, 2, var2))
```

```
    output <- list(boottheta=boottheta,IRLSiter=IRLSiter)
quantile2<-function(x,probs){ quantile(x,probs=probs, na.rm = TRUE)}
 if(dosedep){
output$k3.3 <- t(apply(boottheta, 1,
function(x){theta2k(x, unique(tdymat[,2]), TRUE)[1,]}))
output$k7.3 <- t(apply(boottheta, 1,
function(x){theta2k(x, unique(tdymat[,2]), TRUE)[2,]}))
output$k15.2 <- t(apply(boottheta, 1,
function(x){theta2k(x, unique(tdymat[,2]), TRUE)[3,]}))
output$k33 <- t(apply(boottheta, 1,
function(x){theta2k(x, unique(tdymat[,2]), TRUE)[4,]}))
output$bias <-bias
output$se<-se
output$pctCIbeta <- apply(boottheta, 2, quantile2, probs=probs)
output$centpctCIbeta <- (matrix(2*theta, nrow=2, ncol=6, byrow=T)
                         - output$pctCIbeta)[2:1,]
output$normCIbeta <- rbind(theta-output$bias-qnorm(.975)*output$se
 theta-output$bias+qnorm(.975)*output$se)
output$pctCIk3.3 <- apply(output$k3.3, 2, quantile2, probs=probs)
output$pctCIk7.3 <- apply(output$k7.3, 2, quantile2, probs=probs)
output$pctCIk15.2 <- apply(output$k15.2, 2, quantile2, probs=probs)
output$pctCIk33 <- apply(output$k33, 2, quantile2, probs=probs)
output$centpctCIk3.3<-(matrix(2*theta2k(theta,unique(tdymat[,2]),
TRUE)[1,], nrow=2, ncol=3, byrow=T)- output$pctCIk3.3)[2:1,]
output$centpctCIk7.3<-(matrix(2*theta2k(theta,unique(tdymat[,2]),
TRUE)[2,],  nrow=2, ncol=3, byrow=T)-output$pctCIk7.3)[2:1,]
output$centpctCIk15.2<-(matrix(2*theta2k(theta,unique(tdymat[,2]),
TRUE)[3,],  nrow=2, ncol=3, byrow=T)-output$pctCIk15.2)[2:1,]
output$centpctCIk33<-(matrix(2*theta2k(theta,unique(tdymat[,2]),
TRUE)[4,],  nrow=2, ncol=3, byrow=T)-output$pctCIk33)[2:1,]
    }else{ output$bias <-bias
output$se<-se
output$bootk <- t(apply(output$boottheta, 1, theta2k,
d=unique(tdymat[,2]), dosedep=FALSE))
output$pctCIk <- apply(output$bootk, 2, quantile2, probs=probs)
output$pctCItheta <- apply(output$boottheta, 2,
quantile2, probs=probs)
output$centpctCItheta <- (matrix(2*theta, nrow=2, ncol=3, byrow=T)
              - output$pctCItheta)[2:1,]
k <- theta2k(theta, unique(tdymat[,2]), FALSE)
output$centpctCIk <- rbind(2*k - output$pctCIk[2,], 2*k
- output$pctCIk[1,])    }
  times<-proc.time() - ptm
  output$times<-c(times[3],times[3]/60,times[3]/60/60)
  names(output$times)<-c("secs","Mins","Hrs")
  output$datetimestart<- datetimestart
  output$datetimend<-date()
    output$call <- match.call()
    return(output)
}#End o'function for bootstrapping.
#PBPK model and MCMC#########################
#NEED TO LOAD LIBARIES#################
library('lattice')    #requried for FME
library('coda')      #required for FME
library('deSolve')    #required for simecol
library('MASS')       #required for FME
library('rootSolve')  # required for FME
```

```
library('minpack.lm') # required for FME
library('simecol')   # NECESSARY!
library('FME')        # NECESSARY!
