DRAFT - DO NOT CITE OR QUOTE

EPA/600/R-10/038C
January 2010
Agency/Interagency Review Draft

EPA's Response to "Health Risks from Dioxin and Related Compounds Evaluation of the EPA Reassessment" Published by the National Research Council of the National Academies

NOTICE

THIS DOCUMENT IS AN AGENCY/INTERAGENCY REVIEW DRAFT. It has not been formally released by the U.S. Environmental Protection Agency and should not at this stage be construed to represent Agency policy. It is being circulated for comment on its technical accuracy and policy implications.

National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Cincinnati, OH

DISCLAIMER

This document is distributed solely for the purpose of pre-dissemination peer review under applicable information quality guidelines. It has not been formally disseminated by EPA. It does not represent and should not be construed to represent Agency determination or policy. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

CONTENTS

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

CONTENTS (continued)

LIST OF TABLES

LIST OF TABLES (continued)

LIST OF TABLES (continued)

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

LIST OF ABBREVIATIONS AND ACRONYMS

2,4,5-T	2,4,5-trichlorophenoxyacetic acid
2,4-D	2,4-dichlorophenoxyacetic acid
AA	ascorbic acid
ACOH	acetanilide-4-hydroxylase
AHH	aryl hydrocarbon hydroxylase
AhR	aryl hydrocarbon receptor
AhR-/-	AhR-deficient
AIC	Akaike Information Criterion
ANL	Argonne National Laboratory
ANOVA	analysis of variance
APE	airborne particulate extract
ASAT	aspartate aminotransferase
AUC	area under the curve
bHLH-PAS	basic helix-loop-helix, Per-Arnt-Sim
B_{max}	equilibrium maximum binding capacity
BMD	benchmark dose
BMDL	benchmark dose lower confidence bound
BMDS	Benchmark dose software
BMI	body mass index
BMR	benchmark response
BPS	balanopreputial separation
BROD	benzyloxy resoufin-O-deethylase
b-TSH	blood thyroid-stimulating hormone
BW	body weight
C	cerebellum
CADM	concentration- and age-dependent elimination model
Cc	cerebral cortex
CI	confidence interval
CSAF	chemical-specific adjustment factor
CSLC	cumulative serum lipid concentration
Cx	connexin
CYP	cytochrome P450
D_a:HED	ratio of administered dose to HED
DEN	diethylnitrosamine
df	degrees of freedom
DLC	dioxin-like compound
DRE/XRE	dioxin/xenobiotic response elements
DRL	differential reinforcement of low rate
DSA	delayed spatial alteration
E_2	17β-estradiol
ED_x	effective dose eliciting x percent response
EGFR	epidermal growth factor receptor

EPA	Environmental Protection Agency
ER	estrogen receptor
EROD	7-ethoxyresorufin-O-deethylase
ERα	estrogen receptor alpha
EU	European Union
FFA	free fatty acid
FR	fixed-ratio
FSH	follicle stimulating hormone
FT4	free thyroxine
GD	gestation day
GSH	glutathione stimulating hormone
GSH-Px	glutathione stimulating hormone peroxidase
GST	glutathione-*S*-transferase
H	hippocampus
HCH	hexachlorocyclohexane
HED	human equivalent dose
HQ	hazard quotient
HR	hazard ratio
Hsp90	heat shock protein 90
IARC	International Agency for Research on Cancer
IGF	insulin-like growth factor
IL	interleukin
ILSI	International Life Sciences Institute
i.p.	intraperitoneal
IRIS	Integrated Risk Information System
KABS	oral absorption parameters
LASC	lipid-adjusted serum concentration
LD_{50}	lethal dose eliciting x percent response
LED	lower confidence effective dose
LED_x	lower bound of the 95% confidence interval on the dose that yields an x% effect
LH	luteinizing hormone
LOAEL	lowest-observed-adverse-effect level
$LOAEL_{HED}$	HED estimate based on LOAELs
LOEL	lowest-observed-adverse level
MCH	mean corpuscular hemoglobin
MCMC	Markov Chain Monte Carlo
MCV	mean corpuscular volume
MOA	mode of action
MOE	margin of exposure
MROD	7-methoxyresorufin-O-deethylase
MTD	maximum tolerated dose
NAS	National Academy of Sciences
NIOSH	National Institute for Occupational Safety and Health

NOAEL	no-observed-adverse-effect level
NOEL	no-observed-effect level
NRC	National Research Council
NTP	National Toxicology Program
OR	odds ratio
OSF	oral slope factor
PA	permeability x area
PAI2	plasminogen activator inhibitor 2
PBMC	peripheral blood mononuclear cells
PBPK	physiologically based pharmacokinetic
PCB	polychlorinated biphenyl
PCDD	polychlorinated dibenzo-*p*-dioxin
PCDF	polychlorinated dibenzofuran
PEPCK	phosphoenolpyruvate carboxykinase
PF	adipose tissue:blood partition coefficient
PHAH	polyhalogenated aromatic hydrocarbons
PK	pharmacokinetic
PND	postnatal day
POD	point of departure
pp	phosphotyrosyl protein
PRA	probabilistic risk assessment
PRE	body:blood partition coefficient
PROD	7-pentoxyresorufin-O-deethylase
RAR	retinoic acid receptor
REP	relative potency
RfC	reference concentration
RfD	reference dose
RL	reversal learning
RL	risk level
RR	rate ratios
RR	relative risk
RT-PCR	reverse transcription polymerase chain reaction
RXR	retinoid X receptor
S	saline
SA	superoxide anion
SAhRM	SRM for AhRs
S-D	Sprague-Dawley
SD	standard deviation
SIR	standardized incidence ratio
SMR	standardized mortality ratio
SOD	superoxide dismutase
SRBC	sheep red blood cell
SSB	single-strand break

SWHS	Seveso Women's Health Study
T4	thyroxine
TBARS	thiobarbituric acid-reactive substances
TCB	3,3',4,4'-tetrachlorobiphenyl
TCDD	2,3,7,8-Tetrachlorodibenzo-*p*-dioxin
TCP	2,4,5-trichlorophenol
TEF	toxicity equivalence factor
TEQ	toxicity equivalence
TGFα	transforming growth factor α
TK	toxicokinetic
TNF-α	tumor necrosis factor alpha
TOTTEQ	total toxicity equivalence
TSH	thyroid stimulating hormone
TT4	total thyroxine
TWA	time-weighted average
U.S. NRC	U.S. Nuclear Regulatory Commission
UDP	uridine diphosphate
UDPGT	UDP-glucoronosyl transferase
UED	upper confidence bound for the effective dose
UF	uncertainty factor
UF_A	interspecies extrapolation factor
UF_D	database factor
UF_H	human interindividual variability
UF_L	LOAEL-to-NOAEL UF
UF_S	subchronic-to-chronic UF
UGT	UDP-glucuronosyltransferase
UGT1	uridine diphosphate glucuronosyltransferase I
Vd	volume of distribution
WHO	World Health Organization
ZS@Z	zero slope at zero dose

PREFACE

This document contains the U.S. Environmental Protection Agency's (EPA), Office of Research and Development, National Center for Environmental Assessment's response to the National Academy of Sciences (NAS, 2006) review of *Exposure and Human Health Reassessment of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) and Related Compounds* (U.S. EPA, 2003). Given that the key recommendations of the NAS refer to issues related to TCDD dose-response assessment and quantitative uncertainty analysis, EPA's response focuses on understanding human dose response for TCDD (both cancer and noncancer endpoints), and the feasibility of conducting quantitative uncertainty analysis in TCDD dose-response assessment.

AUTHORS, CONTRIBUTORS, AND REVIEWERS

PRIMARY AUTHORS

National Center for Environmental Assessment, U.S. Environmental Protection Agency, Cincinnati, OH

Belinda Hawkins
Janet Hess-Wilson
Glenn Rice
Jeff Swartout
Linda K. Teuschler

CONTRIBUTING AUTHORS

National Center for Environmental Assessment, U.S. Environmental Protection Agency, Cincinnati, OH

Scott Wesselkamper
Michael Wright
Bette Zwayer

National Health and Environmental Effects Research Laboratory, U.S. Environmental Protection Agency, Research Triangle Park, NC

Hisham El-Masri

Argonne National Laboratory, Argonne, IL

Margaret MacDonell

Emory University, Atlanta, GA

Kyle Steenland

Resources for the Future, Washington, DC

Roger M. Cooke

University of Montreal, Montreal, Canada

Kannan Krishnan

CONTRIBUTORS

National Center for Environmental Assessment, U.S. Environmental Protection Agency, Washington, DC

Karen Hogan
Leonid Kopylev

CONTRIBUTORS continued

Argonne National Laboratory, Argonne, IL
- Maryka H. Bhattacharyya
- Andrew Davidson
- Mary E. Finster
- David P. Peterson

Clark University, Worcester, MA
- Dale Hattis

Colorado State University, Fort Collins, CO
- Raymond Yang

Bruce Allen Consulting, Chapel Hill, NC
- Bruce C. Allen

ICF International, Durham, NC
- Robyn Blain
- Rebecca Boyles
- Patty Chuang
- Cara Henning
- Baxter Jones
- Penelope Kellar
- Mark Lee
- Nikki Maples-Reynolds
- Amalia Marenberg
- Garrett Martin
- Margaret McVey
- Chandrika Moudgal
- Bill Mendez
- Ami Parekh
- Andrew Shapiro
- Audrey Turley

National Toxicology Program, Research Triangle Park, NC
- Michael Devito

Penn State University, University Park, PA
- Jack P. Vanden Heuvel

Risk Sciences International, Ottawa, Ontario
- Jessica Dennis
- Dan Krewski
- Greg Paoli
- Salomon Sand
- Natalia Shilnikova
- Paul Villenueve

University of California-Berkeley, Berkeley, CA
- Brenda Eskenazi

University of California-Irvine, Irvine, CA
- Scott Bartell

University of Montreal; BioSimulation Consulting, Newark, DE
- Claude Emond

INTERNAL REVIEWERS

National Center for Environmental Assessment, U.S. Environmental Protection Agency

Glinda Cooper, Washington, DC
Ila Cote, Research Triangle Park, NC
Lynn Flowers, Washington, DC
Martin Gehlhaus, Washington, DC
Kate Guyton, Washington, DC
Samantha Jones, Washington, DC
Matthew Lorber, Washington, DC
Eva McLanahan, Research Triangle Park, NC
Susan Rieth, Washington, DC
Reeder Sams, Research Triangle Park, NC
Paul Schlosser, Research Triangle Park, NC
Jamie Strong, Washington, DC
John Vandenberg, Research Triangle Park, NC
Paul White, Washington, DC

ACKNOWLEDGMENTS

National Center for Environmental Assessment, U.S. Environmental Protection Agency

Jeff Frithsen, Washington, DC
Annette Gatchett, Cincinnati, OH
Andrew Gillespie, Cincinnati, OH
Marie Nichols-Johnson, Cincinnati, OH
Maureen Johnson, Washington, DC
Peter Preuss, Washington, DC
Linda Tuxen, Washington, DC

Colorado State University, Fort Collins, CO

William H. Farland

ECFlex, Inc., Fairborn, OH

Dan Heing
Heidi Glick
Amy Prues
Lana Wood

IntelliTech Systems, Inc., Fairborn, OH

Cris Broyles
Luella Kessler
Debbie Kleiser
Stacey Lewis
Linda Tackett

National Institute of Environmental Health Sciences, Research Triangle Park, NC

Linda S. Birnbaum
Christopher J. Portier

National Toxicology Program, Research Triangle Park, NC

Nigel Walker

2009 Dioxin Workshop Participants

EXECUTIVE SUMMARY

OVERVIEW

The U.S. Environmental Protection Agency (EPA) is committed to the development of risk assessment information of the highest scientific integrity for use in protecting human health and the environment. Scientific peer review is an integral component of the process EPA uses to generate high quality toxicity and exposure assessments of environmental contaminants. To this end, EPA asked the National Academy of Sciences (NAS) to review its comprehensive human health risk assessment external review draft entitled, *Exposure and Human Health Reassessment of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) and Related Compounds* (U.S. EPA, 2003; "2003 Reassessment"). This current document, *EPA's Response to "Health Risks from Dioxin and Related Compounds: Evaluation of the EPA Reassessment" Published by the National Research Council of the National Academies,* directly and technically responds to key comments and recommendations pertaining to TCDD dose-response assessment published by the NAS in their review (NAS, 2006a). This document only addresses issues pertaining to TCDD dose-response assessment.

In response to the recommendations presented in the 2006 NAS review, EPA Administrator Jackson released EPA's "*Science Plan for Activities Related to Dioxins in the Environment*" ("Science Plan") on May 26, 2009.[1] There are five key components of the Science Plan that pertain to EPA's response to the NAS comments on TCDD dose-response assessment:

1. EPA will release a draft report that responds to the recommendations and comments included in the NAS review of EPA's 2003 Reassessment.
    - EPA's National Center for Environment Assessment (NCEA), in the Office of Research and Development (ORD), will prepare a limited response to key comments and recommendations in the NAS report (draft response to comments report).
    - The draft response will focus on dose-response issues raised by the NAS and will include an analysis of relevant new key studies.
    - The draft response will be provided for public review and comment and independent external peer review.

[1]Available online at http://www.epa.gov/dioxin/scienceplan.

- The draft response will also include an evaluation of some of the significant recommendations that are difficult for EPA to address given the current state of science, and a detailed rationale for these conclusions.
- The peer review will be conducted by EPA's Science Advisory Board (SAB), an independent review body chartered under the Federal Advisory Committee Act.
- The draft response to comments report will be completed and released for public and peer review by December 31, 2009.

2. EPA will provide the draft response to comments report for internal and external review.
3. The SAB will review the science content of the response to comments report.
4. EPA will review impacts of the draft response to comments report on its 2003 Reassessment.
5. EPA will release the final response to comments report and focus on completion of the 2003 Reassessment.

This document responds to and addresses key NAS comments relating to TCDD dose-response assessment. Three separate EPA activities address additional NAS comments pertaining to toxicity equivalence factors (TEFs) and background exposure levels. Information on the application of the dioxin TEFs is published elsewhere by EPA for both ecological (U.S. EPA, 2008a) and human health (U.S, EPA, 2009a) risk assessment. As a consequence, EPA does not directly address TEFs herein, but makes use of the concept of toxicity equivalence (TEQ)[2] as applicable to the analysis of exposure dose in epidemiologic studies. Furthermore, addressing the NAS recommendations pertaining to the assessment of human exposures to TCDD and other dioxins, information on updated background levels of dioxin in the U.S. population has been recently reported by EPA (Lorber et al., 2009).

The NAS identified three key recommendations requiring substantial improvement to support a scientifically robust characterization of human responses to exposures to TCDD. These three key areas are (1) improved transparency and clarity in the selection of key data sets for dose-response analysis, (2) further justification of approaches to dose-response modeling for cancer and noncancer endpoints, and (3) improved transparency, thoroughness, and clarity in quantitative uncertainty analysis. NAS also encouraged EPA to calculate a Reference Dose (RfD), and provided numerous specific comments on various aspects of EPA's 2003

[2]Toxicity equivalence (TEQ) is the product of the concentration of an individual dioxin like compound in an environmental mixture and the corresponding TCDD TEF for that compound. These products are summed to yield the TEQ of the mixture.

Reassessment. The three key recommendations specifically pertain to dose-response assessment and uncertainty analysis. Therefore, EPA's response to the NAS in this document is focused on these issues. EPA thoroughly considered the recommendations of the NAS and responds with scientific and technical evaluation of TCDD dose–response data via:

- an updated literature search that identified new TCDD dose-response studies (see Section 2);
- a kickoff workshop that included the participation of external experts in TCDD health effects, toxicokinetics, dose-response assessment and quantitative uncertainty analysis; these experts discussed potential approaches to TCDD dose-response assessment and considerations for EPA's response to NAS (see Appendix A);
- detailed study inclusion criteria and processes for the selection of key studies (see Section 2.3) and epidemiologic and animal bioassay data for TCDD dose-response assessment (see Section 2.4.1/Appendix B and Section 2.4.2, respectively);
- kinetic modeling to quantify appropriate dose metrics for use in TCDD dose-response assessment (see Section 3);
- dose-response modeling for all appropriate noncancer and cancer data sets (see Section 4.2/Appendix E and Section 5.2.3/Appendix F, respectively);
- thorough and transparent evaluation of the selected TCDD data for use in the derivation of an RfD and an oral slope factor (OSF) (see Sections 4.2 and 5.2.3, respectively);
- the development of an RfD (see Section 4.3);
- the development of a revised OSF (see Section 5.3) with an updated cancer weight of evidence determination for TCDD based on EPA's 2005 *Cancer Guidelines* (U.S. EPA, 2005) (see Section 5.1.2);
- consideration of nonlinear dose-response approaches for cancer, including illustrative RfDs for cancer precursor events and tumors (see Section 5.2.3.4) ; and
- discussion of the feasibility and utility of quantitative uncertainty analysis for TCDD dose-response assessment (see Section 6).

Each of the activities listed above is briefly described in this Executive Summary, and is described in detail in the related sections of this document.

PRELIMINARY ACTIVITIES UNDERTAKEN BY EPA TO ENSURE THAT THIS TECHNICAL RESPONSE REFLECTS THE CURRENT STATE-OF-THE-SCIENCE

As part of the development of this document, EPA undertook two activities that included public involvement: an updated literature search and a scientific expert workshop. The adverse health effects associated with TCDD exposures are documented extensively in epidemiologic and toxicologic studies. As such, the database of relevant information pertaining to the dose-response assessment of TCDD is vast and constantly expanding. Responding directly to the NAS recommendation to use the most current and up-to-date scientific information related to TCDD, EPA, in collaboration with Argonne National Laboratory (ANL), developed an updated literature database of peer-reviewed studies on TCDD toxicity, including in vivo mammalian dose-response studies and epidemiologic studies. An initial literature search for studies published since the 2003 Reassessment was conducted to identify studies published between the year 2000 and October 31, 2008. EPA published the initial literature search results in the Federal Register in November 2008 and invited the public to review the list and submit additional peer-reviewed relevant studies. Additional studies identified by the public and through continued work on this response have been incorporated into the final set of studies for TCDD dose-response assessment (updated through October 2009). EPA believes that the implementation of this rigorous search strategy ensures that the most current and relevant studies were considered for the technical response to NAS and TCDD dose-response assessment included herein.

To assist in responding to the NAS, EPA, in collaboration with ANL, convened a scientific expert workshop ("Dioxin Workshop") in February 2009 that was open to the public. The primary goals of the Dioxin Workshop were to identify and address issues related to the dose-response assessment of TCDD and to ensure that EPA's response to the NAS focused on the key issues while reflecting the most meaningful science. EPA and ANL assembled expert scientists and asked them to identify and discuss the technical challenges involved in addressing the NAS comments, discuss approaches for addressing these key recommendations, and to assist in the identification of important published and peer-reviewed literature on TCDD. The workshop was structured into seven scientific topic sessions as follows: (1) quantitative dose-response modeling issues, (2) immunotoxicity, (3) neurotoxicity and nonreproductive endocrine effects, (4) cardiovascular toxicity and hepatotoxicity, (5) cancer, (6) reproductive and developmental toxicity, and (7) quantitative uncertainty analysis of dose-response. External

co-chairs (i.e., scientists who were not members of EPA or ANL) were asked to facilitate the sessions and then prepare summaries of discussions occurring in each session. The session summaries formed the basis of a final workshop report (U.S. EPA, 2009c, Appendix A of this document). Some of the key outcomes from the workshop include the following recommendations:

- to further develop study selection criteria for evaluating the suitability of developing dose-response models based on animal bioassays and human epidemiologic studies;
- to use kinetic modeling to identify relevant dose metrics and dose conversions between test animal species and humans, and between human internal dose measures and human intakes;
- to consider newer human or animal (e.g., NTP, 2006) publications when evaluating quantitative dose-response models for cancer;
- to consider both linear and nonlinear modeling in the cancer dose-response analysis.

The discussions held during the Dioxin Workshop helped inform, guide, and focus EPA's response to NAS.

EPA'S APPROACH TO CONSIDERING TRANSPARENCY AND CLARITY IN THE SELECTION OF KEY STUDIES AND DATA SETS FOR DOSE-RESPONSE MODELING

One of the key NAS recommendations to EPA was to utilize a clear and transparent process for the selection of key studies and data sets for dose-response assessment. EPA agrees with the NAS and believes that clear delineation of the study selection process and decisions regarding key studies and data sets will facilitate communication of critical decisions made in the TCDD dose-response assessment. EPA developed detailed processes and TCDD-specific criteria for the selection of key dose-response studies. These criteria are based on common practices and current guidance for point of departure (POD) identification and RfD and OSF derivation while also considering issues specifically related to TCDD. Following the selection of key studies, EPA employed additional processes to further select and identify cancer and noncancer datasets from these key studies for use in dose-response analysis of TCDD.

Figure ES-1 presents EPA's study evaluation process for the epidemiologic studies considered for TCDD dose-response assessment, including specific study inclusion criteria (see

Section 2.3). EPA applied TCDD-specific epidemiologic study inclusion criteria to all epidemiologic studies published on TCDD and dioxin-like compounds (DLCs) that had been identified in the TCDD literature database (see Section 2.4.1, Appendix B). The studies were initially evaluated using five considerations (see Figure ES-1) that provide the most relevant kinds of information needed for quantitative human health risk analyses. Additionally, EPA examined whether the human exposures were primarily to TCDD and if the TCDD exposures could be quantified so that dose-response analyses could be conducted. Finally, EPA required that the effective dose be estimable: (1) for cancer, information is required on long-term exposures, and (2) for noncancer, information is required regarding the appropriate time window of exposure that is relevant for a specific, nonfatal health endpoint. Therefore, the study should include an appropriate latency period between TCDD exposure and the onset of the effect. Only studies meeting these criteria were included in EPA's TCDD dose-response analyses (see Section 2.4.3).

Figure ES-2 presents EPA's study evaluation process for mammalian bioassays considered for TCDD dose-response assessment, including the specific study inclusion criteria (see Section 2.3). EPA applied TCDD-specific in vivo mammalian bioassay study inclusion criteria to all of the bioassay studies of TCDD that had been identified in the TCDD literature database (see Section 2.4.2). After ascertaining that a study had been published in the peer-reviewed literature, EPA applied dose requirements to the lowest tested average daily doses in each study, with specific requirements for cancer (≤1 μg/kg-day) and noncancer (≤30 ng/kg-day) studies to ensure that only low-dose TCDD bioassays would be considered. These dose requirements were used to eliminate those studies that would not be selected for development of an RfD or an OSF because the lowest doses tested were too high relative to other TCDD bioassays. EPA also required that the bioassays exposed the animals via the oral route to TCDD only and that the purity of the TCDD was specified. Finally, the studies were evaluated using four considerations (see Figure ES-2) regarded as providing the most relevant information for development of quantitative human health risk analyses from animal bioassay data. Only the bioassay studies meeting these criteria and considerations were included in EPA's TCDD dose-response analyses (see Section 2.4.3).

Applying the study inclusion criteria for both epidemiologic and mammalian bioassay datasets resulted in a list of key noncancer and cancer studies that were considered for

quantitative dose-response analyses of TCDD. Endpoints from these studies that were not considered to be toxicologically relevant were eliminated from consideration (see Section 4.2.1.1, Appendix G). The study/endpoint dataset combinations from the remaining studies were then subjected to dose-response assessment, and PODs for use in developing RfDs or OSFs were identified. PODs included no-observed-adverse-effect levels (NOAELs), lowest-observed-adverse-effect levels (LOAELs) or lower bound benchmark dose levels (BMDLs). The most sensitive PODs were selected as candidates for derivation of the RfD and OSF.

USE OF KINETIC MODELING TO ESTIMATE TCDD DOSES

NAS recommended that EPA utilize state-of-the-science approaches to finalize the 2003 Reassessment. Although NAS concurred with EPA's use of first-order body burden models in the 2003 Reassessment, analyses of recent TCDD literature and comments by experts at the Dioxin Workshop suggested that the understanding of TCDD kinetics had increased significantly since the release of EPA's 2003 Reassessment. These advances led to the development of several pharmacokinetic models for TCDD (e.g., Emond et al., 2004, 2005, 2006; Aylward et al., 2005a) and resulted in EPA's incorporation of TCDD kinetics in the dose-response assessment of TCDD.

The evaluation of internal dose in exposed humans and other species is facilitated by an understanding of pharmacokinetics (i.e., absorption, distribution, metabolism, and excretion). TCDD pharmacokinetics are influenced by three distinctive features: (1) TCDD is highly lipophilic, (2) TCDD is slowly metabolized, and (3) TCDD induces binding proteins in the liver. The overall impact of these factors results in preferential storage of TCDD in adipose tissue, a long half-life of TCDD in blood due to slow metabolism, and sequestration in liver tissue when binding induction becomes significant. As these kinetic features control target tissue levels of dioxin, they become important in relating toxicity in animals to possible effects in humans.

Consideration of pharmacokinetic mechanisms is critical to the selection of the dose metrics of relevance to dose-response modeling of TCDD. Earlier assessments for TCDD, including the 2003 Reassessment, used estimates of body burden as the dose metric for extrapolation between animals and humans. These body burden calculations used a simple one-compartment kinetic model based on the assumption of a first-order decrease in the levels of administered dose as a function of time. However, the assumption of a constant half-life value

for the clearance of TCDD from long-term or chronic exposure is not well-supported biologically given the dose-dependant elimination observed in rodents and humans. The dynamic disposition and redistribution of TCDD between blood, fat, and liver as a function of time and dose is better described using biologically-based models. Additionally, these models provide estimates for other dose metrics (e.g., serum or tissue levels) that are more biologically relevant to response than body burden estimated based on an assumption of first-order elimination over time.

EPA considered the following possible dose metrics for TCDD: administered dose, absorbed dose, first-order body burden, serum or whole blood concentration, tissue concentration, and functional-related metrics of relevance to the mode of action (MOA) (e.g., receptor occupancy) (see Section 3.3.4.1). After careful evaluation of these dose metrics, EPA chose to use TCDD concentration in whole blood as the dose metric for assessing TCDD dose response in this document; blood concentration reflects both the body burden and the dose to target tissues. EPA used the time-weighted average whole-blood concentration over the relevant exposure periods for all continuous dosing protocols, dividing the area under the time-course concentration curve (AUC) by the exposure duration.[3]

Several biologically-based kinetic models for TCDD exist in the literature. The more recent pharmacokinetic models explicitly characterize the concentration-dependent elimination of TCDD (Carrier et al., 1995a, b; Emond et al., 2004, 2005, 2006; Aylward et al., 2005a). The biologically-based pharmacokinetic models describing the concentration-dependent elimination (i.e., the pharmacokinetic models of Aylward et al. [2005a] and Emond et al. [2005, 2006]) are relevant for application to simulate the TCDD dose metrics in humans and animals exposed via the oral route. The rationale for considering the application of the Aylward et al. (2005a) and Emond et al. (2004, 2005, 2006) models was largely based on the fact that both models reflect research results from recent peer-reviewed publications, and both models are formulated with dose-dependent hepatic elimination consistent with the physiological understanding of TCDD kinetics.

[3]For the Seveso cohort, which had a high single exposure followed by low-level background exposures leading to a gradual decline in the internal TCDD concentrations, EPA estimated dose as the mean of the peak exposure and the average exposure over a defined critical exposure window (see Section 4.2.2).

Of the two selected models, the pharmacokinetic model developed by Emond et al. (2004, 2005, 2006) is more physiologically-based, as compared to the Aylward et al. (2005a) model, and models the blood compartment directly in the rat, mouse, and human; there are also gestational and life-time nongestational forms of the Emond PBPK model. Therefore, in this document, EPA chose the Emond rodent physiologically-based pharmacokinetic (PBPK) model to estimate blood TCDD concentrations based on administered doses (see Section 3.3.4, Appendix C). EPA assumed that the same blood TCDD levels that led to effects in animals would also lead to effects in humans; therefore, the Emond human PBPK model was used to estimate the lifetime average daily oral doses (consistent with the chronic RfD and OSF) that would correspond to the blood TCDD concentrations estimated to have occurred during the animal bioassays. EPA used the same Emond human PBPK model to estimate the lifetime average daily doses that would correspond to the TCDD blood or tissue concentrations reported in the epidemiological studies. These estimates are the Human Equivalent Doses (HEDs) that are used to develop candidate RfDs and OSFs for TCDD.

Because TCDD elimination is inducible in the Emond model, ratios of daily averaged intake to long-term blood concentrations are not linear. OSFs based on internal dose measures (i.e., blood or fat concentrations) will be linear only with respect to blood or fat TCDD concentrations; however, ingested TCDD doses are not linear with TCDD fat or blood concentrations in the Emond PBPK model. Thus, an OSF that is linear with TCDD in the fat is not linear with the ingested TCDD dose. A consequence, then, of using the Emond PBPK model is that risk-specific TCDD intake rates corresponding to target risk levels need to be tabulated for use in human health risk assessment.

DERIVATION OF AN RfD FOR TCDD

The NAS specifically recommended that EPA derive an RfD for TCDD. Through a transparent study selection process, EPA identified key studies from both human epidemiologic and animal bioassay studies. To select candidate PODs for its RfD methodology, EPA applied additional processes to the key human epidemiologic studies and animal bioassays. Figure ES-3 (exposure-response array) shows all of the candidate PODs graphically in terms of human-equivalent intake (ng/kg-day). The human study endpoints are shown at the far left of the figure

and, to the right, the rodent endpoints are arranged by the following study categories: less than 1 year, greater than 1 year, reproductive, and developmental.

For each noncancer epidemiologic study that EPA selected as key, EPA evaluated the dose-response information developed by the study authors to determine whether the study provided noncancer effects and TCDD-relevant exposure data for a toxicologically-relevant endpoint. If such data were available, EPA identified a NOAEL or LOAEL as a candidate POD. Then, EPA used the Emond human PBPK model to estimate the continuous oral daily intake (ng/kg-day) that would lead to the relevant blood TCDD concentrations associated with the candidate POD that could be used in the derivation of an RfD. If all of this information was available, then the result was included as a candidate POD.

Through this process, EPA identified health effects from the following four epidemiologic studies to be considered as the basis for the RfD: Eskenazi et al. (2002) (reproductive—increased length of menstrual cycle), Alaluusua et al. (2004) (developmental—tooth development), Mocarelli et al. (2008) (reproductive—decreased sperm concentrations), and Baccarelli et al. (2008) (developmental—increased thyroid-stimulating hormone levels in neonates). All four studies are from the Seveso cohort, whose members were exposed environmentally to high peak concentrations of TCDD as a consequence of an industrial accident. This complicated the estimation of average daily doses associated with these specific endpoints, however EPA was able to calculate candidate PODs for derivation of an RfD from each of these human studies (see Section 4.2.3).

Figure ES-4 summarizes the strategy employed for identifying and selecting candidate PODs from the key animal bioassays EPA identified for use in noncancer dose-response analysis of TCDD (see Section 4.2.4). For each noncancer endpoint, EPA first evaluated the toxicological relevance of each endpoint, rejecting those judged not to be relevant for RfD derivation. Next, initial PODs (NOAELs, LOAELs, and BMDLs) based on the first-order body burden metric, and expressed as continuous human-equivalent oral daily doses (HEDs), were determined for all relevant endpoints. Because there were very few NOAELs, and BMDL modeling was largely unsuccessful due to data limitations, the next stage of evaluation was carried out using LOAELs only. Endpoints not observed at the LOAEL (i.e., reported at higher doses) with BMDLs greater than the LOAEL were eliminated from further analysis, as they would not be considered as candidates for the final POD on either a BMDL or NOAEL/LOAEL

basis (i.e. the POD would be higher than the PODs of other relevant endpoints). In addition, all endpoints with HEDs for LOAELs ($LOAEL_{HED}$s) beyond a 100-fold range of the lowest identified $LOAEL_{HED}$ were eliminated from further consideration, as they would not be potential POD candidates either (i.e. the POD would be higher than the PODs of other relevant endpoints). For the remaining endpoints, EPA then determined final potential PODs (NOAELs, LOAELs, and BMDLs) based on TCDD blood concentrations obtained from the Emond rodent PBPK models. HEDs were then estimated for each of these PODs using the Emond human PBPK model. From these HEDs, a POD_{HED} was selected for each study as the basis for the candidate RfD, to which appropriate uncertainty factors were applied following EPA guidelines. The resulting candidate RfDs were then considered in the final selection process for the RfD. Other endpoints occurring at slightly higher doses representing additional effects associated with TCDD exposure (beyond the 100-fold LOAEL range) were evaluated, modeled, and included in the final candidate RfD array to examine endpoints not evaluated by studies with lower PODs. In addition, BMD modeling based on administered dose was performed on all endpoints for comparison purposes.

For selection of the POD to serve as the basis of the RfD, EPA gave the epidemiologic studies the highest consideration because human data are preferred in the derivation of an RfD, given that the underlying epidemiologic and animal bioassay data all met EPA's selection criteria. This preference for epidemiologic study data also is consistent with reccomendations of panelists at the Dioxin Workshop. The lower end of the candidate POD distribution is dominated by mouse studies, with mouse studies comprising six of the first seven animal bioassays yielding the lowest PODs. EPA has less confidence in the values derived from these mouse bioassays than the values derived from rat and human studies. The primary reason that the mouse $LOAEL_{HED}$s are low is the large toxicokinetic interspecies extrapolation factor applied to mouse data in the Emond PBPK model. In addition, each of these first seven rodent studies has other qualitative limitations and uncertainties that make them poor candidates as the basis for the RfD.

Most of the other rodent studies yielding POD values higher than the first seven animal bioassay studies and lower than the human studies of Mocarelli et al (2008) and Baccarelli et al. (2008) are of small size, using 10 or fewer animals per dose group and are considered too uncertain on which to base the final RfD. However, two of the rat bioassays—Bell et al (2007) and NTP (2006)—were very well designed and conducted, using 30 or more animals per dose

group (see Section 4.3.4). Bell et al (2007) evaluated several reproductive and developmental endpoints initiating TCDD exposures well before mating and continuing them through gestation. NTP (2006) is the most comprehensive evaluation of TCDD chronic toxicity in rodents to date, evaluating dozens of endpoints at several time points in all major tissues. It is EPA's judgement that the toxicokinetic extrapolation of the results of these two studies to humans, however, is still less certain than the use of human data. Despite the overall strength of the Bell and NTP studies, EPA considers the human data to be a better basis for the TCDD RfD.

The most relevant human PODs are based on the Baccarelli et al. (2008) and Mocarelli et al. (2008) studies, which exhibited similar LOAELs of 0.024 and 0.020 ng/kg-day, respectively. For Baccarelli et al. (2008), EPA defined a LOAEL as the group mean of 39 ppt TCDD in neonatal plasma which corresponds to thyroid-stimulating hormone (TSH) values above 5 μU/mL. Baccarelli et al. did not estimate the equivalent oral intake associated with TCDD serum concentrations, rather they provided neonatal serum TCDD concentrations for the groups above and below 5 μU/mL. EPA estimated the maternal intake at the LOAEL from a maternal serum-TCDD/TSH regression model presented in Baccarelli et al. by estimating the maternal TCDD lipid adjusted serum concentration (LASC) at which neonatal TSH exceeded 5 μU/mL. EPA then used the Emond PBPK model to estimate the continuous daily TCDD intake that would result in this TCDD LASC. The resulting predicted maternal daily intake rate established the LOAEL (0.024 ng/kg-day). EPA did not defined a NOAEL because it is not clear what maternal intake should be assigned to the group below 5 μU/mL.

For Mocarelli et al. (2008), EPA defined a LOAEL as the lowest exposed group mean of 68 ppt (1^{st}-quartile) corresponding to decreased sperm concentrations, decreased motile sperm counts, and decreased serum estradiol in men who were 1–9 years old at the time of the Seveso accident (initial TCDD exposure event). TCDD LASC levels were measured within approximately one year of the initial exposure event. Because effects were only observed in men who were under 10 years of age at the time of exposure, EPA assumed a maximum 10-year critical exposure window for elicitation of these effects. EPA has estimated a continuous daily oral intake of 0.020 ng/kg-day associated with the designated LOAEL from the lowest exposure group (68 ppt), (see Section 4.2.3.2). The reference group is not designated as a NOAEL because there is no clear zero-exposure measurement for any of these endpoints, particularly considering the contribution of background exposure to DLCs, which futher complicates the

interpretation of the reference group response as a true "control" response (see discussion in Section 4.4). However, males less than 10 years old can be designated as a sensitive population by comparison to older males who were not affected.

The two human studies, Baccarelli et al. (2008) and Mocarelli et al. (2008), have similar LOAELs of 0.024 and 0.020 ng/kg-day, respectively. Together, these two studies constitute the best foundation for establishing a POD for the RfD, and are designated as co-principal studies. Therefore, increased TSH in neonates (Baccarelli et al., 2008) and male reproductive effects (decreased sperm count and motility, increased estradiol) are designated as cocritical effects. Although the exposure estimate used in determination of the LOAEL for Mocarelli et al. (2008) is more uncertain than the Baccarelli et al. (2008) exposure estimate, the slightly lower LOAEL of 0.020 ng/kg-day from Mocarelli et al. is designated as the POD.

EPA used a composite UF of 30 for both studies. EPA applied a factor of 10 for UF_L to account for lack of a NOAEL. EPA also applied a factor of 3 ($10^{0.5}$) for UF_H to account for human interindividual variability because the effects were elicited in sensitive populations. A further reduction to 1 was not made because the sample sizes in these two epidemiologic studies were relatively small, which, combined with uncertainty in exposure estimation, may not fully capture the range of interindividual variability. The resulting RfD for TCDD in standard units is 7×10^{-10} mg/kg-day (Table 4-7 details the RfD derivation).

WEIGHT-OF-EVIDENCE STATEMENT FOR CARCINOGENICITY

The NAS recommended that EPA update its cancer classification for TCDD and the weight-of-evidence (WOE) statement to reflect the current state of the science and incorporate the latest EPA Cancer Guidelines (U.S. EPA, 2005). Several notable new studies addressing TCDD's carcinogenic potential have been published since the release of EPA's 2003 Reassessment, including several new studies of the Seveso epidemiologic cohort and an NTP 2-year cancer bioassay in female rats (NTP, 2006).

Under the 2005 *Guidelines for Carcinogen Risk Assessment* (U.S. EPA, 2005) TCDD is characterized as *carcinogenic to humans*, based on the available data as of 2009 (see Section 5.1.2). When evaluating the carcinogenic potential of a compound, EPA employs a WOE approach in which all available information is evaluated and considered. In the case of TCDD, EPA based the classification on numerous lines of evidence, including: multiple

occupationally- and accidentally-exposed epidemiologic cohorts showing an association between TCDD exposure and certain cancers or increased mortality from all cancers; extensive evidence of carcinogenicity at multiple tumor sites in both sexes of multiple species of experimental animals; consensus that the mode of TCDD's carcinogenic action in animals involves aryl hydrocarbon receptor (AhR)-dependent key precursor events and proceeds through modification of one or more of a number of cellular processes; the human AhR and rodent AhR are similar in structure and function, and human and rodent tissue and organ cultures respond to TCDD in a similar manner and at similar concentrations; and general scientific consensus that AhR activation is anticipated to occur in humans and to progress to tumors.

DERIVATION OF CANDIDATE OSFs FROM EPIDEMIOLOGIC STUDIES AND ANIMAL BIOASSAYS

There are several well-studied occupationally-exposed epidemiologic cohorts showing an association between TCDD and increased all-cancer mortality, and several epidemiologic cohorts exposed to TCDD as a consequence of industrial accidents showing an association between TCDD and cancer or cancer mortality (see Section 5.2.3.1). The 2003 Reassessment included cancer dose-response analyses based on the following three occupational cohorts: the NIOSH cohort, an occupational cohort subject to chronic TCDD exposures (Steenland et al., 2001); the Hamburg cohort, an occupational cohort also subject to chronic TCDD exposures (Becher et al., 1998); and the BASF cohort, an occupational cohort subject to peak TCDD exposures through clean-up following an industrial accident (Ott and Zober, 1996). In this document, EPA determined that each of these studies met the epidemiologic study inclusion criteria. Thus, after further evaluating the OSFs presented in the 2003 Reassessment for these three studies, EPA accepted those OSF estimates and retained them as candidate OSFs in this document. EPA also determined that two additional studies met the epidemiologic study inclusion criteria: Cheng et al. (2006) (NIOSH cohort) and Warner et al. (2002) (Seveso cohort, a cohort exposed environmentally as a consequence of an industrial accident). EPA was unable to derive a credible OSF from the data presented by Warner et al. (2002) but did derive an OSF from Cheng et al. (2006), as detailed in Text Box ES-1. In Table ES-1, EPA presents estimates of OSFs for specific TCDD intake rates based on target risk levels of 1×10^{-3}, 1×10^{-4}, 1×10^{-5}, 1×10^{-6}, and 1×10^{-7} based on Cheng et al. (2006).

Text Box ES-1
OSF Calculations Using Cheng et al. (2006) Information

Following the estimation of dose using the AUC values developed from the kinetic modeling, Cheng and colleagues derived dose-response estimates for the NIOSH cohort. For exposures lagged 15 years, the regression coefficient of the linear slope derived by Cheng et al. (2006) was 3.3×10^{-6} per ppt-year lipid-adjusted serum TCDD (the standard error of this regression coefficient was 1.4×10^{-6}). The upper 5% of the exposure range (individuals ≥252,950 ppt-year lipid adjusted serum TCDD) was excluded in estimating this slope. Because this exclusion reduces the upper portion of the response where the slope is shallow, this likely better represents the slope in the region of the curve where the fatal cancer risk is increasing with dose, which is the equivalent of dropping the highest dose in an animal bioassay.

To develop an OSF for TCDD, EPA used information from Cheng et al. (2006) in its calculations as follows:

- Upper 95th percentile slope (β_{95}) of 6.0×10^{-6} per ppt-year lipid adjusted serum TCDD.
- Background cancer mortality risk estimate (R_0) of 0.112.
- Risk in the exposed group associated with a 1% extra risk of fatal cancer (R_{exp}) of 0.12088.
- Incremental cancer mortality risk in the exposed population based on a 1% extra risk (R_D) of 8.9×10^{-3} using the equation, $R_D = R_{exp} - R_0$
- Cumulative TCDD concentration in the fat compartment for a 1% extra risk (AUC_{01}) using the following formula corresponding to equations used in Cheng et al. (2006):

  $$AUC_{01} = LN(R_D + R_0/R_0)/\beta_{95}$$

  Cheng et al. (2006) appear to assume that the TCDD concentration in fat is the same as ppt-yr lipid adjusted serum concentration.
- OSF associated with 1% extra risk [$OSF(AUC_{01})$], calculated to be 7.92×10^{-7} by dividing 0.01 by the AUC_{01}. $OSF(AUC_{01})$ is linear with the TCDD concentration in fat. Ingested TCDD doses, however, are not linear with the predicted TCDD fat concentrations in the Emond pharmacokinetic model. Thus, the $OSF(AUC_{01})$ that is linear with TCDD in the fat is not linear with ingested TCDD dose.

EPA calculated estimates of OSFs for specific TCDD intake rates based on target risk levels (RLs) of 1×10^{-3}, 1×10^{-4}, 1×10^{-5}, 1×10^{-6}, and 1×10^{-7}, using the following calculations:

- Area under the TCDD fat concentration curve associated with a target risk level (AUC_{RL}) (ppt-yr) [i.e., $RL/7.92 \times 10^{-7}$ $(ppt\text{-}yr)^{-1}$].
- Lifetime averaged concentration of TCDD in the fat compartment associated with the target risk level (FAT_{RL}) (ng/kg). The AUC_{RL} estimates were then further divided by 70 years to identify (FAT_{RL}) (ng/kg). This step essentially reverses the integration undertaken to calculate AUC_{01}.
- Continuous daily TCDD intake (D_{RL})(ng/kg-day) associated with a target risk level over a lifetime. Using the Emond pharmacokinetic model, EPA estimated the D_{RL} necessary to achieve the FAT_{RL}.
- Oral slope factor at the target risk level (OSF_{RL}) (per mg/kg-day). At the target risk levels, the associated OSF_{RL}s range from 3.7×10^5 to 1.3×10^6 per mg/kg-day. These are calculated as $OSF_{RL} = RL/D_{RL} \times 10^6$.

EPA also identified candidate OSFs for TCDD from key animal bioassays (see Section 5.2.3.2). Based on the inclusion criteria, EPA selected four key rodent cancer bioassays suitable for quantitative dose-response assessment. These included Kociba et al. (1978), NTP (1982), and Toth et al. (1979) that were evaluated in the 2003 Reassessment, and the new NTP (2006) rat chronic bioassay. EPA conducted dose-response modeling for each tumor type separately (individual tumor models) as well as composite tumor incidence dose estimates (multiple tumor models). The tumor types that EPA analyzed are shown in Table ES-2.

For each in vivo animal cancer study that qualified for TCDD dose-response assessment, EPA selected the species/sex/tumor dataset combinations characterized as having statistically significant increases in tumor incidences, then used the Emond rodent PBPK model to estimate blood concentrations corresponding to each study's average daily administered dose for use in dose-response modeling. $BMDL_{01}$s were then estimated for the blood concentration by two different methodologies: (1) using the multistage cancer model for each species/sex/tumor combination within each study, and (2) using a Bayesian Markov Chain Monte Carlo framework that assumes independence of tumors, modeling all tumors together for each species/sex combination within each study. The final selected models were subjected to goodness-of-fit tests and visual inspection of fit to the raw data. Thus, for each sex/species combination within each study, EPA generated a $BMDL_{01}$ for each single tumor type and another $BMDL_{01}$ for the combined tumors. Using the Emond human PBPK model, $BMDL_{HED}$s were then calculated for each of the $BMDL_{01}$s, and using a linear extrapolation, OSFs were calculated by $OSF = 0.01/BMDL_{HED}$ The highest OSF for a species/sex combination for either a single tumor type or all combined tumors was selected as a candidate OSF. The OSF candidates from the key animal bioassays are shown in Table ES-2.

DERIVATION OF TCDD ORAL SLOPE FACTOR AND RISK ESTIMATES

EPA was able to derive OSFs for tumor incidence data from four animal cancer bioassays, as well as cancer mortality data from four epidemiological cohort studies that were selected for TCDD dose-response modeling using the study inclusion criteria (see Section 5.3). For the animal data, OSFs based on individual tumors were developed for 25 study/sex/endpoint combinations, and the results ranged from 1.8×10^4 to 5.9×10^6 (per mg/kg-day). The OSFs based on combined tumors were developed for seven study/sex combinations, and the results

ranged from 3.2×10^5 to 9.4×10^6 (per mg/kg-day). EPA also developed OSFs based on four epidemiologic studies from three cohorts, ranging from 3.75×10^5 to 2.5×10^6 (per mg/kg-day).

EPA has chosen to use the human data over the animal data as recommended by expert panelists at EPA's 2009 Dioxin Workshop (U.S. EPA, 2009c) and in the 2005 Cancer Guidelines (U.S. EPA, 2005). OSFs derived from the human data are consistent with the animal bioassay results; human OSFs fall within the same range as the animal bioassay OSFs.

Among the human studies, the occupational TCDD exposures in the NIOSH and Hamburg cohorts are assumed to be reasonably constant over the duration of occupational exposure. In contrast, the TCDD exposure pattern for the Seveso and BASF accidents is acute, high dose, followed by low-level background exposure. Such exposure patterns similar to those experienced by the BASF and Seveso cohorts have been shown to yield higher estimates of risk when compared to constant exposure scenarios with similar total exposure magnitudes (Kim et al. 2003; Murdoch and Krewski, 1988; Murdoch et al, 1992). Thus, EPA has judged that the NIOSH and Hamburg cohort response data are more relevant than the BASF and Seveso data for assessing cancer risks from continuous ambient TCDD exposure in the general population.

The NIOSH (Steenland et al. 2001; Cheng et al. 2006) and Hamburg (Becher et al. 1998) cohort studies report cumulative TCDD levels in the serum for cohort members. The most significant difference among the Cheng et al. (2006) analysis and those of Steenland et al. (2001) and Becher et al. (1998) is the method used to back-extrapolate exposure concentrations based on serum TCDD measurements. Steenland et al. (2001) and Becher et al. (1998) back-extrapolated exposures and body burdens using a first-order model with a constant half-life. In contrast, Cheng et al. (2006) back-extrapolated body burdens using a kinetic modeling approach that incorporated concentration- and age-dependent elimination kinetics.

Although all three of these are high-quality studies, the kinetic modeling used by Cheng et al. (2006) is judged to better reflect TCDD pharmacokinetics, as currently understood, than the first-order models used by Steenland et al. (2001) and Becher et al. (1998). EPA believes that the representation of physiological processes provided by Cheng et al (2006) is more realistic than the assumption of simple first-order kinetics and this outweighs the attendant modeling uncertainties. Furthermore, the use of kinetic modeling is consistent with recommendations both by the NAS and the Dioxin Workshop panel.

Therefore EPA has selected the results from the Cheng et al. (2006) study for derivation of the TCDD OSF (see Section 5.3). Table ES-1 shows the oral slope factors at specific target risk levels ($OSF_{RL}s$) which range from 3.7×10^5 to 1.3×10^6 per (mg/kg-day). EPA recommends the use of an OSF of 1.3×10^6 per (mg/kg-day) when the target risk range is 10^{-5} to 10^{-7}.

CONSIDERATION OF NONLINEAR DOSE-RESPONSE APPROACHES FOR CANCER

The NAS focused much of its review on EPA's derivation of a cancer slope factor, commenting extensively on the extrapolation of dose-response modeling below the POD. The NAS questioned EPA's choice of a linear, nonthreshold model for extrapolating risk associated with exposure levels below the POD, concluding that the current scientific evidence was sufficient to justify the use of nonlinear methods when extrapolating below the POD for dioxin carcinogenicity.

While, based on the 2005 Cancer Guidelines, EPA deemed linear extrapolation to be most appropriate for TCDD, EPA carefully considered the NAS recommendation to provide risk estimates using both linear and nonlinear methods. In this document, EPA has evaluated the information available for identifying a threshold and for estimating the shape of the dose-response curve below the POD (see Section 5.2.3.4). EPA presents a hypothetical sublinear dose-response modeling example of rodent carcinogenicity. EPA also presents two illustrative examples of RfD development (i.e., nonlinear method) for carcinogenic effects of TCDD, using data derived from animal bioassays. EPA derives illustrative RfDs for cancer based on combined tumor response and also on hypothesized key events in TCDD's MOA for female rat liver and lung tumors. EPA identifies a number of limitations that prevent making strong conclusions based on the nonlinear dose-response modeling exercises.