######PBPK model code################
newPBPK<-new("odeModel",
   main = function (time, init, parms, loadDose) {
      with(as.list(c(parms, init, inputs)), {
######Input:Dose, Dose level, ave body weight####
      Dose      <- Bio*inputs$Dose(time); #
      Doselevel <- inputs$Doselevel(time);
      BWt0 <- inputs$Bwt0(time);
######Body weight over time and Cardiac ouput################
     Vmax<-exp(Va+Vb*Doselevel*5) # mult by 5
     Km<-exp(Ka+ Kb*Doselevel*5)  # mult by 5
      BWtime<-Wint+Wint*((Vmax*time)/(Km+time))    #grams
     Qc <- QCCAR*10080*((BWtime)/1000)^(0.75); # QCCAR is in ml/min/kg
Qli <- Qli0*Qc; #tissue blow flows (fraction of cardiac output) liver=.183
     Qf <-  Qf0*Qc;     #fat = .069
     Qr <-  Qr0*Qc;     #rest o body = .748
     Mf <-  Mf0*Qf;     #Tissue permeabilities.
    Mli <- Mli0*Qli;
    Mr <-  Mr0*Qr;
#######Organ weights as fraction of body weight#####################
      LivW<-WLi*BWtime;   LivbW<-WLib* LivW
      BloodW<-WB*BWtime
      FatW<-  WF*BWtime;  FatbW<- WFb*FatW
      RestW<- WR*BWtime;  RestbW<- WRb*RestW
#######Elimination term from liver...possibly doselevel dependent######
 KLi<-exp(alpha.e+ beta.e*Doselevel*5); #multiply doselevel by 5
      # this is the liver elimination (metabolism and elim combined)
#######Elimination term from blood, urinary..possibly doselevel dependent#
 Kuri<-exp(KuriA+ KuriB*Doselevel*5);  #multiplyby 5; urinary clearance.
#####Gastrointestinal absorption and
          #Distribution of TCDD to the portal lymphatic Circulation.
    #  dAst <-  Dose*bio - Kst*Ast;              #STOMACH
    #  dAlum <- Kst*Ast - Kabs*Alum;             #LUMAN
    #new from closedform solution.
     Alum<- (Dose/Kabs) - (Dose/(Kabs-Kst))*exp(-Kst*time) +
                          ((Dose/(Kabs-Kst))-(Dose/Kabs))*exp(-Kabs*time)
#####BLOOD##########################################################
  dAb<-Qli*((Alib/LivbW) - (Ab/BloodW)) + Qf*((Afb/FatbW) -
                    (Ab/BloodW)) + Qr*((Arb/RestbW)
            - (Ab/BloodW) ) + a*Kabs*Alum - Kuri*(Ab/BloodW)*BloodW   #new
######LIVER#########################################################
    Clif<- Cfree2(Ali,LivW,Pli,Ahli,KAhd,Kcyp1a2d) #Free conc in Liver
    Alif<-Clif*LivW
    if(Alif<0) {Alif<-0}#constraint
dAli <- Mli*((Alib/LivbW)-((Alif/LivW)/Pli)) - KLi*(Alif/LivW)*LivW #LIVER
   dAlib <- Qli*((Ab/BloodW)-(Alib/LivbW)) + (1-a)*Kabs*Alum -
                      Mli*((Alib/LivbW)-((Alif/LivW)/Pli)) #LiverBLood
#####BINDING OF FREE LIVER CONC TO AH receptor and CYP1A2#################
     CAhTCDD <- (Ahli*(Alif/LivW)) /(KAhd + (Alif/LivW)) #   Ah-TCDD
     Acyp1a2<-KACOHCYP*AACOH  #linear relationship between ACOH and CYP1a2
     CcypTCDD <- ( Acyp1a2*(Alif/LivW)) /(Kcyp1a2d + (Alif/LivW))
     denom<-(KAhd + (Alif/LivW))
```

```
########### ACOH Activity in Liver#######################################
    #stimulation function
     Sacoh<- 1 + InACOH*((CAhTCDD^H)/((CAhTCDD^H)+(IcACOH^H)))
     dAACOH<- (K0ACOH*Sacoh - K2ACOH*(AACOH/LivW))*LivW  #ACOH  amount
     SACOHp2<-InACOH*((CAhTCDD^H)/((CAhTCDD^H)+(IcACOH^H)))
########### EROD Activity in Liver#######################################
 #stimulation function
  Serod<- 1 + InEROD*((CAhTCDD^H)/((CAhTCDD^H)+(IcEROD^H)))
  dAEROD<- (K0EROD*Serod - K2EROD*(AEROD/LivW))*LivW         #EROD  amount
    SERODp2<-InEROD*((CAhTCDD^H)/((CAhTCDD^H)+(IcACOH^H)))
#####FAT#################################################################
    dAf <- Mf*((Afb/FatbW)-((Af/FatW)/Pf))                      #tissue
    dAfb<-Qf*((Ab/BloodW) - (Afb/FatbW)) - Mf*((Afb/FatbW) -
           ((Af/FatW)/Pf)) #conc in Fat bloo
#####REST OF BODY########################################################
  dAr <- Mr*((Arb/RestbW)-((Ar/RestW)/Pr))          #REST OF BOdY
  dArb <- Qr*((Ab/BloodW) - (Arb/RestbW)) - Mr*((Arb/RestbW)-
       ((Ar/RestW)/Pr)) # Rest of body blood
########################################################################
#Constraints...
if(Ali<0) {Ali<-0};if(Alib<0) {Alib<-0}
if(Af<0) {Af<-0};if(Afb<0) {Afb<-0}
if(Ar<0) {Ar<-0};if(Arb<0) {Arb<-0};if(Ab<0) {Ab<-0}
########################################################################
     list(c(dAli,dAlib,dAf,dAfb,dAr,dArb,dAACOH,dAEROD,dAb)    #9
       ,c(Alum=Alum,Alif=Alif,AhTcddL=CAhTCDD,Cyp=CcypTCDD,KLi=KLi,
        Dose=Dose,BWt0=BWt0,   #16
          BWt=BWtime,Qc=Qc,Qli=Qli,Qf=Qf,Qr=Qr,
          Acyp1a2= Acyp1a2,LivW=LivW,
          Sacoh=Sacoh,Serod=Serod,Wb=BloodW,#26
          LivW=LivW,LivbW=LivbW,BloodW=BloodW,
          FatW=FatW,FatbW=FatbW, RestW=RestW,#32
          Mli=Mli,Mf=Mf,Mr=Mr,Krui=Kuri)) #36
          })},
 equations = list(
    Cfree2 = function(Ali,LivW,Pli,Ahli,KAhd,Kcyp1a2d){
                        #uses bisection method to solve
                        # non linear eq for Free conc in liver.
     n<-10
     xa<-0
     xb<-Ali/LivW
    lastv<-.15*(Ali/LivW)  #(Alif/(WLi*BWtime))
 for(i in 1:n){
   testv<-(xa-(1/Pli)*((Ali/LivW)-(Ahli*xa/(KAhd+xa))   -
 (Ahli*xa/(Kcyp1a2d+xa))))*(((xa+xb)/2)-(1/Pli)*
 ((Ali/LivW)-(Ahli*((xa+xb)/2)/
 (KAhd+((xa+xb)/2)))-(Ahli*((xa+xb)/2)/(Kcyp1a2d+((xa+xb)/2)))))
   if( is.nan(testv)==TRUE) {midpoint<-lastv}
          else{
if((xa-(1/Pli)*((Ali/LivW)-(Ahli*xa/(KAhd+xa))-(Ahli*xa/(Kcyp1a2d+xa))))*
(((xa+xb)/2)-(1/Pli)*((Ali/LivW)-(Ahli*((xa+xb)/2)/(KAhd+((xa+xb)/2)))-
(Ahli*((xa+xb)/2)/(Kcyp1a2d+((xa+xb)/2)))))<0)  xb<-(xa+xb)/2
    else xa<-(xa+xb)/2; }
          list(left<-xa,right<-xb, midpoint<-(xa+xb)/2)
          midpoint->lastv}
       return(midpoint)}   ),
```

```
  parms =c(Wint=  185.4187,   #BW parameters Now in weekly units.
        Va= -0.2000187,Vb=-0.0006579897,#BW parameters Now in weekly units.
        Ka= 1.861746,Kb=-0.001258892,
        Kst = 60.48,#60.48, #gastric nonabsorption constant
        Kabs= 33.6,#33.6,  # oral absroption constant.
        a = .3,  #fraction of lymph to blood [[(1 - a) goes to liver]]
       alpha.e=2.75,
       beta.e= -0.001, #slope and intercept of liver elimination.