FEASIBILITY OF QUANTITATIVE UNCERTAINTY ANALYSIS

EPA also addresses the third key recommendation of the NAS, specifically, improving transparency, thoroughness, and clarity in *quantitative uncertainty analysis* (see Section 6). In summary, NAS suggested that EPA should

- describe and define (quantitatively to the extent possible) the variability and uncertainty for key assumptions used for each key endpoint-specific risk assessment (choices of data set, POD, model, and dose metric),
- incorporate probabilistic models to the extent possible to represent the range of plausible values,
- clearly state it when quantitation is not possible and explain what would be required to achieve quantitation (NAS, 2006a, p. 9).

Although the NAS summarized the shortfalls in the 2003 Reassessment categorically, the elaborations within their report often contain the qualification "if possible" and do not take a position with regard to the feasibility of many suggestions. With appreciation for the extent of information available for dioxin, EPA's goal herein was to examine the feasibility of a data-driven quantitative uncertainty analysis for TCDD dose-response assessment.

In examining feasibility of quantitative uncertainty analysis, EPA recognized that different kinds of uncertainty require different statistical treatment. *Cognitive uncertainty* concerns uncertainty that can be expressed as probabilities and may be operationalized using either frequentist or Bayesian approaches. For example, classical statistical methods yield distributions on model parameters which reflect sample fluctuations, assuming that the model is true. This type of uncertainty can be taken into account in the BMDL estimation. Also, for TCDD epidemiologic data, the dose reconstruction often involves assumptions that may be amenable to data-driven uncertainty analysis if sufficient data can be retrieved; back-extrapolated TCDD levels, biological half-life, body fat, and background levels are example variables that could be included in such an analysis. In addition, a Monte Carlo analysis has been examined to develop quantitative uncertainty distributions for the RfD (e.g., Swartout et al., 1998). Given a set of animal bioassay data, quantifying dose-response uncertainty may be approached in different ways. The differences reflect different types of uncertainty that are captured. A recent evaluation enumerates the following possible methodologies (Bussard et al., 2009):

> **Benchmark Dose Modeling (BMD)**: Choose the 'best' model, and assess uncertainty assuming this model is true. Supplemental results can compare estimates obtained with different models, and sensitivity analyses can investigate other modeling issues.

Probabilistic Inversion with Isotonic Regression (PI-IR): Define model-independent 'observational' uncertainty, and look for a model that captures this uncertainty by assuming the selected model is true and providing for a distribution over its parameters.

Non-Parametric Bayes (NPB): Choose a prior mean response (potency) curve (potentially a "non-informative prior") and a precision parameter to express prior uncertainty over all increasing dose-response relations, and update this prior distribution with the bioassay data.

Bayesian Model Averaging (BMA) (as considered here): Choose an initial set of models, and then estimate the parameters of each model with maximum likelihood. Use classical methods to estimate parameter uncertainty, given the truth of the model. Determine a probability weight for each model using the Bayes Information Criterion (BIC), and use these weights to average the model results.

The first of the above methods involves standard classical statistical methods and captures sampling uncertainty conditional on the truth of the model used. The other methods are "exotic" in the sense that they attempt to capture uncertainty that is not conditional on the truth of a given model. In this response document, EPA has not applied such methods, but recognizes that quantitative uncertainty analysis is possible in these cases.

In contrast to cognitive uncertainty, *Volitional uncertainty* concerns uncertainty regarding choices on the best course of action to take; volitional uncertainty cannot be analyzed by sampling from a probability distribution and, thus, is not amenable to a complete quantitative uncertainty analysis. Some of the choices made in TCDD dose-response assessment that are volitional include: choice of occupational cohort data set or bioassay data set; choice of PODs (e.g., ED_{01}, ED_{05}, and ED_{10}); choice of species, strain, or sex within an animal bioassay; and choice of dose metric (e.g., administered doses, blood concentrations, lipid-adjusted serum concentrations). These volitional uncertainties cannot be quantified by sampling an input distribution. However, EPA believes that NAS was requesting that dose-response modeling results be shown for specific choices of interest to TCDD assessment. To this end, for the cancer dose-response modeling, BMDLs are reported for 1, 5, and 10% extra risk levels, and a comparison is provided of linear and nonlinear dose-response assessments for cancer. For the noncancer dose-response modeling, different model forms are run and contrasted. Finally, TCDD kinetic doses from the Emond et al. (2005, 2006) PBPK model that is primarily used in

the technical analysis in this document are compared with those predicted by the Aylward et al. (2005a) model.

Uncertainty quantification is an emerging area in science. There are many examples of highly vetted and peer-reviewed uncertainty analyses based on structured expert judgment. Under this process, experts in effect synthesize a wide diversity of information in generating their subjective probability distributions. Where considerable data exist for an environmental pollutant, such as for the well-studied TCDD, it is natural to ask whether these extensive data can be leveraged more directly in uncertainty quantification. This is an area where research could be focused. Additional research topics relevant to dioxin that could further inform health assessments include population variability of biokinetic constants and threshold mechanisms for the mass action model. Further data and improved methodologies in these areas, combined with developments illustrated elsewhere in this report, will help reduce or better quantify uncertainties and strengthen EPA's understanding of potential health implications of environmental TCDD exposures.

Table ES-1. Comparison of lipid-adjusted serum concentrations, fat concentrations, risk specific dose estimates and equivalent oral slope factors based on upper 95th percentile estimate of regression coefficienta of all fatal cancers reported by Cheng et al. (2006) for risk levels of 1×10^{-3}, 1×10^{-4}, 1×10^{-5}, 1×10^{-6}, and 1×10^{-7}

Risk level (RL)	AUC_{RL}, (ppt-yr)	FAT_{RL} (ng/kg)	Risk specific dose (D_{RL}) (ng/kg-day)	Equivalent oral slope factors (OSF_{RL}) per (mg/kg-day)
1×10^{-3}	1.26×10^{-3}	1.803×10^{1}	2.73×10^{-3}	3.7×10^{5}
1×10^{-4}	1.26×10^{2}	1.803×10	1.23×10^{-4}	8.1×10^{5}
1×10^{-5}	1.26×10^{1}	1.803×10^{-1}	8.57×10^{-6}	1.2×10^{6}
1×10^{-6}	1.26×10	1.803×10^{-2}	7.77×10^{-7}	1.3×10^{6}
1×10^{-7}	1.26×10^{-1}	1.803×10^{-3}	7.62×10^{-8}	1.3×10^{6}

[a]Based on regression coefficient of Cheng et al. (2006; Table III), excluding observations in the upper 5% range (≥252,950 ppt-year lipid adjusted serum TCDD) of the exposures; where reported $\beta = 3.3 \times 10^{-6}$ ppt-years and standard error = 1.4×10^{-6}. Upper 95th percentile estimate of regression coefficient (β_{95}) calculated to be: $6.04 \times 10^{-6} = (3.3 \times 10^{-6}) + 1.96 \times (1.4 \times 10^{-6})$.

Table ES-2. Tumor points of departure and oral slope factors using blood concentrations

Study	Sex/species: tumor sites	$BMDL_{HED}$[a] (ng/kg-day)	OSF (per mg/kg-day)
NTP, 1982	Male mice: liver adenoma and carcinoma, lung	1.1E−03	9.4E+6
Toth et al., 1979	Male mice: liver tumors	1.9E−03	5.2E+6
NTP, 1982	Female mice: liver adenoma and carcinoma, thyroid adenoma, subcutaneous fibrosarcoma, all lymphomas	5.3E−03	1.9E+6
NTP, 1982	Female rats: liver neoplasitc nodules, liver adenoma and carcinoma, adrenal cortex adenoma or carcinoma, thyroid follicular cell adenoma	5.7E−03	1.8E+6
Kociba et al., 1978	Female rats: liver adenoma carcinoma, oral cavity, lung	7.3E−03	1.4E+6
NTP, 1982	Male rats: thyroid follicular cell adenoma, adrenal cortex adenoma	9.6E−03	1.0E+6
NTP, 2006	Female rats: liver cholangiocarcinoma, hepatocellular adenoma, oral mucosa squamous cell carcinoma, lung cystic keratinizing epithelioma, pancreas adenoma, carcinoma	2.3E−02	4.4E+5
Kociba et al., 1978	Male rats: adrenal cortex adenoma, tongue carcinoma, nasal/palate carcinoma	3.1E−02	3.2E+5

[a]$BMDL_{HED}$s are from the multiple tumor analyses, with the exception of Toth et al. (1979) which is the result of modeling a single tumor site.

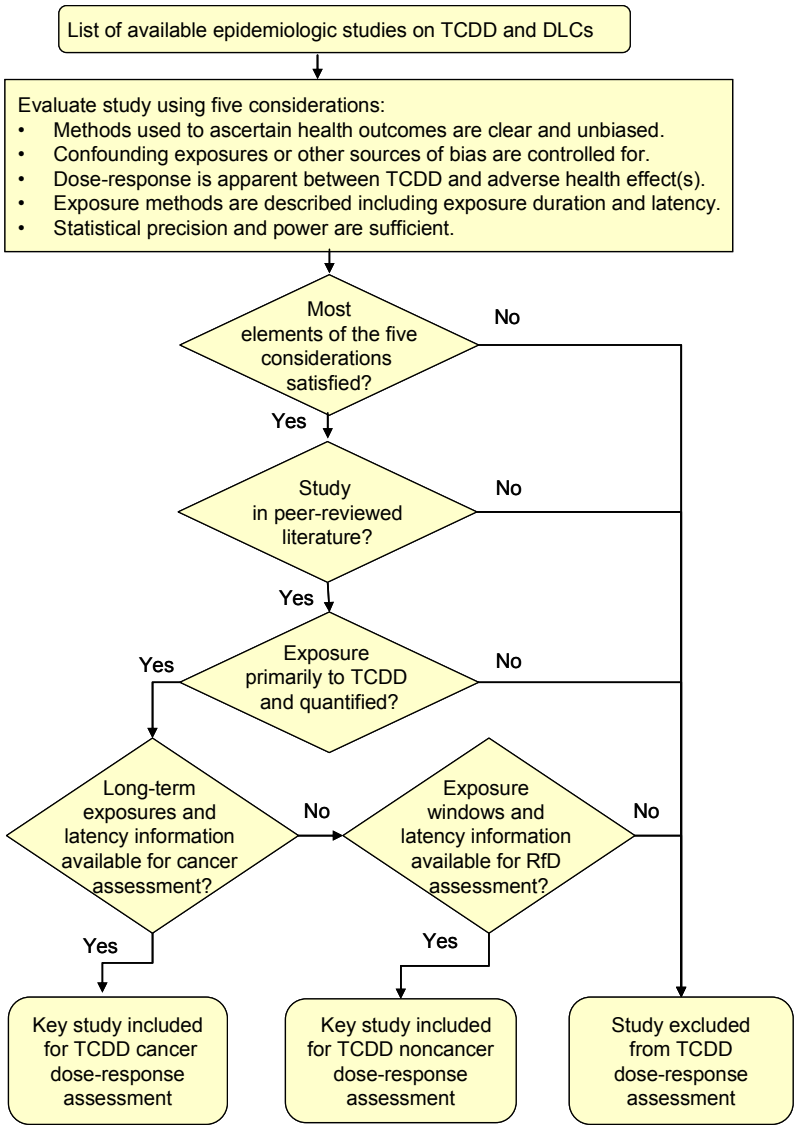


Figure ES-1. EPA's process to evaluate available epidemiologic studies using study inclusion criteria for use in the dose-response analysis of TCDD. EPA applied its TCDD-specific epidemiologic study inclusion criteria to all studies published on TCDD and DLCs. The studies were initially evaluated using five considerations regarded as providing the most relevant kind of information needed for quantitative human health risk analyses. For each study that satisfied most of these considerations and was published in the peer-reviewed literature, EPA then examined whether the exposures were primarily to TCDD and if the TCDD exposures could be quantified so that dose-response analyses could be conducted. Finally, EPA required that the effective dose and oral exposure be estimable: (1) for cancer, information is required on long-term exposures, (2) for noncancer, information is required regarding the appropriate time window of exposure that is relevant for a specific, nonfatal health endpoint, and (3) for all endpoints, the latency period between TCDD exposure and the onset of the effect is needed. Only studies meeting these criteria were included in EPA's TCDD dose-response analysis.

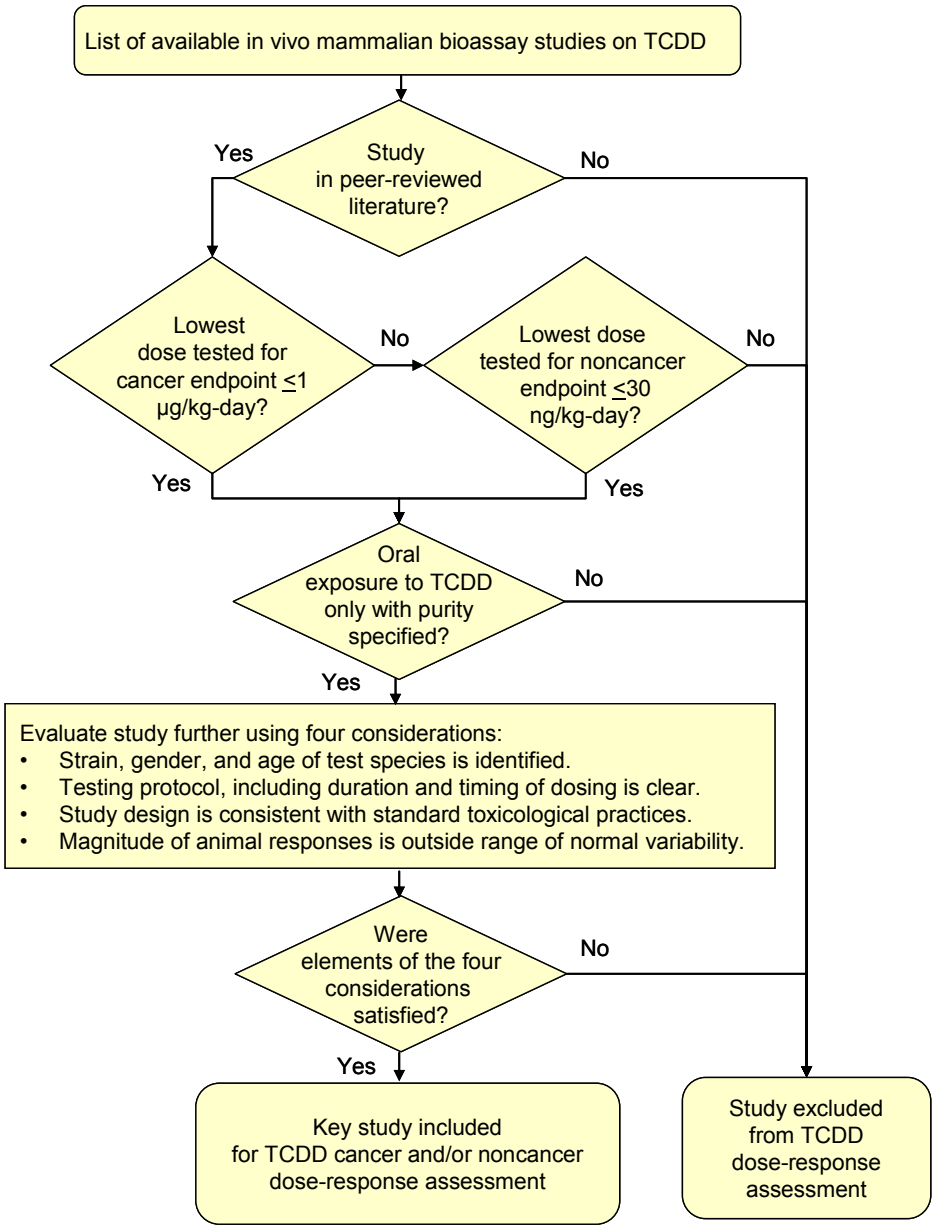


Figure ES-2. EPA's process to evaluate available animal bioassay studies using study inclusion criteria for use in the dose-response analysis of TCDD. EPA evaluated all available in vivo mammalian bioassay studies on TCDD. Studies had to be published in the peer-reviewed literature. Next, to ensure working in the low-dose range for TCDD dose-response analysis, EPA applied dose requirements to the lowest tested average daily doses in each study, with specific requirements for cancer (≤1 µg/kg-day), and noncancer (≤30 ng/kg-day) studies. Third, EPA required that the animals were exposed via the oral route to only TCDD and that the purity of the TCDD was specified. Finally, the studies were evaluated using four considerations regarded as providing the most relevant kind of information needed for quantitative human health risk analyses from animal bioassay data. Only studies meeting all of these criteria and considerations were included in EPA's TCDD dose-response analysis.

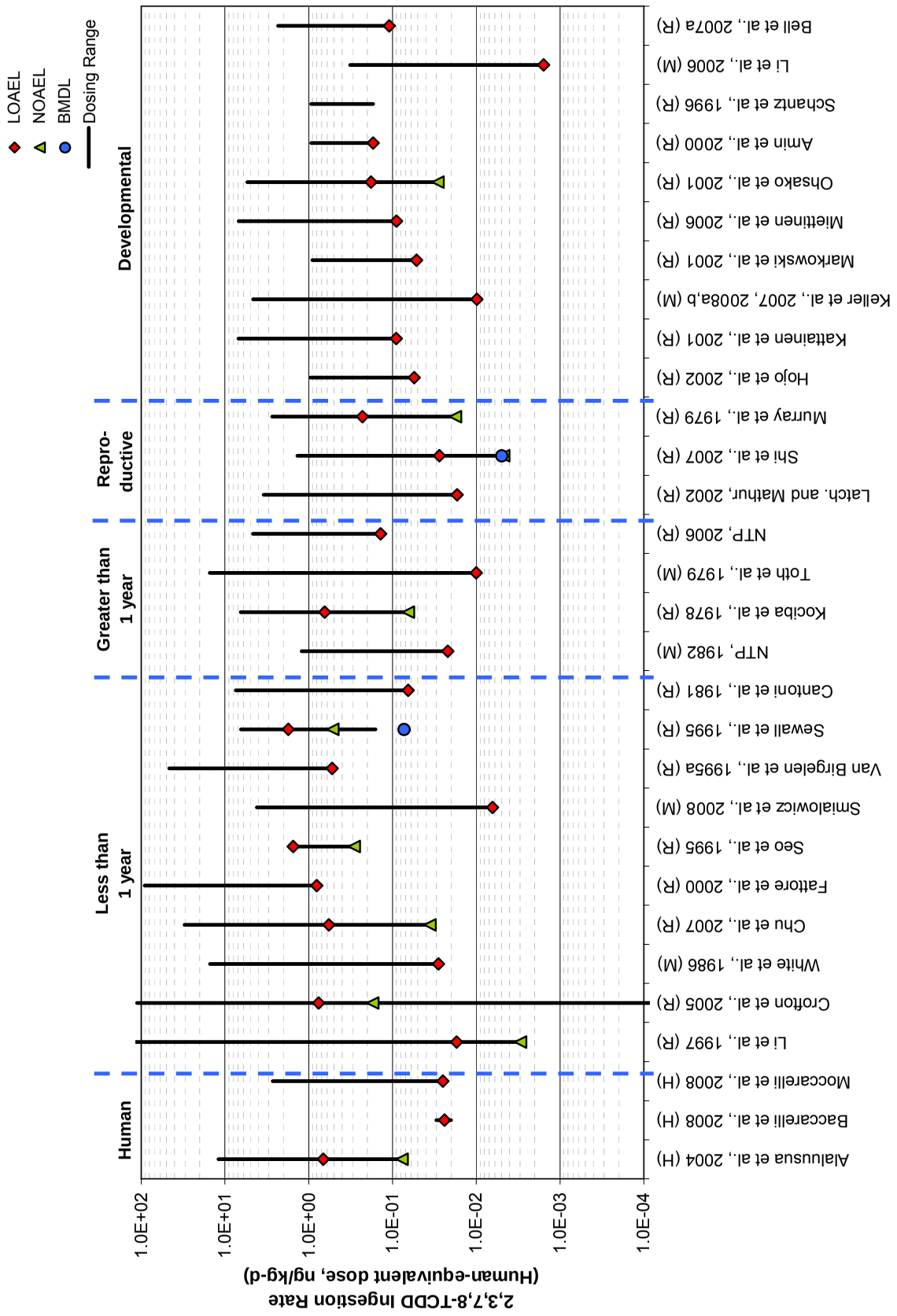


Figure ES-3. Exposure-response array for ingestion exposures to TCDD.

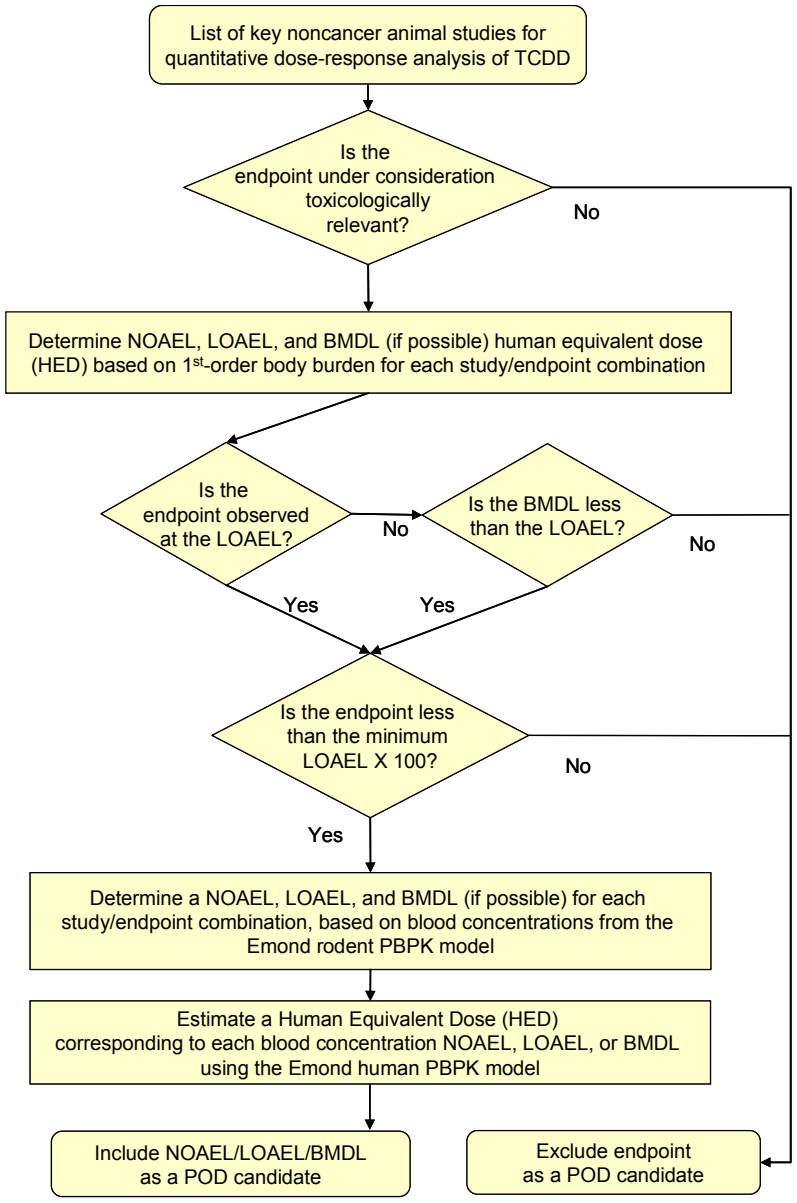


Figure ES-4. EPA's process to select and identify candidate PODs from key animal bioassays for use in noncancer dose-response analysis of TCDD. For each noncancer endpoint found in the studies that qualified for TCDD dose-response assessment using the study inclusion criteria, EPA first determined if the endpoint was toxicologically relevant. If so, EPA determined the NOAEL, LOAEL, and BMDL Human Equivalent Dose (HED) based on 1st-order body burdens for each endpoint. These potential PODs were examined for statistical relevance and included when the endpoint was observed at the LOAEL. If the BMDL was less than the LOAEL, and if the endpoint was less than the minimum LOAEL × 100, EPA then calculated NOAELs, LOAELs, or BMDLs based on blood concentrations from the Emond rodent PBPK model. Then, for all of the candidate PODs, HEDs were estimated using the Emond human PBPK model. Finally, the lowest group of the toxicologically relevant candidate PODs was selected for final use in derivation of an RfD.

1. INTRODUCTION

Dioxins and dioxin-like compounds (DLCs), including polychlorinated dibenzo-dioxins, polychlorinated dibenzofurans, and polychlorinated biphenyls are structurally and toxicologically related halogenated dicyclic aromatic hydrocarbons.[4] Dioxins and DLCs are released into the environment from several industrial sources such as chemical manufacturing, combustion, and metal processing; from individual activities including the burning of household waste; and from natural processes such as forest fires and volcanic activity. Dioxins and DLCs are widely distributed throughout the environment and typically occur as chemical mixtures. Additionally, they do not readily degrade; therefore, levels persist in the environment, build up in the food chain, and accumulate in the tissues of animals. Human exposure to these compounds occurs primarily through the ingestion of contaminated foods (Lorber et al., 2009).

The health effects from exposures to dioxins and DLCs have been documented extensively in epidemiological and toxicological studies. 2,3,7,8-Tetrachlorodibenzo-*p*-dioxin (TCDD) is one of the most toxic members of this class of compounds and has a robust toxicological database. Characterization of TCDD toxicity is critical to the risk assessment of mixtures of dioxins and DLCs because it has been selected repeatedly as the "index chemical" to serve as the basis for standardization of the toxicity of components in a mixture of dioxins and DLCs. The dose-response information for TCDD is used to evaluate risks from exposure to mixtures of DLCs (Van den Berg et al., 1998, 2006; also see the World Health Organization's Web site for the dioxin toxicity equivalence factors [TEFs]),[5] therefore, it is imperative to correctly assess the dose response of TCDD and understand the uncertainties and limitations therein.

In 2003, the U.S. Environmental Protection Agency (EPA) produced an external review draft of the multiyear comprehensive reassessment of dioxin exposure and human health effects entitled, *Exposure and Human Health Reassessment of 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) and Related Compounds* (U.S. EPA, 2003). This draft report, herein called the "2003 Reassessment," consisted of (1) a scientific review of information relating to sources of and exposures to TCDD, other dioxins, and DLCs in the environment; (2) detailed reviews of

[4]For further information on the chemical structures of these compounds, see U.S. EPA (2003, 2008a).
[5]Available at http://www.who.int/ipcs/assessment/tef_update/en/.

scientific information on the health effects of TCDD, other dioxins, and DLCs; and (3) an integrated risk characterization for TCDD and related compounds.

In 2004, EPA asked the National Research Council of the National Academy of Sciences (NAS) to review the 2003 Reassessment. The NAS Statement of Task was as follows

> The National Academies' National Research Council will convene an expert committee that will review EPA's 2003 draft reassessment of the risks of dioxins and dioxin-like compounds to assess whether EPA's risk estimates are scientifically robust and whether there is a clear delineation of all substantial uncertainties and variability. To the extent possible, the review will focus on EPA's modeling assumptions, including those associated with the dose-response curve and points of departure; dose ranges and associated likelihood estimates for identified human health outcomes; EPA's quantitative uncertainty analysis; EPA's selection of studies as a basis for its assessments; and gaps in scientific knowledge. The study will also address the following aspects of EPA's 2003 Reassessment: (1) the scientific evidence for classifying dioxin as a human carcinogen; and (2) the validity of the nonthreshold linear dose-response model and the cancer slope factor calculated by EPA through the use of this model. The committee will also provide scientific judgment regarding the usefulness of toxicity equivalence factors (TEFs) in the risk assessment of complex mixtures of dioxins and the uncertainties associated with the use of TEFs. The committee will also review the uncertainty associated with the 2003 Reassessment's approach regarding the analysis of food sampling and human dietary intake data, and, therefore, human exposures, taking into consideration the Institute of Medicine's report Dioxin and Dioxin-Like Compounds in the Food Supply: Strategies to Decrease Exposure. The committee will focus particularly on the risk characterization section of EPA's 2003 Reassessment report and will endeavor to make the uncertainties in such risk assessments more fully understood by decision makers. The committee will review the breadth of the uncertainty and variability associated with risk assessment decisions and numerical choices, including, for example, modeling assumptions, including those associated with the dose-response curve and points of departure. The committee will also review quantitative uncertainty analyses, as feasible and appropriate. The committee will identify gaps in scientific knowledge that are critical to understanding dioxin reassessment (NAS, 2006a, p. 43, Box 1-1).

In 2006, the NAS published its review of EPA's 2003 Reassessment entitled *Health Risks from Dioxin and Related Compounds: Evaluation of the EPA Reassessment* (NAS, 2006a).

1.1. SUMMARY OF KEY NAS (2006a) COMMENTS ON DOSE-RESPONSE MODELING IN THE 2003 REASSESSMENT

While recognizing the effort that EPA expended to prepare the 2003 Reassessment, the NAS committee identified three key areas that they believe require substantial improvement to support a scientifically robust risk assessment. These three key areas are

- transparency and clarity in selection of key data sets for analysis;
- justification of approaches to dose-response modeling for cancer and noncancer endpoints; and
- transparency, thoroughness, and clarity in quantitative uncertainty analysis.

In their Public Summary, the NAS made the following overall recommendations to aid EPA in addressing their key concerns:

- EPA should compare cancer risks by using nonlinear models consistent with a receptor mediated mechanism of action and by using epidemiological data and the new National Toxicology Program (NTP) animal bioassay data (NTP, 2006). The comparison should include upper and lower bounds, as well as central estimates of risk. EPA should clearly communicate this information as part of its risk characterization (NAS, 2006a, p. 9).
- EPA should identify the most important data sets to be used for quantitative risk assessment for each of the four key end points (cancer, immunotoxicity, reproductive effects, and developmental effects). EPA should specify inclusion criteria for the studies (animal and human) used for derivation of the benchmark dose (BMD) for different noncancer effects and potentially for the development of RfD (reference dose) values and discuss the strengths and limitations of those key studies; describe and define (quantitatively to the extent possible) the variability and uncertainty for key assumptions used for each key end-point-specific risk assessment (choices of data set, POD [point of departure], model, and dose metric); incorporate probabilistic models to the extent possible to represent the range of plausible values; and assess goodness-of-fit of dose-response models for data sets and provide both upper and lower bounds on central estimates for all statistical estimates. When quantitation is not possible, EPA should clearly state it and explain what would be required to achieve quantitation (NAS, 2006a, p. 9).
- When selecting a BMD as a POD, EPA should provide justification for selecting a response level (e.g., at the 10%, 5%, or 1% level). In either case, the effects of this choice on the final risk assessment values should be illustrated by comparing point estimates and lower bounds derived from selected PODs (NAS, 2006a, p. 9).
- EPA should continue to use body burden as the preferred dose metric but should also consider physiologically based pharmacokinetic modeling as a means to adjust for differences in body fat composition and for other differences between rodents and humans (NAS, 2006a, p. 9).
- Although EPA addressed many sources of variability and uncertainty qualitatively, the committee noted that the 2003 Reassessment would be substantially improved if its risk characterization included more quantitative approaches. Failure to characterize variability and uncertainty thoroughly can convey a false sense of precision in the conclusions of the risk assessment (NAS, 2006a, p. 5).

Importantly, the NAS encouraged EPA to calculate an RfD as the 2003 Reassessment does not contain an RfD derivation. The committee suggested that

> …estimating an RfD would provide useful guidance to risk managers to help them (1) assess potential health risks in that portion of the population with intakes above the RfD, (2) assess risks to population subgroups, such as those with occupational exposures, and (3) estimate the contributions to risk from the major food sources and other environmental sources of TCDD, other dioxins, and DLCs for those individuals with high intakes (NAS, 2006a, p. 6).

The NAS made many thoughtful and specific recommendations throughout their review; additional NAS recommendations and comments pertaining to the dose-response assessment of TCDD will be presented and addressed in various sections throughout this document.

1.2. EPA'S SCIENCE PLAN

In response to recommendations presented in the 2006 NAS review, EPA Administrator Jackson released EPA's "*Science Plan for Activities Related to Dioxins in the Environment*" ("Science Plan") on May 26, 2009.[6] There are five key components of the Science Plan that pertain to EPA's response to the NAS comments on TCDD dose-response assessment:

1. EPA will release a draft report that responds to the recommendations and comments included in the NAS review of EPA's 2003 Reassessment.
   - EPA's National Center for Environmental Assessment in the Office of Research and Development, will prepare a limited response to key comments and recommendations in the NAS report (draft response to comments report).
   - The draft response will focus on dose-response issues raised by the NAS and will include an analysis of relevant new key studies.
   - The draft response will be provided for public review and comment and independent external peer review.
   - The draft response will also include an evaluation of some of the significant recommendations that are difficult for EPA to address given the current state of science, and a detailed rationale for these conclusions.
   - The peer review will be conducted by EPA Science Advisory Board, an independent review body chartered under the Federal Advisory Committee Act.

[6]Available at http://www.epa.gov/dioxin/scienceplan.

- The draft response to comments report will be completed and released for public and peer review by December 31, 2009.

2. EPA will provide the draft response to comments report for internal and external review.
3. The Science Advisory Board will review the science content of the response to comments report.
4. EPA will review impacts of the draft response to comments report on its 2003 Reassessment.
5. EPA will release the final response to comments report and focus on completion of the 2003 Reassessment.

This document comprises EPA's report that responds to the recommendations and comments on TCDD dose-response assessment included in the NAS review of EPA's 2003 Reassessment. This document focuses on TCDD only.

1.3. OVERVIEW OF EPA'S RESPONSE TO NAS (2006a) "HEALTH RISKS FROM DIOXIN AND RELATED COMPOUNDS: EVALUATION OF EPA's 2003 REASSESSMENT"

In their key recommendations, the NAS commented that EPA should thoroughly justify and communicate approaches to dose-response modeling, increase transparency in the selection of key data sets, and improve the communication of uncertainty (particularly quantitative uncertainty). They also encouraged EPA to calculate an RfD. These main areas of improvement refer to issues specifically related to TCDD dose-response assessment (and uncertainty analysis); therefore, as noted in the Science Plan, EPA's response to the NAS is particularly focused on these issues.

EPA thoroughly considered the recommendations of the NAS and responds with scientific and technical evaluation of TCDD dose–response data via:

- an updated literature search that identified new TCDD dose-response studies (see Section 2);
- a kickoff workshop that included the participation of external experts in TCDD health effects, toxicokinetics, dose-response assessment and quantitative uncertainty analysis; these experts discussed potential approaches to TCDD dose-response assessment and considerations for EPA's response to NAS (see Appendix A);

- detailed study inclusion criteria and processes for the selection of key studies (see Section 2.3) and epidemiologic and animal bioassay data for TCDD dose-response assessment (see Section 2.4.1/Appendix B and Section 2.4.2, respectively);
- kinetic modeling to quantify appropriate dose metrics for use in TCDD dose-response assessment (see Section 3);
- dose-response modeling for all appropriate noncancer and cancer data sets (see Section 4.2/Appendix E and Section 5.2.3/Appendix F, respectively);
- thorough and transparent evaluation of the selected TCDD data for use in the derivation of an RfD and an oral slope factor (OSF) (see Sections 4.2 and 5.2.3, respectively);
- the development of an RfD (see Section 4.3);
- the development of a revised OSF (see Section 5.3) with an updated cancer weight of evidence determination for TCDD based on EPA's 2005 *Cancer Guidelines* (U.S. EPA, 2005) (see Section 5.1.2);
- consideration of nonlinear dose-response approaches for cancer, including illustrative RfDs for cancer precursor events and tumors (see Section 5.2.3.4) ; and
- discussion of the feasibility and utility of quantitative uncertainty analysis for TCDD dose-response assessment (see Section 6).

Each of these activities is described in detail in subsequent sections of this document.

It should be noted that three separate EPA activities address additional TCDD issues, specifically related to the application of dioxin TEFs and to TCDD and DLC background exposure levels. Information on the application of the dioxin TEFs is published elsewhere by EPA for both ecological (U.S. EPA, 2008a) and human health risk assessment (U.S. EPA, 2009a). As a consequence, EPA does not directly address TEFs herein, but makes use of the concept of toxicity equivalence[7] as applicable to the analysis of exposure dose in epidemiological studies. Furthermore, this document does not address the NAS recommendations pertaining to the assessment of human exposures to TCDD and other dioxins. Information on updated background levels of dioxin in the U.S. population has been recently reported (Lorber et al., 2009).

[7]Toxicity equivalence (TEQ) is the product of the concentration of an individual DLC in an environmental mixture and the corresponding TCDD TEF for that compound. These products are summed to yield the TEQ of the mixture.

1.3.1. TCDD Literature Update

EPA has developed a literature database of peer-reviewed studies on TCDD toxicity, including in vivo mammalian dose-response studies and epidemiological studies. An initial literature search for studies published since the 2003 Reassessment was conducted by the U.S. Department of Energy's Argonne National Laboratory (ANL) through an Interagency Agreement with EPA. ANL used the online National Library of Medicine database (PubMed) and identified studies published between the year 2000 and October 31, 2008. Supporting references published since the release of the 2003 Reassessment were also identified. Supporting studies were classified as studies pertaining to TCDD kinetics, TCDD mode-of-action, in vitro TCDD studies, and TCDD risk assessment approaches. The literature search strategy explicitly excluded studies addressing (1) analytical/detection data and cellular screening assays; (2) environmental fate, transport and concentration data; (3) dioxin-like compounds and toxic equivalents; (4) nonmammalian dose-response data; (5) human exposure analyses only, including body burden data; and (6) combustor or incinerator or other facility-related assessments absent primary dose-response data. EPA published the initial literature search results in the Federal Register on November 24, 2008 (73 FR 70999; November 24, 2008) and invited the public to review the list and submit additional peer-reviewed in vivo mammalian dose-response studies for TCDD, including epidemiological studies that were absent from the list (U.S. EPA, 2008b). The literature search results and subsequent submissions were used during a 2009 scientific workshop, which was open to the public and featured a panel of experts on TCDD toxicity and dose-response modeling (discussed below). Additional studies identified during the workshop and those collected by EPA scientists during the development of this report through October 2009 have been incorporated into the final set of studies for TCDD dose-response assessment.

1.3.2. EPA's 2009 Workshop on TCDD Dose Response

To assist EPA in responding to the NAS, EPA, and ANL convened a scientific workshop (the "Dioxin Workshop") on February 18−20, 2009, in Cincinnati, Ohio. The goals of the Dioxin Workshop were to identify and address issues related to the dose-response assessment of TCDD and to ensure that EPA's response to the NAS focused on the key issues and reflected the most meaningful science. The Dioxin Workshop included seven scientific sessions: quantitative dose-response modeling issues, immunotoxicity, neurotoxicity and nonreproductive endocrine

effects, cardiovascular toxicity and hepatotoxicity, cancer, reproductive and developmental toxicity, and quantitative uncertainty analysis of dose-response. During each session, EPA asked a panel of expert scientists to perform the following tasks:

- Identify and discuss the technical challenges involved in addressing the NAS comments related to the dose-response issues within each specific session topic and the TCDD quantitative dose-response assessment.
- Discuss approaches for addressing the key NAS recommendations.
- Identify important published, independently peer-reviewed literature—particularly studies describing epidemiological studies and in vivo mammalian bioassays expected to be most useful for informing EPA's response.

The sessions were followed by open comment periods during which members of the audience were invited to address the expert panels. The session's Panel Co-chairs were asked to summarize and present the results of the panel discussions—including the open comment periods. The summaries incorporated points of agreement as well as minority opinions. Final session summaries were prepared by the session Panel Co-chairs with input from the panelists, and they formed the basis of a final workshop report (U.S. EPA, 2009c, Appendix A of this report). Because the sessions were not designed to achieve consensus among the panelists, the summaries do not necessarily represent consensus opinions; rather reflect the core of the panel discussions. Some of the key discussion points from the workshop that influenced EPA's development of this document are listed below (see Appendix A for detail):

- In the development of study selection criteria, more relevant exposure-level (i.e., dose) decision points using tissue concentrations could be defined.
- A linear approach to body-burden estimation, which was utilized in the 2003 Reassessment (U.S. EPA, 2003), does not fully consider key toxicokinetic issues related to TCDD—e.g., sequestration in the liver and fat, age-dependent elimination, and changing elimination rates over time. Thus, kinetic/mechanistic modeling could be used to quantify tissue-based metrics. In considering human data, lipid-adjusted serum levels may be preferable over body burden, although the assumptions used in the back calculation of the body burden in epidemiologic cohorts are of concern. In considering rat bioassay data, lipid-adjusted body-burden estimates may be preferable.

- New epidemiological studies on noncancer endpoints have been published since the 2003 Reassessment that may need to be considered (e.g., thyroid dysfunction literature from Wang et al. [2005] and Baccarelli et al. [2008]).
- The 1% of maximal response (ED_{01}) that was utilized in the 2003 Reassessment has not typically been used in dose-response assessment. Some alternative ideas were as follows: (1) the POD should depend on the specific endpoint; (2) for continuous measures, the benchmark response (BMR) could be based on the difference from control and consider the adversity level; and (3) for incidence data, the BMR should be set to a fixed-risk level.
- The quantitative dose-response modeling for cancer could be based on human or animal data. There are new publications in the literature for four epidemiological cohort studies (Dutch cohort, NIOSH cohort, BASF accident cohort, and Hamburg cohort). The increase in total cancers could be considered for modeling human cancer data. However, non-Hodgkin's lymphoma and lung tumors are the main TCDD-related cancer types seen from human exposure. In reviewing the rat data, the NTP (2006) data sets are new and can be modeled. Although the liver and lungs are the main target organs, modeling all cancers, as well as using tumor incidence in lieu of individual rats as a measure, should be considered.
- Both linear and nonlinear model functions should be considered in the cancer dose-response analysis because there are data and rationales to support use of either below the POD.
- For quantitative uncertainty analysis, consider the impacts of choices among plausible alternative data sets, dose metrics, models, and other more qualitative choices. Issues to consider include how much difference these choices make and, also, how much relative credence should be put toward each alternative as a means to gauge and describe the landscape of imperfect knowledge with respect to possibilities for the true dose response. This may be difficult to do quantitatively because the factors are not readily expressed as statistical distributions. However, the rationale for accepting or questioning each alternative in terms of the available supporting evidence, contrary evidence, and needed assumptions, can be delineated.

1.3.3. Overall Organization of EPA's Response to NAS Recommendations

The remainder of this document is divided into five sections that address the three primary areas of concern resulting from the NAS (2006a) review. Section 2 describes EPA's approach to the recommendation for transparency and clarity during selection of key data sets—including criteria for the selection of key dose-response studies, evaluations of the important epidemiologic studies and animal bioassays, and a summary of the key studies used for subsequent dose-response modeling. Sections 3, 4, and 5 present EPA's response to the NAS recommendation to better justify the approaches used in dose-response modeling of TCDD.

Section 3 discusses the toxicokinetic modeling EPA conducted to support the dose-response analyses. Section 4 presents EPA's approach to noncancer data set selection, dose-response modeling, derivation of an RfD for TCDD, and contains a qualitative discussion of the uncertainties associated with the RfD. Section 5 presents an updated cancer weight-of-evidence summary, EPA's approach to cancer data set selection, dose-response modeling, derivation of an OSF for TCDD, and a qualitative discussion of the uncertainties associated with the OSF, including an evaluation of alternative approaches to cancer assessment of TCDD. Finally, Section 6 discusses the feasibility of conducting a quantitative uncertainty analysis of TCDD dose response.

2. TRANSPARENCY AND CLARITY IN THE SELECTION OF KEY DATA SETS FOR DOSE-RESPONSE ANALYSIS

This section focuses on addressing transparency and clarity in the study selection process and on identifying key data sets for 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) dose-response analysis. Section 2.1 summarizes the National Academy of Sciences (NAS) committee's comments specifically regarding this issue. Section 2.2 presents U.S. Environmental Protection Agency's (EPA's) response to those comments and describes EPA's approach to ensuring transparency and clarity in selection of studies for subsequent dose-response analyses. Section 2.3 describes the TCDD-specific study inclusion criteria and evaluation process EPA used in this document for determining the eligibility of both epidemiologic and experimental animal studies for TCDD dose-response analysis. Section 2.4.1 summarizes epidemiologic data and evaluates the suitability of these data for TCDD dose-response analyses. Section 2.4.2 summarizes animal bioassay data that have met the study inclusion criteria for TCDD dose-response assessment. Finally, Section 2.4.3 identifies key TCDD epidemiologic and animal bioassay studies that were determined using the study inclusion criteria. Then it previews the evaluation processes used to further identify study/endpoint combination data sets for developing TCDD toxicity values for noncancer (see Section 4) and cancer (see Section 5) effects.

2.1. SUMMARY OF NAS COMMENTS ON TRANSPARENCY AND CLARITY IN THE SELECTION OF KEY DATA SETS FOR DOSE-RESPONSE ANALYSIS

The NAS committee proposed that EPA develop a clear and readily understandable methodology for evaluating and including epidemiologic and animal bioassay data sets in dose-response evaluations. The NAS committee recommended the development and application of transparent initial criteria to judge whether or not specific epidemiologic or animal bioassay studies be included in TCDD dose-response analysis.

Specific NAS comments on the topic of study evaluation and inclusion criteria include

> EPA should specify inclusion criteria for the studies (animal and human) used for derivation of the benchmark dose (BMD) for different noncancer effects and potentially for the development of RfD values and discuss the strengths and limitations of those key studies (NAS, 2006, p. 27).

> …in its [EPA's] evaluation of the epidemiological literature of carcinogenicity, it did not outline eligibility requirements or otherwise provide the criteria used to assess the methodological quality of other included studies (NAS, 2006, p. 56.)

> With regard to EPA's review of the animal bioassay data, the committee recommends that EPA establish clear criteria for the inclusion of different data sets (NAS, 2006, p. 191).

> …the committee expects that EPA could substantially improve its assessment process if it more rigorously evaluated the quality of each study in the database (NAS, 2006, p. 56).

> EPA could also substantially improve the clarity and presentation of the risk assessment process for TCDD…by using a summary table or a simple summary graphical representation of the key data sets and assumptions…(NAS, 2006, p. 56).

2.2. EPA'S RESPONSE TO NAS COMMENTS ON TRANSPARENCY AND CLARITY IN THE SELECTION OF KEY DATA SETS FOR DOSE-RESPONSE ANALYSIS

EPA agrees with the NAS committee regarding the need for a transparent and clear process for selecting studies and key data sets for TCDD dose-response analyses. The delineation of the study selection process and decisions regarding key data sets will facilitate communication regarding critical decisions made in the TCDD dose-response assessment. In keeping with the NAS committee's recommendation to use a transparent process and improve clarity and presentation of the risk assessment process for TCDD, Figure 2-1 overviews the approach that EPA has used in this document to develop a final list of key cancer and noncancer studies for quantitative dose-response analysis of TCDD. The steps in Figure 2-1 are further explained below.

Literature search for in vivo mammalian and epidemiologic TCDD studies (2000−2008): EPA conducted a literature search to identify peer-reviewed, dose-response studies for TCDD that have been published since the 2003 Reassessment. This search included in vivo mammalian and epidemiological studies of TCDD from 2000 to 2008. Additional details describing the conduct of this literature search are presented in Section 1.3.1 of this document.

Federal Register Notice—Web publication of literature search for public comment: In November 2008, EPA published a list of ~500 citations from results of this literature search (U.S. EPA, 2008b) and invited the public to review this preliminary list of dose-response citations for use in TCDD dose-response assessment. EPA requested that interested parties identify and submit peer-reviewed studies for TCDD that were absent

from this list. All additional studies submitted to EPA were included in the final TCDD literature database considered by EPA for TCDD dose-response analysis.

Initial study inclusion criteria development for TCDD in vivo mammalian bioassays: EPA developed an initial set of draft criteria for evaluating the extensive TCDD database of in vivo mammalian bioassays. These initial inclusion criteria had three purposes. First, they provided a transparent and rigorous evaluation of the scientific quality of each study in EPA's database, a deficiency in the 2003 Reassessment identified by the NAS committee. Second, given the vast TCDD mammalian bioassay database, they provided a transparent method for initially screening studies to be considered for TCDD dose-response analyses. Third, they served as a starting point for discussions of study inclusion criteria by expert panelists who were convened by EPA for its scientific workshop on TCDD dose-response analysis (the Dioxin Workshop), described next (also see the workshop report in Appendix A).

Dioxin Workshop and expert refinement of TCDD in vivo mammalian bioassay inclusion criteria: In February 2009, EPA convened "A Scientific Workshop to Inform EPA's Response to NAS Comments on the Health Effects of Dioxin in EPA's 2003 Dioxin Reassessment." The goals of this 3-day public and scientific workshop were to identify and address issues related to the dose-response assessment of TCDD. Sessions at the workshop examined toxicities associated with TCDD, issues related to developing dose-response estimates based on these data and associated uncertainties. At the workshop, EPA presented the draft set of study inclusion criteria for evaluating the extensive TCDD in vivo mammalian bioassay literature and asked workshop panelists to discuss these criteria and make recommendations for their revision. Further details on this workshop are presented in Section 1.3.2 of this document, and the complete report from this workshop is available in Appendix A, including detailed summaries of the panels' comments on the inclusion criteria in relation to the various toxic endpoints that were discussed.

Final development of inclusion criteria for TCDD in vivo mammalian studies: Based on discussions at the Dioxin Workshop, the initial draft inclusion criteria for evaluating the TCDD mammalian bioassay literature were revised and are presented in Section 2.3.2 (see Figure 2-3). An initial criterion is that studies for consideration must be publically available and published in a peer-reviewed scientific journal. Because the methodology EPA uses to develop reference doses (RfDs) and cancer oral slope factors (OSFs) relies on identification of studies reporting potential adverse effects at low doses (relative to the overall database), another important criterion shown in Section 2.3.2 identifies a maximum value for the lowest TCDD dose tested in a bioassay. This maximum value was used to eliminate those studies that could not be selected for development of an RfD or an oral slope factor because tested doses were too high relative to other TCDD bioassays.