        Bio = .88,   #bioaviablity
        QCCAR= 311.4 , #Cardiac output ml/min/kg
        Mf0=0.091,#Tissue permeamibity (fraction of tissue blood flow)
        Mli0=0.35,#Tissue permeamibity (fraction of tissue blood flow)
        Mr0 = .0298,  #Tissue permeamibity (fraction of tissue blood flow).
        KuriA=.581, KuriB=-0.0004 ,   #dose dependent urinary clearance
        Pf=100 ,Pr=1.5,Pli=6 , #partition coefficients
        WB=0.076,              #weight proportions Blood
        WLi=0.0362,WLib=0.266,   #Liver and liver blood
        WF=0.069,WFb=0.05, #fat and fat blood
        WR=0.729,WRb=.03,  #rest of body , rest of body compartment blood.
        Qli0=.183, #Blood flow rates as fraction of total caridac output.
        Qf0=.069,   #fat
       Qr0=.748,   #rest of body
        K0ACOH=1.1, #BASAL ACOH induction rate
        K2ACOH=16.8, # degradation rate
        InACOH=67.2,    # Induction fold
        IcACOH=130,   # TCDD-Ah-DNA
        KAhd=.1,     # TCDD-AH Dissociation
        Kcyp1a2d=.1, # TCDD-CYP1A2 Dissociation
        KACOHCYP=1,  #linear relationship between  ACOH and CYP1A2
        K0EROD = 1100,#BASAL EROD induction rate
        K2EROD =16.8,# 1.1from Santostefano# degradation rate EROD 1/hr
        InEROD=365,# units fold induction
        IcEROD=1300,#
        H=.6,    #hill coefficient
        Ahli=.35    #receptor level
           ),
  inputs=list(Bwt0 = approxfun(0:103, rep(loadDose[1],104), rule=2),
  #first element of input vector is body weight,
   # SECOND element is time
    Doselevel = approxfun(0:103, rep(loadDose[3],104), rule=2),
       #3rd element of imput vector.
        Dose  = approxfun(0:103, loadDose[4:107], rule=2)),
                    # and 4-107 is the dose.
  times = c(0:103),
  init = c( Ali  = 0, Alib = 0, Af  = 0, Afb = 0, Ar = 0 , Arb = 0,
             AACOH=.5,AEROD=500, Ab=0),
   solver = function (y, times, func, parms, ...) {
       lsoda(y, times, func, parms, hmax=1, verbose =T,
               maxsteps=50000000,rtol=1e-2, atol=1e-2)
                                                  }
       )#END OF MODEL CODE!########
############ WEIGHTED COST FUNCTION ##############################
 weightedcostbyWeek <- function(sim, obs,ttt) {
# print("OBJEcTIVE FUNCTION")
   ###inputs are simulated values and observed values,
   #ttt is a vector of times to repeat the simulated.
  names(obs)<-c("Ali","Af","Ab")    #subtracts simulated from observed
```

```
  #and only uses residuals to compute var covar for the weights.
  nr<-dim(obs)[1]
  Liv<-unlist(obs[c("Ali")])
  Fat<-unlist(obs[c("Af")])
  Blood<-unlist(obs[c("Ab")])
  sim<-sim[c("Ali","Af","Ab")]
  Testdiff<-obs-sim
 YminusYhat<-  matrix(data = unlist(obs) - unlist(sim) ,
                          nrow = nr, ncol = 3, byrow = F)
 obs.df<-data.frame(Liv,Fat,Blood,ttt)
 nt<-sort(unique(ttt))
 lnt<-length(nt)
 savevec<-rep(0,lnt)
 for(i in 1:lnt) {
        temp<-obs.df[obs.df$ttt==nt[i],]
        tempHat<-YminusYhat[obs.df$ttt==nt[i],]
        vcmat<-matrix(   #Variance covariance matrix
        c(var(tempHat[,1]),
        cov(tempHat[,1],tempHat[,2]),
        cov(tempHat[,1],tempHat[,3]),
          cov(tempHat[,2],tempHat[,1]),
          var(tempHat[,2]),
          cov(tempHat[,2],tempHat[,3]),
          cov(tempHat[,3],tempHat[,1]),
          cov(tempHat[,3],tempHat[,2]),
          var(tempHat[,3])),nrow=3);
        testsum<- sum(apply(tempHat,1,function(x){#print("row");
test<-as.numeric(t(x)%*%ginv(vcmat)%*%as.matrix(x)); #solve(vcmat)
        return(test);
        }));        #multivariate sum of squares
        savevec[i]<-testsum
                  }
       return(sum(savevec) )}
   ObjectivefuncALLDOSEbyWeekTEF <- function(par) {
     OUTCOMES<-TEFoutcomes
    DOSES<-FINALavedosebyWeekNg[,49:64]
    fave <- function(x2) {
    inp <- list(Bwt0    = approxfun(0:103, rep(x2[1],104), rule=2),
    #first element of input vector is body weight,
    # SECOND time is 2nd
      Doselevel = approxfun(0:103, rep(x2[3],104), rule=2),
    #3rd element of imput vector.