Development of inclusion criteria for epidemiologic studies: Following the Dioxin Workshop, EPA determined that an evaluation process was also needed for inclusion of epidemiologic studies for TCDD dose-response assessment. These criteria were developed and are detailed in Section 2.3.1 (see Figure 2-2). Analogous to animal bioassay data, epidemiologic studies for consideration must also be publically available

and published in a peer-reviewed scientific journal. In addition to assessing the methodological considerations relative to epidemiologic cohorts and studies (e.g., statistical power and precision of estimates, consideration of latency periods), a key consideration for use of a data set in TCDD dose-response modeling is characterization of the exposure assessment methodology, and specifically, whether this methodology allowed assignment of individual-level exposures within a study.

Final literature collection (October 2009): Additional literature was collected as it was identified by EPA following the Dioxin Workshop through October 2009 to ensure the consideration of all recently published data for this report.

Studies screened using inclusion criteria: The two sets of TCDD-specific study inclusion criteria presented in Section 2.3 were used to evaluate all studies included in the 2003 Reassessment, studies identified in the 2000−2008 literature search, studies identified through public comment and submission, and studies collected in 2009 as identified by EPA during the development of this document. Section 2.4 presents results of EPA's evaluation of epidemiologic and mammalian bioassay literature for both cancer and noncancer endpoints.

Final list of key cancer and noncancer studies for quantitative dose-response analysis of TCDD: Application of the study inclusion criteria concludes in Section 2.4 with development of a list of key noncancer and cancer studies that were considered for quantitative dose-response analyses of TCDD in Sections 4 and 5, respectively. In those sections, points of departure (PODs) are developed and evaluated for all study/endpoint combinations from these final key study lists, and key data sets and PODs for the development of TCDD toxicity values are identified.

2.3. STUDY INCLUSION CRITERIA FOR TCDD DOSE-RESPONSE ANALYSIS

One of the three major recommendations made by the NAS (2006a) committee was that EPA should provide greater clarity and transparency on the selection of studies that were used in the quantitative dose-response modeling of TCDD in the 2003 Reassessment. In this section, EPA describes TCDD-specific study inclusion criteria that have been developed to evaluate epidemiologic studies and animal bioassays for TCDD dose-response assessment. These criteria reflect EPA's goal of developing an RfD and a cancer OSF for TCDD through a transparent study selection process; they are intended to be used by EPA for TCDD dose-response assessment only. These criteria were applied to each of the ~500 studies listed in *Preliminary Literature Search Results and Request for Additional Studies on 2,3,7,8-Tetrachlorodibenzo-p-Dioxin (TCDD) Dose-Response Studies* (U.S. EPA, 2008b); studies identified and submitted by the public and by participants in the Dioxin Workshop (U.S. EPA, 2009c); studies included in

the 2003 Reassessment, and other relevant published studies collected by EPA scientists through October 2009.

EPA has undertaken different approaches for epidemiologic versus in vivo animal bioassay study evaluation and key data set selection. The significant differences between animal and human health effects data and their use in EPA risk assessment support development of separate criteria for study inclusion and different approaches to study evaluation. For the vast majority of compounds on EPA's Integrated Risk Information System (IRIS), cancer and noncancer toxicity values have been derived using animal bioassay data; therefore, approaches to dose-response modeling and POD selection from in vivo mammalian bioassays have been standardized and codified (U.S. EPA, 2008c). The study criteria shown below and in Figure 2-3 for animal bioassay data reflect EPA's preferences for TCDD-specific study inclusion, some of which are based on common practices and guidance for POD selection and RfD and OSF derivation. Far fewer IRIS toxicity values have been derived from human data, although some examples do exist. For example, benzene, beryllium and compounds, chromium IV, and 1,3-butadiene have RfDs, Reference Concentrations, Inhalation Unit Risks and/or OSFs all based on occupational cohort data and the methyl mercury RfD is based on high fish consuming cohorts (U.S. EPA, 2009b). The modeling and interpretation of such human data have been conducted on a case-by-case basis because each cohort is uniquely defined and has its own set of exposure conditions, significant confounders, and biases that may need to be considered in dose-response modeling. For TCDD, not all data are from occupational cohorts, but include cohorts exposed for relatively short time periods to high concentrations as a consequence of industrial accidents, a scenario that has not commonly been used to establish chronic EPA toxicity values.

Because of these differences in data characteristics, divergent selection approaches are used in this document to present and evaluate the epidemiologic studies (see Section 2.3.1) and the in vivo animal bioassays (see Section 2.3.2). In Section 2.4.1, all of the available epidemiologic studies on TCDD are summarized and evaluated for suitability for dose-response modeling using the TCDD-specific study inclusion criteria below and shown in Figure 2-2; only studies meeting the inclusion criteria are presented as key studies in Section 2.4.3 (see Tables 2-4 and 2-5 for the cancer and noncancer endpoints, respectively). In Section 2.4.2, because summarizing and showing the evaluation of the thousands of available animal bioassays on

TCDD was prohibitive, only studies first meeting the in vivo animal bioassays study inclusion criteria below (and shown in Figure 2-3) are summarized. These studies are also presented as key studies in Section 2.4.3 (see Tables 2-6 and 2-7 for cancer and noncancer endpoints, respectively).

2.3.1. Study Inclusion Criteria for TCDD Epidemiologic Studies

This section identifies the process EPA used to select epidemiologic studies for defining candidate PODs for TCDD dose-response modeling. These criteria are based on EPA's approaches for deriving OSFs and RfDs. A discussion of the considerations used in selecting epidemiologic data for quantitative dose-response modeling is valuable, particularly given EPA's preference to use high-quality human studies over animal studies because such human studies are regarded as providing the most relevant information needed for quantitative human health risk analyses (U.S. EPA, 2005). As described by Hertz-Picciotto (1995), key components needed for the use of an epidemiologic study as a basis for quantitative risk assessment include issues regarding exposure assessment (a well-quantified exposure assessment with exposures linked to individuals) and study quality ("strong biases," for example with respect to inclusion criteria for membership in the cohort and follow-up procedures "ruled out or unlikely" and "confounding controlled or likely to be limited"). The strength of the association, either within the full study or within a high exposure subgroup, can also be considered in the evaluation of suitability for dose-response modeling (Hertz-Picciotto, 1995). Stayner et al. (1999), however, note that even weak associations could be useful in terms of providing an estimate of a potential upper bound for a quantitative risk estimate.

EPA's method for applying the TCDD study inclusion criteria to epidemiologic data is detailed below and in Figure 2-2. Based on the framework discussed above, EPA evaluated the available epidemiologic cohorts and studies based on the five following considerations:

1. The methods used to ascertain health outcomes are clearly identified and unbiased, with high sensitivity and specificity.
2. The risk estimates generated from the study are not susceptible to important biases arising from an inability to control for potential confounding exposures or other sources of bias arising from either study design or statistical analysis.

3. The study demonstrates an association between TCDD and an adverse health effect (assuming minimal misclassification of exposure and absence of important biases) with some suggestion of an exposure-response relationship.

4. The exposure assessment methodology is clearly described and can be expected to provide adequate characterization of exposure, with assignment of individual-level exposures within a study (e.g., based on biomarker data, or based on a job-exposure-matrix approach). Limitations and uncertainties in the exposure assessment are considered.

5. The size and follow-up period of a cohort study are large enough and long enough, respectively, to yield sufficiently precise estimates for use in development of quantitative risk estimates and to ensure adequate statistical power to limit the possibility of not detecting an association that might be present (i.e., to avoid Type II Errors due to failing to reject the null hypothesis when the null hypothesis is true). Similar considerations regarding sample size and statistical precision and power apply to case-control studies.

Three specific study inclusion criteria were used to select studies for further evaluation and potential TCDD quantitative dose-response assessment

- The study is published in the peer-reviewed scientific literature and includes an appropriate discussion of strengths and limitations.
- The exposure is primarily to TCDD, rather than dioxin-like compounds (DLCs), and is properly quantified so that dose-response relationships can be assessed.
- The effective dose and oral exposure must be reasonably estimable. The measures of exposure must be consistent with the current biological understanding of dose. For TCDD dose-response assessment, it is critical that reported dose is consistent with a dose that is likely to be toxicologically relevant. The timing of the measurement of effects (i.e., the response) also must be consistent with current biological understanding of the effect and its progression.

  For cancer endpoints, EPA assumes that cumulative TCDD dose estimates are toxicologically relevant measures. Thus, cancer studies must provide information about long-term TCDD exposure levels. Further, EPA reasons that measures of cancer occurrence or death need to allow for examination of issues of latency between the end of effective exposure and cancer detection or death.

  For noncancer endpoints, exposure estimates and analysis must allow for examination of issues of latency and other issues regarding the appropriate time window of exposure relevant for specific endpoints. Also, to be consistent with the RfD methodology, the response must be to a nonfatal endpoint.

Those studies that met the aforementioned considerations and inclusion criteria (see Sections 2.4.1, 2.4.3, and Appendix B) were then subjected to further consideration for quantitative dose-response analyses.

2.3.2. Study Inclusion Criteria for TCDD In Vivo Mammalian Bioassays

This section identifies the criteria EPA applied to select nonhuman in vivo mammalian studies for defining candidate PODs for use in TCDD dose-response modeling. These inclusion criteria are based on EPA's approaches for deriving OSFs and RfDs from bioassay data (U.S. EPA, 2005). EPA agrees with the NAS committee regarding the utility of an oral RfD and the need for reevaluation of the OSF for TCDD, specifically in light of data that have been published since the 2003 Reassessment was released. RfDs and OSFs are generally derived using data sets that demonstrate the occurrence of adverse effects, or their precursors, in low-dose range for that chemical. RfDs and OSFs are derived from a health protective perspective for chronic exposures. Thus, when a group of studies is available on a chemical for which a number of effects are observed at various doses across those studies, the studies using the lowest exposures that show effects will typically drive the RfD and OSF derivations, all other considerations being equal. Studies conducted at higher exposures relative to other available studies are used as supporting evidence for the final RfD or OSF since they were conducted at doses too high to impact the numeric derivations of toxicity values. EPA expresses RfDs and OSFs in terms of average daily doses, usually as mg/kg-day and per mg/kg-day, respectively. Thus, the study inclusion criteria for the animal bioassay data presented in this section include requirements that average daily exposures in the studies are within a low dose range where, relative to other studies, they could be considered for development of a toxicity value. These low-dose requirements do not imply that TCDD studies conducted at higher doses are of poor quality, simply that they are not quantitatively useful in the development of toxicity values because other studies with lower exposures will drive the RfD and OSF derivations under current EPA practice. Because EPA has identified ~2,000 studies on TCDD that may be considered for this purpose, the development and application of these study inclusion criteria has been critical to moving the risk assessment process forward.

EPA's method for applying study inclusion criteria for mammalian bioassays is detailed below and in Figure 2-3. The first study inclusion criterion is that the study is published in the

peer-reviewed scientific literature. Then, two specific study inclusion criteria were used to select studies for further evaluation and potential TCDD quantitative dose-response analyses and identification of candidate PODs:

1. The lowest dose level tested is ≤1 μg/kg-day for cancer studies and ≤30 ng/kg-day for noncancer studies.
2. The study design consists of orally administered TCDD-only doses, and specifies the purity and matrix used to administer the doses.

Then, EPA evaluated the remaining in vivo animal studies based on the following four considerations.

1. The study tests mammalian species, identifying the strain, gender, and age of the tested animals.
2. The study clearly documents testing protocol, including dosing frequency, duration, and timing of dose administration relative to age of the animals.
3. The overall study design is consistent with standard toxicological principles and practices. The control group or groups are appropriate, given the testing protocol, and are well characterized. Clinical and pathological examinations conducted during the study are endpoint-appropriate, particularly for negative findings.
4. The magnitude of animal responses is outside the range of normal variability exhibited by control animals (e.g., greater than or less than one standard deviation).

Those studies that met the aforementioned considerations and inclusion criteria (see Sections 2.4.2 and 2.4.3) were then subjected to dose-response analysis.

The criteria for dose requirements, although somewhat arbitrary, are intended to be reasonable cutoffs that restrict the number of studies that would need to be modeled while ensuring that all study/data set combinations that could be candidates for the cancer slope factor or RfD were modeled. Thus, the dose range under consideration allows for liberal ranges of no-observed-adverse-effect levels (NOAELs), lowest-observed-adverse-effect levels (LOAELs), and benchmark dose lower confidence bound (BMDLs) for assessment of both cancer and noncancer effects.

For cancer studies, the dose requirements were selected based on an initial evaluation of available average daily doses administered in TCDD animal bioassays in which adverse effects

were observed. For example, in cancer studies, a sample of the relatively low ranges of tested average daily doses include 1-1000 ng/kg-day (Toth et al., 1979), 1−100 ng/kg-day (Kociba et al., 1978), 1.43−286 ng/kg-day (NTP, 1982) and 2.14−71.4 ng/kg-day (NTP, 2006) with statistically significant increases in tumor incidence via pair-wise or trend tests found in all of these studies. The entire range of each these studies is ≤1 μg/kg-day. The linearized multi-stage model used by EPA to estimate OSFs is most appropriately applied to studies from which PODs can be estimated as closely as possible to the experimental data. Thus, given the dose ranges in these studies that are available for modeling, the restriction to ≤1 μg/kg-day for cancer was considered to be a reasonable cutoff.

For noncancer studies, dose ranges are more complex and vary according to study endpoint. Examples of the lowest administered doses (not including developmental studies) that might be considered as NOAELs or LOAELs in POD determinations for noncancer endpoints include 1 ng/kg-day (Toth et al., 1979), 1.43 ng/kg-day (Cantoni et al., 1981), 1.07 ng/kg-day (Smialowicz et al., 2008) 1.43 ng/kg-day (NTP, 1982) and 2.14 ng/kg-day (NTP, 2006). Most of the lowest tested doses in the TCDD studies have been designated as LOAELs (see Section 4.1). Given the available database, it is likely that the same composite uncertainty factor (e.g., of 300; 3 for UF_A [interspecies], 10 for UF_H [intraspecies], and 10 for UF_L [LOAEL to NOAEL]) would be applied to any animal noncancer LOAEL used to derive an RfD for TCDD. This implies that any study that has a LOAEL of 30 ng/kg-day or more would result in a candidate RfD that is more than an order of magnitude higher than the example doses of 1−2 ng/kg-day shown here. BMDLs that might be derived from such data also would not be expected to be lower than these example doses of 1−2 ng/kg-day. Thus, a tested dose ≤30 ng/kg-day is considered to be a reasonable cutoff where the lowest tested dose would never be used as a POD to derive an RfD given that much lower tested doses (associated with adverse effects) are available from other studies of acceptable quality.

2.4. EVALUATION OF KEY STUDIES FOR TCDD DOSE RESPONSE

2.4.1. Evaluation of Epidemiological Cohorts for Dose-Response Assessment

2.4.1.1. *Cancer*

In the 2003 Reassessment, EPA selected three cohort studies from which to conduct a quantitative dose-response analysis: the National Institute for Occupational Safety and Health

(NIOSH) cohort (Steenland et al., 2001), the BASF cohort (Ott and Zober, 1996), and the Hamburg cohort (Becher et al., 1998). Although these studies were deemed suitable for quantitative dose-response analysis, the criteria EPA used to reach this conclusion were unclear. In this section, the study selection criteria and methodological considerations presented in Section 2.3 are systematically applied to evaluate a number of studies to determine their suitability for inclusion in dose-response modeling. In addition to the three cohorts used in previous TCDD quantitative risk assessment, considerations are applied to other relevant TCDD epidemiological data sets that were identified through a literature review for epidemiological studies of TCDD and cancer. Study summaries and suitability for quantitative dose-response analysis evaluations are discussed below.

2.4.1.1.1. *Cancer cohorts.*

2.4.1.1.1.1. ***The NIOSH cohort.***

In 1978, the NIOSH undertook research that identified workers employed by U.S. chemical companies that made products contaminated with TCDD between 1942 and 1982. TCDD was generated in the production of 2,4,5-trichlorophenol and subsequent processes. This chemical was used to make 2,4,5-trichlorophenoxyacetic acid (2,4,5-T), which was a major component of the widely-used defoliant, Agent Orange. The NIOSH cohort is the largest cohort of occupational workers studied to date and has been the subject of a series of investigations spanning more than two decades. It is important to note that this cohort consists mostly of male workers that were exposed to TCDD via daily occupational exposure, as compared to an acute accidental exposure scenario seen with other cohorts. The investigations have progressed from a comparison of the mortality patterns of the cohort to the U.S. general population to dose-response modeling using serum-derived estimates of TCDD that have been back-extrapolated several decades. Analyses of cancer data from the NIOSH cohort that are addressed in this section include Fingerhut et al. (1991), Steenland et al. (1999, 2001), Cheng et al. (2006), and Collins et al. (2009).

2.4.1.1.1.1.1. *Fingerhut et al., 1991.*

2.4.1.1.1.1.1.1. *Study summary.*

The investigation of Fingerhut and her colleagues published nearly two decades ago attracted widespread attention (Fingerhut et al., 1991). This retrospective study examined patterns of cancer mortality for 5,172 workers who comprised the NIOSH cohort, which combined workers from the company-specific cohorts of Dow Chemical (Ott et al., 1987; Cook, 1981) and the Monsanto Company (Zack and Gaffey, 1983; Zack and Suskind, 1980). These workers were employed at 12 plants producing chemicals contaminated with TCDD. Almost all workers in the cohort (97%) had production or maintenance jobs with processes involving TCDD contamination. On average, workers were employed for 2.7 years specifically in processes that involved TCDD contamination, and overall, were employed for 12.6 years. The mortality follow-up began in 1940 and extended until the end of 1987. Vital status was determined using records from the Social Security Administration, the Internal Revenue Service, or the National Death Index. The ascertainment of vital status in the cohort was nearly complete, with less than 1% of the cohort not followed up until death or the end of the study period.

Comparisons of mortality were made relative to the U.S. male general population and expressed using the standardized mortality ratio (SMR) metric and 95% confidence intervals (CIs). Life-table methods were used to generate person-years of risk accrued by cohort members at each plant. Person-years and corresponding deaths were tabulated across age, race, and year of death strata, which permitted the SMRs to be examined for potential confounding from these three characteristics. No unadjusted SMRs were presented in the paper. Cross-classification of person-years and deaths was also done across several exposure-related groupings, including duration of employment, years since first exposure, years since last exposure, and duration of exposure. Employment duration was categorized as <5, 5− <10, 10− <15, 15− <20, 20− <25, 25− <30, and ≥30 years. The variable "years since first exposure" (<10, 10− <20, and ≥20 years) was used to evaluate associations in relation to different latency periods. The analysis was jointly stratified by duration of employment and for varying latency intervals to evaluate whether cohort members with higher cumulative TCDD levels had higher cancer mortality rates than those cohort members with lower cumulative levels.

Overall, the cohort of workers had slightly elevated cancer mortality than the general population (SMR = 1.15, 95% CI = 1.02−1.30). Comparisons to the general population,

however, yielded no statistically significant excess for any site-specific cancer. Cancer mortality was examined for the subset of workers that worked for at least one year and had a latency interval of at least 20 years (n = 1,520). The 1-year cut-point was selected based on analyses of serum levels in a subset of 253 workers which revealed that every worker employed for at least one year had a lipid-adjusted serum level that exceeded the mean (7 ppt). Relative to the U.S. general population, statistically significant excesses in cancer mortality were observed for all cancers (SMR = 1.46, 95% CI = 1.21−1.76), cancers of the respiratory system (SMR = 1.42, 95% CI = 1.03−1.92), and for soft tissue sarcoma (SMR = 9.22, 95% CI = 1.90−26.95) among this subset of 1,520 male workers. The elevated SMR for soft tissue sarcoma, however, was based on only three cases in this subset.

SMRs also were generated across joint categories of duration of exposure and period of latency for deaths from all cancer sites (combined), and cancer of the trachea, bronchus, and lung. Increased SMRs were observed in strata defined by longer exposure and latency, but no statistically significant linear trends were found.

2.4.1.1.1.1.1.2. *Study evaluation.*

This cohort was the largest of four the International Agency for Research on Cancer (IARC) considered in its 1997 classification of TCDD as a Group 1 human carcinogen (IARC, 1997). Duration of employment in processes that involved TCDD contamination was used as a surrogate measure of cumulative exposure. In using this exposure metric, Fingerhut et al. (1991) assumed that TCDD exposures were equivalent at all production plants. Doses for individual cohort members were not reconstructed for these analyses, although they were in subsequent analyses of this cohort.

Workers in this cohort also were exposed to other chemicals, which could lead to bias due to confounding if these exposures were associated with both TCDD exposure and the health outcomes being examined. At one plant, workers were exposed to 4-aminobiphenyl. Previous investigators also reported that workers at another plant were exposed to 2,4,5-T and 2,4-dichlorophenoxyacetic acid (2,4-D) (Bond et al., 1988, 1989; Ott et al., 1987). Subsequent analyses revealed only modest correlations between duration of employment (a surrogate measure of exposure to other chemicals) and cumulative TCDD exposure. This suggests that confounding due to other coexposures is unlikely to have biased the results. In addition,

subsequent analyses of this cohort indicate that after excluding workers exposed to pentachlorophenols from analyses, positive associations between cumulative TCDD and all cancer mortality persisted (Steenland et al., 1999). Removal of workers who died from bladder cancer also did not substantially change the dose-response association between TCDD and cancer mortality from all other sites combined. This finding suggests that exposures to 4-aminobiphenyl did not confound the association between cancer mortality and TCDD exposure. Overall, there is little evidence of confounding by these co-exposures among this cohort, however, exposure to other possible confounders, such as dioxin-like compounds, was not examined.

The study collected no information on smoking behavior of the workers, and therefore, the SMRs do not account for any differences in the prevalence of smoking that might have existed between the workers and the general population. For several reasons, however, the inability to take into account smoking is unlikely to have been an important source of bias. First, mortality from other smoking-related causes of death such as nonmalignant respiratory disease were not more common in the cohort than in the general population (SMR = 0.96, 95% CI = 0.54−1.58). Second, stratified analyses of workers with at least a 20-year latency (assuming this subset shared similar smoking habits) revealed that excesses were apparent only among those who were exposed for at least 1 year. Specifically, when compared to the general population, the SMR among workers exposed for at least 1 year with a latency of 20 years was 1.46, (95% CI = 1.21−1.76) while those exposed for less than 1 year had an SMR of 1.02 (95% CI = 0.76−1.36). Third, for comparisons of cancer mortality between blue-collar workers and the general population, smoking is unlikely to explain cancer excesses of greater than 10−20% (Siemiatycki et al., 1988). Finally, the investigators found no substantial changes in the results for lung cancer when risks were adjusted for smoking histories obtained in 1987 from 223 workers employed at two plants. These data were used to adjust for the expected number of lung cancer deaths expected in the entire cohort (Fingerhut et al., 1991). Following this adjustment, the expected number of lung cancer deaths resulted in only a small change in the SMR for lung cancer in the overall cohort from 1.11 (95% CI = 0.89−1.37) to 1.05 (95% CI = 0.85−1.30). Similarly, there was little change in the SMR for lung cancer in higher exposure subcohort (SMR = 1.39, 95% CI = 0.99−1.89; to SMR = 1.37, 95% CI = 0.98−1.87).

The use of death certificate information from the National Death Index is appropriate for identifying cancer mortality outcomes. For site-specific cancers such as soft tissue sarcoma, however, the coding of this underlying cause of death is more prone to misclassification (Percy et al., 1981). Indeed, a review of tissues from four men concluded to have died from soft-tissue sarcoma determined that two deaths had been misclassified (Fingerhut et al., 1991). A review of hospital data revealed that two other individuals had soft tissue sarcomas that were not identified by death certificate information. The use of death certificate information to derive SMRs for cancer as a whole is likely not subject to significant bias; the same might not hold true, however, for some site-specific cancers such as soft tissue sarcoma.

Using the SMR metric to compare an occupational cohort with the general population is subject to what is commonly referred to as the "healthy worker effect" (Li and Sung, 1999; Choi, 1992). The healthy worker effect is a bias that arises because those healthy enough to be employed have lower morbidity and mortality rates than the general population. The healthy worker effect is likely to be larger for occupations that are more physically demanding (Aittomaki et al., 2005; Checkoway et al., 1989), and the healthy worker effect is considered to be of little or no consequence in the interpretation of cancer mortality (McMichael, 1976; Monson, 1986). Few cancers are associated with a prolonged period of poor health that would affect employability long before death. Also recognized is that, as the employed population ages, the magnitude of the healthy worker effect decreases as the absolute reduction in mortality becomes relatively smaller in older age groups (McMichael, 1976). The mortality follow-up of occupational cohorts generally spans several decades, which should minimize the associated healthy worker effect in such studies. Bias could also be introduced in that workers who are healthier might be more likely to stay employed and therefore accrue higher levels of exposure. In the NIOSH cohort, however, mortality was ascertained for those who could have left the workforce or retired by linking subjects to the National Death Index. Although internal cohort comparisons can minimize the potential for the healthy worker effect for the reasons presented above, for cancer outcomes, the SMR statistic is a valuable tool for characterizing whether occupational cohort are more likely to die of cancer than the general population. Moreover, stratified analyses across categories of duration of exposure, or latency periods within a cohort can yield important insights about which workers are at greatest risk. Perhaps most important, subsequent analyses of the NIOSH cohort that presented risk estimates derived from external

comparisons using the SMR were remarkably consistent with rate ratios derived using an internal referent (Steenland et al., 1999).

2.4.1.1.1.1.1.3. *Suitability of data for TCDD dose-response modeling.*

This cohort meets most of the identified considerations for conducting a quantitative dose-response analysis for mortality from all cancer sites combined. The NIOSH cohort is the largest cohort of TCDD-exposed workers, exposure characterization at an individual level is possible but not available in this particular study, the follow-up period is long enough to evaluate latent effects, and the study is not subject to important sources of bias that would materially alter findings. For the purpose of quantitative dose-response modeling, it is important to note that subsequent studies of this cohort adopted methods that greatly improved the characterization of TCDD exposure in this cohort and increased the follow-up interval (Steenland et al., 2001; Cheng et al., 2006). As such, for all practical purposes, due consideration for dose-response modeling should focus on the more recently developed data sets.

For quantitative dose-response modeling for individual cancer sites, the data are much more limited. A statistically significant positive association with TCDD was noted only for soft-tissue sarcoma among those with more than 1 year of exposure and 20 years of latency (SMR = 9.22, 95% CI = 1.90−26.95). However there were only three deaths from soft tissue sarcoma among this exposed component of the cohort, and four deaths in total in the overall cohort. Also, misclassification of outcome for soft-tissue sarcoma through death registries is well recognized and supported with additional review of tissue from two of the men. Specifically, tissues from the four men who died of soft-tissue sarcoma revealed that only two of these cases were coded correctly.

The design of this initial publication of the NIOSH cohort did not allow for exposures to other possible confounders, such as dioxin-like compounds and other occupational exposures, to be examined and accounted for in the analyses. TCDD exposures were based on duration of employment and effective doses could not be estimated based on the TCDD exposure measure (i.e., duration of employment). Therefore, dose-response modeling was not conducted for this study.

2.4.1.1.1.1.2. *Steenland et al., 1999.*

2.4.1.1.1.1.2.1. *Study summary.*

A subsequent analysis of the NIOSH cohort extended the follow-up interval of Fingerhut et al. (1991) by 6 years (i.e., from 1940−1993) and improved characterization of TCDD exposure (Steenland et al., 1999). A key distinction from the work of Fingerhut et al. (1991) was the exclusion of several workers that had been included in the previous mortality analyses. The authors excluded 40 workers who were either female, had never worked in TCDD-exposed departments, or had missing date of birth information. An additional 238 workers were excluded as occupational data for characterizing duration of exposure were lacking, preventing their use in a subcohort dose-response analysis. This subcohort was further reduced by excluding workers from four plants (n = 591) because the information on the degree of TCDD contamination in work histories was limited, preventing the characterization of TCDD levels by job type. Thirty-eight additional workers were excluded from the eight remaining plants because TCDD contamination could not be estimated. Finally, 727 workers were excluded because they had been exposed to pentachlorophenol. In total, exposures were assigned to 3,538 (69%) members of the overall cohort, a cohort substantially reduced from the 5,172 on which Fingerhut et al. (1991) reported. Steenland et al. (1999) also evaluated the mortality experience of a subcohort of 608 workers with chloracne who had no exposure to pentachlorophenol.

For each worker, a quantitative exposure score for each day of work was calculated based on the concentration of TCDD (μg/g) present in process materials, the fraction of the day worked, and a qualitative contact level based on estimates of the amount of TCDD exposure via dermal absorption or inhalation. The authors derived a cumulative measure of TCDD exposure by summing the exposure scores across the working lifetime history for each worker. The authors validated this cumulative exposure metric indirectly by comparing values obtained for workers with and without chloracne. Such a validation is appropriate, given that chloracne is considered a clinical sign of exposure to dioxin. The median exposure score among those with chloracne was 11,546 compared with 77 among those without (Steenland and Deddens, 2003).

Cancer mortality was compared using two approaches. As in Fingerhut et al. (1991), external comparisons were made to the U.S. general population using the SMR statistic. The authors adjusted the SMR statistics for race, age, and calendar time. They also applied life-table methods to characterize risks across the subcohort of 3,538 workers with exposure data by

categorizing the workers into seven cumulative exposure groups. The cut-points for these categories were selected so that the number of deaths in each category was nearly equal to optimize study power. Life-table analyses were extended further to consider a 15-year lag interval, which in a practical sense means that person-years at risk would not begin to accrue until 15 years after the first exposure occurred. The person-years and deaths that occurred in the first 15 years were included in the lowest exposure grouping. The Cox proportional hazards model was used to characterize risk within the cohort. Cox regression was used to provide an estimate of the hazard ratios and the 95% CIs for ischemic heart disease, all cancers combined, lung cancer, smoking related cancers, and all other cancers. The authors also performed Cox regression analyses using the seven categories of exposure, adjusting the regression coefficients for year of birth and age. The regression models were run for both unlagged and lagged (15 years) cumulative exposure scores.

Overall, when compared with the U.S. general population, a slight excess of cancer mortality (from all sites) was noted in the 5,132 cohort study population (SMR = 1.13, 95% CI = 1.02−1.25). This result did not substantially differ from the earlier finding that Fingerhut et al. (1991) published (SMR = 1.15, 95% CI = 1.03−1.30). Site-specific analyses revealed statistically significant excesses relative to the U.S. general population for bladder cancer (SMR = 1.99, 95% CI = 1.13−3.23) and for cancer of the larynx (SMR = 2.22, 95% CI = 1.06−4.08). In the chloracne subcohort (n = 608), a borderline statistically significant excess relative to the general population was found for all cancer sites combined (SMR = 1.25, 95% CI = 0.98−1.57) and for lung cancer (SMR = 1.45, 95% CI = 0.98−2.07). The authors also found statistically significant excesses for connective and soft tissue sarcomas (SMR = 11.32, 95% CI = 2.33−33.10) and for lymphatic and hematopoietic malignancies (SMR = 3.01, 95% CI = 1.43−8.52).

External comparisons made by grouping workers into septiles of cumulative TCDD exposure and generating an SMR for each septile using the U.S. population as the referent group suggested a dose-response relationship. For all cancer sites combined, workers in the highest exposure score category had an SMR of 1.60 (95% CI = 1.15−1.82); increases also were observed in the sixth (SMR = 1.34) and fifth (SMR = 1.15) septiles. The two-sided p-value associated with the test for trend for cumulative TCDD exposure was statistically significant (p = 0.02). A similar approach for lung cancer revealed virtually the same pattern. The

incorporation of a 15-year latency for the analyses of all cancer deaths, in general, produced slightly higher SMRs across the septiles, although a slight attenuation of effect was noted in the highest septile ($SMR_{unlagged}$ = 1.60 vs. SMR_{lagged} = 1.54). For a 15-year lag, the lung cancer SMRs were mixed compared to the unlagged results with some septile exposure categories increasing and others decreasing relative to the lowest exposure group.

For the internal cohort comparisons using Cox regression analyses higher hazard ratios were found among workers in the higher exposure categories than in the lowest septile. The linear test for trend, however, was not statistically significant (p = 0.10). The associations across the septiles for the unlagged exposure for the internal cohort comparisons were not as strong as for the external cohort comparisons. The opposite was true, however, for cumulative exposures lagged 15 years.

Relative to the lowest septile, stratified analyses revealed increased hazard ratios in the upper septiles of the internal cohort comparisons for both smoking- and nonsmoking-related forms of cancer. The test for linear trend was statistically significant for all other cancers (after smoking-related cancers were excluded). These analyses suggest that the overall cancer findings were not limited to an interaction between TCDD and smoking. Additional sensitivity analyses by the authors indicated the findings for smoking-related cancers were largely unaffected by the exclusion of bladder cancer cases. This observation suggests that the exposure to 4-aminobiphenyl, which occurred at one plant and might have contributed to an increased number of bladder cancers, did not substantially bias the dose-response relationship between TCDD and all cancers combined.

The investigators also evaluated the dose-response relationship with a Cox regression model separately for each plant using internal cohort comparisons and found some heterogeneity. This finding is not unexpected particularly given the relatively small number of cancer deaths at each plant, and given that exposures were quite low for one plant at which no positive association was found. The variability among plants was taken into account by modeling plant as a random effect measure in the Cox model, which produced little change in the slope coefficient (β = 0.0422 vs. β = 0.0453).

2.4.1.1.1.1.2.2. *Study evaluation.*

This study represents a valuable extension of that by Fingerhut et al. (1991). Internal comparisons were performed to help minimize potential biases associated with using an external comparison group (e.g., healthy worker effect, and differences in other risk factors between the cohort and the general population). That similar dose-response relationships were found for internal and external comparison populations suggests that the bias due to the health worker effect in the cohort might be minimal for cancer mortality. More importantly, the construction of the cumulative exposure scores provides an improved opportunity to evaluate dose-response relationships compared with the length of exposure and duration of employment metrics that Fingerhut et al. (1991) used.

A potential limitation of the NIOSH study was the inability to account for cigarette smoking. If cigarette smoking did contribute to the increased cancer mortality rates in this and other cohorts, increased cancer mortality from exposure to TCDD would be expected only for smoking-attributable cancers. This study demonstrates associations with TCDD for both smoking- and nonsmoking-related cancers, including a stronger association for nonsmoking-related cancers. Therefore, the data provide evidence that associations between TCDD and cancer mortality are not likely due to cigarette smoking.

The findings regarding latency should be interpreted cautiously as the statistical power in the study to compare differences across latency intervals was limited. Caution also should be heeded, given that latency intervals can vary on an individual basis as they are often dose-dependent (Guess and Hoel, 1977). The evaluation of whether TCDD acts as either an initiating or promoting agent (or both) is severely constrained by the reliance on cancer mortality data rather than incidence data. This constraint is due to the fact that survival time can be quite lengthy and can vary substantially across individual and cancer subtype. For example, the 5-year survival among U.S. males for all cancer sites combined ranged between 45 and 60% (Clegg et al., 2002). When only mortality data are available, evaluating the time between when individuals are first exposed and when they are diagnosed with cancer is nearly impossible.

Starr (2003) suggested that Steenland et al. (1999) focused too heavily on the exposures that incorporated a 15-year period of latency and that those who experienced high exposures would inappropriately contribute person-years to the lowest exposure group "irrespective of how great the workers' actual cumulative exposure scores may have been." Most cancer deaths

would, however, typically occur many years postemployment. Given that the follow-up interval of the cohort was long and the average exposure duration was 2.7 years, at the time of death, person-years for those with high cumulative exposures would be captured appropriately. The median 5-year survival for all cancers is approximately 50% (Clegg et al., 2002), so applying a minimum latency of 5 years when using cancer mortality rather than cancer incidence data is needed to assure that the exposure metric is capturing exposures that occur before diagnoses. Increasing this latency period, for example to 10 or 15 years, would eliminate consideration of exposures that occur in the period between tumor occurrence and tumor detection (diagnosis), and allows for an appropriate focus on exposures that act either early or late in the pathogenic process. The IARC recommendation also underscores the relevance of including a period of latency when estimating cancer risks related to TCDD. Specifically, IARC indicated that if the association of TCDD with cancer is causal, effects might become apparent only at high exposures and with adequate latency; they suggested that a latency interval of 15 years could be too short (IARC, 1997). EPA considers the Steenland et al. (1999) presentation to be balanced in that they provided the range in lifetime excess risk estimated across the various models used. The authors' finding that the models with a 15-year lag provided a statistically significant improvement in fit based on the chi-square test statistic should not be readily dismissed.

2.4.1.1.1.1.2.3. *Suitability of data for TCDD dose-response modeling.*

This study meets most of the epidemiological considerations for conducting a quantitative dose-response analysis for mortality from all cancer sites combined. This study excludes a large number of workers who were exposed to pentachlorophenol, thus eliminating the potential for bias from this exposure and used an improved methodology for assigning TCDD exposures to the workers. However, given that exposures to other dioxin-like compounds were not described, it is unclear if the exposures among this cohort were primarily to TCDD. Therefore, dose-response modeling was not pursued for this study, but was for the subsequent NIOSH study by Steenland et al. (2001), which did examine exposure to dioxin-like compounds.

2.4.1.1.1.1.3. *Steenland et al., 2001.*

2.4.1.1.1.1.3.1. *Study summary.*

In 2001, Steenland et al. published a risk analysis using the NIOSH cohort that for the first time incorporated serum measures in the derivation of TCDD exposures for individual workers. The authors applied the same exclusion criteria to the entire cohort of workers across the 12 plants in the Steenland et al. (1999) study, which left 3,538 workers for which risk estimates could be calculated. Cumulative TCDD serum levels were estimated on an individual basis for all 3,538 workers by developing predictive models that used a subset of 170 workers for which both serum measures and TCDD exposures scores were available (Steenland et al., 2001). Unlike previous analyses of the NIOSH cohort that considered several different mortality outcomes, the analyses presented in Steenland et al. (2001) focused exclusively on mortality from all cancers sites combined. The authors observed 256 cancer deaths in the cohort during the follow-up interval that extended from 1942 until the end of 1993. All risks estimated in the Steenland et al. (2001) study were based on internal cohort comparisons.

Characterization of TCDD exposure levels among the workers was based on serum measures obtained in 1988 from 199 workers who were employed in one of the eight plants. The researchers restricted the development of the model to include only those workers whose measured serum levels were deemed to be greater than the upper range of background levels (10 ppt), which resulted in 170 workers.

The authors developed a regression model that could estimate the level of TCDD at the time of last exposure for the 170 workers. The model was developed based on the estimated half-life of TCDD, the known work history of each worker, a pharmacokinetic model for the storage and excretion of TCDD, and exposure scores for each job held by each worker over time. The resulting equation follows

$$y_{last\ exposure} = y_{1988}\ exp(\lambda \Delta t) \qquad \text{(Eq. 2-1)}$$

The first-order elimination rate constant (λ) was based on a half-life of 8.7 years previously reported for the Ranch Hands cohort (Michalek et al., 1996). The background rate of TCDD exposure was assumed to be 6.1 parts per trillion (ppt), which was based on the median level in a sample of 79 unexposed workers in the NIOSH cohort (Piacitelli et al., 1992). This

value was subtracted when TCDD values were back-extrapolated, and then added again after the back-extrapolation was completed. A background level of 5 ppt also was used in some of the analyses with minimal demonstrable effects on the results. Sensitivity analyses also were incorporated to consider a 7.1-year half-life estimate that had been developed for the earlier Ranch Hands study (Pirkle et al., 1989).

After back-extrapolating to obtain TCDD serums levels at the time of last exposure, the investigators estimated cumulative (or "area under the curve") TCDD serum levels for every cohort member. This estimation procedure was the same method Flesch-Janys et al. (1998) applied to the Hamburg cohort to derive a coefficient for relating serum levels to exposure scores. The "area under the curve" approach integrates time-specific serum levels over the employment histories of the individual workers. The slope coefficient was estimated using a no-intercept linear regression model. This model is based on the assumption that a cumulative score of zero is associated with no serum levels above background.

Cox regression was also used to model the continuous measures of TCDD. A variety of exposure metrics were considered that took into account different lags, nonlinear relationships (e.g., log-transform and cubic spline), as well as threshold and nonthreshold exposure metrics. Categorical analyses were used to evaluate risks across TCDD exposure groups, while different shapes of dose-response curves were evaluated through the use of lagged and unlagged continuous TCDD measures. Categorical analyses of TCDD exposure were conducted using the Cox regression model to derive estimates of relative risk (RR) as described by hazard ratios and 95% CIs. The reference group in this analysis was those workers in the lowest septile cumulative exposure grouping (<335 ppt-years). The septiles were chosen based on cumulative serum levels that considered no lag and also a 15-year lag.

The investigators also conducted dose-response analyses using the toxicity equivalence (TEQ) approach. The TEQ is calculated as the sum of all exposures to dioxins and furans weighted by the potency of each specific compound. In this study, TCDD was assumed to be account for all dioxin exposures in the workplace. For background TEQ levels, the investigators used a value of 50 ppt in the dose-response modeling. This is based on the assumption that TCDD accounted for 10% of the toxicity of all dioxins and furans (WHO, 1988), and is equivalent to using a background level of 5 ppt/yr that was used in the derivation of cumulative serum TCDD levels. A statistically significant dose-response pattern was observed for all cancer

mortality and TCDD exposure based on log of cumulative TEQs with a 15-year lag. A comparison of the overall model chi-square values indicated that the fit of this model was not as good as that for TCDD.

The hazard ratios among workers grouped by categories of cumulative TCDD exposure (lagged 15 years) suggested a dose-response relationship. Steenland et al. (2001) found statistically significant excesses in the higher exposure categories compared to the lowest septile. The RR was 1.82, 95% CI = 1.18−2.82 for the sixth septile (7,568−20,455 ppt-years) and 1.62, 95% CI = 1.03−2.56) for the seventh septile (>20,455 ppt-years). Cox regression indicated that log TCDD serum concentrations (lagged 15 years) was positively associated with cancer mortality ($\beta = 0.097$, standard error (β) = 0.032, $p < 0.003$). A statistically significant improvement in fit was observed when a 15-year lag interval was incorporated into the model compared to a model with no such lag [Model χ^2 with 4 degrees of freedom (df) = 7.5]. Results were similar when using a half-life of 7.1 years rather than 8.7 years. The excess lifetime risk of death from cancer at age 75 for TCDD intake (per 1.0-picogram per kilogram [pg/kg] of body weight (BW) per day) was about 0.05−0.9% above a background lifetime risk of cancer death of 12.4%. The results from the best-fitting models provide lifetime risk estimates within the ranges derived using data from the Hamburg cohort (Becher et al., 1998).

In both categorical and continuous analyses of TCDD based on a linear exposure metric, the dose-response pattern tailed off at high exposures suggesting nonlinear effects. This phenomenon could be due to saturation effects (Stayner et al., 2003) or, alternatively, could have resulted from increased exposure misclassification of higher exposures (Steenland et al., 2001). As the authors highlighted, some of the highest exposures might have been poorly estimated as they occurred in workers exposed to short-term high exposures during the clean-up of a spill. The choice of a linear model to develop data from a single time point can also result in exposure misclassification in those individuals that have differences in the length of exposure (Emond et al., 2005). Misclassification would be less likely at low concentrations where dose-dependent elimination is minimal.

2.4.1.1.1.1.3.2. *Study evaluation.*

An important consideration in the Steenland et al. (2001) study was the use of a small subset of workers ($n = 170$) to infer exposures for the remainder of the cohort. This subset

comprised surviving members of the cohort (in 1988), and therefore, their age distribution would have differed from the rest of the cohort. Furthermore, these workers were employed at a single plant, at which the work histories were less detailed than at other plants; thus, the development of the exposure scores differed between this plant and that of the others. Also, many of the workers at this plant had the same job title and were employed during the same calendar period. The use of serum data from this subset adds a level of uncertainty that is not readily characterized. Despite this limitation, the use of these sera data to derive cumulative measures for all cohort workers has merit given the strong correlation observed between the exposure scores, and TCDD serum levels estimates at the time of last exposure (Spearman $r = 0.90$).

The authors performed an extensive series of sensitivity analyses and considered several alternative exposure metrics to the simple linear model. The lifetime excess risk above background was nearly twice as high for the log cumulative serum measures with a 15-year lag when compared to the piecewise linear models with no lag. An important observation was that the exposure metric based on cumulative serum (lagged 15 years) did not fit the data as well as the cumulative exposure score used in earlier analyses (Steenland et al., 1999). A priori, one would expect that a better fit would be obtained with serum-based measures because serum levels are a better measure of relevant biological dose. As the authors noted, inaccuracies introduced in estimating the external-based exposure scores could have contributed to a poorer fit of the data. Alternatively, exposure misclassification error could be introduced if serum samples based on the 170 workers were not representative of the entire cohort. Although the serum-based measures did not fit the data as well as the exposures scores, the authors regarded them as providing a reasonable fit. Moreover, the serum-based measures enabled better characterization of risk in units (pg/kg-day) that can be used in regulation exposures.

2.4.1.1.1.1.3.3. *Suitability of data for TCDD dose-response modeling.*

This study meets all of the epidemiological considerations for conducting a quantitative dose-response analysis for mortality from all cancer sites combined. As mentioned previously, the NIOSH cohort is the largest assembled to date for which TCDD-related risks of cancer mortality can be estimated. The use of serum-based measures provides an objective measure of TCDD exposure. Repeated measures in other study populations have provided reasonable estimates of the half-life of TCDD, which permitted back-extrapolation of exposures.

The authors have made extensive efforts to evaluate a wide variety of nonlinear and linear models with varying lengths of latency and log transformations. The model chi-square test statistics were fairly similar for the log cumulative serum (15-year lag) (Model $\chi^2_{(4df)} = 11.3$) model and the piecewise linear model (no lag) (Model $\chi^2_{(5df)} = 12.5$). These models, however, produced results with twofold differences in lifetime excess risks. These differences underscore the importance of characterizing uncertainty in modeling approaches when conducting dose-response analysis.

The Steenland et al. (2001) study characterizes risk in terms of pg/kg of body weight per day. Given that tolerable daily intake dioxin levels are typically expressed in pg/kg of body weight (WHO, 1998), the presentation of risks in terms of these units is an important advance from the earlier analyses that used exposure scores (Steenland et al., 1999). Many of the Steenland et al. (2001) findings are consistent with earlier work from this cohort, which is not surprising given that exposures scores were used to derive serum-based levels for the cohort. The findings of excess lifetime risks obtained for the best- fitting model are also consistent with those derived from the Hamburg cohort (Becher et al., 1998). This study meets the epidemiological considerations noted previously as there is no evidence that the study is subject to bias from confounding due to cigarette smoking or other occupational exposures. Given the considerable efforts to measure effective dose to TCDD among the study participants, this study also meets the requisite dose-response modeling criteria and will be used in quantitative dose-response analyses of cancer mortality.

2.4.1.1.1.1.4. *Cheng et al., 2006.*

2.4.1.1.1.1.4.1. *Study summary.*

Cheng et al. (2006) undertook a subsequent quantitative risk assessment of 3,538 workers in the NIOSH cohort using serum-derived estimates of TCDD. This dose-response analysis was published after the 2003 Reassessment document was released. The goal of this study was to examine the relationship between TCDD and cancer mortality (all sites combined) using a new estimate of dose that estimated TCDD as a function of both exposure intensity and age using a kinetic model. This physiologically based pharmacokinetic model has been termed the "concentration- and age-dependent elimination model" (CADM) and was developed by Aylward

et al. (2005b). This model describes the kinetics of TCDD following oral exposure to humans by accounting for key processes affecting kinetics by simulating the total concentration of TCDD based on empirical consideration of hepatic processes (see Section 3.3). An important feature of this kinetic model is that it incorporates concentration- and age-dependent elimination of TCDD from the body; consequently, the effective half-life of TCDD elimination varies based on exposure history, body burden, and age of the exposed individuals. The study was motivated by the reasoning that back-calculations of TCDD using a first-order elimination model and a constant half-life of 7−9 years underestimated exposures to TCDD among workers. This underestimate, in turn, would result in overestimates of the carcinogenic potency of TCDD.

As with the earlier Steenland et al. (2001) analyses, the cohort follow-up period was extended from 1942 until the end of 1993 and work histories were linked to a job exposure matrix to obtain cumulative TCDD scores. Two cumulative serum lipid exposure metrics (in ppt-years) were constructed using the data obtained from the sample of 170 workers. The first replicated the metric used in a previous analysis of the cohort (Steenland et al., 2001) and was based on a first-order elimination model with an 8.7-year half-life (Michalek et al., 1996). The second metric was based on CADM and had two first-order elimination processes (Aylward et al., 2005a). This metric assumes that the elimination of TCDD in humans occurs at a faster rate when body concentrations are high and at slower rates in older individuals (Aylward et al., 2005a, b). The model was optimized using individuals for which serial measures of serum TCDD were available. These measures were obtained from 39 adults with initial serum levels between 130 and 144,000 ppt (Aylward et al., 2005b). This group included 36 individuals who had been exposed in the Seveso accident and 3 exposed in Vienna, Austria. In practice, for serum levels greater than 1,000 ppt, the effective half-life would be less than 3 years, and for serum TCDD levels less than 50 ppt, the effective half-life would be more than 10 years (Aylward et al., 2005b). Results from the model indicate that men eliminate TCDD faster than women do as demonstrated previously by Needham et al. (1994). These age- and concentration-dependent processes were assumed to operate independently on TCDD in hepatic and adipose tissues, and TCDD levels in liver and adipose tissue were assumed to be a nonlinear function of body concentration. Cheng et al. (2006) calibrated CADM using a dose of 156 ng per unit of exposure score and assumed a background exposure rate of 0.01 ng/kg-month. The average TCDD ppt-years derived from CADM with a 15-year lag was 4.5−5.2 times higher than

with the first-order elimination model. The two metrics, however, were highly correlated based on a Pearson correlation coefficient of 0.98 ($p < 0.001$). Comparisons of fit between the CADM and first-order elimination model were made using R^2 values and presented in Aylward et al. (2005b).

Cheng et al. (2006) compared the mortality experience of NIOSH workers to the U.S. general population using the SMR statistic. SMR statistics also were generated separately for each of the 8 plants and for all plants combined. Cox regression models were used to analyze internal cohort dose-response. These models used age as the time variable, and penalized smoothing spline functions of the CADM metric also were considered. The possible confounding effects of other occupational exposures and other regional population differences were assessed by repeating analyses after excluding one plant at a time. Lagged and unlagged TCDD exposures were analyzed separately, and stratified analyses compared risk estimates for smoking- and nonsmoking-related cancers. Cheng et al. (2006) adjusted the slope estimates derived from the Cox model for potential confounding effects of race and year of birth.

Overall, a statistically significant excess in all cancer mortality in the cohort occurred relative to the general population (SMR = 1.17, 95% CI = 1.03−1.32). The plant-specific SMRs ranged from 0.62−1.87, with a statistically significant excess evident only for plant 10 (SMR = 1.87, 95% CI = 1.35−2.52). For lung cancer mortality, the overall SMR was not statistically significant (SMR = 1.11, 95% CI = 0.89−1.37). A statistically significant excess for lung cancer also was found for plant 10 (SMR = 2.35, 95% CI = 1.44−3.64). The SMRs between smoking- (SMR = 1.22, 95% CI = 1.01−1.45) and nonsmoking-related cancers (SMR = 1.12, 95% CI = 0.94−1.33) were comparable.

For the internal cohort analyses of serum-derived measures, the authors were able to replicate the one-compartmental model used previously (Steenland et al., 2001). As had been noted by Steenland et al. (2001), an inverse-dose-response pattern was seen for individuals with high exposures (above 95th percentile); this type of pattern is often seen in occupational studies (Stayner et al., 2003). Excluding these data produced a stronger association between TCDD and all-cause mortality. In fact, only when the upper 2.5% or 5% of observations was removed did a statistically significant positive association become evident with the untransformed data. Similarly, when the model incorporated a lag of 15 years, a statistically significant association was noted only for the untransformed TCDD ppt-years with the upper 5% of observations

removed. Stratified analyses revealed little difference between smoking- and nonsmoking-related cancers, and the removal of one plant at a time from the analyses of TCDD ppt-years changes did not substantially change the slope.

2.4.1.1.1.1.4.2. *Study evaluation.*

The authors reported that CADM provided an improved fit over the one-compartmental model, but presented no evidence regarding any formal test of statistical significance. A comparison of R^2 values presented in Aylward et al. (2005b), however, does reveal that the R^2 value increased from 0.27 (first-order compartmental model with an 8.7-year half-life) to 0.40 for CADM. TCDD exposures estimated using CADM were approximately fivefold higher than the one-compartmental model estimates among cohort members with higher levels of exposure. Differences in exposure estimates between the two metrics were less striking among individuals with lower TCDD exposures. The net effect was that CADM produced a 6- to 10-fold decrease in estimated risks compared to estimates previously reported (Steenland et al., 2001). Nonetheless, the estimates produced by CADM span more than two orders of magnitude under various assumptions. Further uncertainties arise from between-worker variability of TCDD elimination rates, possible residual confounding, and the variability associated with the use of data obtained from other cohorts. Nevertheless, the use of the CADM model to estimate TCDD exposure is considered a significant advantage over the previous first-order body burden calculations.

2.4.1.1.1.1.4.3. *Suitability of data for TCDD dose-response modeling.*

The value of including the NIOSH cohort data has already been established based on investigations published by Steenland et al. (1999, 2001). The decision to include data from the quantitative dose-response analysis that Cheng et al. (2006) conducted relates to the added value that the CADM exposure estimates would provide. The earlier modeling work of Aylward et al. (2005b) provided some support for a modest improvement of the fit of CADM over the first-order compartmental model, and they also confirmed previous studies that found that TCDD elimination rates varied by age and sex. Recent work by Kerger et al. (2006) also demonstrates that the half-life for TCDD is shorter among Seveso children than the corresponding half-life for adults, and that body burdens influence the elimination of TCDD in humans. That estimates of

half-lives among men have been remarkably consistent, with mean estimates ranging between 6.9 and 8.7 years (Flesch-Janys et al., 1996; Pirkle et al., 1989; Michalek et al., 2002; Needham et al., 2005), however, is noteworthy. Based on the underlying strengths of the NIOSH cohort data and efforts by Cheng et al. (2006) to improve estimates of effective dose, these data support further dose-response modeling.

2.4.1.1.1.1.5. *Collins et al., 2009.*

2.4.1.1.1.1.5.1. *Study summary.*

In a recent study, Collins et al. (2009) investigated the relationship between serum TCDD levels and mortality rates in a cohort of trichlorophenol workers exposed to TCDD. These workers were part of the NIOSH cohort having accounted for approximately 45% of the person-years in an earlier analysis (Bodner et al., 2003). The investigators completed an extensive dioxin serum evaluation of workers employed by the Dow Chemical plant in Midland, Michigan, that made 2,4,5-trichlorophenol (TCP) from 1942 to 1979 and 2,4,5-T from 1948 to 1982. Collins et al. (2009) developed historical TCDD exposure estimates for all TCP and 2,4,5-T workers. This study represents the largest group of workers from a single plant ever studied for the health effects of TCDD. Little information on how vital status was ascertained, either in this paper or in the Bodner et al. (2003) report of mortality in this cohort. Although the authors indicate that death certificates were obtained from the states in which the employees died, whether vital status was ascertained from company records or through record linkage to the National Death Index is unclear.

The follow-up interval for these workers covered the period between 1942 and 2003. Thus, the study included 10 more years of follow-up than earlier investigations of the entire NIOSH cohort. Serum samples were obtained from 280 former workers collected during 2004−2005. A simple one-compartment first-order pharmacokinetic model and elimination rates as estimated from the BASF cohort were used (Flesch-Janys et al., 1996). The "area under the curve" approach was used to characterize workers' exposures over the course of their working careers and provided a cumulative measure of exposure. Analyses were performed with and without 165 of the 1,615 workers exposed to pentachlorophenol to evaluate the impact of these exposures.

External comparisons of cancer mortality rates to the general U.S. population were made using SMRs. Internal cohort comparisons of exposure-response relationships were made using the Cox regression model. This model used age as the time variable, and was adjusted for year of hire and birth year. Only those causes of death for which an excess was found based on the external comparisons or for which previous studies had identified a positive association were selected for dose-response analyses.

A total of 177 cancer deaths were observed in the cohort. For the external comparison with the U.S. general population, overall, no statistically significant differences were observed in all cancer mortality among all workers (SMR = 1.0, 95% CI = 0.8−1.1). Results obtained after excluding workers exposed to pentachlorophenol were similar (SMR = 0.9, 95% CI = 0.8−1.1). Statistically significant excesses in the cohort were found for leukemia (SMR = 1.9, 95% CI = 1.0−3.2) and soft tissue sarcoma (SMR = 4.1, 95% CI = 1.1−10.5). Although not statistically significant SMRs for other lymphohemopoietic cancers included non-Hodgkin's lymphoma SMR = 1.3; 95%CI = 0.6, 2.5) and Hodgkin's disease (SMR = 2.2; 95% CI = 0.2, 6.4).

Internal cohort comparisons using the Cox regression model were performed for all cancers combined, lung cancer, prostate cancer, leukemia, non-Hodgkin's lymphoma, and soft-tissue sarcoma. Whether the internal comparisons excluded those workers exposed to pentachlorophenol is not entirely clear from the text or accompanying table, but presumably they do not. The RR was 1.002 (95% CI = 0.991−1.013) for all cancer mortality per 1 ppb-year increase in cumulative TCDD exposure was not statistically significant. Except for soft tissue sarcomas, no statistically significant exposure-response trends were observed for any cancer site. For soft tissue sarcoma, analyses were based on only four deaths.

2.4.1.1.1.1.5.2. *Study evaluation.*

A key limitation of this study is that SMRs were not derived for different periods of latency for the external comparison group analysis. The original publication on the NIOSH cohort found that SMRs increased when a 20-year latency period was incorporated (Fingerhut et al., 1991), and similar patterns have been observed in other occupational cohorts (Manz et al., 1991; Ott and Zober, 1996) and among Seveso residents (Consonni et al., 2008). Additionally, dose-response analyses showed marked increases in slopes with a 15-year latency period

(Steenland and Deddens, 2003; Cheng et al., 2006). In this context, the absence of an elevated SMR for cancer mortality is consistent with previous findings of the NIOSH cohort. While the cohort did have sufficient follow-up, no evaluation of possible latent effects was presented and this is a major limitation of this study. Further, the evaluation of the exposure metrics should be expanded from what was presented in Collins et al. (2009) due to the previous analyses of the same workers finding positive associations between cancer mortality and TCDD (Steenland et al., 2001).

Unfortunately, the Collins et al. (2009) study did not include a categorical analysis of TCDD exposure and cancer mortality. This categorical analysis would have enabled an evaluation of whether a nonlinear association exists between TCDD exposure and cancer risk. The analyses of both Cheng et al. (2006) and Steenland et al. (2001) suggest an attenuation of effects at higher doses, and several investigations have considered log-transformed associations as a means to address nonlinearity. Also, the earlier plant-specific dose-response analyses of Steenland et al. (2001) are not consistent with the findings for the Midland plant that Collins et al. (2009) presented. These differences could be due to differences in the construction of exposure metrics, additional follow-up, or lagging of exposures.

2.4.1.1.1.1.5.3. *Suitability of data for dose-response modeling.*

The Collins et al. (2009) study uses serum levels to derive TCDD exposure estimates and does not appear to be subject to important biases. The reliance on data from one plant offers some advantages over the multiplant analyses, as heterogeneity in exposure to other occupational agents would be lower. The number of individuals who provided serum samples (n = 280) is greater than the 170 individuals used to derive TCDD estimates for the NIOSH cohort. This study's main limitation is that it lacks rigor in sensitivity analyses to explore the impact that other exposure metrics and model assumptions had on the study findings. The data used in quantitative dose-response modeling should retain flexibility to account for latency effects. The data, as structured for this study, are inadequate for quantitative dose-response analysis given that no dose-response effects were detected.

2.4.1.1.1.2. *The BASF cohort.*

In 1953, dioxin contamination occurred as a result of an autoclave accident during the production of trichlorophenol at the BASF plant in Ludwigshafen, Germany. A second dioxin incident occurred in 1988 that was attributed to the blending of thermoplastic polyesters with brominated flame retardants. Of the two events, the one on November 13, 1953, was associated with more severe acute health effects, including chloracne that resulted in immediate hospitalizations for seven workers. These adverse events were not linked to TCDD until 1957 when TCDD was identified as a byproduct of the production of trichlorophenol and was shown to induce chloracne (Zober et al., 1994). Zober and colleagues (1998) noted that with the 1988 accident, affected individuals did not exhibit clinical symptoms or chloracne, but rather were identified through "analytical measures." In both instances, efforts were made to limit the potential for exposure to employees.

2.4.1.1.1.2.1. *Thiess and Frentzel-Beyme, 1977 and Thiess et al., 1982.*

2.4.1.1.1.2.1.1. *Study summary.*

A study of the mortality of workers employed at the BASF plant was first presented in 1977 (Thiess and Frentzel-Beyme, 1977) with subsequent updates in both 1982 (Thiess et al., 1982), and in 1990 (Zober et al., 1990). In the first published paper (Thiess et al., 1982), 74 employees involved in the 1953 accident were traced and their death certificate information extracted. Of these, 66 suffered chloracne or severe dermatitis. Observed deaths were compared to the expected number using three external reference groups: the town of Ludwigshafen (n = 180,000), the district of Rhinehessia-Palatinate (n = 1.8 million), and the Federal Republic of Germany (n = 60.5 million). Another comparison group was assembled by selecting age-matched employees taken from other cohorts under study. This additional comparison was aimed at avoiding potential biases associated with healthy worker effect when using an external referent.

During a follow-up interval of up to 26 years (1953−1979), 21 individuals died. Of these, seven deaths were from cancer. The expected number of cancer deaths derived for the three external comparison groups ranged between 4.1 and 4.2, producing an SMR of 1.7 (p-values ranged between 0.12 and 0.14). Excess mortality was found for stomach cancer based on the external comparisons ($p < 0.05$); however, this was based on only three cases. No other

statistically significant excesses were found with the external comparisons made to the other cohorts of workers.

2.4.1.1.1.2.1.2. *Study evaluation.*

In the Thiess et al. (1982) study, no TCDD exposures were derived for the workers, thus no dose-reconstruction was performed. The findings from this study are limited by the small size of the cohort. The 74 workers followed in this cohort represent the smallest number of workers across the occupational cohorts (Fingerhut et al., 1991; Steenland et al., 2001; Becher et al., 1998; Hooiveld et al., 1998; Michalek and Pavuk, 2008; McBride et al., 2009a, b) that have investigated TCDD exposures and cancer mortality. Mechanisms of follow-up were excellent as all individuals were traced, and death certificates were obtained from all deceased workers.

Although the study does compare the mortality experience to other occupational cohorts, the paper provides insufficient information to adequately interpret the associated findings. For example, a description of these occupations is lacking making it impossible to determine whether these cohorts were exposed to other occupational carcinogens that might have confounded the associations between TCDD exposure and cancer mortality.

2.4.1.1.1.2.1.3. *Suitability of data for TCDD dose-response modeling.*

Subsequent data assembled for the BASF cohort provide more detailed exposure characterization and also include information for 243 male workers employed at the plant. As such, this study did not meet the considerations for further dose-response analysis.

2.4.1.1.1.2.2. *Zober et al., 1990.*

2.4.1.1.1.2.2.1. *Study summary.*

Zober et al. (1990) also examined the mortality patterns of 247 individuals involved in the 1953 accident at the BASF plant. As detailed in their paper, the size of the original cohort was expanded by efforts to locate all individuals who were exposed in the accident or during the clean-up. Three approaches were followed in assembling the cohort. Sixty-nine cohort members were identified from the company physician's list of employees exposed as a result of the accident (Subcohort C1). Sixty-six of these workers were included in the original study population of workers Thiess et al. (1982) examined. Eighty-four other workers who were

potentially exposed to TCDD due to their involvement in demolitions or operations were added to the cohort. This group included 43 firemen, 18 plant workers, 7 bricklayers, 5 whitewashers, 4 mechanics, 2 roofers, and 5 individuals in other occupations (Subcohort C2). The cohort was further augmented through the Dioxin Investigation Program, which sought to locate those who were involved in the 1953 accident and were still alive in 1986. Current and former workers enrolled in the study were asked to identify other current or former coworkers (including deceased or retired) who might have been exposed from the accident. This third component of 94 workers (Subcohort C3) included 27 plant workers, 16 plumbers, 10 scaffolders, 10 professionals, 7 mechanics, 6 transportation workers, 5 bricklayers, 5 laboratory assistant, 3 insulators, and 5 individuals in other occupations. A medical examination was performed for those identified through the Dioxin Investigation Program, and blood measures were obtained for 28 of these workers.

External comparisons of the workers' mortality experience to the general population of the Federal Republic of West Germany were made using SMRs. Person-years were tabulated across strata defined by calendar period, sex, and age group. Sixty-nine deaths including twenty-three from cancer were detected among the workers during the 34-year follow-up period (November 17, 1953 through December 31, 1987). Cause-specific death rates for these same strata were available for the Federal Republic of West Germany. Stratified analyses were conducted to examine variations in the SMRs according to years since first exposure (0−9, 10−19, and ≥20 years) for each of the three subcohorts, as well as 114 workers with chloracne.

Although it was consistent in magnitude with findings from the NIOSH cohort, a statistically significant SMR for all cancer mortality was not observed (SMR = 1.17, 90% CI = 0.80−1.66). The SMRs for each of the three subcohorts varied substantially. For Subcohorts C1, C2, and C3, the SMRs were 1.30 (90% CI = 0.68−2.26), 1.71 (90% CI = 0.96−2.83), and 0.48 (90% CI = 0.13−1.23), respectively. The SMRs increased dramatically when analyses were restricted to those with 20 or more years since first exposure in Subcohort C1 (SMR = 1.67, 90% CI = 0.78−3.13) and Subcohort C2 (SMR = 2.38, 90% CI = 1.18−4.29). Meanwhile, in a subgroup analysis of those with chloracne, for the period of 20 or more years after first exposure, a statistically significant excess in cancer mortality was noted (SMR = 2.01; 90% CI = 1.22−3.15).

2.4.1.1.1.2.2.2. *Study evaluation.*

An important limitation of the study is the manner in which the cohort was constructed. Subcohort C3 was constructed by identifying individuals who were alive in 1986. This resulted in 97 active and retired employees who participated in the program, with 94 included in the analysis. Although these individuals did identify other workers who might have also retired or died, inevitably, some individuals who had died were not included in the cohort. This would serve to underestimate the SMRs that were generated with external comparisons to the German population. Indeed, cancer mortality rates in this subcohort were about half of what would have been expected based on general population rates (SMR = 0.48, 90% CI = 0.13−1.23). Additionally, more than half of Subcohort C2 were firemen (43 of 84), who would likely have been exposed to other carcinogens as a consequence of their employment. Quantitative analyses of epidemiologic data for firefighters have demonstrated increased cancer risk for several different forms of cancer (Youakim, 2006). Therefore, potential confounding from other occupational exposures of the firefighters could have contributed to the higher SMR in Subcohort C2 cohort and is a concern. Data on cigarette smoking were not available either. No excess for nonmalignant respiratory disease was found, however, suggesting this might not be an important source of bias.

2.4.1.1.1.2.2.3. *Suitability of data for TCDD dose-response modeling.*

As with the Thiess et al. (1982) publication, worker exposure was not estimated. Lack of exposure estimates precludes a quantitative dose-response analysis using these data. Also, the study design is not well suited to characterization of risk using the SMR statistic. Mortality is also likely under-ascertained in the large component of the cohort that was constructed through the identification of surviving members of the cohort.

2.4.1.1.1.2.3. *Ott and Zober, 1996.*

2.4.1.1.1.2.3.1. *Study summary.*

Ott and Zober (1996) extended the analyses of the BASF cohort to include estimates of individual-level measures of TCDD. The researchers also investigated associations with cancer mortality and identified incident cancer cases. The cohort follow-up period of 39 years extended until December 31, 1992, adding 5 years to a previous study (Zober et al., 1990). Ott and Zober

(1996) identified incident cases of cancer using occupational medical records, death certificates, doctor's letters, necropsy reports, and information from self-reported surveys sent to all surviving cohort members. Self-reported cancer diagnoses were confirmed by contacting the attending physician.

This study characterized exposure by two methods: (1) determining chloracne status of the cohort members and (2) estimating cumulative TCDD (μg/kg) levels. In 1989, serum measures were sought for all surviving members of the 1953 accident, and serum TCDD levels were quantified for 138 individuals. These serum levels were used to estimate cumulative TCDD concentrations for all 254 members of the accident cohort. Ott et al. (1993) published a description of the exposure estimation procedure, which was a regression model that accounted for the circumstances and duration of individual exposure. The average internal half-life of TCDD was estimated to be 5.8 years based on repeated serum sampling of 29 individuals. The regression model allowed for this half-life to vary according to the percentage of body fat, and yielded half-lives of 5.1 and 8.9 years among those with 20% and 30% body fat, respectively. Previous analyses of this cohort had used a half-life of 7.0 years (Ott et al., 1993).

TCDD half-life has been reported to increase with percentage of body fat in both laboratory mammals (Geyer et al., 1990) and humans (Zober and Papke, 1993). Ott and Zober (1996) contend that observed correlations with chloracne severity and cumulative estimates of TCDD exposure indirectly validated this exposure metric. Specifically, the mean TCDD concentration for those without chloracne was 38.4 ppt; for those with moderate and severe forms of chloracne, the mean was 420.8 ppt and 1,008 ppt, respectively.

Unlike for the NIOSH cohort, individual-level data were collected for other cancer risk factors. These factors included body mass index at time of first exposure, history of occupational exposure to β-naphthylamine and asbestos, and history of smoking. Smoking data were available for 86% of the cohort. SMRs were based on the external referent population of West Germany. For cancer incidence, Ott and Zober (1996) generated standardized incidence ratios (SIRs) using incidence rates for the state of Saarland (1970−1991) as the external referent. They calculated SMRs (and SIRs) for three categories of cumulative TCDD levels: <0.1 μg/kg, 0.1−0.99 μg/kg and ≥1 μg/kg. The Cox regression model was used to characterize risk within the cohort using a continuous measure of TCDD. These analyses considered the potential confounding influence of age, smoking, and body mass index using a stepwise regression

modeling approach. The Cox modeling employed a stratified approach using the date of first exposure to minimize possible confounding between calendar period and exposure. The three first exposure groups were exposure within the first year of the accident, exposure between 1 year after the accident and before 1960, and exposure after 1959. The Cox regression estimates were presented in terms of conditional risk ratios (i.e., hazard ratios adjusted for body mass index, smoking and age).

Although no statistically significant excesses relative to the general population were detected for all cancer mortality, there was some suggestion of an exposure-response relationship. In the 0.1−0.99 μg/kg and ≥1 μg/kg exposure groups, the all cancer SMRs were 1.2 (95% CI = 0.5−2.3) and 1.6 (95% CI = 0.9−2.6), respectively. Higher SMRs for cancer (all sites combined) were also found with an increased interval since exposure first occurred. Specifically, when observed versus expected counts of cancer were compared in the time interval 20 years after first exposure, the SMR in the highest exposure group (≥1 μg/kg) was 1.97 (95% CI = 1.05−5.36). An excess in lung cancer also was noted with the same lag in this exposure group (SMR = 3.06, 95% CI = 1.12−6.66). For cancer incidence, a statistically significant increased SIR for lung cancer was observed in the highest exposure category (SIR = 2.2, 95% CI = 1.0−4.3), but no other statistically significant associations were detected for any other cancer site. No cases of soft-tissue sarcoma were found among the cohort members in this analysis.

Based on internal cohort comparisons, Cox regression models also were used to generate hazard ratios as measures of relative risk for TCDD exposures following adjustment for smoking, age and body mass index. A statistically significant association between TCDD dose (per μg/kg) and cancer mortality was detected (RR = 1.22, 95% CI = 1.00−1.50), but not for cancer incidence (RR = 1.11, 95% CI = 0.91−1.35). Statistically significant findings were observed for stomach cancer mortality (RR = 1.46, 95% CI = 1.13−1.89) and incidence (RR = 1.39, 95% CI = 1.07−1.69).

The Ott and Zober (1996) study also compared the relationship between TCDD exposure categories and cancer mortality from all sites combined according to smoking status. Associations were noted between increased exposure to TCDD and mortality from cancer among smokers, but not among nonsmokers or former smokers.

2.4.1.1.1.2.3.2. *Study evaluation.*

The Ott and Zober (1996) study characterizes exposure to TCDD at an individual level. Therefore, unlike in past studies involving this cohort, these data can provide an opportunity for conducting quantitative dose-response modeling. As with the more recent studies involving the NIOSH cohort, serum samples were obtained from surviving cohort members and then used to back-extrapolate TCDD values for all cohort members. In the BASF cohort, however, serum data were available for a much higher percentage of cohort members (54%) than in the NIOSH cohort (5%). An additional study strength was the collection of questionnaire data, which allowed for the potential confounding from cigarette smoking and body mass index to be examined.

The Ott and Zober (1996) study also evaluates the relationship between TCDD and cancer incidence. Most cohort studies of TCDD-exposed workers have relied solely on mortality outcomes. The availability of incidence data better allows for period of latency to be described, and moreover, to characterize risks associated with cancers that typically have long survival periods. The authors provide few details on the expected completeness of ascertainment for incident cancer cases, which makes determining any associated bias difficult. They do, however, suggest that nonfatal cancers are more likely to have been missed in the earlier part of the follow-up. The net result of differential case ascertainment over time makes evaluating differences in risk estimates across different periods of latency impossible.

The small sample size of the cohort (n = 243 men) likely limited the statistical power to detect small associations for some of the exposure measures. This also effectively limited the ability to analyze dose-response relationships quantitatively, particularly across strata such as time since exposure. For site-specific analyses, the cancer site with the most cancer deaths was the respiratory system (n = 11). Thus, quantitative dose-response analysis using these cohort data would be limited to the evaluation of all cancer sites combined.

The most important limitation of this study is related to the construction of the third component of the cohort. As mentioned earlier, this cohort was assembled by actively seeking out surviving members of the cohort in the mid-1980s. The mortality experience of this cohort is much lower than that of the general population over the entire follow-up, a result that is expected given that the individuals were known to be alive as of 1986. The net result is likely an underestimate of the SMR.

2.4.1.1.1.2.3.3. *Suitability of data for TCDD dose-response modeling.*

This study was included in the quantitative dose-response modeling for the 2003 Reassessment (U.S. EPA, 2003). The characterization of exposure data and availability of other risk factor data at an individual level are appropriate for use in quantitative dose-response analyses.

2.4.1.1.1.3. ***The Hamburg cohort.***

The Hamburg cohort has been the subject of several cancer risk assessments. As with the NIOSH and BASF cohorts, analyses have progressed from basic comparisons of mortality experience to general population rates to more sophisticated internal cohort analyses involving the reconstruction of TCDD exposures using serum measures. This cohort consists of approximately 1,600 workers who were employed in the production of herbicides at a plant in Hamburg, Germany during 1950−1984 (Flesch-Janys et al., 1995; Becher et al., 1998). The herbicides produced included 2,4,5-T, β-hexachlorocyclohexane and lindane. The production of TCP and 2,4,5-T was halted in 1954 following a chloracne outbreak. The plant ceased operations in 1984. Approximately 20 different working areas were identified, which, in turn, were grouped into five main areas based on putative TCDD exposure levels. One working area was deemed to be extremely contaminated, having TCDD exposures at least 20-fold higher than in other areas. In this section, the studies undertaken in this cohort that have examined cancer mortality are summarized.

2.4.1.1.1.3.1. *Manz et al., 1991.*

2.4.1.1.1.3.1.1. *Study summary.*

Manz et al. (1991) investigated patterns of mortality in the Hamburg cohort. The study population consisted of 1,583 workers (1,184 men, 399 women) who were employed for at least three months between 1952 and 1989. Casual workers were excluded as they lack sufficient personal identifying information thereby not allowing for associations with mortality outcomes to be examined. Vital status was determined using community-based registries of inhabitants throughout West Germany. Cause of death until the end of 1989 was determined from medical records for all cancer deaths and classified based on the ninth revision of the International Classification of Diseases (WHO, 1978). Although Manz et al. (1991) present some data on

cancer incidence for the cohort, the data are incomplete as information was available on only 12 cases; 93 cancer deaths were observed in the cohort.

In this study, the authors used information on production processes to group workers into categories of low, medium, or high exposure to TCDD. This information was based on TCDD concentrations in precursor materials, products, waste, and soil from the plant grounds, measured after the plant closed in 1984. The distribution of workers into the low, medium, and high exposure groups was 186, 901, and 496, respectively. The authors examined the validity of the three exposure categories using a separate group of 48 workers who provided adipose tissue samples. The median exposure of the 37 volunteers in the high group was 137 and 60 ng/kg in the remaining 11. Information about chloracne in the cohort was incomplete, and, therefore, was not used as a marker of TCDD exposure. Other surrogate measures of exposure were considered in this study, including duration of exposure and year of first employment. For the latter measure, employment that began after 1954 was assumed to result in much lower exposures given that production of 2,4,5-T and TCP stopped in 1954.

External comparisons of cancer mortality were made by calculating SMRs using the general population of West Germany as a referent. Comparisons of mortality in the cohort also were made to a separate cohort of 3,417 gas supply workers to avoid bias from a healthy worker effect. Vital status and cause of death in the gas supply workers were determined using the same methods as used in the Hamburg cohort. SMRs were calculated relative to both referent populations (West Germany and gas supply workers) across low, medium, and high TCDD exposure groups. The comparison of mortality to the gas supply workers, however, extended only until the end of 1985, whereas, comparisons to the general population extended until 1989. Stratified analyses were undertaken to calculate SMRs for each of the three exposure groups for categories of duration of employment (<20 versus ≥20 years) and date of entry into the cohort (≤1954 versus >1954).

When compared to the general population, overall cancer mortality was elevated in male cohort members (SMR = 1.24, 95% CI = 1.00−1.52) but not in females (SMR = 0.80, 95% CI = 0.60−1.05). A two-fold increase in female breast cancer mortality was noted but was of borderline statistical significance (SMR = 2.15, 95% CI = 0.98−4.09). The SMR among men was further increased when analyses were restricted to workers who were employed for at least 20 years (SMR = 1.87, 95% CI = 1.11−2.95). Analyses restricted to those in the highest

exposure group produced an even higher SMR for those with at least 20 years of employment (SMR = 2.54, 95% CI = 1.10−5.00). Statistically significant excesses in risk were detected among those who first worked before 1954, but not afterward. Furthermore, a dose-response trend was observed across increasing exposure categories in the subset of workers employed before 1954. The SMRs using the cohort of gas supply workers as the referent group for the low, medium, and high groups in this subset were 1.41 (95% CI = 0.46−3.28), 1.61 (95% CI = 1.10−2.44), and 2.77 (95% CI = 1.59−4.53), respectively. This finding is consistent with what was known about TCDD exposures levels at the plant, namely, that TCDD concentrations were much higher between 1951 and 1954, with subsequent declining levels after 1954.

Generally speaking, patterns of excess mortality were similar when the cohort of gas workers was used as a reference group. The overall SMR for men was 1.39 (95% CI = 1.10−1.75); and was 1.82 (95% CI = 0.97−3.11) when analyses were restricted to workers with 20 or more years of employment. A dose-response trend also was observed across exposure categories when analyses were restricted to those employed for at least 20 years. In particular, with these analyses, no cancer deaths were observed among those in the lowest exposure group, while the SMRs in the middle and high exposure groups were 1.36 (95% CI = 0.50−2.96) and 3.07 (95% CI = 1.24−6.33).

SMRs also were generated for several site-specific cancers relative to the West German general population and the gas worker cohort. No statistically significant excesses were observed using the general population reference. In contrast, statistically significant excesses were observed for lung cancer (SMR = 1.67, 95% CI = 1.09−2.44) and hematopoietic system cancer (SMR = 2.65, 95% CI = 1.21−5.03) relative to the gas workers cohort.

2.4.1.1.1.3.1.2. *Study evaluation.*

The Manz et al. (1991) findings indicate an excess of all cancer mortality among the workers with the highest exposures, particularly those who worked for at least 20 years and were employed before 1954. The findings across categories of exposure within the subsets of workers employed for at least 20 years and before 1954, particularly using the cohort of gas supply workers, are consistent with a dose-response relationship. These elevated cancer mortality rates found among those employed before 1954 were likely due to higher TCDD exposures. Other

carcinogenic coexposures, such as benzene, asbestos, and dimethyl sulfate, could have occurred among this population. Given that no substantial changes in the production processes at the Hamburg plant occurred after 1954, comparable levels of these coexposures would be expected before and after 1954. Therefore, confounding due to these coexposures is unlikely to explain the dose-response relationship demonstrated between all cancer mortality and TCDD exposures. No information, however, was presented on potential exposure to other dioxin-like compounds which may confound the associations that were detected.

Detailed information on workers' smoking behaviors was not collected. Limited evidence indicated, however, that smoking prevalence between the Hamburg cohort and the gas supply workers cohort was quite similar. A nonrepresentative sample of 361 workers in the Hamburg cohort and the sample of 2,860 workers in the gas supply cohort indicated that the self-reported smoking prevalence was 73% and 76%, respectively. This suggests that the two cohorts are comprised predominantly of smokers. The similarity in overall smoking prevalence indicates that comparisons of cancer mortality between the two groups are not unduly influenced by an inability to adjust for smoking.

2.4.1.1.1.3.1.3. *Suitability of data for TCDD dose-response modeling.*

The data compiled for the Manz et al. (1991) study do satisfy many of the considerations for conducting quantitative dose-response analysis; health outcomes appear to be ascertained in an unbiased manner, and exposure was characterized on an individual-level basis. However, as demonstrated in later studies, there was a large dioxin-like compound component that was not quantified or assessed in this study. Dose-response associations between TCDD and cancer mortality were detected, with stronger associations observed with increased periods of latency and for those who first worked when TCDD was at higher levels.

The size of the cohort, although not as large as the NIOSH cohort, does offer sufficient statistical power to evaluate TCDD-related risk for cancers from all cancer sites. The data are limited, however, for characterizing cancer risks among women; only 20 cancer deaths occurred in the 399 women included in the cohort. It is unlikely that the findings are biased by confounding due to cigarette smoking since dose-response patterns were strengthened when comparisons were made to the cohort of gas supply workers rather the general population referent where smoking rates were likely lower. The inability to account for other occupational

exposure when TCDD exposures were much higher (pre-1955) could result in confounding if these other exposures were related to TCDD and the health outcomes under consideration. This data set would be suitable for quantitative dose-response modeling if the exposure characterization of the cohort could be improved using biological measures of dose.

2.4.1.1.1.3.2. *Flesch-Janys et al., 1995.*

2.4.1.1.1.3.2.1. *Study summary.*

In 1995, Flesch-Janys et al. published an analysis of the male employees from the Hamburg cohort that extended the follow-up to 40 years (1952−1992). Inclusion of these three additional years of follow-up resulted in a sample size of 1,189 male workers.

The authors estimated a quantitative exposure variable for concentrations of TCDD in blood at the end of exposure (i.e., when employment in a department ended) and above German median background TCDD levels. The TCDD exposure assessment defined 14 production departments according to TCDD levels in various products in the plant, in waste products, and in various buildings. The time (in years) each worker spent in each department then was calculated. Concentrations of TCDD were determined in 190 male workers using serum (n = 142) and adipose tissue samples (n = 48). The authors used a first-order kinetic model to calculate TCDD levels at the end of exposure for the 190 workers with available polychlorinated dibenzo-p-dioxin (PCDD) and -furan (PCDF) at various time points. Half-lives were calculated from an elimination study of 48 workers from this cohort, and the median TCDD background level was estimated at 3.4 ng/kg blood fat from the German population (Papke et al., 1994; Flesch-Janys et al., 1994). Using the one-compartment, first-order kinetic model, the half-life of TCDD was estimated to be 6.9 years (Flesch-Janys, 1997). Increased age and higher body fat percentage were associated with increased TCDD half-life, while smoking was associated with a higher decay rate for most of the congeners examined (Flesch-Janys et al., 1996). Cumulative TCDD exposures were estimated by summing exposures over the time spent in all production departments and were expressed in terms of ng/kg of blood fat. The authors also applied a metric of total toxicity equivalence (TOTTEQ) as the weighted sum of all congeners where weights were TEQs that denoted the toxicity of each congener relative to TCDD.

Similar to previous analyses on this cohort, comparisons were made using an external referent group of workers from a gas supply company (Manz et al., 1991). In contrast to

previous analyses where SMR statistics were generated using this "external" reference, however, Flesch-Janys et al. (1995) used Cox regression. The Cox regression models treated the gas worker cohort as the referent group, and six exposure groups were defined by serum-derived cumulative TCDD estimates. The groups were determined by using the first four quintiles with the upper two exposure categories corresponding to the ninth and tenth deciles of the cumulative TCDD. Internal cohort comparisons used those workers in the lowest quintile as the referent group, as opposed to the cohort of gas workers. A similar approach was used to model TEQs. No known TCDD exposures occurred in the gas workers, so they were assigned exposures based on the median background levels in the general population. RRs were calculated based on exposure above background levels; in other words, background levels were assumed to be equivalent across all workers and also for those employed by the gas supply company. The RRs derived using the Cox model were adjusted for total duration of employment, age, and year when employment began.

The Cox regression with the cohort of gas workers as the referent exposure group yielded a linear dose-response relationship between cumulative TCDD exposure and cancer mortality for all sites combined ($p < 0.01$). The RRs for all-cancer mortality were 1.59, 1.29, 1.66, 1.60, 1.70, and 3.30. For four of the six categories (excluding the referent group), the RRs were statistically significant ($p < 0.05$); in the highest TCDD exposure category (344.7−3890.2 ng/kg) the RR was 3.30 (95% CI = 2.05−5.31). Similar findings were evident with TOTTEQ. A dose-response pattern for all cancer mortality ($p < 0.01$) based on the internal cohort comparisons was also detected.

The authors performed an additional analysis to evaluate the potential confounding role of dimethylsulfate. Although no direct measures of dimethylsulfate were available, the investigators repeated analyses by excluding 149 workers who were employed in the department where dimethylsulfate was present. A dose-response pattern persisted for TCDD ($p < 0.01$), and those in the highest exposure group (344.7−3,890.2 ng/kg of blood fat) had a RR of 2.28 (95% CI = 1.14−4.59).

2.4.1.1.1.3.2.2. *Study evaluation.*

The Flesch-Janys et al. (1995) study used serum-based measures to determine cumulative exposure to TCDD at the end of employment for all cohort members. They used the standard

one-compartment, first-order kinetic model and samples obtained from 190 male workers. This quantitative measure of exposure permits an estimation of a dose-response relationship.

Confounding for other occupational exposures is unlikely to have biased the results. A dose-response relationship persisted after excluding workers exposed to dimethylsulfate. Other potential exposures of interest included benzene and isomers of hexachlorocyclohexane. Exposure to these agents, however, was highest in the hexachlorocyclohexane and lindane department, where TCDD exposures were lower. As outlined earlier, the study findings are unlikely to be biased for cigarette smoking as cigarette smoking in the cohort was similar to that in the comparison population. Moreover, more recent analyses of serum-based TCDD exposure measures found no correlation with smoking status in this cohort (Flesch-Janys et al., 1995)—a necessary condition for confounding.

The authors used an exposure metric that described cumulative TCDD exposure of workers at the time they were last exposed. As a result, the authors were unable to characterize risks associated with this metric for different periods of latency despite a sufficient follow-up period. Subsequent analyses constructed time-dependent measures of cumulative TCDD and accounted for excretion of TCDD during follow-up.

In contrast to most risk assessments of TCDD exposure, this study modeled the relationship between other dioxin-like compounds and the risk of cancer mortality using the TOTTEQ metric.

2.4.1.1.1.3.2.3. *Suitability of data for TCDD dose-response modeling.*

The data used in this study satisfy most of the considerations developed for performing a quantitative dose-response analysis. However, latency period was not examined in this study. Dose-response analyses were, therefore, limited to a subsequent study of this cohort (Becher et al., 1998), which did examine latency.

2.4.1.1.1.3.3. *Flesch-Janys et al., 1998.*

2.4.1.1.1.3.3.1. *Study summary.*

Flesch-Janys et al. (1998) undertook another analysis on this cohort that incorporated additional sera data for 275 workers (39 females and 236 males). The follow-up period was the same as that used in the 1995 analyses, with mortality follow-up extending until December 31,

1992. Analyses were based on 1,189 males who were employed for at least 3 months from January 1, 1952 onward. The authors continued this dose-response analysis to address limitations in their previous work. One limitation was that the previous method did not account for the elimination of TCDD while exposures were being accrued during follow-up. A second limitation was that the amount of time workers spent in different departments was not considered. In the 1998 study, the "area under the curve" approach was used because it accounts for variations in concentrations over time and reflects cumulative exposure to TCDD. The authors used a first-order kinetic model to link blood levels and working histories to derive department-specific dose rates for TCDD. The TCDD background level of 3.4 ng/kg blood fat for the German population was used (Papke et al., 1994). The dose rates were applied to estimate the concentration of TCDD at every point in time for all cohort members. A cumulative measure expressed as ng/kg blood fat multiplied by years was calculated and used in the SMR analysis. SMRs were calculated using general population mortality rates for the German population between 1952 and 1992. No lag period was incorporated into the derivation of the SMRs. The SMRs were estimated for the entire cohort and for exposure groups based on quartiles obtained from the area under the curve. Linear trend tests were also performed. The overall SMR for cancer mortality in the cohort was 1.41 (95% CI = 1.17−1.68). This SMR value was higher than the SMR of 1.21 reported for this same cohort with 3 fewer years of follow-up (Manz et al., 1991). In terms of site-specific cancer mortality, excesses were found for respiratory cancer (SMR = 1.71, 95% CI = 1.24−2.29) and rectal cancer (SMR = 2.30, 95% CI = 1.05−2.47). Increased risk for lymphatic and hematopoietic cancer (SMR = 2.16, 95% CI = 1.11−3.17) were also noted largely attributable (SMR = 3.73, 95% CI = 1.20−8.71) to lymphosarcoma (i.e., non-Hodgkin's lymphoma). A dose-response relationship was observed across quartiles of cumulative TCDD for all-cancer mortality ($p < 0.01$). The SMRs for these quartiles were 1.24, 1.34, 1.34, and 1.73. Dose-response relationships were not observed for lung cancer or hematopoietic cancers using this same metric. Dose-response relationships were not observed with cumulative TEQ for any of the cancer sites examined (i.e., all cancers, lung cancer, hematopoietic cancer).

2.4.1.1.1.3.3.2. *Study evaluation.*

The approach used in the Flesch-Janys et al. (1998) study offers a distinct advantage over earlier analyses involving the same cohort. Three more years of follow-up were available, and the characterization of exposure using the "area under the curve" better captures changes in cumulative exposure using a person-years approach rather than cumulative TCDD at the time of last exposure. As noted previously, other occupational exposures or cigarette smoking are unlikely to have biased the study findings. A sufficient length of follow-up had accrued, and dose-response associations were evident. Dioxin-like compounds were evaluated in this study. For TCDD, the mean concentration was 101.3 ng/kg at the time of measurement. For other higher chlorinated congeners, the corresponding mean (without TCDD) was 89.3 ng/kg.

2.4.1.1.1.3.3.3. *Suitability of data for TCDD dose-response modeling.*

The data used in this study satisfy most of the considerations developed for performing a quantitative dose-response analysis. However, latency was not examined in this study. Dose-response analyses were, therefore, limited to a subsequent study of this cohort (Becher et al., 1998) which did examine latency and supersedes the Flesch-Janys et al. (1998) study.

2.4.1.1.1.3.4. *Becher et al., 1998.*

2.4.1.1.1.3.4.1. *Study summary.*

The Becher et al. (1998) quantitative cancer risk assessment for the Hamburg cohort was highlighted in the 2003 Reassessment as being appropriate for conducting dose-response analysis. The integrated TCDD concentration over time, as estimated in the Flesch-Janys et al. (1998) study, was used as the exposure variable. Estimates of the half-life of TCDD based on the sample of 48 individuals with repeated measures were incorporated into the model that back-calculated TCDD exposures to the end of the employment (Flesch-Janys et al., 1996). This method took into account the age and body fat percentage of the workers. In Becher et al. (1998), the analysis used the estimate of cumulative dose (integrated dose or area under the curve) as a time-dependent variable.

Poisson and Cox regression models were used to characterize dose-response relationships. Both models were applied to internal comparisons where a person-years offset was used and to an external comparison where an offset of expected number of deaths was used.

The person-years offset was used to account for varying person-time accrued by workers across exposure categories. The use of the expected number of deaths as an offset allows risks to be described in relation to that expected in the general population. Within each classification cell of deaths and person-years, a continuous value TCDD and TEQ levels based on the geometric mean were entered into the Poisson model. For the Cox model, accumulated dose was estimated based on area under the curve for TCDD, TEQ, TEQ without TCDD, and β-hexachlorocyclohexane. These other coexposure metrics were adjusted for in the Cox regression analyses. Other covariates considered included in the models were year of entry, year of birth, and age at entry into the cohort. A background level of 3.4 ng/kg blood fat for the German population was used (Papke et al., 1994). A variety of latencies was evaluated (0, 5, 10, 15, and 20 years), and attributable risk and absolute risk were estimated. The unexposed cohort of gas workers was used for most internal analyses.

Internal and external comparisons using the Poisson model found positive associations with TCDD exposure and mortality from all cancers combined. The slope associated with the continuous measure of TCDD (μg/kg blood fat × years) for the internal comparison was 0.027 ($p < 0.001$), which decreased to 0.0156 ($p = 0.07$) after adjusting for age and calendar period. The slope for the external comparison was 0.0163 ($p = 0.055$); this estimate was not adjusted for other covariates. For TEQ, the slopes based on the internal comparisons were 0.0274 ($p < 0.001$) in the univariate model and 0.0107 ($p = 0.175$) in the multivariate model after adjusting for age and calendar period. The external estimate of slope for TEQ was 0.0109 ($p = 0.164$). Cox regression of TCDD across six exposure categories, with a lag of 0 years, found a statistically significant linear trend ($p = 0.03$) and those in the upper exposure group had a RR of 2.19 (95% CI = 0.76−6.29). These estimates were adjusted for year of entry, age at entry, and duration of employment. A similar pattern was observed with the Cox regression analysis of TEQ; the linear test for trend, however, was marginally statistically significant ($p = 0.06$).

Cox regression models that included both TCDD and TEQ (excluding TCDD) were applied. In this model, the slope parameter for TCDD was marginally statistically significant ($\beta = 0.0089$, $p = 0.058$), while the coefficient for TEQ (excluding TCDD) was not statistically significant ($\beta = -0.024$, $p = 0.70$). This suggests that confounding by other dioxin-like compounds was unlikely and the increased risk of cancer was due to TCDD exposure. For all

TEQs combined, the parameter estimate was generated from another Cox regression model and was borderline statistically significant ($\beta = 0.0078$, $p = 0.066$).

The authors used multiple Cox models to evaluate the effect of latency. The slope estimates for both TCDD and TEQ increased dramatically with increasing latency. The slope estimates for TCDD increased from 0.0096 to 0.0160 ($p < 0.05$) when latency was increased from 0 to 20 years. Similar changes in the TEQ slopes were noted (0.0093 to 0.0157). Evaluations of dose-response curves found that the best-fitting curve was concave in shape, thereby yielding higher risk at low exposure. Differences between the fit of the class of models considered [i.e., $RR(x,\beta) = \exp(\beta \log(kx = 1))$], however, were small.

Attributable risks were generated only for TCDD, as the data suggested no effects with other TEQs. The additional lifetime risk of cancer assuming a daily intake of 1 pg TCDD/kg body weight/day was estimated to range between 0.001 and 0.01.

2.4.1.1.1.3.4.2. *Study evaluation.*

The Becher et al. (1998) study represent perhaps the most detailed analyses performed on any cohort to date. The findings were robust, as similar patterns were found with and without using the gas supply worker cohort as the referent group. Exposures to other potential confounding coexposures, such as dioxin-like compounds, were taken into account, and workers with exposure to other carcinogens (e.g., lindane) were excluded. Furthermore, latency was examined in this study, unlike earlier studies of this cohort.

2.4.1.1.1.3.4.3. *Suitability of data for TCDD dose-response modeling.*

This study was included in the quantitative dose-response modeling for the 2003 Reassessment (U.S. EPA, 2003). The data in the Becher et al. (1998) study are suitable for conducting quantitative dose-response modeling. The exposure data capture cumulative exposure to TCDD as well as exposures to other dioxin-like compounds. The length of the follow-up is sufficient, and the study appears to not be subject to confounding or other types of biases. Therefore, this study is utilized in quantitative dose-response analysis.

2.4.1.1.1.4. ***The Seveso cohort.***

Several studies have evaluated the morbidity and mortality effects of residents exposed to TCDD following a July 10, 1976, accidental release through an exhaust pipe at a chemical plant in the town of Meda near Seveso, Italy. The released fluid mixture contained 2,4,5-T, sodium trichlorophenate, ethylene glycol, and sodium hydroxide. Vegetation in the area showed immediate signs of damage, and in the days following the accident, residents developed nausea, headaches, eye irritation, and dermal lesions, particularly children.

This accident transported TCDD up to 6 km from the plant. Soil samples taken near the plant revealed average levels of TCDD that ranged from 15.5 $\mu g/m^2$ to 580.4 $\mu g/m^2$ in the most contaminated area near the plant (referred to as Zone A) (Bertazzi et al., 2001). Zone A covered 87 hectares and extended 2,200 m south from the plant. Another, more distant contaminated zone (Zone B) covering 270 hectares also had contaminated soil levels, but the TCDD concentration range was much lower (1.7−4.3 $\mu g/m^3$). A reference zone (Zone R), which surrounded the two contaminated areas, had lower TCDD soil levels (range: 0.9−1.4 $\mu g/m^3$) and included approximately 30,000 residents. Following the accident, most residents in Zone A left the area. Although residents in Zone B remained, they were under strict regulations to avoid consuming homegrown products. In total, 736, 4,737, and 31,800 individuals lived in Zones A, B, and R, respectively. Within days of the accident, 3,300 animals (mostly poultry and rabbits) were found dead. Emergency slaughtering was undertaken to prevent TCDD from entering the food chain, and within 2 years more than 80,000 animals had been slaughtered. Mechanisms were put into place for long-term follow-up of these residents. Unlike the other studies based on occupational cohorts, the follow-up of this population allows for risks to be characterized for females.

The mortality studies from Seveso published to date have not incorporated serum TCDD levels that were measured in individuals. Needham et al. (1997) describe the collection of serum samples from a sample of the exposed population and control subjects in 1976. In 1988, human exposure to TCDD was assessed by measuring small volumes of serum remaining from medical examinations done in 1976. An examination of these data revealed some of the highest serum TCDD levels ever reported, that the half-life of TCDD in this population was between 7 and 8 years, and that half-life varied between women and men. The half-life of TCDD in serum was longer in women (~9 years) than in men (~7 years) (Needham et al., 1994). In this report, the

findings of studies that characterized cancer risks in relation to exposure to TCDD from the 1976 accident are highlighted. These studies include comparisons of cancer mortality rates to the general population based on zone of residence at the time of accident (Bertazzi et al., 2001; Consonni et al., 2008). More recent work done by Warner et al. (2002) investigated the relationship between serum-based measures of TCDD and breast cancer among participants in the Seveso Women's Health Study (SWHS).

2.4.1.1.1.4.1. *Bertazzi et al., 2001.*

2.4.1.1.1.4.1.1. *Study summary.*

Several studies have reported on the mortality experience of Seveso residents. The more recent publications having a longer follow-up of the cohort are evaluated here. In 2001, the findings from a 20-year mortality study of Seveso residents was published (Bertazzi et al., 2001). The Bertazzi et al. (2001) study was an extension of the 10- and 15-year follow-ups for mortality (Bertazzi et al., 1989, 1997; Pesatori et al., 1998) and the 10-year follow-up for cancer incidence (Bertazzi et al., 1993).

In this cohort, TCDD exposures were assigned to the population using a three-level categorical variable representative of the individual's place of residence (Zones A, B, or R) at the time of the accident or when the person first became a resident of the zone, if that was after 1976. An external comparison to the province of Lombardy was made by generating rate ratios (RR) using Poisson regression techniques. Person-years of follow-up were tabulated across strata defined by age, zone of residence, duration of residence, gender, calendar time, and number of years that had elapsed since the time of exposure. Mortality rates during the preaccident period also were compared to evaluate potential changes in rates due to the accident and to evaluate whether patterns were consistent before and after the accident.