      Dose      = approxfun(0:103, x2[4:107], rule=2))
    # and 4-731 is the dose.
        inputs(newPBPK) <- inp
      #cat("DOSE:= ", x2[3], "\n")
          if(x2[2]==13)  {
       times(newPBPK) <-c(0,12);               };
 #not 13,30,52, and 104 but minus one!dosing starts at time = 0.
       if(x2[2]==30)  {times(newPBPK) <-c(0,29);};
       if(x2[2]==52)  { times(newPBPK) <-c(0,51); };
        if(x2[2]==104) {times(newPBPK) <-c(0,103);};
 info<-out(sim(newPBPK,hmax=1,maxsteps=5000000,rtol=1e-2, atol=1e-2))
      info[info<0]<-0.0000001
```

```
        dim(info)[1]->rownumber
       return(info[rownumber,]   )      }
       parms(newPBPK)[names(par)] <- par
  res <- lapply(DOSES , fave) # Applies simulation to each
                  # dose vector for each time
    print(parms(newPBPK))
 es.df<-do.call(rbind,lapply(res,data.frame)) #combine all in a data set.
  res.df0<-res.df[c("Ali","Alib","Alif","Af","Afb","Ab")]#choosing cols
  res.df <-data.frame(Ali=res.df0$Ali+res.df0$Alib+res.df0$Alif,
                   Af=res.df0$Af+res.df0$Afb,Ab=res.df0$Ab)
 res.df2<-data.frame(     #repeat simulated data
 rep(res.df$Ali,times=c(10,10,8,10,10,10,8,10,10,10,8,10,10,10,8,8)) ,
 rep(res.df$Af,times=c(10,10,8,10,10,10,8,10,10,10,8,10,10,10,8,8)),
 rep(res.df$Ab,times=c(10,10,8,10,10,10,8,10,10,10,8,10,10,10,8,8)))
 names(res.df2)<-c("Ali","Af","Ab")
 obtime<-rep(c(13,30,52,104,13,30,52,104,13,30,52,104,13,30,52,104),
             times=c(10,10,8,10,10,10,8,10,10,10,8,10,10,10,8,8))
  return(weightedcostbyWeek(res.df2,
  OUTCOMES[,c("LIVERamount", "FATamount", "BLOODamount")],obtime))
  }
modMCMC(f=ObjectivefuncALLDOSEbyWeekTEF,
    p= c(alpha.e=-2.0616721,beta.e= -0.13229725,KuriA=8.3557425,
          KuriB=-0.0095490400,Pf= 193.56113,Pli=47.025481),
           jump =c(.005, .005   ,.05    ,.001   ,3.3   ,.3),
           lower =c(-5 , -1     , 0     ,-4    , 50   ,20),
           upper =c( 5, .05     , 9     , 4     ,450   ,180),
           prior = Prior ,
           var0 = NULL, wvar0 = NULL, n0 = NULL,
           niter = 1000,
           outputlength =100,
           burninlength =0,     #updatecov = 2, covscale = (2.4^2)/5,
           verbose = TRUE)
```

About the Author

Zachary is trained in mathematics and statistics. He holds a bachelor's degree in mathematics from Youngstown State University and a master's degree in Applied Statistics from the University of North Carolina at Charlotte. He also plays guitar. He lives with his beautiful wife Rebecca and loyal canine companion Snickers (A.K.A Doodle).