No overall excess in mortality rates from all cancer sites combined was observed in Zones A or B (combined) when compared to the reference population of Lombardy (n = 9 million residents) (RR = 1.0, 95% CI = 0.9−1.2). Analyses of site-specific cancer mortality revealed statistically significant excesses among residents in Zones A or B (combined) for cancer of the rectum (RR = 1.8, 95% CI = 1.0−3.3) and lymphatic and hematopoietic malignancies (RR = 1.7, 95% CI = 1.2−2.5). Lymphatic and hematopoietic malignancies were elevated in women (RR = 1.8, 95% CI = 1.1−3.2) and in men (RR = 1.7, 95% CI = 1.0−2.8).

Analyses stratified by the number of years since first exposure (i.e., 1976) revealed higher risk among men with an increased number of years elapsed. Similar to other studies, the RR for all cancers (combined) was 1.3 (95% CI = 1.0-1.7) among men 15−20 years after first exposure. No such increase after 15 years postexposure, however, was noted in women (RR = 0.8, 95% CI = 0.6−1.2).

2.4.1.1.1.4.1.2. *Study evaluation.*

Ascertainment of mortality appears to be excellent. Vital status was established using similar methods for both the exposed and reference populations. No individual data were collected and, therefore, the possibility that confounding by individual characteristics such as cigarette smoking cannot be entirely dismissed. Bertazzi et al. (2001) do note that the sociodemographic characteristics of residents in the three zones were similar based on independently conducted surveys, and no differences in chronic respiratory disease were found across the different zones. If excess mortality was attributable to cigarette smoking, such excesses would be expected to be evident during the entire study period. Latency analyses revealed elevated risks 15−20 years postaccident. Finally, no excesses were observed for other smoking-related cancers of the larynx, esophagus, pancreas, and bladder. The observed excesses in all cancer mortality do not appear to be attributed to differential smoking rates between the two populations.

To examine potential for bias due to noncomparability in the two study populations, a comparison of cancer mortality rates between the Seveso regions and the reference population of Lombardy was conducted. Elevated rates for brain cancer mortality were noted in Seveso relative to Lombardy, but the higher rates of leukemia mortality were found in Lombardy relative to Seveso. That no excess was reported for all cancer sites combined lends credence to the hypothesis that the exposure to TCDD from the accident increased rates of cancer after a sufficient period of latency.

Stratified analyses were performed across several categorical variables including gender and time since exposure. The numbers of cancer site-specific deaths are quite small in many of the 5-year increments since first exposure. The study, therefore, has limited statistical power to detect differences in mortality rates among the comparison groups for many cancer sites.

Bertazzi et al. (2001) assigned exposures based on zone of residence. Soil sampling within each zone revealed considerable variability in TCDD soil levels within each zone. Moreover, some individuals would have left the area shortly after the accident, and determining the extent to which individuals in Zone B who were subject to the recommendations near the time of the accident adhered to them is difficult. As a result, exposure misclassification is possible, and the use of individual measures of TCDD level in serum is preferred over zone of residence for determining exposure. As noted by the authors, the study is better suited to "hazard identification" than to quantitative dose-response analysis.

2.4.1.1.1.4.1.3. *Suitability of data for TCDD dose-response modeling.*

Given the variability in soil TCDD levels within each zone and the lack of individual level, no effective dose can be estimated for quantitative dose-response analyses. Uncertainty in identifying the critical exposure window for the Seveso cohort is a key limitation. The evaluation of this study indicates that this study is not suitable for quantitative dose-response analysis.

2.4.1.1.1.4.2. *Pesatori et al., 2003.*

2.4.1.1.1.4.2.1. *Study summary.*

Pesatori et al. (2003) published a review of the short- and long-term studies of morbidity and mortality outcomes in the Seveso cohort in 2003. This paper presented external comparisons of cancer incidence from 1977 to 1991 for Seveso males and females separately. As in the Bertazzi et al. (2001) study, RRs were estimated using Poisson regression. For males who lived in Zones A and B, the only statistically significantly elevated RR was for lymphatic and hematopoietic cancers (RR = 1.9, 95% CI = 1.1−3.1). This excess was due primarily to non-Hodgkin's lymphoma, which accounted for 8 of the 15 incidence cases (RR = 2.6, 95% CI = 1.3−5.3). Among females living in Zones A and B, higher rates were observed for multiple myeloma (RR = 4.9, 95% CI = 1.5−16.1), cancer of the vagina (RR = 5.5, 95% CI = 1.3−23.8), and cancer of the biliary tract (RR = 3.0, 95% CI = 1.1−8.2).

2.4.1.1.1.4.2.2. *Study evaluation.*

All RRs presented in the Pesatori et al. (2003) study were based on fewer than five incident cases.

2.4.1.1.1.4.2.3. *Suitability of data for TCDD dose-response modeling.*

As with the studies of mortality among Seveso residents, the Pesatori et al. (2003) study does not capture TCDD exposure on an individual basis, and soil TCDD levels considerably vary within each zone. Therefore, the quality of the exposure data is insufficient for estimating the effective dose needed for quantitative dose-response analysis.

2.4.1.1.1.4.3. *Consonni et al., 2008.*

2.4.1.1.1.4.3.1. *Study summary.*

Similar analytic methods were applied with 25 years of follow-up of the Seveso cohort (Consonni et al., 2008). An important addition in this paper was the presentation of RRs for Zone R, which had the lowest TCDD levels. Poisson regression models were used to calculate RRs of mortality using Seregno as the reference population. Cancer deaths observed in Zones A and B were 42 and 244, respectively.

No statistically significant differences in all cancer mortality relative to the reference population were noted in any of the zones (Zone A: RR = 1.03, 95% CI = 0.76−1.39; Zone B: RR = 0.92, 95% CI = 0.81−1.05; Zone R: RR = 0.97, 95% CI = 0.92−1.02). Statistically significant excesses in mortality from non-Hodgkin's lymphoma (RR = 3.35, 95% CI = 1.07−10.46) and multiple myeloma (RR = 4.34, 95% CI = 1.07−17.52) were observed in the area with the highest TCDD levels (Zone A). No other statistically significant increases in cancer mortality relative to the reference population were apparent. The absence of elevated breast cancer mortality among women in this study was noteworthy, as this finding differs from the results of a study of Seveso women for which TCDD exposures were estimated using serum samples (Warner et al., 2002). A more detailed description of this study is provided later in this section.

2.4.1.1.1.4.3.2. *Study evaluation.*

Although no individual-level data on smoking were available, the potential for confounding is likely minimal. Independent smoking surveys found that the smoking prevalence rates in Desio, one of cities affected by the accident, were similar to those in districts just outside the study area (Cesana et al., 1995). As mentioned earlier, one would expect elevated RRs over the entire study period if smoking had biased the study results, and not just after 15−20 years since exposure to TCDD.

2.4.1.1.1.4.3.3. *Suitability of data for TCDD dose-response modeling.*

The lack of individual-level exposure data precludes quantitative dose-response modeling using these data.

2.4.1.1.1.4.4. *Baccarelli et al., 2006.*

2.4.1.1.1.4.4.1. *Study summary.*

Given previous findings from Seveso, Baccarelli et al. (2006) examined t(14;18) translocations in the DNA of circulating lymphocytes of healthy dioxin-exposed individuals. These translocations are associated with the development of cancer, namely follicular lymphomas. The study included 211 healthy subjects of the Seveso area, and 101 who had developed chloracne. The investigators analyzed data from 72 high-TCDD plasma level individuals (≥10 ppt) and 72 low-TCDD plasma levels (<10 ppt). A three-level categorical variable was used to evaluate dose-response. This variable was developed by dividing those with exposures ≥10 ppt into two groups: 10– <50 ppt, and 50−475.0 ppt. Trained interviewers administered a questionnaire that collected data on demographic characteristics, diet, and residential and occupational history.

The prevalence of t(14;18) was estimated as those individuals having a t(14;18) positive blood sample divided by the t(14;18) frequency (number of copies per million lymphocytes). Baccarelli et al. (2006) found that the frequency of t(14;18) was associated with plasma TCDD levels, but no association between TCDD and the prevalence of t(14;18) was detected.

2.4.1.1.1.4.4.2. *Study evaluation.*

Whether the frequency of t(14;18) associated with plasma TCDD levels translates into an increased risk of lymphoma is uncertain as prospective data of TCDD on those who developed non-Hodgkin's lymphoma are lacking. Moreover, the t(14;18) translocation could be an important event in the pre-B stage cell that contributes to tumorigenicity, however subsequent exposure to carcinogenic agents might be necessary for t(14;18) cells to develop into a malignancy (Hoglund et al., 2004).

2.4.1.1.1.4.4.3. *Suitability of data for TCDD dose-response modeling.*

Given that current TCDD plasma levels were measured for this study, it is unclear if the effects of lymphocyte translocations may be due to initial high exposure or are a function of the cumulative exposure for a longer exposure window. Additionally, whether the frequency of t(14;18) associated with plasma TCDD levels translates into an increased risk of lymphoma is unknown. Dose-response analysis for this outcome, therefore, was not conducted.

2.4.1.1.1.4.5. *Warner et al., 2002.*

2.4.1.1.1.4.5.1. *Study summary.*

To date, Warner et al. (2002) is the only published investigation of the relationship between serum-based measures of TCDD and cancer in Seveso. Eligible participants from the SWHS (see Section 2.4.1.2.1.4 for details) were women who, at the time of the accident in 1976, were 40 years of age or younger, had lived in one of the most highly contaminated zones (A or B), and had adequate sera collected soon after the explosion. Enrollment in SWHS was begun in March 1996 and lasted until July 1998. Of the total 1,271 eligible women, 981 agreed to participate in the study. Cancer cases were identified during interview and confirmed through review of medical records. Information on other risk factors including reproductive history and cigarette smoking was obtained through interview.

Serum volumes greater than 0.5 mL collected between 1976 and 1981 volume were analyzed. Most sera were collected in 1976/77 (n = 899); samples were collected in 1978−1981 for 54 women, and in 1996/97 for 28 women. For most samples collected after 1977, serum TCDD levels were back-extrapolated using a first-order kinetic model with a 9-year half-life

(Pirkle et al. 1989). For 96 women with undetectable values, a serum level that was equal to one-half the detection level was used.

Analyses were based only on women who provided serum samples; no extrapolation of values to a larger population was done. Risks were therefore generated using data collected at an individual level. Serum TCDD was analyzed as both a continuous variable and a categorical variable. The distribution of serum TCDD levels of the 15 cases of breast cancer was examined in relation to the distribution of all women in the SWHS. The median exposure was slightly higher among with the 15 cases of breast cancer (71.8 ppt) compared to those without (55.1 ppt), and the exposure distribution among breast cancer cases appeared to be shifted to the right (i.e., the exposures were higher but followed the same distribution); however, no formal test of significance was conducted.

Warner et al. (2002) used Cox proportional hazards modeling techniques to evaluate risk of breast cancer in relation to TCDD serum levels while controlling for a variety of potential risk factors. In all, 21 women had been diagnosed with cancer, and of these, 15 cases were cancer of the breast. The analysis revealed that for every 10-fold increase in TCDD log-serum levels (e.g., from 10 to 100 ppt) the risk of breast cancer increased by 2.1 (95% CI = 1.0−4.6). Risk estimates also were generated across four categories (<20, 20.1−44, 44.1−100, >100 ppt), with the lowest category used as the reference. The RRs estimated in the third and fourth highest exposure categories were 4.5 (95% CI = 0.6−36.8) and 3.3 (95% CI = 0.4−28.0). Although statistical significance was not achieved for either category, likely because of the small number of cases, the greater than threefold risk evident in both categories is worth noting. Given that the reference category had only one incident case underscores the limited inferences that can be drawn from these analyses. The authors adjusted for numerous potential confounders, but observed no differences between the crude and adjusted results; the authors, therefore, presented unadjusted risks.

2.4.1.1.1.4.5.2. *Study evaluation.*

The findings from the Warner et al. (2002) study differ from reports in earlier studies in which mortality outcomes noted the absence of an SMR association. The design of this study is much stronger than earlier ones, given the improved characterization of exposure, the ability to compare incidence rates within the cohort, the ability to control for potential confounding

variables at an individual level, and the availability of incident outcomes. The use of incident cases (versus mortality data) should also help minimize potential bias due to disease survival. Another important advantage was the ability to measure TCDD near the time of the accident, thereby reducing the potential for exposure measurement error.

A potentially important limitation of the Warner et al. (2002) study was that information was collected only from those who were alive as of March 1996. Therefore, TCDD and other relevant risk factor data could not be collected for those who had previously died of breast cancer. Thirty-three women could not participate because they were either too ill or had died. Of these, three died of breast cancer. Given that there were only 15 breast cancer cases, the exclusion of these 3 cases could have dramatically impacted the findings in either direction.

Another limitation was that, at the time of the follow-up, most women were still premenopausal and therefore, most of the cohort had not yet attained the age of greater risk of breast cancer. An ongoing follow-up of the cohort should be completed by 2010, which should allow for increased number of incident breast cancers to be identified. Given that the current analyses were based only on 15 incident cases, this will substantially improve the statistical power of the study. A secondary benefit is that the increased follow-up will allow for an investigation of possible differential effects according to the age the women were at the time of exposure.

2.4.1.1.1.4.5.3. *Suitability of data for TCDD dose-response modeling.*

Several aspects of the Warner et al. (2002) study are weaknesses in the consideration of this study for further dose-response modeling. Only 15 cases of breast cancer were available, and no increases in risk were found with serum TCDD exposures between 20.1 and 44 ppt ($n = 2$) when compared to those with <20 ppt ($n = 1$). The average age at the time of enrollment was 40.8 years while the average age at diagnosis among the cases was 45.2 years. As most women had not yet reached the age when breast cancer cases are typically diagnosed, additional follow-up of the cohort would improve the quantitative dose-response analysis and strengthen this study. A key strength of this study, however, is that Warner et al. (2002) includes an investigation of the relationship between individual serum-based measures of TCDD and cancer in Seveso. Despite the weaknesses, this study meets the evaluation considerations and criteria for inclusion and will be analyzed for quantitative dose-response modeling.

2.4.1.1.1.5. ***Chapaevsk study.***

Industrial contamination of dioxin in the Chapaevsk region of Russia has been the focus of research on the environmentally-induced cancer and other adverse health effects. The Chapaevsk region is located in the Samara region of Russia and has a population of 83,000. The region is home to a chemical plant that produced lindane and its derivatives between 1967 and 1987, which are believed to be responsible for local dioxin contamination. Soil sampling has demonstrated a strong gradient of increased TCDD concentrations with decreased proximity to the chemical plant (Revich et al., 2001).

2.4.1.1.1.5.1. *Revich et al., 2001.*

2.4.1.1.1.5.1.1. *Study summary.*

Revich et al. (2001) used a cross-sectional study to compare mortality rates of Chapaevsk residents to two external populations of Russia and the region of Samara. Mortality rates for all cancers combined among males in Chapaevsk were found to be 1.2 times higher when compared to the Samara region as a whole and 1.3 times higher than Russia. Similar to other studies, statistically significant excess was noted in men (SMR = 1.8, 95% CI = 1.6−1.9) but not in women (SMR = 0.9, 95% CI = 0.8−1.1). Among men, the excess was highest for the smoking-related cancers of the lung (SMR = 3.1, 95% CI = 2.6−3.5) and larynx (SMR = 2.3, 95% CI = 1.2−3.8) and urinary organs (SMR = 2.6, 95% CI = 1.7−3.6). Among females, there was no increased SMR for all cancer sites combined, but excesses for breast cancer (SMR = 2.1, 95% CI = 1.6−2.7) and cancer of the cervix (SMR = 1.5, 95% CI = 1.0−3.1) were statistically significant.

Revich et al. (2001) also compared age-standardized cancer incidence rates in Chapaevsk to those in Samara. Although statistical tests examining these differences were not reported, higher incidence rates were observed for all cancers combined, cancer of the lip, cancer of the oral cavity, and lung and bladder cancer among males in Chapaevsk. Considerably lower cancer incidence rates also were observed for prostate cancer, cancer of the esophagus, and leukemia/lymphoma among males from Chapaevsk. Among females, incidence rates were higher in 1998 for all cancers in Chapaevsk when compared to Russia and the Samara region, an observation that appears somewhat counter to the presented SMR of 0.9 for all cancer mortality from 1995−1998. Like mortality, rates of breast cancer incidence among women in Chapaevsk

were higher than in Russia, as were rates of cervical cancer. Leukemia/lymphoma rates were higher among women in Chapaevsk than in those who lived in the reference populations of Samara and Russia. This finding is contrary to the finding for males who had lower rates of leukemia/lymphoma in Chapaevsk.

2.4.1.1.1.5.1.2. *Study evaluation.*

Although the Revich et al. (2001) findings suggest TCDD exposures in Chapaevsk are quite high relative to other parts of the world (Akhmedkhanov et al., 2002), evaluation of health outcomes to date have been based on ecological data only. This analysis did not adjust for the influence of other risk factors (e.g., smoking, reproductive characteristics) that could contribute to increased cancer rates for lung cancer in men and breast cancer in women. Given that both the SMRs and SIRs for cancer outcomes vary considerably between men and women, this suggests the possibility that occupational exposures might be a contributing factor in these adverse health outcomes.

Future research in Chapaevsk includes plans to conduct a breast cancer case-control study. Women who were born from 1940 onward and who have been diagnosed with breast cancer before the age of 55 were included in the study, although the plan to characterize TCDD using serum is uncertain (Revich et al., 2005).

2.4.1.1.1.5.1.3. *Suitability of data for TCDD dose-response modeling.*

This study did not meet the considerations and criteria for inclusion in a quantitative dose-response assessment. Given the lack of exposure data on an individual basis, no effective dose can be estimated for this study population. As such, no dose-response modeling was conducted.

2.4.1.1.1.6. *The Air Force Health ("Ranch Hands" cohort) study.*

Between 1962 and 1971, the U.S. military sprayed herbicides over Vietnam to destroy crops that opposition forces depended upon, to clear vegetation from the perimeter of U.S. bases, and to reduce the ability of opposition forces to hide. These herbicides were predominantly a mixture of 2,4-D, 2,4,5-T, picloram, and cacodylic acid (Institute of Medicine, 2006). A main chemical sprayed was Agent Orange, which was a 50% mixture of 2,4-D and 2,4,5-T. TCDD

was produced as a contaminant of 2,4,5-T and had levels ranging from 0.05 to 50 ppm (Institute of Medicine, 1994). A series of studies have investigated cancer outcomes among Vietnam veterans. A review of military records to characterize exposure to Agent Orange led Stellman and Stellman (1986) to conclude that assignment of herbicide levels should not be based solely on self-reports or a crude measure such as military branch or area of service within Vietnam. Investigations have been performed on the Ranch Hands cohort, which consisted of those who were involved in the aerial spraying of Agent Orange between 1962 and 1971. More elaborate methods were used to characterize exposures among these individuals, and these studies are summarized below.

2.4.1.1.1.6.1. *Akhtar et al., 2004.*

2.4.1.1.1.6.1.1. *Study summary.*

Akhtar et al. (2004) investigated the incidence of cancer in the Ranch Hand cohort, which was published after the release of the 2003 Reassessment document (U.S. EPA, 2003). The Ranch Hand Unit was responsible for aerial spraying of herbicides, including Agent Orange, in Vietnam from 1962 to 1971. Cancer incidence in the Ranch Hand cohort were compared to a cohort that included other Air Force personnel who served in Southeast Asia during the same period but were not involved in the spraying of pesticides. Health outcomes were identified during the postservice period that extended from the time each veteran left Southeast Asia until December 31, 1999. In contrast to previous analyses of this cohort, the Akhtar et al. (2004) study took into account concerns that both the comparison and spraying cohorts had increased risks of cancer, and addressed the possibility that workers with service in Vietnam or Southeast Asia might have increased cancer risk. The authors addressed the latter concern by adjusting risk estimates for the time spent in Southeast Asia and for the proportion of time spent in Vietnam.

The Ranch Hand cohort comprised 1,196 individuals, and the comparison cohort had 1,785 individuals. The comparison cohort was selected by matching date of birth, race, and occupation (i.e., officer pilot, officer navigator, nonflying officer, enlisted flyer, or enlisted ground personnel). TCDD levels were determined using serum levels collected from veterans who completed a medical examination in 1987. For those who did not have a serum measure taken in 1987, but provided one in subsequent years, TCDD levels were back-extrapolated to 1987 using a first-order kinetic model that assumed a half-life of 7.6 years. Those with

nonquantifiable levels were assigned a value of the limit of detection divided by the square root of 2. A total of 1,009 and 1,429 individuals in the Ranch Hand and comparison cohorts, respectively, provided serum measures that were used in the risk assessment. Veterans also were categorized according to the time their tours ended. This date corresponded to changes in herbicide use. These categories were before 1962 or after 1972 (no herbicides were used), 1962−1965 (before Agent Orange was used), 1966−1970 (when Agent Orange use was greatest), and 1971−1972 (after Agent Orange was used). Information on incident cases of cancer in the cohort was determined from physical examinations and medical records. Some malignancies were discovered at death and coded from the underlying causes of death as detailed on the death certificate. A total of 134 and 163 incident cases of cancer were identified in the Ranch Hand and comparison cohort, respectively. Akhtar et al. (2004) describe case ascertainment verified by record review as being complete.

External comparisons were made based on the expected cancer experience derived from U.S. national rates using SIRs and the corresponding 95% confidence interval. Person-years and events were tabulated by 5-year calendar and age intervals.

When compared to the general population, no statistically significant excesses in all cancer incidence were observed for either the Ranch Hand (SIR = 1.09, 95% CI = 0.91−1.28) or the comparison cohort (SIR = 0.94, 95% CI = 0.81−1.10). Statistically significant differences were found for three site-specific cancers in the Ranch Hands cohort relative to the general population. Excesses were noted for malignant melanoma (SIR = 2.33, 95% CI = 1.40−3.65) and prostate cancer (SIR = 1.46, 95% CI = 1.04−2.00). In contrast, a reduced SIR was found for cancers of the digestive system (SIR = 0.61, 95% CI = 0.36−0.96). The excess in prostate cancer was also noted in the comparison cohort (SIR = 1.62, 95% CI = 1.23−2.10) relative to the general population. External comparisons were repeated by restricting the cohorts to the period when Agent Orange was used (1966−1970). Again, no statistically significant excesses in all cancer incidence were noted in the Ranch Hand (SIR = 1.14, 95% CI = 0.95−1.37) or comparison cohort (SIR = 0.94, 95% CI = 0.80−1.11). Statistically significant excesses continued to be observed for malignant melanoma (SIR = 2.57, 95% CI = 1.52−4.09) and prostate cancer (SIR = 1.68, 95% CI = 1.19−2.33) in the Ranch Hand component of the cohort. No other statistically significant differences were found among Ranch Hands personnel.

For internal cohort analyses, veterans were assigned to one of four exposure categories. Those in the comparison cohort were assigned to the "comparison category." Ranch Hand veterans that had TCDD serum levels <10 ppt were assigned to the "background" category. Those with a TCDD levels >10 ppt had their TCDD level estimated at the end of their Vietnam service with a first-order kinetic model that used a half-life of 7.6 years. These back-extrapolated values that were less than 118.5 ppt were assigned to a "low" exposure group, while those with values above 118.5 ppt were classified as "high" exposure. Akhtar et al. (2004) used Cox regression models to describe risks across the exposure groups using the comparison category as the reference. Risks were adjusted for age at tour, military occupation, smoking history, skin reaction to sun exposure, and eye color. Internal cohort analyses were restricted to those who spent no more than 2 years in Southeast Asia and Ranch Hand workers who served exclusively in Vietnam, and the comparison cohort who served exclusively outside of Vietnam.

Statistically significant excesses of cancer incidence (all sites combined) were observed in the highest two exposure groups. A statistically significant trend test ($p = 0.04$) was detected based on the RRs for the background-, low-, and high- exposure groups: 1.44 (95% CI = 0.82−2.53); 2.23 (95% CI = 1.24−4.00), and 2.02 (95% CI = 1.03−3.95). For malignant melanoma, the RRs across the three increasing exposure categories were 2.99, 7.42, and 7.51. The corresponding risk estimates for prostate cancer were 1.50, 2.17, and 6.04.

2.4.1.1.1.6.1.2. *Study evaluation.*

An important strength of this study is the manner in which TCDD exposure was estimated. Serum data were available for most veterans, and therefore, generalizing exposure from a small sample of cohort members is not a concern as was the case with the NIOSH and Hamburg cohorts. Back-extrapolating to derive past exposures was based on a methodology that has been applied in many of the cohorts, thereby facilitating risk comparisons. An additional strength of the study is the examination of incidence as a measure of disease occurrence rather than mortality.

In contrast to the previous analysis (Ketchum et al., 1999) the analysis by Akhtar et al. (2004) was restricted to individuals who spent no more than 2 years in Southeast Asia. Previous research had demonstrated that increased time spent in Southeast Asia was associated with an increased risk of cancer. Confounding might have been introduced given that the comparison

cohort spent much more time in Southeast Asia than the Ranch Hands. To illustrate, the median number of days spent in Southeast Asia was 790 for comparison cohort members, and the median days for the Ranch Hand cohort in the background, low, and high exposure groups were 426, 457, and 397, respectively. After restricting to those who spent at most 2 years, statistically significant associations were observed for all cancer sites combined, prostate cancer, and malignant melanoma using the internal cohort comparisons.

An important issue in the study is the high correlation between 2,4,5-T and 2,4-D, given that both were used in equal concentrations in Agent Orange. As a result, distinguishing the effects of each is impossible. This point is relevant, given that 2,4-D has been associated with prostate cancer in several studies. As a result, the dose-response association with prostate cancer might be due to 2,4-D exposure and not TCDD. This issue also has implications for the interpretation of the dose-response pattern for all cancer sites combined, given that incident prostate cancers accounted for 4 of the 12 incident cases in the high-exposure group.

2.4.1.1.1.6.1.3. *Suitability of data for TCDD dose-response modeling.*

The ascertainment of incident cases and characterization of exposure to TCDD based on serum measures are strong features of the cohort. Confounding by 2,4-D is a major concern. Since delineating the independent effects of other Agent Orange contaminants is not possible, quantitative dose-response analysis was not conducted on this study.

2.4.1.1.1.6.2. *Michalek and Pavuk, 2008.*

2.4.1.1.1.6.2.1. *Study summary.*

Michalek and Pavuk (2008) recently published an updated analysis of the incidence of cancer and diabetes in the cohort of Ranch Hand veterans. As with the Akhtar et al. (2004) analysis, the study included a comparison cohort of other Air Force veterans who served in Southeast Asia at the same time but were not involved with the spraying of herbicides. This study extended previous analyses (Henriksen et al., 1997; Ketchum et al., 1999) by addressing the number of days of herbicide spraying, calendar period of service, and the time spent in Southeast Asia. Veterans who attended at least one of five examinations were eligible for inclusion. Incident cancer cases also were identified from medical records.

The methods used to determine TCDD exposures were as described above in the review of the Akhtar et al. (2004) study. Blood measures also were taken in 1992, 1997, and 2002 for subjects with no quantifiable TCDD levels in 1987, those who refused in 1987, and those new to the study. TCDD dose at the end of service in Vietnam was assigned to Ranch Hands that had TCDD levels above background using a constant half-life of 7.6 years. Each veteran was then assigned to one of four dose categories: comparison, background (i.e., Ranch Hands with 1987 TCDD ≤10 ppt), low (Ranch Hands with 1987 levels 10.1−91 ppt), and high (Ranch Hands with 1987 ≥91 ppt). Serum TCDD estimates are available for 1,597 veterans in the comparison cohort, and 986 veterans in the Ranch Hand cohort. The comparison cohort was selected by matching on date of birth, race, and occupation of the Ranch Hands.

Michalek and Pavuk (2008) used Cox regression to characterize risks of cancer incidence across the three upper exposure categories using the comparison category as the referent group. Risk estimates were adjusted for year of birth, race, smoking, body mass index at the qualifying tour, military occupation, and skin reaction to sun exposure. Tests for trend for increased risk of cancer were conducted by testing the continuous covariate $\log_{10}$TCDD.

Overall, no association between the TCDD exposure categories and RR of all-site cancer was observed. Those in the highest exposure group had an RR of 0.9 (95% CI = 0.6−1.4). Stratified analyses by calendar period of service showed more pronounced risk for those who served before 1986 (when higher amounts of Agent Orange were used). A statistically significant dose-response trend ($p < 0.01$) was observed for cancer risk and $\log_{10}$TCDD exposure. The RRs for the background, low, and high groups used in these comparisons were 0.7 (95% CI = 0.4−1.3), 1.7 (95% CI = 1.0−2.9), and 1.5 (95% CI = 0.9−2.6). A statistically significant increase, however, was noted when analyses were restricted to those who had sprayed for at least 30 days before 1967 and spent time in Southeast Asia (RR = 2.2, 95% CI = 1.1−4.4).

2.4.1.1.1.6.2.2. *Study evaluation.*

Michalek and Pavuk (2008) used the same study population that Akhtar et al. (2004), and so it has the same strengths and limitations as noted above. The follow-up, however, extends an additional 5 years (until the end of 2004). The findings for the dose-response analyses were not as compelling as the earlier Akhtar et al. (2004) findings.

2.4.1.1.1.6.2.3. *Suitability of data for TCDD dose-response modeling.*

The key limitation precluding dose-response analysis for the Michalek and Pavuk (2008) study is the possible confounding from the inability to control for 2,4-D and other agents used in Agent Orange. As such, quantitative dose-response analysis was not conducted on this study.

2.4.1.1.1.7. *Other studies of potential relevance to dose-response modeling.*

2.4.1.1.1.7.1. *t' Mannetje et al., 2005—New Zealand herbicide sprayers.*

2.4.1.1.1.7.1.1. *Study summary.*

t'Mannetje et al. (2005) described the mortality experience of a cohort of New Zealand workers who were employed in a plant located in New Plymouth. The plant produced phenoxy herbicides and pentachlorophenol between 1950 and the mid-1980s. This study population also was included in the international cohort of producers and sprayers of herbicides that was analyzed by IARC (Saracci et al., 1991; Kogevinas et al., 1997). In this 2005 study, analyses were restricted to those who had worked at least 1 month; clerical, kitchen, and field research staff were excluded. The authors followed up 1,025 herbicide producers and 703 sprayers from 1969 and 1973, respectively, until the end of 2000.

The cohort consisted of two components: those involved with the production of herbicides and those who were sprayers. For the herbicide producers, exposures were determined by consulting occupational history records; no direct measures of exposure were available. Each department of employment was assigned to one of 21 codes as in the IARC international cohort (Saracci et al., 1991). Industrial hygienists and factory personnel with knowledge of potential exposures in this workforce classified each job according to potential to be exposed to TCDD, other chlorinated dioxins, and phenoxy herbicides. Exposure was defined as a dichotomous variable (i.e., exposed and unexposed). Among producers, 813 were classified as exposed, with the remaining 212 considered unexposed.

The "sprayer" component of the cohort includes those who were registered in the national registry of applicators at any time from January 1973 until the end of 1984. For the sprayers, detailed occupational information was lacking. Exposure was, therefore, based on an exposure history questionnaire completed in a previous study of congenital malformations (Smith et al., 1982). This questionnaire, administered to 548 applicators in 1980 and 232 applicators in 1982, achieved a high response rate (89%). Participants were asked to provide information about

2,4,5-T-containing product use on an annual basis from 1969 up to the year the survey was completed. As the use of 2,4,5-T ceased in the mid-1980s, data on occupational exposure to TCDD among these workers are fairly complete. Virtually all sprayers (699 of 703) were exposed to TCDD, higher chlorinated dioxins, and phenoxy herbicides.

Deaths among workers were identified through record linkage to death registrations in the New Zealand Health Information Service. Electoral rolls, drivers' licenses, and social security records also were consulted to confirm identified deaths. External comparisons of mortality were made to the New Zealand population using the SMR statistic. The mortality follow-up for the producers began on January 1, 1969 and extended until December 31, 2000. For the sprayers, the follow-up period extended from January 1, 1973 until December 31, 2000. A total of 43 cancer deaths occurred in the producer group and 35 cancer deaths occurred in the sprayer group in the cohort. Where possible, stratified analyses by duration of employment and department were conducted. The departments examined for producers included synthesis, formulation and lab, maintenance and waste, packing and transport, other, and unexposed. SMRs were generated using the New Zealand population as an external referent. A linear test for trend was applied to evaluate dose-response trends according to categories of duration of employment. Stratified analyses also were also done for sprayers who started working before 1973, as TCDD levels in 2,4,5-T produced at the New Zealand plant dropped dramatically after 1973. Although an SMR was presented for female producers, given that only one cancer death was observed, this study can provide no insight on differential risks between the sexes.

Among TCDD-exposed producers, for all cancers combined, no statistically significant excess mortality was found when compared to the general population (SMR = 1.24, 95% CI = 0.90−1.67). No dose-response trend in the SMRs for all cancers was observed with duration of employment ($p = 0.44$). No statistically significant elevated SMR was observed in any of the duration of employment categories for any of the six specific departments examined. A statistically significant positive linear trend, however, was noted among synthesis workers ($p = 0.04$). There was some suggestion of reduced mortality in the upper exposure levels for workers in the formulation and lab departments. For sprayers, the SMR for all cancer sites combined was not elevated relative to the New Zealand general population (SMR = 0.82, 95% CI = 0.57−1.14), nor was a dose-response pattern observed with increasing duration of employment ($p = 0.86$). Additionally, no statistically significant excess in cancer mortality for

all sites combined was evident in workers who were first employed either before 1973 (SMR = 0.75, 95% CI = 0.50−1.07) or from 1973 on (SMR = 1.81, 95% CI = 0.59−4.22). For site-specific analyses of cancer mortality, an excess of multiple myeloma was observed among production workers relative to the general population (SMR = 5.51, 95% CI = 1.14−16.1). This SMR was based on three deaths. No statistically significant excess (or deficit) of mortality was found for any other cancer site examined in either the sprayers or the producers.

2.4.1.1.1.7.1.2. *Study evaluation.*

The physical activity demands of spraying contribute to a healthy worker effect that manifests itself in a lower SMR based on both external comparisons to the general population as a referent, and the SMR generated for the producers in the cohort. The analyses conducted using a simple dichotomy of exposure and duration of employment are limited, as nearly all of the sprayers were unexposed.

The dose-response pattern with duration of employment coupled with the observation that higher levels of exposure to TCDD occurred among workers in the synthesis department is an important finding. These workers were also exposed to several other contaminants, however, that include processing chemicals, technical products, intermediates, and byproducts (Kauppinen et al., 1993). These included phenoxy herbicides and dioxin-like compounds such as chlorinated dioxins. Since the dichotomous exposure measure was based on exposure to TCDD, chlorinated dioxins and phenoxy herbicides, the associated dose-response analyses presented in this study should be interpreted cautiously in light of the inability to either characterize or control for these potential confounders. As such, these co-exposures might have contributed to the dose-response pattern observed with increased duration of employment in the synthesis workers.

2.4.1.1.1.7.1.3. *Suitability of data for TCDD dose-response modeling.*

Although the study authors completed a subsequent analysis of this cohort using serum-derived TCDD (McBride et al., 2009a), the lack of individual-level TCDD exposures precludes dose-response modeling.

2.4.1.1.1.7.2. *McBride et al., 2009b—New Zealand herbicide sprayers.*

2.4.1.1.1.7.2.1. *Study summary.*

McBride et al. (2009b) published an updated analysis of the mortality of the New Zealand cohort. The follow-up period was from January 1, 1969 to December 31, 2004 extending the previous study by an additional 4 years. In contrast to the previous study where the cohort comprised individuals employed for at least 1 month prior to 1982 (or 1984) (t'Mannetje et al., 2005), the cohort in this study consisted of all those who worked at least one day between January 1, 1969 and October 1, 2003. This resulted in a cohort of 1,754 workers, of which 247 died in the follow-up interval. Seventeen percent of the cohort members were lost to follow-up, which could be a source of selection bias if loss to follow-up was related to both the exposure metrics and the health outcome of interest. Previous data from this cohort (t'Mannetje et al., 2005), however, showed fairly comparable loss to follow-up rates among the unexposed (23%) and the exposed populations (17%).

Comparisons to the New Zealand general population were made using the SMR statistic. Stratified analyses were conducted by duration of employment (<3 months, ≥3 months), sex, latency (<15 years, ≥15 years), and period of hire (<1976, ≥1976). The authors defined latency as the period between the day last worked and the earliest of date of death, date of emigration or loss to follow-up, or December 31, 2004.

The overall SMR for mortality from all cancer sites combined relative to the New Zealand population was 1.01 (95% CI = 0.85−1.10). Although not statistically significant there was suggestion of an increased risk of rectal cancer (SMR = 2.03; 95%CI = 0.88–4.01) among the employees. SMRs for lymphatic and hematopoietic cancers (overall SMR = 1.21, 95% CI = 0.52–2.39) included 3.12 (95% CI = 0.08−17.37) for Hodgkin's disease, 1.59 (95% CI = 0.43–4.07) for non-Hodgkin's lymphoma and 3.73, 95% CI = 1.20−8.71), and 1.66 (95% CI = 0.20–5.99) for multiple myeloma. No statistically significant excess of cancer mortality was noted among workers employed for <3 months (SMR = 1.19, 95% CI = 0.65−2.00), or for ≥3 months (SMR = 0.98, 95% CI = 0.75−1.26). A statistically significant excess of digestive cancers was found for those who worked fewer than 3 months relative to the New Zealand population (SMR = 2.52, 95% CI = 1.15−4.78). No excesses were observed for any site-specific cancers when analyses were restricted to those who worked for 3 or more months. No statistically significant elevated SMRs were found for all cancers

(combined) either for a latency period of fewer than 15 years (SMR = 1.14, 95% CI = 0.72−1.71) or a latency period of ≥15 years (SMR = 0.96, 95% CI = 0.72−1.26). Similarly, no statistically significant excess in cancer mortality was observed for all cancer sites combined, or any site-specific cancer when analyses were stratified by date of hire (<1976, ≥1976) or by sex. The SMR among women who were employed at the site was 0.68 (95% CI = 0.45−1.00).

2.4.1.1.1.7.2.2. *Study evaluation.*

High rates of emigration in New Zealand (9% among workers in the cohort) contributed to a fairly high loss to follow-up (22% among workers) during the study period. The loss to follow-up would reduce the overall mortality estimates among the workers, which could underestimate the SMRs if loss to follow-up (and health status) was not comparable in the general population. For example, it is unclear if workers and the general population who emigrated were sicker than those remaining in the cohort. Previous data from the cohort workers suggests that loss to follow-up rates were slightly higher among the low and unexposed populations (t'Mannetje et al., 2005; McBride et al., 2009a) worker population, so presumably the highly exposed workers were not lost to follow-up more so than other workers.

2.4.1.1.1.7.2.3. *Suitability of data for TCDD dose-response modeling.*

This study extended the mortality follow-up and included stratified analyses to investigate effect modification by period of latency, sex, and date of hire. A key limitation was the lack of direct measures of exposure for study participants which precluded estimating effective dose needed for dose-response modeling. This study did not meet the considerations and criteria for inclusion in quantitative dose-response analysis.

2.4.1.1.1.7.3. *McBride et al., 2009a—New Zealand herbicide sprayers.*

2.4.1.1.1.7.3.1. *Study summary.*

McBride et al. (2009a) recently published the mortality experience of the New Zealand cohort in relation to serum estimates of TCDD levels. This study included 1,599 workers who were employed between 1969 and November 1, 1989, which was the date that 2,4,5-T was last used. As in their study published earlier in the same year (McBride et al., 2009b), the follow-up period extended from the first day of employment until December 31, 2004. Vital status was

ascertained through record linkage to the New Zealand Health Information Service Mortality Collection and the Registrar General's Index to Deaths for deaths up to 1990.

All current and former workers who lived within 75 km of the plant were invited to provide serum samples. A total of 346 of the eligible workers (68%) provided samples, which represented 22% of the overall study population (346/1599). Based on the serum measures, 70% (241/346) had been exposed to TCDD. This percentage is similar to the estimated 71% of workers who were deemed to have been exposed based on a review of occupational records. The mean serum TCDD value was 9.9 ppt. The highest exposures were observed for those employed in the trichlorophenol operation (23.4 ppt). Values among unexposed workers averaged 4.9 ppt, which is close to the background level of 3.9 ppt among individuals of similar age in the New Zealand general population (Bates et al., 2004). Details on smoking histories of individuals were also collected for the 346 individuals who provided serum, allowing for an examination of the potential confounding role that smoking might have on derived risk estimates for TCDD.

Cumulative exposure to TCDD as a time-dependent metric was estimated for each worker. A detailed description of the methods used to derive TCDD exposure was described in Aylward et al. (2009). The qualitative TCDD scores available for those with serum measures were used to estimate the cumulative exposures based on a half-life of approximately 7 years. A time-dependent estimate of TCDD exposure was derived and the area under the curve was used to obtain cumulative workplace TCDD exposure above background levels. Model performance appears modest as the model explained only 30% of the variance (adjusted R^2) when these TCDD exposure estimates were compared with actual serum levels (Aylward et al., 2009).

As with previous analyses of the cohort (t' Mannetje et al., 2005; McBride et al., 2009b), external comparisons to the New Zealand general population were made using the SMR statistic. The SMR statistic also was used to compare mortality across four exposure groups relative to the general population, as defined by the serum TCDD estimates: 0−68.3, 68.4−475.0, 475.1−2085.7, and ≥2085.8 ppt-month. The proportional hazards model also was used to conduct internal cohort comparisons across these same four exposure groups. In these analyses, age was used as the time variable, and the covariates of date of hire, sex, and birth year were included in the proportional hazards model. The cut-points for these four exposure categories were chosen so that approximately equal numbers of deaths were included in each category.

Consistent with earlier SMR analyses of the same cohort, no increased cancer mortality was observed among "ever" exposed workers in this cohort when compared to the general population (SMR = 1.1, 95% CI = 0.9−1.4). No statistically significant excess was noted for any of the site-specific cancers, although there was some suggestion of increased risk of soft tissue sarcoma (SMR = 3.4, 95% CI = 0.1–19.5), multiple myeloma (SMR = 2.2, 95% CI = 0.2−8.1), non-Hodgkin's lymphoma (SMR = 1.6, 95% CI = 0.3–4.7), and cancer of the rectum (SMR = 2.0, 95% CI = 0.7−4.4). No statistically significant increases in cancer mortality (all sites combined) was found in any of the four exposure categories as measured by the SMR statistic, nor was a dose-response trend noted with increasing exposure categories. No dose-response trend (based on SMR analyses) was noted for five site-specific cancers examined (i.e., digestive organs, bronchus, trachea and lung, soft tissue sarcomas, lymphatic and hematopoietic tissue, and non-Hodgkin's lymphoma), although SMRs for three of the four exposure categories exceeded 2.0.

In contrast to the external cohort comparisons, the RRs generated with the proportional hazards model supported a dose-response trend, as rate ratios increased across increasing TCDD exposure categories. The RRs and their 95% confidence intervals relative to the lowest of the four groups were 1.05 (95% CI = 0.48−2.26), 1.38 (95% CI = 0.64−2.97) and 1.58 (95% CI = 0.71−3.52). Neither the linear ($p = 0.29$) or quadratic ($p = 0.82$) test for trend, however, was statistically significant. An increased risk of lung cancer mortality was observed in the highest TCDD exposure category relative to the lowest (RR = 5.75, 95% CI = 0.76−42.24). The tests for trend for lung cancer, however, also were not statistically significant.

A smoking survey was administered to a sample of surviving workers of this cohort, and smoking prevalence was found to be slightly higher among those with higher cumulative exposure (61%) compared to lower exposures (51−56%). These minor differences in smoking prevalence unlikely was a strong enough confounder to explain the fivefold increase in risk of lung cancer found in the highest exposure category. Although the smoking data assessment was a strength of the study, it was limited to only sample of workers and was not available for those who died of lung cancer.

2.4.1.1.1.7.3.2. *Study evaluation.*

Given high rates of emigration, loss to follow-up (22%) was a potential concern in this study. If comparable emigration rates did occur among the general population then the SMRs would be underestimated. It is unclear to what extent emigration occurred among the general population and whether emigration in both the worker and general populations was dependent on health status. If emigration rates were comparable among these two populations, the associated bias from the under-ascertainment of mortality in the lost to follow-up group would likely attenuate a positive association between TCDD and cancer mortality. Among the worker population, there was not much evidence of differential loss to follow-up with respect to exposure as average exposures were lower (3.2 ppt) among those loss to follow up compared to those with complete follow-up (5.7 ppt). Previous studies among this population also found slightly higher loss to follow-up rates among the unexposed (23%) compared to the exposed (17%) workers (t'Mannetje et al., 2005).

McBride et al. (2009a) did not present results using a continuous measure of TCDD exposure (lagged or unlagged) as was done in most other occupational cohorts. Additionally, the modeling did not consider the use of different periods of latency.

2.4.1.1.1.7.3.3. *Suitability of data for TCDD dose-response modeling.*

There is no evidence that the authors considered exposure metrics that are consistent with environmental cancer-causing agents such as exposure modeling that takes latency into account. Given that past occupational cohort studies of TCDD-exposed workers have consistently demonstrated stronger association with lag interval of 15 years, such an approach should be applied to this cohort. This precludes this study from consideration for quantitative dose-response modeling.

2.4.1.1.1.7.4. *Hooiveld et al., 1998—Netherlands workers.*

2.4.1.1.1.7.4.1. *Study summary.*

Hooiveld et al. (1998) re-analyzed the mortality experience of a cohort of workers employed in two chemical plants in the Netherlands using 6 additional years of follow-up from an earlier study (Bueno de Mesquita et al., 1993). The cohort consisted of those employed between 1955 and June 30, 1985, and vital status was ascertained until December 31, 1991 (i.e.,

36 years of follow-up). These cohort members were involved in the synthesis and formulation of phenoxy herbicides, of which the main product was 2,4,5-trichlorophenoxyacetic acid and monochloroacetic acid. This cohort, with a shorter follow-up interval than the original study (t'Mannetje et al., 2005), was included in the IARC international cohort. The cohort consisted of 1,167 workers, of which 906 were known to be alive at the end of the follow-up. The average length of follow-up was 22.3 years, and only 10 individuals were lost to follow-up.

The authors used detailed occupational histories to assign exposures. Workers were classified as exposed to phenoxy herbicides or chlorophenols and contaminants if they worked in selected departments (i.e., synthesis, finishing, formulation, packing, maintenance/repair, laboratory, chemical effluent waste, cleaning, shipping-transport, or plant supervision); were exposed to the accident in 1963; or were exposed by proximity (i.e., if they entered an exposed department at least once a week). The 1963 accident was the result of an uncontrolled reaction in the autoclave in which 2,4,5-trichlorophenol was synthesized; an explosion resulted, with subsequent release of PCDDs that included TCDD. Based on these methods of exposure assignment, 562 workers were deemed to be exposed to phenoxy herbicides or chlorophenols, and 567 were unexposed. Due to limited information, 27 workers were classified as having unknown exposure.

TCDD exposures also were assigned using serum measured on a sample of workers who were employed for at least 1 year and first started working before 1975. Dioxin-like compounds including PCDDs were also measured in the serum samples, but not analyzed for this study. Of the 144 subjects who were invited to provide samples, 94 agreed. TCDD levels were back-extrapolated to the time of maximum exposure using a one-compartment, first-order kinetic model that used a half-life estimate of 7.1 years. The mathematical model used was $\ln(\text{TCDDmax}) = \ln(\text{TCDD}) + \text{lag} \times \ln(2)/7.1$. The lag was defined as the number of years since last exposure for those exposed by virtue of their normal job duties. For those exposed as a result of the accident in 1963, the lag was defined as the number of years since the accident occurred.

The authors made external comparisons of cohort mortality to the Netherlands population using the SMR statistics. Poisson regression was used to perform internal cohort comparisons using unexposed workers as the referent. RRs (measured using rate ratios) generated from the Poisson model also were used to compare mortality based on low, medium, and high TCDD

serum-derived categories. The Poisson model included the following covariates as adjustment factors: age, calendar period at end of follow-up, and time since first exposure.

When compared to the general population, workers had an excess mortality from cancer (SMR = 1.5, 95% CI = 1.1−1.9), based on 51 cancer deaths. Generally, no excesses were observed for site-specific cancers. The exception included eight deaths from cancers of the urinary organs (SMR = 3.9, 95% CI = 1.7−7.6). Although not statistically significant, SMRs comparable in magnitude to other studies were detected for non-Hodgkin's lymphoma (SMR = 3.8, 95% CI = 0.8−11.0**)** and Hodgkin's disease (SMR = 3.2, 95% CI = 0.1−17.6). A statistically significant excess of cancer mortality (*n*=20 deaths among occupational workers) also was also observed relative to the general population when analyses were restricted to those exposed as a result of the 1963 accident (SMR = 1.7, 95% CI = 1.1−2.7). Three deaths from prostate cancer were also noted among these workers (SMR = 5.2, 95% CI = 1.1−15.3), but no excess was observed with any other cancer site.

Internal cohort comparison also demonstrated an increased risk of all cancer mortality among those exposed to phenoxy herbicides, chlorophenols, and contaminants relative to those unexposed (RR = 4.1, 95% CI = 1.8−9.0). A statistically significant increased risk was also noted for respiratory cancer mortality (RR = 7.5, 95% CI = 1.0−56.1). Analyses across categories of TCDD exposure revealed excesses in cancer mortality for all cancer sites combined; however, no dose-response trend was apparent.

2.4.1.1.1.7.4.2. *Study evaluation.*

Several other studies that have characterized cohorts by TCDD levels have used the area under the curve approach and thus have derived an exposure metric that is time dependent. Hooiveld et al. (1998) instead created an exposure metric to capture the maximum exposure attained during the worker's employment. Characterizing risks using this metric assumes that other TCDD exposures accrued during a workers' lifetime are not relevant predictors of cancer risk.

2.4.1.1.1.7.4.3. *Suitability of data for TCDD dose-response modeling.*

One study limitation is that although dioxin-like compounds were measured in the serum samples, Hooiveld et al. (1998) reported associations with mortality for TCDD only. There is

some utility to examining dose-response analyses using alternative exposure metrics as those constructed in this cohort. However, the small number of identified cancer deaths, limitations in terms of the exposure assignment (based on nonrepresentative sample, and maximum exposure level) and concern over potential confounding by co-exposures preclude using these data for a dose-response analysis.

2.4.1.1.2. *Key characteristics of epidemiologic cancer studies*

See Table 2-1 at the end of the chapter for a comparison of the length of follow-up, latency period used, the half-life for TCDD used, and the fraction of TEQs accounted for by TCDD (when applicable) for each study.

2.4.1.1.3. *Feasibility of TCDD cancer dose-response modeling—summary discussion by cohort.*

2.4.1.1.3.1. *Using the NIOSH cohort in dose-response modeling.*

It is important to evaluate the NIOSH cohort in cancer dose-response modeling of TCDD. This cohort is the largest assembled to date, direct measures of TCDD based on sampling are available, and the lengthy follow-up interval allows for latent effects to be taken into account. Further, although this cohort consists mostly of male workers, these workers were occupationally exposed to TCDD daily, as compared to the acute accidental exposures of other occupational cohorts. Although the most recent analyses of a subset of the NIOSH cohort showed no association between serum TCDD levels and cancer mortality, the study authors did not examine latency effects (Collins et al., 2009). Incorporation of latency intervals is important in light of the stronger dose-response relationships that consistently have been observed with a 15−20 year latency interval in previous investigations of the NIOSH and other cohorts (Steenland et al., 2001).

Most published studies of the NIOSH cohort did not evaluate exposures to dioxin-like compounds. An exception is the analysis by Steenland et al. (2001). Although Steenland et al. (2001) did not incorporate individual-level data on dioxin-like compounds, based on their previous work (Piacitelli et al., 1992) they assumed that TEQ occupational exposures occurred as a result of TCDD alone in this population. TCDD exposures provided a better fit to the data than the TEQ-based metric, and 15-year latencies improved the fit for both metrics (relative to

unlagged exposures). The lifetime risk estimates for an increase in 10 TEQs (pg/kg of body weight/day/sex) ranged from 0.05−0.18%. The value added for this measure is the incorporation of the contribution of other dioxin-like compounds to the background rates.

Blue collar workers, such as those in the NIOSH cohort, typically have higher rates of smoking than the general population (Lee et al., 2007; Bang and Kim, 2001). This potential source of confounding would be expected to produce a higher SMR for lung cancer mortality, and could contribute to the excess noted in the cohort with longer lag intervals. This bias, however, likely is not large as no statistically significant excess of nonmalignant respiratory mortality was found in these workers. Any associated bias from smoking would be expected to be smaller for comparisons conducted within the cohort, as fellow workers would be expected to be more homogeneous with respect to their risk factor profile than with an external general population referent group. Stratified analyses using both internal and external comparison groups also did not identify important differences in associations with TCDD exposure between smoking and nonsmoking cancers. Thus, fatal cancer risk estimates reported for workers in the NIOSH cohort appear to provide a reasonable estimate of the carcinogenic potency of TCDD.

Although the Steenland et al. (2001) study did not directly account for the possible confounding effects of other occupational exposure, the authors did address this source of potential bias. No known occupational exposures to carcinogens occurred, with the exception of 4-aminobiphenyl, which occurred at one plant. Two deaths from mesothelioma also occurred in the cohort, so some exposure to asbestos might also have occurred in the cohort (Fingerhut et al., 1991). The statistical analyses suggested that the inability to control for other occupational exposures would not have unduly affected risk estimates generated from internal cohort comparisons. For instance, the removal of one plant at a time from the analysis did not materially change dose-response estimates generated from the Cox model (Cheng et al., 2006). Moreover, adding a variable to represent plant in the Cox regression had little impact on the risk estimates. Given that other occupational exposures varied by plant, a change in risk estimates would be expected if such exposures were strong confounders.

The Cheng et al. (2006) analysis provides important information about the impact of applying kinetic models to the data. The CADM TCDD kinetic model resulted in dramatic decreases in the TCDD cancer mortality risk estimates when compared to the one-stage compartmental model that had been applied. Although Cheng et al. (2006) suggested that the

CADM model provides a better fit to the data than the typically used simple one-compartmental model, statistical comparisons of model fit were not reported. Therefore, there is value in presenting the range in risk estimates across different models when characterizing dose-response relationships.

Finally, the half-life of TCDD is generally recognized to vary according to body fat percentage, data that were not available for the NIOSH workers. The inability to account for between-worker variability in body fat would introduce exposure measurement error. That body fat percentage would not be expected to correlate with cumulative exposure to TCDD exposure, however, would limit the potential for misclassification bias. The effect of any nondifferential exposure measurement error likely would serve to attenuate the risk estimates of the study.

2.4.1.1.3.2. *Using the BASF cohort in dose-response modeling.*

The availability of blood lipid data for TCDD allows for characterization of cumulative TCDD exposures in the BASF cohort. TCDD blood lipid data were collected for 90% of the surviving members of the cohort (138 of 154) and these serum measures were used to generate TCDD exposure estimates for all 254 cohort members. Therefore, the potential for misclassification from extrapolating these exposures to the entire cohort may not be as likely as for the NIOSH cohort where sera data were available for only a small fraction of workers. These data were, however, collected long after the accident (36 years) and had to be back-extrapolated to derive the initial exposures.

The data on this cohort included several risk factors such as cigarette smoking and body mass index. One advantage is that cumulative TCDD levels by body mass index can be estimates on an individual-level basis. As expected, the derived cumulative measures appear to compare well with severity scores of chloracne. The finding that more pronounced risks are found 15−20 years after first exposure are also consistent with findings from several other cohorts (Fingerhut et al., 1991; Manz et al., 1991; Bertazzi et al., 2001).

One key limitation of the BASF cohort is its relatively small sample size (n = 243), which limits the ability to evaluate dose-response relationships for site-specific cancers. Also, the quality of the ascertainment of cancer incidence cannot be readily evaluated as the geographic area of the cohort is not covered by a tumor registry. Ott and Zober (1996) state that nonfatal cancers could have been more likely to be missed in early years, which could partially contribute

to the larger standardized incidence ratio found for cancer with longer latencies. Commenting on risk differences derived from incident and decedent cancer outcomes is difficult. Among those comprising the cohort, the ascertainment of incident outcomes was recognized to be less complete in early years. Although the ascertainment of mortality outcomes was generally regarded to be good among the 243 workers, some workers who died or moved likely were missed when the cohort was constructed. These deaths would have been more likely to have occurred several years before the second component of the cohort was assembled.

The use of the SMR statistic for this study population is associated with important sources of uncertainties. Deaths were surely missed, particularly for the third component of the cohort that accounts for approximately 38% (94/247) of the entire cohort; this factor would serve to underestimate the overall SMR. As mentioned before, this component of the cohort was assembled through the recruitment of workers known to be alive in 1986. Despite this limitation, the characterization of exposure data and availability of other risk factor data at an individual level allow the development of quantitative dose-response analyses.

2.4.1.1.3.3. *Using the Hamburg cohort in dose-response modeling.*

The Hamburg cohort lacked data on cigarette smoking, and, therefore, effect estimates could not be adjusted for this covariate. Additional analyses that excluded lung cancers resulted in an even stronger dose-response relationship between all cancer mortality and TCDD. Serum levels of TCDD also were also not associated with smoking status in a subgroup of these workers (Flesch-Janys et al., 1995) suggesting that smoking is not likely a confounder of the association between all cancer mortality and TCDD.

An important limitation of the cohort is the reliance on blood and tissue measurements of 190 workers that likely represent a highly selective component of the cohort. This subset of workers was identified at the end of the observation period, and therefore, excludes workers who died or could not be traced. There are uncertainties in deriving department- and period-specific estimates for a period that extends over three decades using this number of workers. Additionally, the criteria applied to the reference population could have introduced some bias. Workers were included only in the reference group if they had been employed for at least 10 years in a gas supply industry. The criteria were much different for the workers who were exposed to TCDD (only 3 months of employment). As a result, the reference group likely would

be more susceptible to the healthy worker effect. Internal cohort comparisons, which should be void of such bias, however, generally produced results similar to those based on the external comparison population. Therefore, the Becher et al. (1998) study meets the criteria and additional epidemiological considerations for dose-response modeling.

2.4.1.1.3.4. ***Using the Seveso cohort in dose-response modeling.***

Unlike many of the occupational cohorts that were examined, data from the Seveso cohort are representative of a residential population whose primary exposure was from a single TCDD release. A notable exception is the BASF cohort where workers were exposed primarily through two accidents that occurred in the plant. The Seveso data, therefore, might permit cancer dose-response investigations in women and children.

Uncertainty in identifying the critical exposure window for most of the outcomes related to the Seveso cohort is a key limitation. An important feature of the Seveso cohort, however, is that TCDD levels were much lower among those in the highest exposure zones in Seveso (medians range from 56−136 ng/kg) (Eskenazi et al., 2004) than those in the occupational cohorts who had TCDD exposures that were sometimes more than 1,000 ng/kg. Given these dramatic differences in exposures, the standardized mortality ratios (after incorporating a 15−20 year latency period) for all cancer sites combined are remarkably similar between the Seveso and the occupational cohort analyses. Perhaps more importantly, the data from Seveso might be more relevant for extrapolating to lower levels, given that exposures to TCDD are two orders of magnitude higher than background levels (Smith and Lopipero, 2001).

The Warner et al. (2002) study found a positive association between serum levels of TCDD and breast cancer. As noted previously, ascertainment of incident cases for all cancers would allow for a dose-response relationship to be evaluated. Moreover, future breast cancer analyses in this cohort should strengthen the quantitative dose response analyses of this specific cancer site. The strengths of the Warner et al. (2002) study outlined earlier suggest that this study should be considered for cancer dose-response modeling.

Earlier Seveso studies likely are unsuitable for conducting quantitative risk assessment. These previous studies used an indirect measure of TCDD exposure, namely, zone of residence. Soil concentrations of TCDD varied widely in these three zones (Zone A: 15.5−580.4 ppt; Zone B: 1.7−4.3 ppt; and Zone R: 0.9−1.4 ppt), which could have resulted in considerable

exposure misclassification. The Warner et al. (2002) study greatly improved the characterization of TCDD exposure using serum measures, and also allowed for control of salient risk factors that may have resulted in bias due to confounding.

At this time it is unclear whether any study has examined the relationship between cancer and serum estimates of TCDD among Seveso males exposed from the 1976 accident.

2.4.1.1.3.5. *Using the Chapaevsk related data in dose-response modeling.*

Currently, individual-level exposure data are lacking for residents of this area and there is no established cohort for which cancer outcomes can be ascertained. These limitations, therefore, preclude the inclusion of Chapaevsk data in a quantitative dose-response analysis.

2.4.1.1.3.6. *Using the Ranch Hands cohort in dose-response modeling.*

An important limitation of the Ranch Hands cohort for TCDD and cancer dose-response modeling is an inability to isolate TCDD effects from the effects of other agents found in the associated herbicides. Exposure to other dioxin-like compounds was not estimated in this study and could confound the previously reported associations. As such, dose-response analyses on this population were not conducted.

2.4.1.1.4. *Discussion of general issues related to dose-response modeling*

2.4.1.1.4.1. *Ascertainment of exposures.*

Several series of epidemiological data have used serum measures to estimate TCDD levels. Serum data offer a distinct advantage in that they provide an objective means to characterize TCDD exposure at the individual level. The serum measures in the occupational cohorts, however, are limited in two important ways. First, these samples are generally collected from small subsets of the larger cohorts; therefore, using these measures to extrapolate to the remainder of the cohort could introduce bias due to exposure misclassification. The second limitation is related to estimating the half-life of TCDD. As noted previously, exposures to TCDD were back-extrapolated several decades from serum samples collected among surviving members of several cohorts. This approach was used in the NIOSH, Ranch Hands, BASF, New Zealand, and Hamburg cohorts. The reported half-life of TCDD among these populations was reported between 7.1 to 9.0 years and shown to vary with several individual characteristics

including age, body fat composition, and smoking. The derivation of half-lives from a sample of workers, and application of these estimates to retrospectively characterize exposure can introduce uncertainty into the lifetime exposure estimates. It is important to note, however, that sensitivity analyses results in several studies have been fairly consistent when evaluating the impact of half-life of TCDD (Steenland et al., 2001; Flesch-Janys et al., 1995).

A unique advantage of the Seveso study is that serum measures were taken shortly after the accident, and therefore characterization of TCDD exposure in this population does not depend on assumptions needed to back-extrapolate exposures several decades.

2.4.1.1.4.2. *Latency intervals.*

Many of the epidemiological studies indicate stronger associations between TCDD and cancer outcomes once a latency period has been considered. Generally, risks are higher when a lag period of 15−20 years is included. As noted previously, this observation is consistent with many other environmental carcinogens such as radon, radiation, and cigarette smoking. That recent exposures do not contribute to increased cancer risk provides some support that the initiation and promotion phases might occur many years before death making recent exposures irrelevant for these analyses. The ability to discriminate between models of varying latency, however, was limited in many studies. The application of biologically based modeling could provide additional important insights on which phase(s) of carcinogenesis TCDD exerts an influence. Such modeling, however, would necessitate having data on an individual-level basis. Ideally, this modeling would use cancer incident data rather than mortality outcomes, given that for many cancers, the median survival time exceeds 5 years.

2.4.1.1.4.3. *Use of the SMR metric.*

The occupational cohorts and the studies in Seveso and Chapaevsk have made inferences regarding the effects of TCDD on mortality using the SMR. When compared to the general population, the healthy worker effect may result in a downward bias in the SMR. This often can manifest as SMRs less than 1 for several causes of mortality. The effect of this bias is, however, generally lower for cancer outcomes. Cancer outcomes, whether incidence or death, typically occur later in life and do not generally affect an individual's ability to work at earlier ages.

There are several approaches that can be taken to minimize potential biases introduced by the healthy worker effect, which would account for workers being healthier than the general population. Comparisons of mortality (or cancer incidence) can be made to other cohorts of similar workers. If done properly, this can allow for some control of characteristics such as sociodemographic characteristics and smoking as the two populations can be matched by these factors. However, it may be the case that other working populations are exposed to other harmful exposures, thereby making it difficult to estimate risk associated with a specific agent (such as TCDD) in the cohort of interest. A second and preferred approach to control for the healthy worker effect, should it prove feasible, is to conduct comparisons of health outcomes in relation to exposure within the cohort. These comparisons are less likely to be influenced by other potential confounding variables such as smoking, socioeconomic status, and other occupational exposures that are generally more homogeneous within the cohort relative to external populations. Moreover, the mechanisms used to identify health outcomes and follow individuals over time are generally applied in the same manner to all cohort members. Taken together, where different comparisons have been made to generate risk estimates, those that have been conducted using internal cohort comparisons are preferable.

In addition to potential bias from the health worker effect, the comparison of SMRs between studies is not always straightforward and is not recommended by some (Rothman, 1986; Myers and Thompson, 1998). The SMR is the ratio of the observed number of deaths to the expected number of deaths and is often referred to as the method of indirect standardization. The expected number of deaths is estimated by multiplying the number of person-years tabulated across individuals in the cohort, stratified by age, by rates from a reference population that are available for the same strata. Therefore, each population cohort will have an estimated number of cases derived using a different underlying age structure. As outlined by Rothman (1986), the mortality rates might not be directly comparable to each other, although the impact of such bias will be much less if the age-distribution of the cohorts is similar. While it might be reasoned that the TCDD exposed workers would have similar age distributions this is in fact not the case (Becher et al., 1998; Ott et al., 1993; Thiess et al., 1982). This may be due to exposure occurring both chronically, as well as from acute exposures due to accidental releases that happened at various times at different plants. This is evident with the Hamburg and the BASF cohorts, as most individuals comprising the BASF cohort were employed at the time of the accident

(1953/1954), while most of the Hamburg cohort (852/1048) was employed after 1954; the follow-up of these cohorts ended at approximately the same time.

The method of direct standardization allows for a more meaningful comparison of mortality rates to be made between cohorts. With this approach, weights (usually based on age and sex) are drawn from a standard population and are, in turn, applied to disease rates for the same strata observed in the cohort of interest. A comparison of weighted rates between different cohorts would then be based on the same population standard.

Despite these limitations in comparing SMRs between studies, Armstrong (1995) argues that the comparisons are valid if the underlying stratum specific rates in each exposure grouping are in constant proportion to external rates. Comparisons of the SMRs between studies will be biased only if there is an interaction between age and TCDD (i.e., the RR of disease due to exposure differs by age). For cancer outcomes, the finding that associations become stronger after a period of latency is incorporated into the analyses suggests that this assumption does not hold true. That is, risk estimates would be lower among young workers. Similarly, for noncancer outcomes, some of the data from the Seveso cohort suggests differential effects according to the age at exposure.

The use of the SMR might also be biased in that workers exposed to TCDD could be subject to more intensive follow-up than the general population, and as a result, differential coding biases with cause of death might occur. Moreover, some cohorts (e.g., the BASF cohort) have been assembled, in part, by actively seeking out survivors exposed to accidental releases of dioxins. As such, they would not include persons who have died or who were lost to follow-up. This would result in underascertainment of deaths and SMRs developed from these data. The use of an internal cohort comparison offers distinct advantages to overcome potential sources of selection bias. Given these uncertainty about comparability across the different studies, conducting a meta-analysis of cancer outcomes for TCDD using the SMR statistic is not warranted for this analysis.

2.4.1.1.4.4. ***All cancers versus site-specific.***

An important consideration for quantitative dose-response modeling is the application of models for all cancers combined, or for site-specific cancers. Consistency is often lacking for site-specific cancers, which might be due in large part to the relatively small number of cases

identified for site-specific cancers in the cohorts. Although the risk estimates produced for all cancer sites have important limitations and uncertainties, the data are far more consistent in terms of the magnitude of an association and latency intervals. The IARC evaluation has put forth the possibility of a pleuripotential mode of action between TCDD and the occurrence of cancer. Despite the criticism of this assertion by some (Cole et al., 2003), the general consistency of an increased risk for all-cancer mortality across the occupational cohorts when latency intervals have been incorporated, provides adequate justification for dose-response quantification of all cancer sites combined.

2.4.1.1.4.5. *Summary of epidemiologic cancer study evaluations for dose-response modeling.*

All epidemiologic cancer studies summarized above were evaluated for suitability of quantitative dose-response assessment using the TCDD-specific considerations and study inclusion criteria. The results of this evaluation are summarized in a matrix style array (see Table 2-2) at the end of this section, and descriptively in Appendix B. Table 2-4 summarizes the key epidemiologic cancer studies suitable for further TCDD dose-response analyses.

2.4.1.2. *Noncancer*

In this section, the available epidemiological data that could be used in a dose-response analysis for noncancer endpoints are evaluated. Because many of the key studies also evaluated cancer outcomes, the noncancer studies are presented in the same order as presented in Section 2.4.1.1. Generally, the strengths and limitations of the cancer studies also apply to the noncancer outcomes. In this section, key features of these studies that have direct relevance to modeling of noncancer outcomes in particular are highlighted. To reduce redundancy, a detailed overview of many of these studies is not provided here. Instead, the reader should refer to Section 2.4.1.1.1.

2.4.1.2.1. *Noncancer cohorts.*

2.4.1.2.1.1. *The NIOSH cohort.*

2.4.1.2.1.1.1. *Steenland et al., 1999.*

2.4.1.2.1.1.1.1. *Study summary.*

The 1999 published report of NIOSH workers exposed to TCDD also conducted external cohort comparisons to the U.S. general population using SMRs for mortality outcomes other than cancer (Steenland et al., 1999). Analyses are based on 3,538 workers employed at 8 plants from 1942 to 1984. SMRs were based on a mortality follow-up that was extended until the end of 1993. Cox regression analyses were used to compare mortality risk in relation to TCDD exposure within the cohort.

2.4.1.2.1.1.1.2. *Study evaluation.*

Overall, no statistically significant differences in all-cause mortality (SMR = 1.03, 95% CI = 0.97−1.08) were observed. Mortality from ischemic heart disease (SMR = 1.09, 95% CI = 1.00−1.20) and accidents (SMR = 1.25, 95% CI = 1.03−1.50) was slightly elevated. The dose-response relationship for ischemic heart disease observed with the SMRs calculated across septile exposure categories, however, was not statistically significant ($p = 0.14$). Overall, excess risk was not evident for diabetes, cerebrovascular disease, or nonmalignant respiratory disease using the external population comparisons. Internal cohort comparisons using the Cox regression model were performed using 0 and 15-year lag intervals. A dose-response trend was observed for the derived ratios across the septiles for ischemic heart disease ($p = 0.05$) and diabetes ($p = 0.02$). For ischemic heart disease mortality, those in the upper two septiles had rate ratios of 1.57 (95% CI = 0.96−2.56) and 1.75 (95% CI = 1.07−2.87), respectively, relative to those in the lowest septile. In contrast, an inverse dose-response relationship was observed for diabetes mortality. The inverse association found for diabetes is inconsistent with the positive association reported in the Ranch Hands study (Michalek and Pavuk, 2008). However, previous reports have questioned the use of death certificates as the means to ascertain outcome as diabetes may be under-reported especially among descendents with diabetes who die from cancer (McEwen et al., 2006).

2.4.1.2.1.1.1.3. *Suitability of data for TCDD dose-response modeling.*

The inverse association with diabetes precludes dose-response analysis for this outcome. Although a dose-response pattern was observed for ischemic heart disease mortality, it was borderline statistically significant, and this association was not observed in other cohorts. Furthermore, fatal outcomes are not a suitable basis for development of an RfD. For these reasons, dose-response analysis for this outcome is precluded.

2.4.1.2.1.1.2. *Collins et al., 2009.*

2.4.1.2.1.1.2.1. *Study summary.*

Collins et al. (2009) recently described the mortality experience of Dow employees who worked in Midland, Michigan. This plant produced 2,4,5-trichlorophenol between 1942 and 1979, and 2,4,5-T between 1948 and 1982. The cohort consisted of 1,615 workers exposed to TCDD from as early as 1942; the follow-up of the cohort extended until 2003.

TCDD exposures were derived using serum samples obtained from 280 surviving individuals. A simple one-compartment, first-order pharmacokinetic model was used to estimate time-dependent TCDD measures. The area under the curve approach was then applied to estimate cumulative TCDD exposure above background. A half-life of 7.2 years for TCDD based on earlier work was incorporated into the exposure estimation (Flesch-Janys et al., 1996).

Collins et al. (2009) made an external comparison of the mortality rates of the cohort to the U.S. general population using the SMR statistic. Noncancer causes of death included all causes, diabetes, cerebrovascular disease, nonmalignant respiratory disease, cirrhosis of the liver, and accidents. Overall, no statistically significant difference in all-cause mortality of these workers was detected when compared to the general population (SMR = 0.9, 95% CI = 0.9−1.0). Except for cirrhosis of the liver (SMR = 0.4, 95% CI = 0.1−0.8), no differences were found for any of the noncancer causes of death relative to the general population.

Internal cohort analyses based on cumulative measures of TCDD were conducted for mortality from diabetes, ischemic heart disease, and nonmalignant respiratory disease using the Cox regression model. These models adjusted for possible confounders such as year of hire and birth year. No statistically significant association was found between continuous measure of TCDD and these causes of death.

2.4.1.2.1.1.2.2. *Study evaluation.*

Given that the external comparisons may result in bias from the healthy worker effect, results from the internal cohort comparisons using the Cox regression model are preferred. These analyses were performed for diabetes, ischemic heart disease, and nonmalignant respiratory disease. TCDD levels for these workers were estimated using a simple one-compartment pharmacokinetic model (Aylward et al., 2007). The hazard ratios generated from the Cox regression model were not statistically significant for any of the three noncancer outcomes modeled.

2.4.1.2.1.1.2.3. *Suitability of data for TCDD dose-response modeling.*

No association of an increased risk for an adverse effect was observed with any of the noncancer outcomes. Therefore, dose-response modeling based on this population was not conducted.

2.4.1.2.1.2. ***The BASF cohort.***

2.4.1.2.1.2.1. *Ott and Zober, 1996.*

2.4.1.2.1.2.1.1. *Study summary.*

In 1996, Ott and Zober published a report on the mortality experience of the cohort of 243 BASF male workers who were accidentally exposed to 2,3,7,8-TCDD in 1954 or in the clean up that followed. The mortality follow-up of this cohort extended until the end of 1992. External comparisons of mortality were made to the German population using the SMR statistic. Internal cohort comparisons were also made by estimating cumulative TCDD for the cohort using serum measures that were obtained from 138 workers. Ott et al. (1993) provided a detailed account of the methodology to estimate TCDD. Briefly, a cumulative measure of TCDD expressed in µg/kg was derived, by first estimating the half-life of TCDD using individuals who had repeated serum measures; the half-life was estimated to be 5.8 years. Individual-level data on body fat were used to account for the influence of body fat on decay rates. Half-life estimates of TCDD varied (range: 5.1−8.9 years) and were dependent on body fat composition (20% and 30%, respectively). This approach differed from previous analysis of this cohort that used a constant 7-year half-life (Ott et al., 1993). TCDD levels at the time of serum sampling were then estimated as the product of TCDD concentration in blood lipid and the total lipid weight for each

worker. Nonlinear models then were applied to estimate the contribution of duration of exposure to TCDD dose extrapolated to the time of exposure.

External comparisons to the German population using the SMR statistic also were examined across dose categories. The noncancer causes of death examined by Ott and Zober (1996) included all-cause mortality, diseases of the circulatory system, ischemic heart disease, diseases of the digestive system, external causes, suicide, and residual causes of death. Overall, no statistically significant differences in the SMR with the general population for all-causes of death (SMR = 0.9, 95% CI = 0.7−1.1) were found. No statistically significant differences were noted for any of the other causes of death examined.

Ott and Zober (1996) performed internal cohort comparisons using the Cox regression model. These analyses found no dose-response patterns when cause-specific mortality was examined across increasing cumulative TCDD exposure categories. Although an inverse association for diseases of the respiratory system (SMR = 0.1, 95% CI = 0.0−0.8) was detected, it was based only on 1 reported case. Many of these comparisons are limited by small sample sizes as 92 deaths occurred in the cohort, and of these, 31 were from cancer. Also, the third component of the cohort was identified primarily from former employees who were alive in 1986. As a result, the SMR based on the general population might be underestimated by the exclusion of deceased workers.

2.4.1.2.1.2.1.2. *Study evaluation.*

As noted previously, caution should be exercised in the interpretation of SMR values of noncancer outcomes as they could be influenced by the healthy worker effect. Although the mechanism of identifying vital status appears to be excellent and unbiased, SMRs might be underestimated for the cohort due to the manner in which they were constructed. Specifically, a large component of the cohort was assembled by actively seeking out former workers who were known to be alive in 1986.

2.4.1.2.1.2.1.3. *Suitability of data for TCDD dose-response modeling.*

No dose-response patterns were observed between TCDD and the noncancer outcomes in the Ott and Zober (1996) study. Therefore, dose-response modeling was not conducted.

2.4.1.2.1.3. *The Hamburg cohort.*

2.4.1.2.1.3.1. *Flesch-Janys et al., 1995.*

2.4.1.2.1.3.1.1. *Study summary.*

Flesch-Janys et al. (1995) reported on the mortality experience of a cohort of individuals employed by an herbicide-producing plant in Hamburg, Germany, covering the period 1952 to 1992. As described in more detail in Section 2.4.1.1.1.3, the authors developed a cumulative measure of TCDD using serum measures from 190 workers. This study also examined the relationship between total TEQ and mortality. In the study population, the mean TEQ without TCDD was 155 ng/kg, and for the mean TEQ including TCDD was 296.5 ng/kg.

Risks relative to the unexposed referent group of gas workers were estimated using Cox regression across six exposed TCDD groups (i.e., the first four quintiles, and the ninth and tenth deciles). A linear dose-response relationship was found with all causes of mortality and cardiovascular mortality ($p < 0.01$). The RR for all cardiovascular deaths in the upper exposure category was 1.96 (95% CI = 1.15−3.34), although there was no evidence of a linear dose-response trend ($p = 0.27$). The dose-response relationship was most marked for ischemic heart disease, with a RR of 2.48 (95% CI = 1.32−4.66) in the highest exposure group. A dose-response relationship was also observed across TEQ groupings for all cause mortality, cardiovascular disease mortality, and ischemic heart disease mortality. The authors did not perform joint modeling of TEQ (without TCDD) and TCDD, so determining the extent that dioxin-like compounds contributed to an increased risk of mortality is not possible.

2.4.1.2.1.3.1.2. *Study evaluation.*

The Flesch-Janys et al. (1995) study lacks information on other potential risk factors for cardiovascular disease, which could result in confounding if those risk factors are also related to TCDD exposure. Dose-response patterns were strong, however, and persisted across numerous TCDD (and TEQ) exposure categories based on the use of an external reference group (i.e., gas workers) or based on the internal comparison. The findings based on the internal comparison are noteworthy in that these groups should be more homogenous with respect to confounding factors. As noted previously, the poor correlation between TCDD and smoking among workers and similar smoking prevalence between the workers and the external gas company workers suggest that smoking was not likely a confounder of the TCDD and cardiovascular disease

relationship. No other evaluation of noncancer mortality outcomes has been undertaken in this cohort since 1995.

A strength of the Flesch-Janys et al. (1995) study was that it included the collection of blood serum measures, which provided an objective measure of TCDD exposure. Blood serum data, however, were obtained only for 16% of the cohort. The assumption of the first-order kinetic elimination model is critical, given that measures were taken at the end of follow-up. The model also assumed the half-life of TCDD was 6.9 years. If the kinetics are not first order, or if the half-life estimate is inaccurate, estimates of TCDD levels during exposure would be biased, particularly for workers having longer periods between exposure and PCDD and PCDF assays. Sensitivity analyses completed by the authors suggest that such bias is not likely to present because the results were unaffected when different model assumptions regarding kinetic and half-lives were examined. The lack of an impact on RR estimates with varying half-life estimates was similar to findings by Steenland et al. (2001).

2.4.1.2.1.3.1.3. *Suitability of data for TCDD dose-response modeling.*

Despite the aforementioned study strengths, the study focused on fatal outcomes such as all cause mortality, cardiovascular disease mortality, and ischemic heart disease mortality. As such, dose-response analysis was not conducted since these outcomes are not suitable for development of an RfD.

2.4.1.2.1.4. *The Seveso Women's Health Study (SWHS).*

Eskenazi et al. (2000) presented an overview of the SWHS. The SWHS is the first comprehensive epidemiologic study of the reproductive health of a female population exposed to TCDD. The primary objective of the SWHS is to investigate the relationship of TCDD and several reproductive endpoints, including endometriosis, menstrual cycle characteristics, birth outcomes, infertility, and age at menopause. A second phase of follow-up that focuses on osteoporosis, thyroid hormone, breast cancer, diabetes, and metabolic syndrome is expected to be completed in 2010.

Women were eligible for participation in the SWHS if they resided in Zones A and B (the most contaminated areas) at the time of the explosion, were 40 years of age or younger at the time of the explosion in 1976, and samples of their blood were collected and stored between

1976 and 1980. The enrollment of women in the SWHS began in March 1996 and continued until July 1998. Of the 1,271 eligible women, 17 could not be found, 21 had died, and 12 were too ill to participate. Of the 96% of the remaining women, 80% (n = 981) participated in the study. Participation in the SWHS included a blood draw and an interview by a trained nurse who was blind to subjects' TCDD level and zones of residence at the time of the accident. The interview included detailed information on potential confounders including occupational, medical, and reproductive, and pregnancy history. Also, women who were premenopausal were asked to undergo a vaginal ultrasound and pelvic exam and to complete a daily diary on menstruation.

Depending on the health outcome under study, TCDD exposures were characterized for the women at different times. For example, TCDD exposure levels were estimated at the time of the accident for some studies and at the time of conception for others. The SWHS study population has been used to investigate associations between maternal TCDD levels and the following health outcomes: menstrual cycle characteristics (Eskenazi et al., 2002a); endometriosis (Eskenazi et al., 2002b); birth outcomes (Eskenazi et al., 2003); age at menarche (Warner et al., 2004); age at menopause (Eskenazi et al., 2005); uterine leiomyomas (Eskenazi et al., 2007); and ovarian function (Warner et al., 2007). An evaluation of the studies in chronological order is presented in this section.

2.4.1.2.1.4.1. *Eskenazi et al., 2002a—Menstrual cycle characteristics.*

2.4.1.2.1.4.1.1. *Study summary.*

Eskenazi et al. (2002a) evaluated serum TCDD exposures in relation to several menstrual cycle characteristics in the SWHS. A total of 981 women who were 40 years of age or younger at the time of the accident comprised the SWHS. The following exclusion criteria was applied 44 years of age or older, women with surgical or natural menopause, those with Turner's syndrome, and those who in the past year had been pregnant, breastfed, or used an intrauterine device or oral contraceptives.

A trained interviewer collected data on menstrual cycle characteristics using a questionnaire. Women were asked to indicate how long their cycles were, whether the cycles were regular (e.g., irregular cycle defined as length varied by more than 4 days), how many days the menstrual flow lasted, and whether this flow was "scanty, moderate, or heavy." Information

was also collected on obstetric and gynecological conditions. TCDD exposures were derived from serum samples collected in 1976−1985. The authors selected the earliest available serum sample, and back-extrapolated to 1976 values using either the Filser model (Kreuzer et al., 1997) for women aged 16 years or younger in 1976 (n = 20) or the first-order kinetic model (n = 6) (Pirkle et al., 1989).

Serum TCDD levels were transformed using the log10 scale, and the relationships between these levels and length of menstrual cycle and days of menstrual flow were examined using linear regression. The authors applied logistic regression to characterize the risk between $\log_{10}$TCDD and heaviness of flow or regularity of cycle. In these analyses, moderate or heavy flow and regular cycle were used as the reference categories. Stratified analysis was performed by menarcheal status at the time of the accident.

Overall, the association with TCDD exposure (per 10-fold increase) and length of menstrual cycle was of marginal statistical significance for premenarcheal (β = 0.93, 95% CI = −0.01, 1.86) women, but not postmenarcheal women (β = −0.03, 95% CI = −0.61, 0.54). The corresponding estimates found for days of menstrual flow were β = 0.18 (95% CI = −0.15, 0.51) and β = 0.16 (95% CI = −0.18, 0.50), respectively. Reduced flow was not associated with TCDD when compared to moderate or heavy flow (odds ratio [OR] = 0.84, 95% CI = 0.44, 1.61); effect modification by menarcheal status, however, was evident (p = 0.03). Specifically, women exposed to TCDD who were premenarcheal had lower odds of reduced flow, while those exposed to TCDD who were postmenarcheal did not. These findings counter the hypothesis that TCDD exposure is related to ovarian dysfunction. Finally, statistically significant ORs were found between serum TCDD levels (per 10-fold increase) and having an irregular cycle (OR = 0.46, 95% CI = 0.23, 0.95). This inverse association was evident in both premenarcheal women (OR = 0.50, 95% CI = 0.18, 1.38) and postmenarcheal women (OR = 0.41, 95% CI = 0.15, 1.16).

2.4.1.2.1.4.1.2. *Study evaluation.*

Overall, the findings from the Eskenazi et al. (2002a) study suggest that exposures to TCDD can affect menstrual cycle characteristics among women who were exposed before menarche. Exposures to TCDD were well characterized using serum samples available on an individual-level basis, and the design allowed for the influence of other risk factor data to be

controlled for in regression analyses. Analysis of TCDD levels and the length of menstrual cycle in premenarcheal women produced associations that were of marginal statistical significance. However, it is unclear whether the endpoints that were measured constitute adverse health outcomes as they are not definitive markers of ovarian dysfunction. Another source of uncertainty is measurement error due to the subjective nature of menstrual flow reporting. Any resulting misclassification of the outcome should be nondifferential, as the measurement error is unlikely to be dependent on TCDD exposure.

2.4.1.2.1.4.1.3. *Suitability of data for TCDD dose-response modeling.*

The lack of a clear adverse health outcome related to TCDD exposure is a weakness of this study. Although it is difficult to define the critical window of exposure for quantitative exposure calculations, it can be estimated for the women that were premenarcheal at the time of the accident as 13 years. Therefore, this study is suitable for further consideration for quantitative dose-response modeling.

2.4.1.2.1.4.2. *Eskenazi et al., 2002b—Endometriosis.*

2.4.1.2.1.4.2.1. *Study summary.*

The SWHS provided the opportunity to investigate the association between serum TCDD levels and endometriosis (Eskenazi et al., 2002b). The rationale the authors provided for undertaking this study was the experimental animal studies that suggested an association, the high prevalence of endometriosis among infertile women where breast milk concentrations of dioxin are high, and the unknown etiology of endometriosis. The study consisted of 601 women who were younger than 30 years at the time of the Seveso accident. Stored sera that had been collected between 1976 and 1980 were also available for these women.

Given that laparoscopy could not be performed on women unless clinically indicated, no "gold" standard was available for endometriosis diagnosis. Based on the results of a validation study they conducted in a clinical population, the researchers classified women as having endometriosis based on symptom report, gynecologic exam results, and vaginal ultrasound.

TCDD was measured in sera in 1976 for 93% of the women. Values for women whose serum TCDD levels were collected after 1977 and had values exceeding 10 ppt were back-extrapolated to 1976 using either the Filser model (<16 years of age) (Kreuzer et al., 1997)

or a first-order kinetic model (≥16 years) (Pirkle et al., 1989). These estimates of TCDD were then modeled as both continuous (on a log scale) and categorical (≤20, 20.1−100, and >100 ppt) exposures.

Polytomous logistic regression was applied within the cohort used to generate RRs. In relation to women in the lowest exposure category, the RR for endometriosis among women in the middle and upper categories was 1.2 (95% CI = 0.3−4.5) and 2.1 (95% CI = 0.5−8.0), respectively. The trend tests were not statistically significant for either the categorical (p = 0.25) and continuous measures of TCDD (p = 0.84).

2.4.1.2.1.4.2.2. *Study evaluation.*

It is important to note that disease misclassification could have led to an underestimate of the true risk of endometriosis if this misclassification was not differential with respect to TCDD exposure. Also, younger women were likely to be under-represented as those who had never been sexually active could not be examined due to cultural reasons. Other dioxin-like compounds (PCDD, PCDFs, or polychlorinated biphenyls [PCBs]) were not considered because of small serum volumes, but any potential TEQ exposures occurring in the population were thought to be mostly attributable to TCDD in the exposed women.

2.4.1.2.1.4.2.3. *Suitability of data for TCDD dose-response modeling.*

Given that no statistically significant dose-response patterns were observed with either log-transformed or across TCDD exposure categories, and that the elevated risks among those with higher exposures had very wide confidence intervals (that included unity) quantitative dose-response analyses were not recommended for this outcome.

2.4.1.2.1.4.3. *Eskenazi et al., 2003—Adverse birth outcomes.*

2.4.1.2.1.4.3.1. *Study summary.*

Eskenazi et al. (2003) examined the relationship between serum TCDD levels and birth outcome measures. Analyses were based on 745 of the 981 women enrolled in the SWHS who reported having been pregnant (n = 1,822). Most of these pregnancies (888 pregnancies among 510 women) occurred after the accident. Analysis of spontaneous abortions was restricted to 769 pregnancies among 476 women that did not end in abortion or in ectopic or molar

pregnancy. Congenital anomalies were evaluated for the 672 pregnancies that did not end in spontaneous abortion. For the birth outcomes of fetal growth and gestational age, analysis was performed using 608 singleton births from women without hypertensive pregnancy disorders.

TCDD exposures were based on serum measures, most of which were taken shortly after the accident. Serum was collected in 1976−1977 for 413 women, between 1978 and 1981 for 12 women, and in 1996 for 19 women. TCDD exposures based on serum samples collected from 1977 onward were back-extrapolated to 1976.

Statistical analyses were performed on pregnancies that ended between 1976 and the time of interview. A continuous measure of $\log_{10}$TCDD (base 10 scale) was used to investigate associations with adverse birth outcomes. Logistic regression was used to characterize the relationship between TCDD exposure spontaneous abortions, small for gestational age, and preterm birth (<37 weeks gestation). Linear regression was used to describe the relationship between TCDD and birth weight (in grams) and gestational age (in weeks).

The risk estimates were adjusted for a series of characteristics that included sex of infant, history of low birthweight child, maternal height, maternal body mass index, maternal education, maternal smoking during pregnancy, and parity. No association was evident between TCDD serum levels and spontaneous abortion for pregnancies between 1976 and 1998 (OR = 0.8, 95% CI = 0.6−1.2), or those between 1976 and 1984 (OR = 1.0, 95% CI = 0.6−1.6). No statistically significant associations were found for birth weight or small gestational age, though the association with birth weight for pregnancies between 1976 and 1984 associated with a 10-fold increase in TCDD was fairly large and marginally statistically significant ($\beta = -92$, 95% CI = −204 to 0). No association was noted for preterm delivery in relation to $\log_{10}$TCCD levels.

2.4.1.2.1.4.3.2. *Study evaluation.*

This study was well-designed with well characterized exposures. Statistically significant associations were not evident, although the birth-weight findings should be pursued with further follow-up of the cohort. As the authors point out, those who were most vulnerable at the time of the accident (the youngest) had not yet completed their childbearing years. While the study lacked exposure data for the fathers, the authors indicated that only a small proportion were believed to have high exposures to TCDD. The key limitation of the study was a reliance on

self-reported measures of pregnancy history, which may lead to some misclassification of the birth outcomes. The observation that a large proportion of Seveso women had a voluntary abortion because of fears of possible birth defects due to exposures from the accident suggest an awareness bias is possible as a result of differential reporting of birth outcomes according to exposure status.

2.4.1.2.1.4.3.3. *Suitability of data for TCDD dose-response modeling.*

No statistically significant associations were found in the study; in addition, possible awareness bias could have influenced the self-reported measures of birth outcomes. Therefore, quantitative dose-response assessment was not considered for this study.

2.4.1.2.1.4.4. *Warner et al., 2004—Age at menarche.*

2.4.1.2.1.4.4.1. *Study summary.*

Warner et al. (2004) examined the relationship between TCDD and age at menarche in the SWHS cohort. As described earlier in this report, the SWHS comprised 981 participants. This study was restricted only to those who were premenarcheal at the time of the accident (n = 282). The proportional hazards model was used to model TCDD exposures and age at menarche. Age at menarche was determined by questionnaire administered by a trained interviewer. Covariates examined as potential confounders included height, weight, body mass index, athletic training at the time of interview, smoking, and alcohol consumption.

TCDD exposures were determined using serum samples collected from 257 of these women between 1976 and 1977. For the remaining women, TCDD levels were quantified from measures collected between 1978 and 1981 (n = 23) and in 1996 (n = 2). TCDD levels were back-extrapolated to the time of the explosion in 1976. TCDD was modeled as both a continuous variable ($\log_{10}$TCDD) and a categorical variable based on quartile values (≤55.9, 56−140.2, 140.3–300, >300 ppt). The lowest group was further subdivided into those with levels ≤20, and >20 ppt; this cut-point represented background levels found in a sample of women living in an unexposed area.

No association was found between the continuous measure of TCDD and age at menarche (hazard ration [HR] = 0.95, 95% CI = 0.83−1.09). Analyses restricted to those who were younger than 8 in 1976 produced similar results (HR = 1.08, 95% CI = 0.89−1.30).

Additionally, no dose-response trend was observed with categorical measures of TCDD among all women, as well as those under the age of 8. A marginally statistically significant association with earlier menarche was found when analyses were limited to 84 women under the age of 5 at the time of the accident (HR = 1.20, 95% CI = 0.98−1.60).

2.4.1.2.1.4.4.2. *Study evaluation.*

An important strength of the Warner et al. (2004) study is the ability to characterize TCDD exposures using serum samples that were collected shortly after the accident occurred. The outcome of interest, age at menarche, was determined by asking women "At what age did you get your first menstrual period?" Recent work suggests that self-reported measures of age at menarche decades later have modest agreement with responses provided during adolescence with recall varying by education and by history of an adverse birth outcome (Cooper et al., 2005). In the Seveso study, bias would be introduced if recall varied according to exposure levels.

2.4.1.2.1.4.4.3. *Suitability of data for TCDD dose-response modeling.*

Although the TCDD exposure characterization of study subjects was based on serum data, and no major biases were introduced from the study design, the analyses produced largely null associations. Therefore, quantitative dose-response assessment was not considered for this study.

2.4.1.2.1.4.5. *Eskenazi et al., 2005—Age at menopause.*

2.4.1.2.1.4.5.1. *Study summary.*

Eskenazi et al. (2005) evaluated the relationship between age at onset of menopause and serum levels of TCDD among women in the SWHS. Of the 981 women who agreed to participate in SWHS, this analysis was restricted to those who had not reached natural menopause before the time of the accident and who were at least 35 years of age at the time of the interview. The recruitment and interview of women occurred approximately 20 to 22 years after the accident (March 1996−July 1998).

The population was divided into quintiles of serum TCDD levels for the categorical analysis. For most women (n = 564), TCDD levels were estimated from samples provided in 1976−1977. For the remaining women included in these analyses, TCDD levels were estimated

from samples collected between 1978 and 1982 (n = 28) and between 1996 and 1997 (n = 24). As noted previously, exposure levels for women with post-1977 detectable levels of TCDD were back-extrapolated to 1976 using either the first-order kinetic model (Pirkle et al., 1989) (>16 years at time of accident) or the Filser model (<16 years at time of accident) (Kreuzer et al., 1997). Women were classified as premenopausal if they were still menstruating or if they had amenorrhea as a result of pregnancy or lactation (at the time of interview) with an indication of subsequent menstruation based on maintained diaries or further examination. Subjects for which amenorrhea had persisted for at least 1 year with no apparent medical explanation were classified into a natural menopause category. The category, surgical menopause, pertained to women with a medically confirmed hysterectomy or an oophorectomy. Finally, impending menopause was defined for subjects in which menstruation had been absent for 2 months, but who provided evidence of subsequent menstruation, or had a secretory endometrial lining, or indicated less predictable cycles in the previous 2−5 years. If participants' menopausal status could not be determined, they were grouped into the "other" category. This category included those for whom status could not be determined due to current use of oral contraceptives, hormone replacement therapy, or previous cancer chemotherapy.

Statistical analysis was based on both a continuous measure of log-transformed TCDD exposures and categories based on quintiles (<20.4 ppt; 20.4−34.2 ppt; 34.3−54.1 ppt; 54.2−118.0 ppt; >118.0 ppt). The Cox model was used to generate hazard ratios as estimates of relative risks and their 95% confidence intervals examining natural menopause as the outcome. Several covariates previously identified as associated with menopausal status in the literature were considered as potential confounders. These covariates included body mass index, physical activity, premenopausal smoking, education, marital status, history of heart disease and other medical conditions, and other reproductive characteristics.

The RRs were found to increase across the second through fourth quintiles (RRs = 1.1, 1.4, and 1.6, respectively) of serum TCDD categories in relation to those in the lowest category, but not in the upper quintile (RR = 1.0, 95% CI = 0.6−1.8). A statistically significant test of trend was detected across the first four quartiles (p = 0.04) but not across all five quintiles (p = 0.44). A statistically significant association with onset of menopause was not detected (RR = 1.02, 95% CI = 0.8−1.3) based on the logTCDD continuous measure.

2.4.1.2.1.4.5.2. *Study evaluation.*

The categorical exposure results from this study support a non-monotonic dose-related-association for earlier menopause with increased serum TCDD levels up to approximately 100-ppt TCDD serum, but not above. Eskenazi et al. (2005) speculated that the inverse "U" shape of the dose-response relationship is explained by the mimicking of hormones at lower doses of a chemical, while at higher levels the toxic effect of a chemical does not have the capacity to either inhibit or stimulate hormonal effects.

A study limitation is the potential for residual confounding due to adjustment based on current smoking status and not at the time of onset of menopause. It is unclear to what extent smoking status may differ between these two time periods and whether smoking is related to TCDD exposures in this cohort. Exposures to other dioxin-like compounds were not considered in this study because of small serum volumes, but any potential TEQ exposures occurring in the exposed population were thought to be mostly attributable to TCDD in the exposed women.

2.4.1.2.1.4.5.3. *Suitability of data for TCDD dose-response modeling.*

To date, this study is the only one that has examined the relationship between TCDD levels and onset of menopause. Although the findings suggest the possibility of a nonlinear dose-response function, the $\log_{10}$TCDD exposure metric was not statistically significant, nor were any category-specific hazard ratios statistically significant relative to the lowest category. Therefore, a quantitative dose-response analysis was not undertaken.

2.4.1.2.1.4.6. *Warner et al., 2007—Ovarian function.*

2.4.1.2.1.4.6.1. *Study summary.*

Warner et al. (2007) investigated the association between serum TCDD levels and ovarian function in subjects in the SWHS who were younger than 40 in 1976 and for whom sera collected after the accident had been stored. These women were recruited from March 1996 until July 1998. Ovarian function analysis was limited to 363 women between 20 and 40 years of age and who were not using oral contraceptives. Of these, 310 underwent transvaginal ultrasound and were included in the functional ovarian cyst analysis. Ninety-six women were in the preovulatory stage of their menstrual cycles and were included in the follicle analysis. For the hormone analysis, 126 women who were in the last 2 weeks of their cycle were included.

The authors used logistic regression to examine the relationship between TCDD and the prevalence of ovarian follicles greater than 10 mm. Linear regression models examined the continuous outcome variables: number of ovarian follicles >10 mm and diameter of dominant ovarian follicle. Covariates considered for inclusion in the model were age at ultrasound, age at accident, age at menarche, marital status, parity, gravidity, lactation history, current body mass index, age at last birth, and smoking history. For the serum hormone analyses, estradiol and progesterone were measured in blood at the time of interview. Ovulation status was defined as a dichotomous variable (yes/no) based on a serum progesterone cut-point value of 3 ng/mL.

The adjusted ORs across categories of TCDD exhibited no dose-response trend for the presence of follicles in relation to TCDD in the follicular phase; also, no statistically significant differences were noted in any of the upper exposure categories relative to those in the lowest. The adjusted OR for the continuous measure of $\log_{10}$TCDD was 0.99 (95% CI = 0.4−2.2). A similar nonstatistically significant finding was found for $\log_{10}$TCDD in relation to ovulation in both the luteal (OR = 0.99, 95% CI = 0.5−1.9) and mid-luteal phases (OR = 1.03, 95% CI = 0.4−2.7). Analyses of progesterone and estradiol also were not related to serum TCDD levels for either the luteal or mid-luteal phases (p = 0.51 and p = 0.47).

2.4.1.2.1.4.6.2. *Study evaluation.*

The investigators found no relationship between serum TCDD levels and serum progesterone and estradiol levels among women who were in the luteal phase at the time of blood draw. No association with number of ovarian follicles detected from ultrasound. Although no association was found, the authors suggested that the lack of significant results could be because the women in SWHS were all exposed postnatally and the relevant and critical time period for an effect might be in utero (animal studies support relevance of in utero exposures).

2.4.1.2.1.4.6.3. *Suitability of data for TCDD dose-response modeling.*

One limitation of the study was the lack of examination of confounding by dioxin-like compounds. The absence of associations between TCDD and adverse health effects in this study precludes conducting quantitative dose-response analyses.

2.4.1.2.1.4.7. *Eskenazi et al., 2007—Uterine leiomyoma.*

2.4.1.2.1.4.7.1. *Study summary.*

Associations between TCDD exposures and uterine leiomyoma (i.e., fibroids) were also examined among 956 women in the SWHS (Eskenazi et al., 2007). The sample population was based on the on the original 981 SWHS participants excluding 25 women diagnosed with fibroids before the date of the accident (July 10, 1976). Women who previously had fibroids were identified both through the administered questionnaire and the review of medical records. Transvaginal ultrasounds were performed for 634 women to determine if they had fibroids at the time of follow-up. Similar to other SWHS studies, exposure to TCDD was estimated using serum collected from women shortly after the time of the accident, between 1978 and 1981 and in 1996. TCDD levels were back-extrapolated to 1976 levels.

The study authors performed statistical analyses using two definitions of fibroids as outcome measures. The first was fibroids detected before the study, and the second was fibroids detected via ultrasound. A proportional odds method Dunson and Baird (2001) developed was used to model the cumulative odds of onset of fibroids. This method combines historical and current information of diagnoses of fibroids. Continuous and categorical measures of TCDD were modeled. Regression models were adjusted for known or suspected risk factors of fibroids including parity, family history of fibroids, age at menarche, body mass index, smoking, alcohol use, and education.

2.4.1.2.1.4.7.2. *Study evaluation.*

Categorical measures of TCDD suggested an inverse dose-response relationship with the onset of fibroids. Relative to those with TCDD levels less than 20 ppt, those having TCDD exposures between 20.1 and 75.0 ppt and greater than 75.0 ppt had RRs of 0.58 (95% CI = 0.41−0.81), and 0.62 (95% CI = 0.44−0.89), respectively. The continuous measure of $\log_{10}$TCDD produced a hazard ratio of 0.83 (95% CI = 0.65−1.07).

2.4.1.2.1.4.7.3. *Suitability of data for TCDD dose-response modeling.*

The inverse association between TCDD and uterine fibroids supports the possibility of an anti-estrogenic effect of TCDD. The observed direction of the reported associations precludes quantitative dose-response modeling.

2.4.1.2.1.5. *Other Seveso noncancer studies.*

2.4.1.2.1.5.1. *Mocarelli et al., 2008—Semen quality.*

2.4.1.2.1.5.1.1. *Study summary.*

Mocarelli et al. (2008) examined the relationship between TCDD and endocrine disruption and semen quality in a cohort of Seveso men. A total of 397 subjects of the eligible 417 males (<26 years old in 1976) from Zone A and nearby contaminated areas were invited to participate. Frozen serum samples were used to derive TCCD exposures. Also, 372 healthy blood donors not living in the TCCD-contaminated area were invited to participate. The researchers collected a health questionnaire and semen samples from participants. Analyses were based on 257 individuals in the exposed group and 372 in the comparison group.

Semen samples were collected postmasturbatory at home. Ejaculate volume, sperm motility, and sperm concentration were measured on these samples. Fasting blood samples also were collected from the subjects for reproductive hormone analyses, including 17β-estradiol (E_2), follicle stimulating hormone (FSH), inhibin B, luteinizing hormone (LH), and testosterone.

The researchers estimated serum concentrations of TCDD from samples provided in 1976−1977, and also in 1997−1998 for individuals whose earlier samples had TCDD values that exceeded 15 ppt. Serum concentrations for the comparison group were assumed to be less than 15 ppt in 1976 and 1977 and <6 ppt in 1998/2002 on the basis of serum results for residents in uncontaminated areas. The exposed and comparison groups were divided into three groups based on their age in 1976: 1−9, 10−17, and 18−26 years. Mocarelli et al. (2008) applied a general linear model to the sperm and hormone data and included exposure status, age, smoking status, body mass index, and occupational exposures as covariates.

Men exposed between the ages of 1 and 9 had reduced semen quality 22 years later. Reduced sperm quality included decreases in sperm count ($p = 0.025$), progressive sperm motility ($p = 0.001$), and total number of motile sperm ($p = 0.01$) relative to the comparison group. The opposite pattern was observed for several indices of semen quality among those aged 10−17 at the time of the accident. For the hormone analyses, those in the exposed group had lower serum E_2 levels, and higher follicle stimulating hormone concentrations. Neither testosterone levels nor inhibin B concentrations were associated with TCDD exposure.

2.4.1.2.1.5.1.2. *Study evaluation.*

The findings of the Mocarelli et al. (2008) study support the hypothesis that exposure to TCDD in infancy/prepuberty reduces sperm quality and could contribute to reported decrease in sperm quality in young men in the industrialized world. Although most semen analysis studies have low compliance rates in general population samples (20−40%) (Jorgenson et al., 2001; Muller et al., 2004), the compliance rate in this study was much higher (60%). Given that the compliance rates were similar between the exposed and comparison groups and the strong differences detected across the two age groups, selection bias appears unlikely in this study.

2.4.1.2.1.5.1.3. *Suitability of data for TCDD dose-response modeling.*

Health outcomes are well defined in the Mocarelli et al. (2008) study, and exposures are well characterized using serum data. Because the men exposed to elevated TCDD levels between the ages of 1 and 9 had reduced semen quality 22 years later, it is difficult to identify the relevant time interval over which TCDD dose should be considered. Specifically, it is difficult to discern whether this effect is a consequence of the initial high exposure between 1 and 9 years of age or a function of the cumulative exposure for this entire exposure window beginning at the early age. However, the differences between these two dose estimates (the initial high exposure versus the cumulative exposure for the 9 year window) are minimal (i.e. within an order of magnitude). Despite the uncertainty in estimating the critical window of exposure, dose-response analysis for this outcome was conducted.

2.4.1.2.1.5.2. *Mocarelli et al., 1996, 2000—Sex ratio.*

2.4.1.2.1.5.2.1. *Study summary.*

A letter to the editor was the first report of a possible change in the sex ratio from dioxin among Seveso residents following the July 10, 1976 accident (Mocarelli et al., 1996). The authors reported that 65% (n = 48) of the 74 total births that had occurred from April 1977 to December 1984 were females. This male to female ratio of 26:48 (35%) is significantly different from the worldwide birth ratio of 106 males to 100 females (51%) (James, 1995). Between 1985 and 1994, the Seveso male to female ratio leveled out at 60:64 (48%). The authors suggested that the finding supported the hypothesis that dioxin might alter the sex ratio through several possible mechanistic pathways.

Mocarelli et al. (2000) later reported on an investigation between serum-based TCDD measures in parents and the sex ratio of offspring. In this study, serum samples were collected from mothers and fathers who lived in the areas at the time of the explosion, were between the ages of 3 and 45 at the time of the explosion, and produced offspring between April 1, 1977 and December 31, 1996. The study population included 452 families and 674 offspring, and serum measures were available for 296 mothers and 239 fathers. An estimate of TCDD at the time of conception was also examined in relation to male to female birth ratios. TCDD exposure estimates between the years of 1976 and 1996 were estimated using Filser's model (Kreuzer et al., 1997).

Mocarelli et al. (2000) used chi-square test statistics to compare observed sex ratio to an expected value of 0.51 in this Seveso population. Concentrations of TCDD were modeled as categorical variables in several ways. First, a dichotomous variable was used whereby unexposed parents were defined as those who lived outside Zones A, B, and R or had a serum TCDD concentration of less than 15 ppt; parents with exposures of 15 ppt or higher were considered exposed. Second, a trichotomous exposure variable was created that consisted of parents who (1) lived outside Zones A, B, and R or had serum concentrations of less than 15 ppt, (2) had serum concentrations of 15−80 ppt, and (3) had serum concentrations that exceeded 80 ppt. These cut-points were chosen as they represented tertiles based on the distribution of TCDD among parents. Analyses were conducted separately for paternal and maternal TCDD levels.

The overall proportion of 0.49 male births (based on male to female ratio of 328:346) was not significantly different from the expected proportion of 0.51 ($p > 0.05$). Statistically significant differences were found, however, if both parents had TCDD levels >15 ppt (sex ratio = 0.44) or just the father had serum TCDD levels >15 ppt (sex ratio = 0.44). No statistically significant differences were found when the fathers had TCDD levels less than 15 ppt, irrespective of the maternal levels. A dose-response pattern in the sex ratio was found across the paternal exposure categories. That is, the sex ratio decreased with increased paternal TCDD levels (linear test for trend, $p = 0.008$). In the unexposed group, the sex ratio (male to female) was 0.56 (95% CI = 0.49−0.61), while in the highest exposure group (281.0−26,400.0 ppt) the corresponding sex ratio was 0.38 (95% CI = 0.28−0.49).

Stratified analyses by age at paternal exposure revealed that the sex ratio was altered to a greater degree among fathers who were younger than 19 at the time of the explosion. The male to female ratio among the unexposed fathers was 0.56 (95% CI = 0.50−0.62), while it was 0.38 (95% CI = 0.30−0.47) for those younger than 19 when exposed and 0.47 (95% CI = 0.41−0.53) for those exposed after 19. Regardless of the age at the time of exposure, however, fathers who were exposed had a statistically significantly different birth ratio (they were more likely to father girls) than those who were unexposed ($p < 0.05$).

Separate analysis of birth ratios based on paternal TCDD exposure estimated at the time of conception did not show the same dose-response pattern but did show strong evidence of consistently decreased male births relative to females. More specifically, the male to female birth ratios among the four successive quartiles (first through fourth) were 0.41, 0.33, 0.33, and 0.46.

2.4.1.2.1.5.2.2. *Study evaluation.*

Mocarelli et al. (2000) based the characterization of TCDD exposure on serum samples, which is an objective method for characterizing dose. Unlike for the occupational cohorts, serum measures for this study were taken close to the time of the accident, and therefore, back-extrapolation of TCDD exposures is unnecessary. Exposure received before the age of 19 at the time of the explosion were more strongly associated with a reduced male to female ratio than those received after the age of 19. The cut off age of 19 seems to be somewhat arbitrary, resulting in a highly uncertain critical exposure window. TCDD levels at the time of conception did not demonstrate a dose-response relationship, but paternal exposures resulted in consistently reduced male to female birth ratios (range: 0.33−0.46).

The study findings are unlikely to be influenced by age at conception as these values were found, on average, to be similar across calendar years. This suggests that age at conception was not an important confounder and that the birth ratio findings may be related to paternal exposures.

The methods used to identify births appear to be appropriate. Even if some under-ascertainment of births occurred, there is no reason to believe that ascertainment would be related to TCDD exposure and the sex of the baby. Therefore, no bias is suspected due to incomplete birth ascertainment.

2.4.1.2.1.5.2.3. *Suitability of data for TCDD dose-response modeling.*

TCDD exposures were well-characterized, and internal cohort analyses demonstrate association between paternal TCDD levels at the time of the accident and birth ratio. However, the change in sex ratio was only statistically significant when exposure occurred before 19 years of age. It is impossible to identify the relevant time interval over which TCDD dose should be considered for dose-response analysis; specifically, it is difficult to discern whether the different sex ratio is a consequence of the initial peak exposure before 19 years of age or a function of the average cumulative exposure over this entire exposure window. Assuming the initial high exposure is the correct exposure window, using the initial exposures in a dose-response model would yield LOAELs that are too high to be relevant to factor into the RfD calculation. The differences between the two dose estimates are quite large. Dose-response analysis for this outcome, therefore, was not conducted.

2.4.1.2.1.5.3. *Baccarelli et al., 2008—Neonatal thyroid function.*

2.4.1.2.1.5.3.1. *Study summary.*

Baccarelli et al. (2008) investigated the relationship between thyroid function and TCDD among offspring of women of reproductive age who were exposed in the 1976 accident. This health endpoint is relevant because thyroid function is important for energy metabolism and nutrients and for stimulating growth and development of tissues. Neonatal thyroid function at birth is evaluated through blood thyroid-stimulating hormone (b-TSH). Apart from iodine deficiency, no other environmental exposure has been associated consistently with reduced neonatal thyroid functioning.

The study population was drawn from 1,772 women who were identified as having lived in the highly contaminated areas (Zones A or B) at the time of the accident or between July 10, 1976 and December 31, 1947; were of fertile age (born after 1947); and were alive as of January 1, 1994. A random sample of 1,772 unexposed women who lived in the reference area was selected using frequency matching by year of birth to the exposed women, and residency in the reference area at the time of the accident. The reference area represents the noncontaminated areas that surround the three zones of decreasing exposure (Zones A, B and R). In total, 55,576 women had lived in the reference area. Population registry offices (n = 472) were contacted to detect children born to these women. Records could be traced for virtually all

subjects (1761/1772 exposed; 1762/1772 unexposed). Children born outside the Lombardy area were excluded as b-TSH could not be obtained for them. This accounted for 156 of the 1,170 children identified. The analyses were based on the remaining 56, 425, and 533 singletons born between January 1, 1994, and June 30, 2005 in Zone A, B, and from the reference area, respectively.

Thyroid function is tested in all newborns by b-TSH measures in the region of Lombardy where Seveso is located. These measures are obtained from blood samples taken 72 hours after birth using a standardized protocol. The b-TSH levels were log transformed to approximate a normal distribution. Linear regression analysis was used to conduct test for trends in mean b-TSH levels across different covariates. Logistic regression was used to assess associations between elevated b-TSH levels defined by the cutpoint of 5 μU/mL and residence in particular zones of contamination. Generalized estimating equations were used to adjust the standard errors of the ORs for correlation between siblings.

The mean levels of b-TSH were positively associated with average soil TCDD concentrations in the three areas (Zone A: 1.66 μU/mL; Zone B: 1.35 μU/mL; and Zone R: 0.98 μU/mL) ($p < 0.001$). Plasma TCDD levels also were shown to be much higher in a group of 51 newborns that had b-TSH levels >5 μU/mL. Compared to the reference population, adjusted ORs were elevated for Zone B (OR = 1.90, 95% CI = 0.94−3.86) and Zone A (OR = 6.63, 95% CI = 2.36−18.6). These ORs were adjusted for gender, birth weight, birth order, maternal age at delivery, hospital, and type of delivery. The adjusted ORs however differed only slightly from those that were unadjusted (Zone B, OR = 1.79, 95% CI = 0.92−3.50; Zone A OR = 6.60, 95% CI = 2.45−17.8). Of the risk factors considered, both gender and birth weights were associated with neonatal b-TSH.

The paper also included an analysis of children born to 109 women who were part of the Seveso Chloracne Study (Baccarelli et al., 2005). A total of 51 children were born to 38 of these women, of these 12 lived in Zone A, 10 in Zone B, 20 in Zone R, and 9 from the reference population. Several congeners including TCDD were measured in maternal plasma. TCDD levels were extrapolated to the date of delivery using a first-order pharmacokinetic model (Michalek et al., 1996). The elimination rate used was 9.8 years based on the mean half-life estimate from a previous study of women in the Seveso region (Michalek et al., 2002). TEQs were calculated for a mixture of dioxin-like compounds by multiplying the concentration of each

congener by its toxicity equivalence factor. The maternal average TEQ was 44.8 ppt (range: 11.6−330.4) among 51 mothers. The measurement of noncoplanar PCBs occurred only later in the study (1996) and, therefore, total mean TEQs (i.e., including the sum of PCDDs, PCDFs, coplanar PCBs, and noncoplanar PCBs) are available only on a subset (n = 37) of the population. Dioxin-like congeners were examined in this study as several studies suggest associations between the sum of PCBs, or individual congeners having decreased thyroxine (T4; Longnecker et al., 2000; Sandau et al., 2002), and increased TSH (Alvarez-Pedrerol et al., 2008; Chevrier et al., 2007).

The authors used a linear model to examine the association between maternal TCDD levels and b-TSH. The standardized regression coefficient obtained from this model was 0.47 ($p < 0.001$). For the evaluation of TEQs, a similar association was noted for PCDDs, PCDFs, and coplanar PCBs (n = 51, $\beta = 0.45$, $p = 0.005$) but not with non-coplanar PCBs (n = 37, $\beta = 0.16$, $p = 0.45$). Multivariate regression models that were adjusted for several covariates (i.e., gender, birth weight, birth order, maternal age at delivery, hospital, and type of delivery) found statistically significant associations with plasma TCDD, PCDDs, PCDFs, and coplanar PCBs, but not with non-coplanar PCBs. The sum of all total TEQs from the measured compounds was not statistically significant (n = 37, $\beta = 0.31$, $p = 0.14$).

2.4.1.2.1.5.3.2. *Study evaluation.*

The Baccarelli et al. (2008) study satisfies the epidemiological considerations and criteria for determining whether dose-response modeling should be pursued. The outcome is well defined, and a dose-response pattern was observed. The study also contained a substudy that characterized TCDD and exposures to other dioxin-like congeners and used serum measures for a sample of mothers. Results were consistent among the zone of residence analysis and the substudy based on serum measures.

2.4.1.2.1.5.3.3. *Suitability of data for TCDD dose-response modeling.*

Given the potential for exposure misclassification due to variability in TCDD soil levels within each zone, modeling should rely on individual-level TCDD exposures derived from the serum sampling substudy. The study data provide an opportunity for quantitative dose-response

analyses as the critical exposure window of 9 months can be used for exposure assessment purposes.

2.4.1.2.1.5.4. *Alaluusua et al., 2004—Oral hygiene.*

2.4.1.2.1.5.4.1. *Study summary.*

Alaluusua et al. (2004) examined the relationship between TCDD and dental defects, dental caries, and periodontal disease among Seveso residents who were children at the time of the accident. Subjects were randomly selected from those individuals who had previously provided serum samples in 1976, which was shortly after the accident. A total of 65 subjects who were less than 9.5 years of age at the time of the accident, and who lived in Zones A, B, or R were invited to participate. Recruitment was initiated 25 years after the time of the Seveso accident. An additional 130 subjects from the surrounding area (outside Zones A, B, or R or "non-ABR zone") having the same age restriction were recruited. Subjects were frequency matched for age, sex, and education. Questionnaires were administered to these individuals to collect detailed information on dental and medical histories, education, and smoking behaviors. Ten subjects who had completed at least high school were randomly excluded from the non-ABR zone to create groups with similar educational profiles. Participation rates for the ABR and non-ABR zones were 74% and 58%, respectively.

One dentist who was blind to the patients' TCDD exposure levels assessed dental aberrations. Dental caries was assessed using recommendations of the World Health Organization. Periodontal status was described following a detailed evaluation of the surfaces of the teeth. A radiographic examination was done to identify missing teeth, alveolar bone loss, deformities in the roots, and jaw cysts.

Comparisons of the presence of dental enamel defects according to exposure status were performed using logistic regression. Chi-square test statistics were applied to compare the distributions in the prevalence of dental defects across several categorical covariates (i.e., education, age, and serum TCDD level). For those who were younger than 5 at the time of the accident, dental defects were more prevalent among patients in zone ABR (42%) than those in the non-ABR zone (26%) ($p = 0.14$). Zone ABR is characterized by higher levels of soil TCDD levels relative to non-ABR. Serum levels permitted an improved characterization of risk as they were available at an individual level, rather than using a zone of residence. Defect prevalence

was highest among those in the upper serum TCDD category (700−26,000 ng/kg) with 60% of subjects having dental defects. The continuous measure of serum TCDD was associated with developmental defects ($p = 0.007$) and hypodontia ($p = 0.05$).

2.4.1.2.1.5.4.2. *Study evaluation.*

Although the subjects with serum measures were selected randomly, no direct measures of TCDD were made in subjects from the unexposed area (i.e., non-ABR zones). That those who resided in the non-ABR areas had lower TCDD exposures would be a reasonable assumption. Alaluusua et al. (2004), however, provide few details about the sampling frame used to identify these participants. Despite this fact, it is important to note that a dose-response pattern was observed between TCDD exposure and presence of developmental defects in the ABR population alone ($p = 0.016$). This finding is based on 27 subjects with developmental defects. This positive association provides support for a quantitative dose-response modeling of dental aberrations. The numbers of such subjects are small, however, with one, five, and nine subjects having defects in the exposure groups of 31−226, 238−592, and 700−26,000 ng/kg TCDD, respectively.

TCDD exposures were characterized using serum measures for those who resided in zone ABR in 1976 (near the time of the accident). The authors could not account for additional exposure to TCDD across subjects that might have occurred since the time of the accident, so there is considerable uncertainty in delineating the critical exposure window for the reported effects. In addition, the lack of exposure data for those in the non-ABR zone, however, makes interpretation of the findings difficult. This difficulty is particularly evident, given that the prevalence of dental defects was less among those in the low exposure category of zone ABR (31−226 ng/kg TCDD) (10%) when compared to those in the non-ABR zone (26%).

2.4.1.2.1.5.4.3. *Suitability of data for TCDD dose-response modeling.*

Most of the considerations for conducting a dose-response analysis have been satisfied with the study population, although, exposure assessment uncertainties are a limitation of this study. For example, it is difficult to discern whether these health effects are a consequence of the initial high exposure during childhood or a function of the cumulative exposure for this entire exposure window beginning at the early age. If the latter is true, averaging exposure over the

critical window would add considerable uncertainty to effective dose estimates given the large difference between initial TCDD body burden and body burden at the end of the critical exposure window. Despite the uncertainty in defining the critical window of exposure, dose-response analysis was conducted for this outcome.

2.4.1.2.1.5.5. *Bertazzi et al., 1989; Consonni et al., 2008—Mortality outcomes.*

2.4.1.2.1.5.5.1. *Study summary.*

Several studies have evaluated the mortality of Seveso residents exposed to TCDD following the 1976 accident. The earlier section of this report described the designs of these studies and discussed their findings as they relate to cancer mortality. In this section, some of the findings for other causes of death are described. A key feature of these studies is that patterns of mortality among Seveso residents were investigated according to their zone of residence at the time of explosion relative to general population rates.

A 10-year mortality follow-up of residents of Seveso was published in 1989 (Bertazzi et al., 1989). Poisson regression was used to derive RRs for those who had lived in Zone A at the time of explosion using a referent group consisting of inhabitants who had lived in the uncontaminated study area. Between 1976 and 1986, no statistically significant difference was observed in all-cause mortality relative to the general population among those who lived in the most highly exposed area (Zone A) at the time of the accident. This finding was evident in both males (RR = 0.86, 95% CI = 0.5−1.4) and females (RR = 1.14, 95% CI = 0.6−2.1). A statistically significant excess in circulatory disease mortality was found among males relative to those in the referent population (RR = 1.75, 95% CI = 1.0−3.2); this increased risk was more pronounced when the follow-up period was restricted to the first 5 years after the accident (1976−1981) (RR = 2.04, 95% CI = 1.04−4.2). Between 1982 and 1986, the RR decreased substantially and was not statistically significant (RR = 1.19, 95% CI = 0.4−3.5). Among females, a risk similar in magnitude was detected for circulatory disease mortality although it was not statistically significant (RR = 1.89, 95% CI = 0.8−4.2). Contrary to the calendar period-specific findings for males, the excess of circulatory mortality among females occurred between 1982 and 1986 (RR = 2.91, 95% CI = 1.1−7.8) and not between 1976 and 1981 (RR = 1.12, 95% CI = 0.3−4.5). The number of deaths in this cohort with the 10 years of

follow-up was relatively small; in Zone A, 16 deaths were observed among males and 11 among females.

The most recently published account of the mortality experience of Seveso residents provides further information on follow-up of these residents until the end of 2001 (25 years after the accident) (Consonni et al., 2008). Three exposure groups were considered: Zone A (very high contamination), Zone B (high contamination), and Zone R (low contamination). The reference population consisted of those residents who lived in unaffected surrounding areas, as well as residents of five nearby towns. The authors used Poisson regression to compare mortality rates for each zone relative to the reference population.

For all causes of death, no excess was found in Zone A, B, or R relative to the reference population. Statistically significant excesses were noted for those who lived in Zone A relative to the reference population for chronic rheumatic heart disease (RR = 5.74, 95% CI = 1.83−17.99) and chronic obstructive pulmonary disease (RR = 2.53, 95% CI = 1.20−5.32). These risks, however, were based on only 3 and 7 deaths, respectively. For those in Zone A, no statistically significant excesses in mortality were noted for diabetes, accidents, digestive diseases, ischemic heart disease, or stroke. Among Zone A residents, stratified analysis by time since accident showed increased rates of circulatory disease 5−9 years since the accident (RR = 1.84, 95% CI = 1.09−3.12). Increased mortality from diabetes relative to the reference population was noted among females who lived in Zone B (RR = 1.78, 95% CI = 1.14−2.77).

2.4.1.2.1.5.5.2. *Study evaluation.*

The ascertainment of mortality in this cohort is nearly complete. Misclassification of some health outcomes, such as diabetes, may occur due to use of death certificate data.

The characterization of exposure is based on zone of residence. Soil sampling indicated considerable variability in TCDD soil levels, and therefore, the generation of risks based on zone of residence likely does not accurately reflect individual exposure. Exposure misclassification might also occur because residency in the areas does not necessarily reflect whether the individual would have been present in the area at the time the accident occurred. Any exposure misclassification would likely be nondifferential which would tend to bias the risk estimates towards the null.

Although some excess of circulatory disease mortality was found, the finding was not consistent between men and women. Moreover, excess circulatory disease mortality was more pronounced among men within the first 5 years of exposure, while, for women, the excess was more pronounced in years 5−10. Numerous other risk factors for circulatory disease were not controlled for in these analyses and may be confounders if related to TCDD exposure. Taken together, the possibility that TCDD increased circulatory disease mortality based on these data is tenuous at best.

2.4.1.2.1.5.5.3. *Suitability of data for TCDD dose-response modeling.*

There is considerable uncertainty in these data due to the potential for outcome and exposure misclassification. The lack of the individual-level TCDD levels and the examination of fatal outcomes reported in this study are not a suitable basis for development of an RfD. For these reasons, dose-response analysis for this outcome is not conducted.

2.4.1.2.1.5.6. *Baccarelli et al., 2005—Chloracne.*

2.4.1.2.1.5.6.1. *Study summary.*

Baccarelli et al. (2005) published findings from a case-control study of 110 chloracne cases and 211 controls. The authors collected information on pigment characteristics and an extensive list of diseases. This study was performed to yield information about the health status of chloracne cases, TCDD-chloracne exposure response, and factors that could modify TCDD toxicity. TCDD was measured from plasma. Following adjustment for confounding, TCDD was associated with chloracne (OR = 3.7, 95% CI = 1.5−8.8), and the risk of chloracne was considerably higher in subjects younger than 8 at the time of the accidents (OR = 7.4, 95% CI = 1.8−30.3). Among individuals with lighter hair, the association between TCDD and chloracne was stronger than among those with darker hair.

2.4.1.2.1.5.6.2. *Study evaluation.*

Although a dose-response association was observed, chloracne is a rare health outcome likely only to occur among those highly exposed.

2.4.1.2.1.5.6.3. *Suitability of data for TCDD dose-response modeling.*

Given the very high TCDD levels needed to cause chloracne, quantitative dose-response modeling to characterize risks for the general population with much lower TCDD exposures would be of little value. Therefore, quantitative dose-response assessment for the Baccarelli et al. (2005) study was not conducted.

2.4.1.2.1.5.7. *Baccarelli et al., 2002, 2004—Immunologic effects.*

2.4.1.2.1.5.7.1. *Study summary.*

The relationship between TCDD and immunological effects was evaluated in a sample of Seveso residents (Baccarelli et al., 2002, 2004). Both studies were based on findings from 62 individuals who were randomly selected from Zones A and B. An additional 59 subjects were chosen from the surrounding noncontaminated areas. Residency was based on where subjects lived at the time of the accident (July 10, 1976) (Landi et al., 1998). Frequency matching ensured that the two groups of subjects were similar with respect to age, sex, and cigarette smoking status.

TCDD levels were determined by mass spectrometric analysis of plasma samples. TCDD levels at the time of sampling were obtained, and estimates of levels at the time of the accident also were estimated by assuming an 8.2-year half-life (Landi et al., 1998). The plasma was also used to characterize levels of the immunoglobulins (Ig) IgG and IgM and the complement components C3 and C4. One subject was excluded due to lack of an immunological evaluation. Analyses are, therefore, based on 58 subjects in the noncontaminated areas and 62 individuals from the contaminated areas.

Nonparametric tests were applied to test for differences between the two groups. Multiple regression also was used to describe the relationship between the variables. Adjustment was made for several potentially confounding variables that were collected via a questionnaire.

An inverse association was noted with increasing TCDD levels and plasma IgG levels; this result remained statistically significant after adjusting for other potential confounding variables in the regression models. Specifically, the slope coefficient and *p*-value for the unadjusted model were −0.35 ($p = 0.0002$) and for the adjusted model the *p*-value was 0.0004. The authors did not present the slope coefficient for the adjusted model in either paper but noted minimal differences between the adjusted and unadjusted results. In the 2004 analysis, the

authors present IgG, IgM, IgA, C3, and C4 median and interquartile values across TCDD exposure quintiles. Decreased levels of IgG were observed in the highest exposure groups. Specifically, the median values across the five quintiles (for lowest to highest) were 1,526; 1,422; 1,363; 1,302; and 1,163. The Kruskal-Wallis test for differences across the TCDD categories was statistically significant ($p = 0.002$), which is consistent with the findings for the continuous measures of TCDD. This finding persisted after excluding those subjects with inflammatory diseases and those who used antibiotics or nonsteroidal anti-inflammatory drugs. For the other plasma measures, no dose-response relationship was apparent based on median values for IgM, IgA, C3, or C4 across TCDD quintiles. The authors highlight the need for additional research, particularly given the excess of lymphatic tumors noted in the area.

Exposure to other dioxin-like compounds for both the TCDD and nonexposed areas were reported to be at background levels.

2.4.1.2.1.5.7.2. *Study evaluation.*

Both TCDD exposure and health outcome measures are well characterized. TCDD exposures, in particular, are based on current serum measures and, therefore, are not dependent on assumptions needed to back-extrapolate to earlier time periods of exposure.

A dose-response relationship between TCDD and IgG is well documented for the unadjusted model, but no details are provided on the change in the slope coefficient when other covariates were added to the model.

Interpreting the inverse association between TCDD exposure and IgG in terms of clinical significance is not possible. The IgG values reported are much higher than those subjects with antibody immunodeficiency disorders.

2.4.1.2.1.5.7.3. *Suitability of data for TCDD dose-response modeling.*

Although the data support an inverse dose-response association between IgG and TCDD, because the relationship cannot be described in terms of clinical relevance with respect to a specific adverse health outcome, these data were not suitable for qunatitative dose-response modeling.

2.4.1.2.1.6. ***The Chapaevsk study.***

2.4.1.2.1.6.1. *Revich et al. (2001)—Mortality and reproductive health.*

2.4.1.2.1.6.1.1. *Study summary.*

Revich et al. (2001) describe a series of investigations that have evaluated adverse health outcomes among residents of Chapaevsk where ecological measures of TCDD have been noted to be higher than expected. In the earlier cancer section of this report, the cross-sectional comparisons of mortality that the authors carried out between Chapaevsk residents and a general population reference were described. Although the general focus of this paper is on cancer, the authors examined other adverse health outcomes.

For all-cause mortality, rates were found to be higher in Chapaevsk relative to the Samara region and other nearby towns. The magnitude of this increase, however, was not quantified in the review by Revich. Cardiovascular mortality accounted for nearly two-thirds of women's deaths and almost half of those among men. The rates of cardiovascular mortality among Chapaevsk men have been reported to be 1.14 times higher than those in Russia.

Revich et al. (2001) also reported on the occurrence of adverse reproductive events. Although the authors indicated that official medical information was used to make comparisons between regions, no details were provided about data quality, completeness, or surveillance differences across areas. The presented rates for reproductive health outcomes should be interpreted cautiously. A higher rate of spontaneous abortions (24.4 per 100 pregnancies finished by delivery) was found in Chapaevsk women relative to rates that ranged between 10.6 and 15.2 found in five other areas. The frequency of preeclampsia also was found to be higher in Chapaevsk women (44.1/100) relative to other towns, as was the proportion of low birth-weight babies and preterm births. The percentage of newborns with low birth weight was slightly larger in Chapaevsk (7.1%) when compared to other towns in Samara (5.1−6.2%); observed differences, however, were not statistically significant. The authors also reported on the sex ratio of newborns born between 1983 and 1997. These ratios (boys:girls) were highly variable and ranged between 0.79 and 1.29. Given the annual variability of this ratio on a year-to-year basis, it is unclear if this is largely due to natural fluctuations and to what extent this may result from prior TCDD (or other contaminants) exposure TCDD and other contaminants.

2.4.1.2.1.6.1.2. *Study evaluation.*

The review by Revich et al. (2001) highlights analyses that have been undertaken using largely cross-sectional data. Although soil sampling measures appear to demonstrate decreasing levels of TCDD in the soil with increasing distance from the plant, at this time, no individual-level TCDD exposure data are available. Increased rates of mortality relative to the Samara region in Russia were observed among Chapaevsk men for all cancer sites combined; this excess risk however, was not observed among women. Although the authors provide compelling evidence of increased adverse events among residents of Chapaevsk, the study lacks a discussion about the validity of comparing health data across regions, and suffers from inherent limitations from ecological studies such as exposure misclassification.

2.4.1.2.1.6.1.3. *Suitability of data for TCDD dose-response modeling.*

As with the cancer outcomes presented in this study, the data for noncancer outcomes are limited by the absence of TCDD levels on an individual-level basis and information on other potential confounding variables that could have biased the comparisons. Additional studies are being undertaken to evaluate the relationship between TCDD and the sexual and physical development of boys. The cross-sectional nature of the data that were presented does not provide the necessary level of detail needed to estimate effective dose given the lack of individual-level exposure data. Therefore, a quantitative dose-response analysis was not conducted.

2.4.1.2.1.7. *The Air Force Health ("Ranch Hands" cohort) study.*

2.4.1.2.1.7.1. *Michalek and Pavuk (2008)—Diabetes.*

2.4.1.2.1.7.1.1. *Study summary.*

Michalek and Pavuk (2008) examined both the incidence of cancer and the prevalence of diabetes in the cohort of Ranch Hand workers exposed to TCDD. As noted previously, these veterans were responsible for aerial spraying of Agent Orange in Vietnam between 1962 and 1971. Exposure to TCDD was estimated using serum collected from participants in 1987 and assayed for TCDD. Exposure to TCDD was estimated using a first-order pharmacokinetic model with a half-life of 7.6 years and provided an estimate of TCDD at the end of the tour of duty in Vietnam. Veterans were grouped into four categories: comparison, background, low, and high.

Diabetes was identified from diagnoses during the post-Vietnam era from medical records. Overall, no differences were shown in the RR of diabetes between the Ranch Hand unit and the reference group (RR = 1.21, $p = 0.16$). Stratified analyses by days of spraying (<90 days, ≥90 days), however, revealed a significant increase in risk of diabetes (RR = 1.32, p = 0.04) among those who sprayed for at least 90 days. A dose-response relationship was also evident when log10TCDD was modeled in the combined cohort. Also, stratification by calendar period showed a dose-response relationship for those whose last year of service was during or before 1969.

2.4.1.2.1.7.1.2. *Study evaluation.*

The Michalek and Pavuk (2008) study provides an opportunity to characterize risks of diabetes as the study is not subject to some of the potential bias of case ascertainment based on death certificates (D'Amico et al., 1999). The quality of the TCDD exposure estimates is high, given that serum data were available at an individual-level basis for all Ranch Hand and comparison veterans used in the cohort. Although disentangling the effects of 2,4-D and TCDD is not possible because their concentrations in Agent Orange are equivalent, 2,4-D has not been associated with diabetes.

2.4.1.2.1.7.1.3. *Suitability of data for TCDD dose-response modeling.*

The reported dose-response relationship between TCDD and diabetes is supported by study strengths including the use of the individual-level level TCDD serum measures and the identification of diabetes through medical records are important strengths of the Michalek and Pavuk (2008) study. Nonetheless, the possible confounding from the inability to control for 2,4-D and other agents used in Agent Orange precludes a quantitative dose-response analysis.

2.4.1.2.1.8. *Other noncancer studies of TCDD.*

2.4.1.2.1.8.1. *McBride et al., 2009a—Noncancer mortality.*

2.4.1.2.1.8.1.1. *Study summary.*

The McBride et al. (2009a) mortality study of New Zealand workers employed as producer or sprayers with potential exposure to TCDD was described earlier in this report. These individuals were employed at a plant that manufactured 2,4,-dichlorophenoxyacetic acid,

and later 2,4,5-T and 4-chloro 2-methyphenoxyacetic acid. In 1987, the plant closed and 2,4,5-T production ceased in 1988.

The cohort consisted of 1,754 individuals who were employed for at least one day at the New Plymouth site between January 1, 1969, and October 1, 2003. Vital status was determined until the end of 2004. Comparisons of mortality were made to the New Zealand general population using the SMR statistic. Exposure was characterized by duration of employment. Person-years of follow-up were tabulated across strata defined by age, calendar period, duration of employment, sex, latency, and period of hire. Analyses were stratified to compare risks by duration of employment (<3 or ≥3 months), latency (<15 or ≥15 years), and period of hire (<1976, ≥1976).

Overall, no statistically significant differences in all-cause mortality relative to the general population were found among those who worked for at least 3 months (SMR = 0.92, 95% CI = 0.80−1.06) or for less than 3 months (SMR = 1.23, 95% CI = 0.91−1.62). No statistically significant excesses were found for mortality from diabetes, cerebrovascular disease, heart diseases, or accidents. The incorporation of a latency period of 15 years revealed no statistically significant excesses for these same causes of death. Similarly, no excesses for any cause of death were noted among those who were hired either before or after 1976.

In subsequent analyses of the same cohort that used estimated TCDD levels from serum samples, McBride et al. (2009b) found no excesses for all-cause mortality or mortality from diabetes or heart disease.

2.4.1.2.1.8.1.2. *Study evaluation.*

For the McBride et al. (2009a) study, the size of the cohort is large enough to characterize mortality risks relative to the general population for most common causes of deaths. An important limitation of this study is the loss to follow-up of a substantial percentage of workers (22%). This would have impacted statistical power by reducing the number of deaths among the workers. If this incomplete ascertainment of mortality outcomes did not occur in a similar fashion with the general population then the SMR may also be biased.

For noncancer causes of death, the use of the SMR statistic is more likely to be influenced by the healthy-worker effect. Therefore, the findings obtained for these outcomes should be interpreted with caution. Subsequent analyses published by the same authors

(McBride et al., 2009b) provide improved characterization of TCDD exposure using serum samples.

2.4.1.2.1.8.1.3. *Suitability of data for dose-response analysis.*

Overall, no associations were evident between surrogate measures of TCDD (duration of employment, year of hire) and noncancer mortality outcomes. Further, the use of mortality endpoints is inconsistent with EPA RfD methodology. As such, these data do not support further use in a quantitative dose-response analysis.

2.4.1.2.1.8.2. *McBride et al., 2009b—Noncancer mortality.*

2.4.1.2.1.8.2.1. *Study summary.*

McBride et al. (2009b) further analyzed the cohort of New Zealand workers to include estimates of TCDD exposure based on serum samples. Current and former employees who were still alive and living within 75 km of the site were asked to provide serum samples. Samples were collected from 346 workers representing 22% (346/1599) of the entire study population. These serum measures were used to estimate cumulative TCDD levels for all workers. The exposure assessment approach by Flesch-Janys et al. (1996) was used to estimate time-dependent exposures based on area under the curve models. This was based on a one-compartment first-order kinetic model with a half-life of 7.2 years.

Comparisons of mortality were made to the general population using the SMR statistic. The Cox proportional hazards model was used to conduct an internal cohort analysis across four categories of cumulative TCDD levels for diabetes and ischemic heart disease mortality. The RRs generated from these models were adjusted for sex, hire year, and birth year. No diabetes deaths were observed among women, and therefore, analysis of this outcome was limited to men.

Relative to the general population, no difference in the all-cause mortality experience was observed in exposed cohort members (SMR = 1.0, 95% CI = 0.9−1.2). Similarly, no excess in these workers was observed for heart disease (SMR = 1.1, 95% CI = 0.9−1.5); cerebrovascular disease (SMR = 1.1, 95% CI = 0.6−1.9); diabetes (SMR = 0.7, 95% CI = 0.2−2.2); or nonmalignant respiratory disease (SMR = 0.8, 95% CI = 0.4−1.4). For the internal cohort analysis, the RR associated with cumulative categorical TCDD measure was 1.0 for both diabetes and ischemic heart disease.

2.4.1.2.1.8.2.2. *Study evaluation.*

The McBride et al. (2009b) study extends the earlier work the same authors completed in two ways. First, serum measures were used to estimate cumulative TCDD with methodology that has been applied to several other cohorts of workers exposed to TCDD. Second, the authors used regression analyses that examined individual-level TCDD exposures in relation to various outcomes as part of the internal cohort comparisons. For noncancer outcomes, no dose-response associations with TCDD were observed with the internal comparisons. Also, as found with earlier analyses of this same cohort, no excess noncancer mortality relative to the New Zealand general population was observed.

Associations between TCDD and diabetes have been found previously in TCDD-exposed populations, most notably in the Ranch Hands cohort (Michalek and Pavuk, 2008). In this cohort, only five deaths from diabetes were identified, and of these, only three occurred among those who were exposed to TCDD. The study, therefore, has limited statistical power to characterize associations between TCDD and mortality from diabetes. Further, the identification of diabetes deaths is subject to misclassification errors due to under-reporting (McEwen et al., 2006).

2.4.1.2.1.8.2.3. *Suitability of data for TCDD dose-response modeling.*

McBride et al. (2009b) found no statistically significant associations in any of the noncancer causes of death. Furthermore, the use of mortality endpoints is inconsistent with EPA RfD methodology. Therefore, the data were not suitable for quantitative dose-response analysis for these outcomes.

2.4.1.2.1.8.3. *Ryan et al., 2002—Sex ratio.*

2.4.1.2.1.8.3.1. *Study summary.*

Ryan et al. (2002) conducted an investigation on the sex ratio in offspring of children of pesticide workers who were involved with the production of trichlorophenol and the herbicide 2,4,5-T in Ufa, Bashkortostan, Russia. Ufa was the site of a state agrochemical plant that has been in operation since the 1940s. Between 1961 and 1988, the plant employed more than 600 workers, most in their early 20s. Females, however, accounted for about 15% of the workforce that produced 2,4,5-T and 30% for 2,4,5-trichlorophenol.

Serum samples previously taken in 1992 among 60 men, women, and children from the factory and city of Ufa showed TCDD exposures that were approximately 30 times higher than background levels (Ryan and Schecter, 2000). Blood data were subsequently measured on a sample of 20 workers between 1997 and 2000, and on 23 2,4,5-trichlorophenol workers between 1997 and 2001. In all, 84 individuals who provided blood samples formed the basis of the analysis in this study. Of these, 55 were exposed to 2,4,5-T and 29 were exposed to 2,4,5-trichlorophenol.

Ryan et al. (2002) reviewed company records for these workers to determine the number, sex, and date of birth of any children; birth data were available for 198 workers. Awareness of the study led other workers who had not provided serum to provide information on births that occurred 9 months after the time of first employment in the factory.

The authors calculated descriptive statistics for the 198 workers and compared them to values for the city of Ufa between 1959 and 1996. Tests of statistical significance were made using the z-test, and the chi-square test. The observed proportion of male births (0.40) among the factory workers was much lower than that for the city of Ufa (0.51) ($p < 0.001$). Stratified analyses revealed that this lower ratio was observed only among those paternally exposed to TCDD. Specifically, the proportion of male births among exposed fathers was 0.38 and among exposed mothers was 0.51. This pattern was observed in both the workers exposed to 2,4,5-T (proportion of male births = 0.40) and 2,4,5-trichlorophenol (proportion of male births = 0.35).

2.4.1.2.1.8.3.2. *Study evaluation.*

The Ryan et al. (2002) findings are consistent with earlier work completed for Seveso residents (Mocarelli et al., 2000). Although serum measures were available for 84 individuals, no dose-response of birth ratios was performed using exposure quantified at an individual-level basis. This approach would have been preferred and consistent with that which Mocarelli et al. (2000) used. All comparisons were made using an external comparison group, namely the sex ratio observed in Ufa between 1959 and 1996.

Although serum measures were used to describe TCDD exposure for a sample of the workers, individual-level dose estimates were not calculated for the study population. Specifically, exposures were characterized many years after exposure, and no attempt was made to back-extrapolate to the time of conception. The two groups of workers in the study also

reportedly had high exposure levels of 1,2,3,7,8-pentachlorodibenzo-*p*-dioxin. So, the group level exposure classification (by plant) did not allow consideration of confounding due to other dioxin-like compounds. Another limitation of the study is that the study population is likely nonrepresentative of all workers employed at the plant. Participants included only those willing to provide serum samples and those who volunteered to participate in the study after learning about it in a public forum. If participation was dependent on TCDD exposures and the reproductive health of these subjects, then bias may have occurred.

2.4.1.2.1.8.3.3. *Suitability of data for TCDD dose-response modeling.*

The findings are notable in their consistency with those found in Seveso residents by Mocarelli et al. (2000). For the Ryan et al. (2002) study, serum data were quantified at an individual-level basis. Risk estimates, however, were not derived in relation to these exposures but instead in two separate subgroups (2,4,5-T and 2,4,5-trichlorophenol workers). This important limitation precludes the use of these data for quantitative dose-response modeling.

2.4.1.2.2. ***Feasibility of dose-response modeling for noncancer.***

Relatively few study populations permit quantitative dose-response modeling to be performed for noncancer outcomes. The serum collected among Seveso men and women provide an opportunity to characterize risks for several health conditions in relation to TCDD exposure. The collection of these serum samples, shortly after the accident does not require the back-extrapolation of TCDD levels as in the occupational cohorts, which should reduce the exposure assessment uncertainty and minimize the potential for exposure misclassification.

An added feature of the SWHS is the detailed collection of other risk factor data from trained interviewers. These data allow for risk estimates to be adjusted for potential confounding variables. For the evaluations of reproductive health outcomes, this adjustment is critical given there are various documented risk factors for the different outcomes that were examined. For some health outcomes, continued follow-up of the cohort is needed, given that several of the Seveso studies suggest that those exposed at a very young age might be more susceptible to subsequent adverse health effects.

The findings of positive associations and dose-response relationships with serum-based measures of TCDD suggest several noncancer health outcomes could be associated with TCDD

exposure. These health outcomes include neonatal thyroid function, sex ratio, diabetes, and semen quality. Although findings have suggested an association between TCDD and age at menopause, they were not statistically significant and no dose-response trend was observed. Weak or nonstatistically significant associations have been noted for endometriosis and menstrual cycle characteristics and do not support quantitative dose-response analyses.

Associations between TCDD exposure and cardiovascular disease have been noted in some, but not all, of the occupational cohorts, and also shortly after the accident among Seveso residents. Findings from the cohort studies based on external comparisons using the SMR statistic should be interpreted cautiously due to potential bias from the healthy worker effect. Because the magnitude of the healthy worker bias is recognized to be larger for cardiovascular diseases than for cancer outcomes, risk estimates in some occupational cohorts might be underestimated for cardiovascular outcomes. Information on cardiovascular risk factors generally was not captured in these studies, and sensitivity analyses were generally designed to examine risk estimates generated for cancer outcomes.

2.4.1.2.3. *Summary of epidemiologic noncancer study evaluations for dose-response modeling.*

All epidemiologic noncancer studies summarized above were evaluated for suitability of quantitative dose-response assessment using the TCDD-specific considerations and study inclusion criteria. The results of this evaluation are summarized in a matrix style array (see Table 2-3) at the end of the chapter, and descriptively in Appendix B. The key epidemiologic noncancer studies suitable for further TCDD dose-response assessment are presented in Table 2-5.

2.4.2. Summary of Animal Bioassay Studies

This section summarizes studies that have met the in vivo animal bioassay TCDD study inclusion criteria (see Section 2.3.2) and are considered in the dose-response modeling conducted later in this document (see Sections 4 and 5). The sections that follow summarize the experimental protocol, the results, and the NOAELs and LOAELs identified in reproductive studies, developmental studies, and general toxicity studies (subdivided by duration).

To evaluate and discuss studies consistently, doses were converted to nanograms per kilogram body weight per day (ng/kg-day) and were also adjusted for continuous exposure. Some doses were adjusted based on daily dietary intake and body weight. For these studies, EPA uses 10% of an animal's body weight as the daily feed rate. More commonly, doses were adjusted from 5 days/week to a 7 days/week standard adjustment, in which case administered doses were multiplied by 5 and divided by 7 to obtain continuous doses. To adjust for weekly dosing, the weekly administered doses were multiplied by the administration frequency per week (in days) and divided by 7 to give continuous doses.

Other exposure protocols used a single loading dose followed by weekly maintenance doses. To adjust these doses, the loading dose was added to the maintenance doses multiplied by the administration frequency, and this sum was divided by the exposure duration to give a continuous dosing rate. The doses administered in single dose studies were not averaged over the observation period.

2.4.2.1. *Reproductive Studies*

2.4.2.1.1. *Bowman et al., 1989a, b (and related Schantz and Bowman, 1989; Schantz et al., 1986).*

Female rhesus monkeys (6 to 10 years old; 8 per treatment) were exposed to 0 or 5 ppt (for 3.5 years), or 25 ppt (for 4 years) TCDD (purity not specified) (Bowman et al., 1989a, b; Schantz and Bowman, 1989; Schantz et al., 1986). Female monkeys were mated to unexposed males after 7 months (Cohort I) and 27 months (Cohort II) of exposure, then again 10 months post-exposure (Cohort III). The average daily doses to mothers were equivalent to 0, 0.15, and 0.67 ng/kg-day. The 0.67 ng/kg-day dose group had reduced reproductive rates in both Cohorts I ($p < 0.001$) and II ($p < 0.025$; Bowman et al., 1989a). The mean number of days of offspring survival ($p < 0.023$) also decreased. No effects on birth weight or growth, or physical evidence of toxicity (Bowman et al., 1989b) were observed. Behavioral effects were observed in the offspring (Cohort I: 7, 6, and 0 offspring, respectively; Cohort II: 3, 5, and 0 offspring, respectively; Cohort III: 6, 7, and 3, respectively). In the 0.67 ng/kg-day dose group, the number of offspring was insufficient to form a group in either Cohorts I or II. Offspring in the 0.15 ng/kg-day dose group had alterations in social behavior of the mother-infant pairs (mothers had increased care giving, which appeared to be an effect of the infants and not due to the

treatment of the mother) and peer group of the offspring after weaning (Cohort I offspring were more dominant or aggressive and exhibited more self-directed behavior; Bowman et al., 1989b). The performance of learning tasks was inversely related to the level of TCDD in the body fat. Schantz and Bowman (1989) examined effects using discrimination-reversal learning (RL) and delayed spatial alteration (DSA). RL detected effects in the 0.15 ng/kg-day group as measured by retarded learning of the shape reversal ($p < 0.05$), but DSA did not. Schantz et al. (1986) combined the cohorts and looked at 5, 5, and 3 mother-infant pairs in the 0, 0.15, and 0.67 ng/kg-day groups, respectively. They found that TCDD-exposed mother-infant pairs spent more time in close, social contact compared to the controls (mutual ventral contact, $p < 0.025$; nipple contact, $p < 0.01$) and infants had reduced locomotor activity ($p < 0.05$), but the dose-effect was complex. Of note is that the control groups contained fewer males than did the TCDD-exposed groups.

In a follow-up study, Rier et al. (2001) examined the DLC levels of sera collected from some monkeys in this study. They reported that animals in this study had elevated serum PCB77 and PCB126 levels and an increased serum TEQ. In fact, the fractional contribution of serum TCDD levels to total serum TEQ was 30% in treated animals. In this study, it is not possible to determine the contribution of TCDD alone to the developmental effect due to the background contamination; thus, EPA has not developed a TCDD LOAEL from the study.

2.4.2.1.2. *Hochstein et al., 2001.*

Adult female mink (12/treatment group) were administered dietary concentrations of 0.0006 (control), 0.016, 0.053, 0.180, or 1.40 ppb TCDD (purity >99.8%) for 132 days (Hochstein et al., 2001). This dose is estimated to be equivalent to 0.03 (control), 0.8, 2.65, 9, and 70 ng/kg-day assuming a food consumption of 5% of body weight per day. Females were mated with unexposed males beginning on treatment day 35. Females were allowed to mate every fourth day during a 29-day mating period or until a confirmed mating. Mated females were presented with a second male either the day after initial mating or 8 days later. In the 70 ng/kg-day group, the treated animals were lethargic after 4 to 5 weeks, with several having bloody (tarry) stools near the end of the trial. Two animals in the 70 ng/kg-day dose group died prior to study termination. These animals had lost a large percentage of their body weight (24−43%), and had pale yellow livers and intestinal hemorrhages. Histopathology from both

mink indicated marked diffuse hepatocellular vacuolation. The mean body weight decreased in all treatment groups including the control (losing an average of 3.29% of initial body weight), compared to a dose-dependent loss of up to 26% in the 70 ng/kg-day group. Mating and reproduction were considered subnormal in all groups. The number of females that gave birth in the 0.03 (control), 0.8, 2.65, 9, and 70 ng/kg-day dose groups were 5/12, 0/12, 3/12, 8/12, and 0/11, respectively. The study authors speculated that the subnormal breeding and reproductive performances in the control females likely were due to the indoor environment in which the mink were housed. In the three groups that gave birth, there was a dose-dependent decrease in kit body weight at birth, which was significant ($p < 0.05$) in the 9 mg/kg-day group compared to the controls. The body weight in the kits was not significantly different at 3 or 6 weeks after birth. Three-week survival rates of 71, 47, and 11% were recorded for kits in the 0.03 (control), 2.65, and 9 ng/kg-day dose groups, respectively. Six-week kit survival rates were 62, 29, and 11% in the 0.03 (control), 2.65, and 9 ng/kg-day dose groups, respectively.

In the adult females, clinical signs of toxicity were noted in the 70 ng/kg-day group near the end of the study and included alopecia and notably thickened, deformed, and elongated toenails. There was a dose-dependent decrease in plasma total solids, total protein, and osmolality that reached statistical significance ($p < 0.05$) in the two highest exposure groups. Anion gap was significantly decreased ($p < 0.05$) and alanine aminotranferase was significantly increased in the 70 ng/kg-day group compared to the controls. At terminal sacrifice, there was a dose-related decrease in body weight. There was a dose-related increase in liver weight that reached statistical significance ($p < 0.05$) in the 70 ng/kg-day dose group. The brains of 42% of the animals in the 70 ng/kg-day dose group had localized accumulation of lymphatic cells within the meninges with mild extension into the adjacent neuropil and mild gliosis. Of the 10 mink surviving to study termination in the 70 ng/kg-day group, 3 had periportal hepatocellular vacuolation. These same brain and liver lesions were not observed in the control mink.

As there were no litters produced in the low-dose group and pregnancy outcomes were not dose related, the 0.8 ng/kg-day exposure level does not inform the choice of NOAEL or LOAEL. Thus, the LOAEL for this study is 2.65 ng/kg-day (132-day maternal exposure duration) based on reduced kit survival (47% of control at 6 weeks). A NOAEL cannot be determined for this study.

2.4.2.1.3. *Ikeda et al., 2005.*

Ikeda et al. (2005) studied the effect of repeated TCDD exposure to F0 dams on the male gonads of F1 generation and sex ratio in the F2 generation. Twelve female Holtzman rats were treated with a single dose of 400 ng/kg TCDD (≥98% purity) orally, via gavage, followed by weekly treatment doses of 80 ng/kg TCDD (16.5 ng/kg-day adjusted for continuous exposure of 10 weeks; specified 2 weeks premating, assumed 1 week for successful mating, 3 weeks of gestation, and specified 4 weeks to weaning) during mating, pregnancy, and lactational periods (total exposure duration approximately 10 weeks). Corn oil served as the control in another group of 12 dams. Four dams were sacrificed on gestation day (GD) 20 to evaluate the in utero toxicity of TCDD. Litter sizes from the remaining eight dams were examined on postnatal day (PND) 2, and some of the F1 offspring were sacrificed to estimate TCDD tissue concentrations. The remaining offspring were weaned on PND 28. Some of the F1 (number not specified) offspring were mated with untreated females on PND 98, following which, litter size, sex ratio, weight, and anogenital distance of F2 pups were examined on PND 2. Mated and unmated F1 males were sacrificed and the testes, epididymis, seminal vesicle, and the ventral prostate were weighed; the cauda epididymis was weighed and examined for sperm count.

All fetuses in the control and TCDD group as a result of in utero exposure in the F0 generation survived. Litter size, sex ratio, and anogenital distance in the F1 generation on PND 2 were not altered as a result of in utero TCDD exposure. Pup weight was significantly ($p < 0.05$) lower in the TCDD-treated group than in controls. TCDD concentration in the adipose tissue of the F0 dams on GD 20 was significantly ($p < 0.05$) higher than in the liver. Adipose TCDD was significantly ($p < 0.01$) reduced at weaning, however, compared to concentrations on GD 20. F1 pup liver TCDD concentration increased significantly ($p < 0.01$) and was higher on PND 28 than PND2. The liver weight in F1 males increased by 14-fold at PND 28 compared to PND 2, implying a transfer of approximately 850 pg of TCDD from the dam to the F1 pup livers during lactation. TCDD also was detected in pup adipose tissue on PND 28. Body weight of TCDD-exposed F1 males was significantly ($p < 0.001$) lower than control males at weaning (PND 28). No significant differences in testis and cauda epididymis weights were observed between the control and treated groups. Ventral prostate weight in the F1 males exposed to TCDD, however, was approximately 60% lower than controls. No change in

weight of the body, brain, testes, cauda epididymis, or seminal vesicle was observed at PND 120. Ventral prostate weight, however, was 16% lower than that of the control group ($p < 0.001$). Sperm count in the cauda epididymis of the F1 males was not affected by TCDD exposure.

Examination of F2 generation litters indicated no significant differences in litter size, pup body weight, and anogenital distance between TCDD-treated or vehicle control groups. The percentage of male F2 pups born to maternally and lactationally TCDD-exposed males was significantly ($p < 0.05$) lower (38%) than those sired by control group males (52%). Every female mated with maternally TCDD-exposed F1 males delivered more female than male pups.

A LOAEL for TCDD of 16.5 ng/kg-day for an estimated 10 week exposure duration in F0 rat dams is identified in this study for decreased development of the ventral prostate in the F1 generation (60% lower than controls) and for significantly ($p < 0.05$) altered sex ratio (decreased percentage of males) in the F2 generation. A NOAEL cannot be determined for this study.

2.4.2.1.4. *Ishihara et al., 2007.*

Ishihara et al. (2007) examined the effect of repeated TCDD exposure of F0 males on the sex ratio of F1 offspring. Seven-week-old male ICR mice ($n = 127$) were divided into three groups and treated via gastric intubation with an initial loading dose of either 2 or 2,000 ng TCDD/kg BW or an equivalent volume of sesame oil (vehicle) as control, followed by a weekly maintenance doses of 0, 0.4, or 400 ng/kg until the animals were 12 weeks old. One week after the last exposure, the animals were mated with untreated female mice. On the day a vaginal plug was identified, F0 male mice were sacrificed and major organs including testes, epididymis, and liver were removed and weighed. Organ tissues also were examined for histopathological and immunohistochemical changes. Treatment levels, averaged over the 6 week period from start of treatment to mating (five maintenance doses), were 0, 0.095, and 950 ng/kg-day for the control, low dose and high dose groups, respectively.

All TCDD-treated males successfully impregnated untreated females and yielded viable offspring. Mortality, pup weights, and mating and fertility indices were not affected by TCDD exposure. There were no significant differences in body weights or in relative weights of testes, epididymis, or livers in the TCDD-treated F0 males compared to the control group. The livers of some animals (number not specified) in the high-dose group, however, were larger and heavier

than in the controls or the low-dose group. Hence, tissues from the high-dose animals were selected for detailed immunohistochemical examination.

General histopathological findings in the TCDD-treated groups showed no changes in cell morphology in germ, Sertoli, and Leydig cells of the testes. Arrangement of the germ cells was normal and there was no difference in the epididymis spermatozoon number in either of the TCDD-treated groups compared to controls. Livers of some of the animals in the high-dose group however, showed enlarged and vacuolated areas in the centrilobular area when compared to the low-dose group and the control group. Immunohistochemical and quantitative immunohistological findings showed a marked increase in staining intensity for cytochrome P450 (CYP)1A1 in the cytoplasm of the hepatocytes in the centrilobular area of the high-dose TCDD group compared to the cells in the low-dose and the control groups. In addition, proportions of immunoreactive CYP1A1 areas in the liver sections of the high-dose group were higher than in the low-dose and control groups. The proportions of immunoreactive CYP1A1 also varied across animals (n = 33) in the high-dose group.

In addition to the above findings, there was a dose-related decrease in the male/female sex ratio. The proportion of male offspring of the high-dose group was significantly lower ($p < 0.05$) than that observed in controls (46.2% versus 53.1%, respectively). Hepatic immunoreactive CYP1A1 staining levels in individual F0 males were strongly correlated with the sex ratio of their offspring.

A LOAEL for TCDD of 950 ng/kg-day for a 6 week exposure duration of F0 male mice is identified for significantly ($p < 0.05$) decreased male/female sex ratio (i.e., higher proportion of female offspring) in the F1 generation. The NOAEL is 0.095 ng/kg-day.

2.4.2.1.5. *Latchoumycandane and Mathur, 2002 (and related: Latchoumycandane et al., 2002a, b, 2003).*

Latchoumycandane and Mathur (2002) conducted a study to determine whether treatment with vitamin E protected rat testes from TCDD-induced oxidative stress. Groups of albino male Wistar rats (n = 6) were administered an oral dose of 0 (vehicle alone) 1, 10, or 100 ng TCDD/kg-day for 45 days, while another group of animals (n = 6) was co-administered TCDD at the same doses, along with vitamin E at a therapeutic dose of 20 mg/kg-day for 45 days. At study termination, animals were fasted overnight, weighed, and sacrificed. Testis, epididymis,

seminal vesicles, and ventral prostate were removed, weighed, and preserved for further examination. The left testis was used to determine daily sperm production, while the right testis was used for biochemical studies. Superoxide dismutase, catalase, glutathione reductase, and glutathione peroxidase activity were measured in the testes, along with production of hydrogen peroxide and lipid peroxidation.

Body weights of TCDD-treated rats did not differ significantly from the control group. Testis, epididymis, seminal vesicle, and ventral prostate weights in the TCDD-treated groups, however, decreased significantly ($p < 0.05$) when compared to controls. None of these changes were observed in the TCDD-exposed groups receiving vitamin E. There was a dose-related decrease in daily sperm production ($p < 0.05$) in all three TCDD-treated groups when compared to the control group. In contrast, the TCDD treatment groups that also received vitamin E did not show any significant changes in daily sperm production compared to the controls. The TCDD-treated groups also showed significantly ($p < 0.05$) lower activities of the antioxidant enzymes (superoxide dismutase, catalase, glutathione reductase, and glutathione peroxidase) than the control group. Levels of hydrogen peroxide and lipid peroxidation increased significantly ($p < 0.05$) in the testes of the rats treated with TCDD compared to the corresponding controls. The TCDD-treated groups that had been co-administered vitamin E show no difference in antioxidant enzyme activities or in reactive oxygen species production when compared with controls.

A LOAEL for TCDD of 1.0 ng/kg-day for a 45-day exposure duration in rats is identified in this study for significantly ($p < 0.05$) reduced sperm production and significantly ($p < 0.05$) decreased reproductive organ weights. A NOAEL cannot be determined for this study.

2.4.2.1.6. *Murray et al, 1979.*

Male (10−16 per treatment) and female (20−32 per treatment) Sprague-Dawley rats were administered diets containing TCDD (purity >99%) to achieve daily concentrations of 1, 10, or 100 ng/kg-day through three generations. After 90 days of treatment, F0 rats were mated to produce F1a offspring. Thirty-three days after weaning of the last F1a litter, the F0 rats were mated again to produce F1b offspring. Some F0 rats were mated a third time for a cross-mating study. The F1b and F2 rats were mated at about 130 days of age to produce the F2 and F3 generations. No clinical signs of toxicity or changes in body weight and food consumption were

observed in F0 rats during the 90 days of treatment before mating. The 100 ng/kg-day group was discontinued due to the lack of offspring. In the three surviving offspring (all males), no changes in appearance, body weight, or food consumption occurred. A dose of 10 ng/kg-day caused a consistent decreased body weight in both sexes of F1 and F2 rats, which was associated with decreased food consumption. A significant ($p < 0.05$) decrease in fertility in F1 and F2 rats occurred, but not in F0 rats, administered 10 ng/kg-day. The number of live pups and gestational survival index were significantly ($p < 0.05$) decreased in the 100 ng/kg-day F0 rats and in the 10 ng/kg-day F1 and F2 rats. The gestational survival index also was significantly ($p < 0.05$) decreased in F2 rats administered 1 ng/kg-day. Postnatal survival was significantly ($p < 0.05$) reduced only in F2 rats administered 10 ng/kg-day. Growth (as measured by body weight) was affected at 10 ng/kg-day only in the third generation. In the 10 ng/kg-day group, a significant ($p < 0.05$) decrease in relative thymus weight and increase in liver weight also occurred in F_3 rats (weights were not measured in F2 rats). Additionally, mating 100 ng/kg-day TCDD-treated females with untreated males increased the percent of implants resorbed as assessed by uterine histopathology.

The reproductive LOAEL is 10 ng/kg-day, based on a significant ($p < 0.05$) decrease in fertility (33−37% lower than controls); decrease in the number of live pups (18−27% lower than controls); decrease in gestational survival (10−11% lower than controls); decrease in postnatal survival (32% lower than controls); and decreased postnatal body weight (14−19% lower than controls at weaning) in one or more generations. The reproductive NOAEL is 1 ng/kg-day.

2.4.2.1.7. *Rier et al., 1993, 1995.*

Reir et al. (1993, 1995) examined the impact of chronic TCDD exposure on endometriosis in monkeys. Female rhesus monkeys (8 animals per treatment group) were exposed to 0, 5, or 25 ppt TCDD (purity not specified) in feed for 4 years. Previously, Bowman et al. (1989b) determined that these dietary concentrations were equivalent to 0, 0.15, and 0.67 ng/kg-day, respectively. Ten years after termination of TCDD treatment, the presence of endometriosis was determined via laparoscopic surgical procedure, and the severity of the disease was assessed. The study authors reported that three monkeys in the 0.67 ng/kg-day exposure group died at 7, 9, and 10 years after termination of TCDD treatment. Autopsy results attributed the deaths to widespread and severe peritoneal endometriosis (all three monkeys)

along with obstruction of the colon (one monkey) and blockage of the jejunum (one monkey). Other deaths also occurred in the control group (1 death from birthing complications and another from an unknown cause); in the 0.15 ng/kg-day dose group (1 death due to natural causes with no endometriosis), and in the 0.67 ng/kg-day dose group (1 death due to a breeding fight with no incidence of endometriosis). At study termination, 17 live animals plus the 3 that had previously died of endometriosis were evaluated (total $n = 20$).

Incidence of endometriosis was significantly ($p < 0.05$) higher than in the control group with 71 and 86 % incidence rates in the 0.15 and 0.67 ng/kg-day dose groups, respectively, compared to 33% in the control group. Severity of endometriosis was also significantly ($p < 0.001$) correlated with TCDD dose. Staging by rAFS indicated that untreated control animals had either minimal or no incidence of endometriosis. In comparison, endometriosis was absent in 2 of the 7 monkeys in the 0.15 ng/kg-day dose group, while only 1 of the 7 animals in the high dose group was disease free. Moderate-to-severe disease was observed in 3 of the 7 animals in the 0.15 ng/kg-day dose group and 5 of the 7 animals in the 0.67 ng/kg-day dose group. Moderate-to-severe disease was not observed in the control group. The authors also compared the incidence and severity of endometriosis in TCDD-exposed animals with 304 normal, non-neutered females with no dioxin exposure and reported that the disease was not present in monkeys that were less than 13 years of age, while the disease rate was 30% among animals 13 years of age or older. The study authors report that these findings are in agreement with human and rhesus studies demonstrating that the prevalence of detectable endometriosis can increase with advanced age.

As noted previously, in a follow-up study, Rier et al. (2001) examined the DLC levels of sera collected from some monkeys in this study. They reported that animals in this study had elevated serum PCB77 and PCB126 levels and an increased serum TEQ; the fractional contribution of serum TCDD levels to total serum TEQ was 30% in treated animals. They also reported that the severity of the endometriosis corresponded to the serum PCB77 concentrations rather than total TCDD. In this study, it is not possible to determine the contribution of TCDD alone to the edometriosis due to the background contamination; thus, EPA has not developed a TCDD LOAEL from the study.

2.4.2.1.8. *Shi et al., 2007.*

Pregnant Sprague-Dawley rat dams (3 per treatment group) were administered 0, 1, 5, 50, or 200 ng/kg TCDD (purity >99%) in corn oil by gavage on GD 14 and GD 21 and on PND 7 and PND 14 for lactational exposure to pups (Shi et al., 2007). Ten female pups per treatment were selected and administered TCDD weekly at the same dose levels through their reproductive lifespan (approximately 11 months). The corresponding equivalent daily TCDD doses are 0, 0.14, 0.71, 7.14, and 28.6 ng/kg-day. Vaginal opening was slightly but significantly ($p < 0.05$) delayed in 28.6 ng/kg-day females. Vaginal opening was also delayed, but not significantly, in the 0.14 and 7.14 ng/kg-day groups. Reproductive senescence with normal cyclicity was significantly ($p < 0.05$) accelerated beginning at 9 months in 7.14 and 28.6 ng/kg-day females. Serum estradiol concentrations were decreased at all time points across the estrous cycle in a dose-dependent manner with a statistically significant decrease ($p < 0.05$) in all but the lowest dose group. TCDD exposure, however, did not affect the number or size distribution of ovarian follicles; responsiveness of the pituitary gland to gonadotropin-releasing hormone, or serum profiles of FSH, LH, or progesterone.

A LOAEL for TCDD of 0.71 ng/kg-day for an 11-month exposure duration was identified in this study based on significantly ($p < 0.05$) decreased estradiol levels in offspring. The NOAEL for this study is 0.14 ng/kg-day.

2.4.2.1.9. *Yang et al., 2000.*

Yang et al. (2000) studied the impact of TCDD exposure on the incidence and severity of endometriosis in female rhesus monkeys. Groups of 7- to 10-year old nulliparous cynomolgus monkeys were treated with 0 ($n = 5$), 1, 5, or 25 ($n = 6$ per group) ng/kg BW TCDD 5 days per week via gelatin capsules for 12 months. Because the monkeys received one capsule 5 days per week, the doses adjusted for continuous exposure were 0, 0.71, 3.57, and 17.86 ng/kg-day. Prior to TCDD administration, all animals had endometriosis induced during days 12–14 of the menstrual cycle by auto-transplantation of endometrial-strips in multiple abdominal sites. All TCDD-treated and control groups were laparoscopically examined during months 1, 3, and 6 to monitor the survival of endometrial implantations and to obtain peritoneal fluid to determine the concentration and immunotype of endometrial growth regulator cytokines interleukin-6 (IL-6) and interleukin-6 soluble receptor (IL-6sR). Because insufficient peritoneal fluids were present

in the treated and control monkeys, however, the study authors collected blood samples at 6 and 12 months during laparoscopy for routine hematology and to assess the circulating levels of IL-6 and IL-6sR. All animals were sacrificed at 12 months, and circulating levels of gonadal steroids also were measured at the time of necropsy.

No changes were observed among treatment levels in general toxicological endpoints such as body weight changes, food consumption, hematological endpoints, general activity levels, and caretaker interaction. In addition, TCDD did not impact circulating levels of gonadal steroids measured during necropsy. Similarly, there were no differences in the number of menstrual cycles, the length of the menstrual cycle, and bleeding intervals. Endometrial implants were found in at least one site in all TCDD-treated and control monkeys during the first laparoscopic examination. Follow-up laparoscopies revealed that there was a continuous loss of endometrial implants over time in each dose group. At the 1-, 3-, and 6-month examination, the number of endometrial losses was not significantly different among different dose groups. At the 12-month examination, however, a significantly ($p < 0.05$) higher rate of survival of endometrial implants was observed in the 3.57 and 17.86 ng/kg-day dose groups compared to the control group. The highest rate of endometrial implant survival was observed in the ovaries regardless of the dose group. In contrast, all lesions disappeared from the left broad ligament, whereas two on the right broad ligament and one on the uterine fundus survived. There was a dose-dependent divergence in the growth response of endometrial implants following TCDD exposure. Both the maximum and minimum implant diameters in the 17.86 ng/kg-day dose group were significantly ($p < 0.05$) larger compared to controls. In contrast, the maximum and minimum implant diameters in the 0.71 ng/kg-day dose group were significantly ($p < 0.05$) smaller compared to controls. TCDD did not impact implant diameters in the 3.57 ng/kg-day dose group when compared to controls. Histological examinations revealed that endometrial glands and stromal cells were present in all surviving implants. Sections examined in the 17.86 ng/kg-day of TCDD possessed cystic endometrial glands that were more frequently observed in this dose group compared to other groups including controls. In addition, circulating levels of IL-6 were significantly ($p < 0.05$) lower in monkeys exposed to 17.86 ng/kg-day TCDD both at 6 and 12 months compared to the control group. In contrast, circulating levels of IL-6sR were significantly ($p < 0.05$) higher in animals treated with 3.57 and 17.86 ng/kg-day TCDD at 6 months, while the levels were higher only in the 17.86 ng/kg-day TCDD group at 12 months.

A LOAEL for TCDD of 17.86 ng/kg-day for a 1 year exposure duration was identified in this study for significantly ($p < 0.05$) increased endometriosis induced by endometrial implant survival, significantly ($p < 0.05$) increased maximum and minimum implant diameters, and growth regulatory cytokine dysregulation (as assessed by significantly decreased IL-6 levels, $p < 0.05$). A NOAEL of 3.57 ng/kg-day is identified in this study.

2.4.2.2. *Developmental Studies*

2.4.2.2.1. *Amin et al., 2000.*

Amin et al. (2000) studied the impact of in-utero TCDD exposure on the reproductive behavior in male pups. Groups of pregnant Harlan Sprague-Dawley rats (n = 108 divided into 4 cohorts; number of animals in the TCDD treatment group is ~3 per dose group) were dosed via gavage with 0, 25, or 100 ng/kg-day TCDD (purity >98%) in corn oil on GDs 10−16. On the day of birth (PND 0), pups were examined for gross abnormalities and the number of live pups, their weights, and sex were recorded from each litter. Litters consisting of more than eight pups were reduced to eight, comprised of four males and four females when possible. Litters consisting of fewer than five pups were excluded from the study to minimize between-litter differences in growth rate, maternal behavior, and lactational exposure. After this exclusion, approximately 10 to 11 litters per exposure group remained. All pups were weaned on day 21 and one male and one female were retained to assess reproductive development, play behavior, reproductive behavior, and saccharin preference behavior. Both male and female pups were tested for saccharin preference between 189 and 234 days of age. A saccharin preference test was conducted for 8 days. For the first 4 days, rats were provided bottles containing tap water, and on days 5 and 6 the animals were provided a bottle containing water and a bottle containing 0.25% saccharin solution. On days 7 and 8, the animals were provided water and a bottle containing 0.50% of saccharin solution. A 0.50% saccharin solution was used because previous studies have reported that male rats exhibited a greater reduction in preference for this saccharin concentration compared to females, hence the sex difference in preference is more marked at this saccharine dose.

None of the treated dams exhibited any signs of toxicity as a result of exposure to TCDD. Gestational body weight, liver weight, litter size and percent live births were all comparable to the corresponding control group. Birth rate and weaning weight of the pups also were not

affected by TCDD exposure. Sex-related water consumption, however, was significantly ($p < 0.001$) affected during the first 4 days with female pups drinking more water per 100 g of body weight compared to the respective male counterparts. Saccharin consumption was significantly ($p < 0.001$) affected, with females consuming greater amounts of saccharin solution per 100 g body weight compared to the corresponding males. Additionally, both male and female pups drank significantly ($p < 0.001$) more of the 0.25% saccharin solution compared to the 0.50% saccharin solution. Females of all exposure groups consumed less of both the 0.25 and 0.50% saccharin solution compared to the same-sex control group. Comparisons of each exposure group to the control group indicated that only the high TCDD exposure group (100 ng/kg-day) different significantly ($p < 0.05$) compared to control in the consumption of 0.25% saccharin solution. In contrast, for the 0.50% saccharin solution, both the low and high TCDD dose groups differed significantly ($p < 0.05$ and $p < 0.01$, respectively) compared to the control group. The saccharin preference of TCDD-exposed male rats did not differ from that of the male control group. The TCDD-exposed females' preference for saccharin solution, however, was significantly reduced in both the 25 ($p < 0.05$) and the 100 ng/kg-day ($p < 0.005$) dose group compared to that of the female controls. The study authors state that the reduction in saccharin consumption and preference in females could be due to the anti-estrogenic action of TCDD and that recent research reports suggest that TCDD can decrease the level of estrogen receptor (ER) mRNA by blocking the ability of ER to transactivate from the estrogen response element.

A LOAEL for TCDD of 25 ng/kg-day for 7 days of gestational exposure is identified for significantly ($p < 0.05$) decreased preference in the consumption of 0.25% saccharin solution. A NOAEL cannot be determined for this study.

2.4.2.2.2. *Bell et al., 2007a.*

Bell et al. (2007a) examined the reproductive effects of TCDD in rats exposed during development. Female CRL:WI (Han) rats were treated with TCDD (99% purity; dissolved in acetone) in the diet at concentrations of 0 (acetone alone; $n = 75$), 28, 93, or 530 ($n = 65$/group) ng TCDD/kg diet, which provided average doses of 0, 2.4, 8, or 46 ng/kg-day, respectively. Rats were exposed to TCDD 12 weeks prior to mating, during mating, and through pregnancy. Dams were switched to the control diet after parturition. Litters from pregnant dams

were reduced to a maximum size of eight on PND 4 and to five males (if possible) on PND 21. These males were left untreated until sacrificed (25/group, one/litter) on PND 70, while all remaining animals were sacrificed on PND 120. All sacrificed animals were necropsied and received a seminology examination. Prior to sacrifice, during weeks 12 and 13, 20 animals from each dose group were tested for learning ability and motor activity, and were also administered a functional observation battery. During postnatal week 16, groups of 20 male F1 rats from each treatment group were paired with untreated virgin females for 7 days, and mated females were killed on GD 16 and examined for terminal body weights, pregnancy status, number of corpora lutea, and number of intrauterine implantations.

The study authors found no evidence of direct maternal toxicity from exposure to TCDD. In the high-dose groups, 8 of 27 dams suffered complete litter loss compared to 3 dams in the control group, but the difference was not statistically significant. Pup survival at PND 4 was also lower in the high-dose group, but the difference again was not statistically significant.

A dose-related decrease in mean pup body weight was observed on PND 1, and this trend continued throughout the lactation period. High-dose male pups had lower body weights when compared to controls at PND 21, with this trend continuing over the course of the study. Balanopreputial separation (BPS) was significantly ($p < 0.05$) delayed compared to controls in all three treatment groups by 1.8, 1.9, and 4.4 days in the low-, medium-, and high-dose groups, respectively. The study authors reported that adjustment for lower body weights observed at PND 21 and PND 42 did not affect the estimate of delay in BPS. No adverse effects from maternal treatment were observed on learning or in functional observational battery performance. Offspring in the high-dose group exhibited less activity when compared to controls ($p < 0.05$) when they were subjected to a test of motor activity for 30 minutes.

The median precoital time was 2−3 days for all 20 F1 males that were mated during postnatal week 16. The uterine and implantation data were similar in all dose groups and there were no significant differences in the proportion of male offspring between groups. Epididymal sperm counts and sperm motility did not differ significantly between dose groups in animals sacrificed during postnatal week 10. The mean number of spermatids was significantly lower (14%; $p < 0.05$) and the proportion of abnormal sperm was significantly ($p < 0.05$) higher in the high-dose group when compared to controls on PND 70. These effects, however, were not seen in animals sacrificed on PND 120.

Terminal body weights were significantly ($p < 0.05$) decreased in the high-dose group (6.9 %) compared to controls on PND 120, while the depression in body weight in the medium-dose group (5.5%) was not statistically significant. At PND 70, the relative and absolute testis weight of the high-dose group was less than the controls (12 and 18%, respectively). Absolute spleen weight in the high-dose group was significantly higher (8%) on PND 70, and increased significantly ($p < 0.05$) by 1−3% on PND 120 in all dose groups compared to controls. Kidney weight in the low and medium-dose groups was significantly ($p < 0.05$) greater than in controls (~2%) at PND 120. In addition to these organs, ventral prostate (9.4%) and relative liver (~4.5%) weights were significantly ($p < 0.05$) higher than controls on PND 120 in the medium- and low- and high-dose groups, respectively. On PND 120, absolute brain weight was significantly ($p < 0.05$) less than the control in the medium-dose group, while relative brain weight was significantly ($p < 0.05$) higher than the control in the low- and high-dose group. Histological examination revealed no unusual findings.

A LOAEL for TCDD of 2.4 ng/kg-day following an estimated 17 week exposure duration of dams was identified in this study for significantly ($p < 0.05$) delayed BPS. A NOAEL was not identified in this study.

2.4.2.2.3. *Franczak et al., 2006.*

Franczak et al. (2006) examined the impact of chronic TCDD exposure on the onset of reproductive senescence in female rats. Pregnant Sprague-Dawley rats (n = 2–3/dose group) were fed 50 or 200 ng/kg TCDD (>99% purity) or corn oil vehicle (4 mL/kg) orally on GD 14 and 21 and PND 7 and 14 to provide in utero and lactational exposure to TCDD. On PND 21, female pups (n = 7/dose group) were weaned and were subsequently given weekly doses of 50 or 200 ng/kg-week TCDD by gavage (7.14 or 28.6 ng/kg-day adjusted for continuous exposure; administered doses divided by 7) or corn oil vehicle. Exposure continued for up to 8 months, and animals were observed for changes in estrus cycle at 4, 6, and 8 months. Rats were sacrificed at 8 months of age when the TCDD-treated animals had entered the transition to reproductive senescence. Following sacrifice, diestrus concentrations of serum LH, FSH, progesterone, and estradiol were measured, and the ovaries were collected for examination.

Estrus cycles at 4 months exhibited normal cyclicity in both TCDD-exposed groups and did not differ significantly from the control group. At 6 months, however, there was a tendency

($p < 0.1$) toward loss of normal estrus cyclicity in animals treated with TCDD. At the 8 month observation, estrus cyclicity was significantly ($p < 0.05$) different in both dioxin-exposed groups compared to controls (cumulative TCDD exposure is reported as 1.7 and 8 μg/kg for the 50 and 200 ng/kg dose groups, respectively). The study authors noted that although the low-dose animals showed an increased prevalence of prolonged cycles, persistent estrus or diestrus was observed in only 10% of the rats. Conversely, approximately 50% of the rats exhibited loss of cyclicity in the high-dose group. There were no changes in the number and size distribution of ovarian follicles or the number of corpora lutea at either dose. Progesterone levels at 8 months tended to be higher ($p < 0.08$) in animals receiving either 7.14 or 28.6 ng/kg-day TCDD compared to controls, while serum estradiol concentrations were significantly ($p < 0.03$) lower at diestrus. Serum LH levels in TCDD-treated animals were comparable to those in the control group, while FSH levels were elevated in rats receiving 7.14 ng/kg-day TCDD, but not in the 28.6 ng/kg-day dose group.

A LOAEL for TCDD of 7.14 ng/kg-day for an 8-month exposure duration was identified for significantly ($p < 0.03$) decreased serum estradiol levels. A NOAEL cannot be determined for this study.

2.4.2.2.4. *Hojo et al., 2002 (and related: Zareba et al., 2002).*

Hojo et al. (2002) studied the impact of prenatal exposure to TCDD on sexually dimorphic behavior in rats. Thirty-six pregnant Sprague-Dawley rats were assigned according to a randomized block design to groups receiving 0, 20, 60, or 180 ng/kg TCDD (98% purity) on GD 8. Litters from pregnant dams were culled to 5 females and 5 males on PND 4 and allowed to wean normally, at which time 5, 5, 6, and 5 litters from the 0, 20, 60, and 180 ng/kg TCDD treatment groups, respectively, were maintained for examination of behavioral response. Offspring were exposed to TCDD (from a single maternal exposure) for about 35 days through gestation and lactation. After weaning at PND 21, offspring were fed ad libitum until PND 80, at which time a fixed amount of food was supplied daily to maintain constant body weights. At 90 days old, the rats in these treatment groups were trained to press a lever to obtain food pellets using two operant behavior procedures. Initially, each lever press was reinforced. The fixed-ratio (FR) requirement was then increased every fourth session from the initial setting of 1 to values between 6 and 71. The responses for 30 days were studied under a multiple schedule

combining FR 11 and another schedule requiring a pause of at least 10 sec between responses (differential reinforcement of low rate, or DRL 10-sec)

Pup and dam body weights were not affected by TCDD exposure, and all pups were successfully trained in the lever-press response within 3–4 days. Analyses of the FR procedure data indicated that the male pups responded at a lower rate at all TCDD doses when compared to the control group. In case of female pups, all TCDD-treated groups responded at a higher rate than controls. None of these results were, by themselves, however, statistically significant. Examination of the FR 11 and DRL 10-second data indicated that when considering the FR component of this multiple procedure, males from all three treatment groups responded at lower rates when compared to the controls. Conversely, all female pups responded at a higher rate than controls. In addition, the treatment-by-sex interaction was significant ($p = 0.036$), with the 60 ng/kg female pups responding at a higher rate than the 60-ng/kg male pups. Examination of the delayed response component in the multiple FR 11 and DRL 10-sec procedures indicated that almost all TCDD treatment groups were affected. Like the FR component, male pups at all TCDD dose groups responded at a lower rate compared to controls, while female pups at all dose groups responded at a higher rate than controls. There was also a significant ($p = 0.001$) sex-by-treatment interaction for the DRL 10-sec similar to the FR component. Following behavioral testing, the animals were sacrificed and cortical depth measurements were taken in selected right and left brain regions. Reduced cortical thickness and altered brain morphometry were observed in both male and female offspring in the 180-ng/kg exposure group when compared to controls (reported in a separate article; Zareba et al., 2002).

A nominal LOAEL for TCDD of 20 ng/kg for a single exposure on GD 8 is established for this study based on abrogation of sexually dimorphic neurobehavioral responses. A NOAEL cannot be derived for this study.

2.4.2.2.5. *Kattainen et al., 2001.*

Pregnant Line A, B, and C rats derived from Han/Wistar and Long-Evans rats (4−8 pregnant dams/strain/treatment group) were administered a single gavage dose of 0, 30, 100, 300, or 1,000 ng/kg TCDD (purity >99%) in corn oil on GD 15 (Kattainen et al., 2001). On PND 1, the litters were culled to three males and three females. Offspring were weaned on PND 28. Female pups were sacrificed on PND 35 and male pups were sacrificed on PND 70.

TCDD treatment did not affect body weight or cause clinical signs of toxicity in the dams. In Line B offspring, body weights in the 1,000 ng/kg group were slightly decreased during PND 1−7, while Line C offspring had slightly decreased body weights throughout the study period (data were not provided). The development of the third molar was affected the most in Line C offspring. In 5 of 10 Line C females and 6 of 10 Line C males treated with 1,000 ng/kg TCDD, the lower third molar did not develop. In comparison, 1 of 19 Line A females and 1 of 18 Line B females administered 1,000 ng/kg TCDD lacked the third molar at sacrifice. Third molars were present in all the controls and all male Line A and B offspring administered 1,000 ng/kg. Due to the lack of eruption of the third molar in the majority of Line B and C control females (only 30% erupted), however, the effects of TCDD on third molar eruption could only be evaluated in Line A female offspring (with 94% eruption). There was a dose-dependent decrease in the eruption of the lower third molar in Line A female offspring with a significant ($p < 0.05$) decrease observed in the 300 and 1,000 ng/kg dose groups. In the male offspring, any third molar that developed erupted by PND 70. The mesiodistal length of the existing lower third molar was reduced in a dose-dependent manner in both genders of all three rat lines. In Line A and C females, the decrease was significant ($p < 0.05$) at all doses. The size of the second molars was also significantly decreased with 1,000 ng/kg ($p < 0.05$) in all but Line C males.

A developmental LOAEL for TCDD of 30 ng/kg for maternal exposure on GD 15 is established for this study, based on impaired tooth development (significantly reduced mesiodistal length of the lower third molar by approximately 12% to 38% [$p < 0.05$]). A NOAEL could not be determined.

2.4.2.2.6. *Keller et al., 2007, 2008a, b.*

Keller et al. (2007, 2008a, b) conducted three separate experiments to assess the impact of TCDD on molar tooth development using different mouse strains. In Experiment 1, Keller et al. (2007) used six inbred mouse strains (C57BL/6J, BALB/cByJ, A/J, CBA/J, C3H/HeJ, and C57BL/10J) known to possess high affinity ligand-binding aryl hydrocarbon receptor alleles (*b*), two with *b1* alleles (C57BL/6J and CBA/J), and four with *b2* alleles (BALB/cByJ, A/J, C3H/HeJ, and CBA/J). Females (number not specified) from each strain were mated with males of the same strain. On GD 13, each pregnant female was assigned to one of the four dose groups

and treated with 0, 10, 100, or 1,000 ng TCDD/kg BW via oral gavage. The control group received corn oil. GD 13 was chosen for dosing because the first morphological signs of tooth development occur on GD 11. The first visible signs of the M1 (molar) occur on GDs 13–14 followed by final cuspal morphology, which is determined on GD 15. The F1 offspring of females from each strain were weaned and separated by sex at PND 28 and were euthanized at PND 70. Each F1 mouse was examined for the presence or absence of both maxillary (M^3) and mandibular third molars (M_3) on both the left and right sides. In addition, all mice were scored as either normal or variant in M_1 morphology for both molar rows.

In Experiment 2 (Keller et al., 2008a), dams from six inbred mouse strains (C57BL/6J, BALB/cByJ, A/J, CBA/J, C3H/HeJ, and C57BL/10J) were orally dosed on GD 13 with 0, 10, 100, or 1,000 ng TCDD/kg BW in corn oil. GD 13 was used as the dosing day because it coincided with the formation of Meckel's cartilage (a major signal center) in the mouse mandible that is followed shortly by intramembranous bone formation on GD 15. The A/J mouse strain was abandoned because the authors had difficulty rearing the offspring from this strain. All offspring (n = 4 or 5 per treatment group) from the remaining strains were euthanized at 70 days of age. Mandible size and shape from all selected offspring were examined using geometric morphometric methods to assess the impact of TCDD exposure.

In Experiment 3 (Keller et al., 2008b), dams from six inbred mouse strains (C57BL/6J, BALB/cByJ, A/J, C3H/HeJ, CBA/J, and C57BL/10J) were treated with a single oral dose of 0, 10, 100, or 1,000 ng TCDD/kg-BW in corn oil. GD 13 was chosen as the dosing day because the first visible signs of the first molar (M1) occurs on GDs 13–14 and the final cuspal morphology (the pattern of projections on the chewing surface of the tooth) is not determined until after GD 15. Similar to Experiment 2, the A/J mouse strain was abandoned due to difficulty in rearing offspring. All offspring (n = 107–110 in each of the five strains for all treatment groups) were euthanized at 70 days of age and their molar size, shape, and asymmetry traits were examined using geometric morphometric methods.

In Experiment 1, all four M_3s were present in all dose groups in mice from C57BL/6J, BALB/cByJ, and C57BL/10J strains. A similar response was observed in the A/J strain mice with only 3 of 51 F1 mice exhibiting missing third molars. Approximately one-third of the mice from the CBA/J and C3H/HeJ strains, however, were missing at least one M^3 or M_3 molar. The numbers of CBA/J mice missing one or both M_3 or M^3 molars were 0/29, 2/21, 6/29, and 30/30

in the 0, 10, 100, and 1,000 ng/kg groups, respectively. In the C3H/Hej animals, the numbers missing one or both molars were 1/24, 3/28, 1/26, and 30/36, respectively.

Maternal TCDD exposure was also found to affect the frequency of M_1 variants, but only in the C57BL/10J strain, and the dose-response relationship was nonmonotonic. The proportions of variants observed in the 0, 10, 100, and 1,000 ng/kg dose groups were 33, 68, 59, and 58%, respectively.

A LOAEL for TCDD of 10 ng/kg maternal exposure on GD 13 is identified for this study for increased incidence (33%) of the M_1 variant in the C57BL/10J mouse strain. A NOAEL cannot be determined in this study.

In Experiment 2 TCDD exposure of dams did not affect offspring survival or 10-week body weight in any of the inbred mouse strains used. Analysis of variance (ANOVA) indicated that although mandible size in both male and female offspring varied significantly ($p < 0.0001$) among strains, it was not affected by TCDD exposure. In contrast, analysis of covariance indicated that TCDD exposure significantly ($p = 0.0033$) decreased the mandible size in male offspring in the C3H/HeJ strain at all treatment groups. The mean mandible size was similar across all treatment groups in both sexes in all strains with male offspring exhibiting larger mandibles compared to females. Males in the C3H/HeJ strain exhibited a significant (level not reported) downward trend in mandible size throughout all treatment groups. Females in the C3H strain also showed a similar trend in mandible size, but the trend was not significant. ANOVA on mandible shape indicated that males had significantly ($p < 0.0001$) different mandible shape in strain × treatment groups. In contrast, in female offspring, although the mandible shape was significantly ($p < 0.0001$) different due to strains, treatment groups, and litter, the strain × treatment interaction was not significant. Male offspring from the C3H/HeJ and C57BL/6J mouse strains appear to be more sensitive to TCDD than BALB/cByJ or CBA/J mice, with the C57BL/10J strain exhibiting intermediate sensitivity. In addition to these analyses, Procrustes distance analysis also indicated that C3H/HeJ mice had the greatest response to the highest dose of TCDD, followed by the C57BL/6J strain. Female offspring in the C3H/HeJ and C57BL/6J strains also exhibited the largest change in Procrustes distance with TCDD exposure. This trend, however, was not statistically significant ($p = 0.29$).

A LOAEL for TCDD of 10 ng/kg maternal exposure on GD 13 was identified for this study for significantly ($p = 0.0033$) decreased mandible shape and size in male C3H/HeJ mice. A NOAEL cannot be determined in this study.

In Experiment 3, effect of TCDD exposure on offspring survival or body weight was not reported. Three-way ANOVA results showed significant ($p < 0.0001$) differences in molar size among strains, sexes, and litters, but not between treatment groups. Molar size difference in sex × strain interaction was significant ($p = 0.03$), whereas differences in sex × treatment and sex × strain × treatment were not significant. Additionally, molar size in treatment × strain interaction also was not statistically significant. Based on these results, the authors reported that molar size varied significantly ($p < 0.0001$) among all five strains tested, with all strains exhibiting similar trends in all four treatment groups. Strain differences in molar size were more apparent in male offspring. A hormesis-like trend in molar size was observed in all strains (except in BALBc/ByJ) and sexes with an increase at the 100 ng/kg dose and a decrease in the 1,000 ng/kg dose. In addition to lack of difference in molar size for all treatment groups in all strains, fluctuating asymmetry in molar size also did not increase with increasing doses of TCDD.

In contrast to these results on molar size, the Procrustes ANOVA indicated that molar shape was significantly ($p < 0.0001$) affected by strain, sex, treatment, and litter size. Molar shape in sex × strain and sex × strain × treatment interactions was also highly significant ($p < 0.0001$). Based on these results, the authors concluded that differences between males and females varied based on the strain, and that the effect of TCDD exposure on each strain also differed for male and female offspring. Because molar shape in treatment × strain interaction was significant ($p < 0.0001$), differences in molar shape between the three treatment groups and the control group were analyzed for each strain using nonorthogonal contrasts. In male offspring, contrasts between the control group and 1,000 ng/kg were statistically significant only in the C3H/HeJ ($p < 0.0001$) and CBA/J ($p < 0.03$) strains. These results suggest that these two strains are most susceptible to TCDD effect on molar shape, and similar results were observed in female offspring of these two strains. The contrast in molar shape between the control and the 100 ng/kg treatment group for the female C57BL/6J mice also was statistically significant ($p = 0.0096$). On the whole, when considering Procrustes distance results for molar shape, the C3H/HeJ male offspring had the largest response at the low and high doses, while the female

offspring had the largest response at low and mid doses. This observation in male C3H/HeJ mice is consistent with that of TCDD-induced changes in mandible size from Keller et al. (2008a).

A LOAEL for TCDD of 10 ng/kg maternal exposure on GD 13 is identified for this study for significant ($p < 0.0001$) differences in molar shape in male C3H/HeJ mice. A NOAEL cannot be determined in this study.

2.4.2.2.7. *Kuchiiwa et al., 2002.*

Kuchiiwa et al. (2002) studied the impact of in utero and lactational TCDD exposure on serotonin-immunoreactive neurons in raphae nuclei on F1 male mouse offspring. Twenty-one adult female ddY mice (seven per treatment group) were administered TCDD (99.1% purity) by oral gavage once a week for 8 weeks at doses of 0, 4.9, or 490 ng/kg (0, 0.7, or 70 ng/kg-day average daily dose; administered doses divided by 7) or an equivalent volume of olive oil vehicle (6.7 mL/kg) by gavage. Immediately following the final treatment, the mice were housed with untreated male mice for mating. At approximately 20–21 days after mating, 3 female mice from each dose group, including the control group gave birth to 10–12 offspring. One day after birth, each litter was culled to 10 offspring to accommodate similar lactational TCDD exposure. On PND 28, the offspring were weaned, and three offspring from each TCDD exposed group and the control group were selected for an immunocytochemical examination at 42 days of age. Following sacrifice of these offspring, the brain of each animal was removed and every second serial section of the brain was processed for immunocytochemistry. In addition to the serial sections of the brain, cells from 18 offspring (6 males per treatment group) were used to assess the number of cells in the dorsal and median raphe nucleus, the supralemniscal area, and the Nucleus raphe magnus.

Examination of external morphology, birth, and postnatal body weights indicated that there were no differences between the male TCDD-exposed offspring and the control male offspring. TCDD-exposed males, however, were aggressive toward other normal mice and were also hypersensitive to soft touch.

Serotonin-immunoreactive neurons were found to be distributed throughout the entire brainstem in 42-day-old males, and the general pattern in the TCDD-exposed animals was consistent with those observed in control male offspring. Serotonergic neurons were identified

and counted in the caudal linear nucleus, the median and dorsal raphe nucleus, Nucleus raphe pontis, interpeduncular nucleus, supralemniscal area, pedunculopontine segmental nuclei, deep mensencephalic nucleus, Nucleus raphe magnus, pallidus, and obscurus, dorsal and medial to the facial nucleus and the ventrolateral medulla. Results from computerized cell counts ($n = 6$) showed an average of 1,573.3 immunoreactive neurons in the raphe nuclei from the control group versus 716.3 and 419.8 neurons in the low- and high-dose offspring, respectively. The numbers of immunoreactive neurons in the individual raphe nuclei (dorsalis, medianus, magnus, and B9) from the TCDD-exposed offspring were significantly ($p < 0.01$) lower than control values, with the degree of reduction being dose-related.

In the absence of other relevant neurotoxicity endpoints, reduced serotonin is not an adverse endpoint of toxicological significance in and of itself, thus, neither a NOAEL nor a LOAEL can be established for this study. A lowest-observed-effect level (LOEL) of 0.7 ng/kg-day for an 8-week exposure duration is identified in this study for a significantly ($p < 0.01$) lower number of serotonin-immunoreactive neurons in the raphe nuclei of male offspring. A no-observed-effect level (NOEL) cannot be determined for this study.

2.4.2.2.8. *Li et al., 2006.*

Pregnant and pseudopregnant (obtained by mating normal estrous female mice with vasectomized male mice) NIH mice (10 per treatment group) were exposed to 0, 2, 50, or 100 ng/kg-day of TCDD (purity 99%) during early gestation (GDs 1–8), preimplantation (GDs 1−3), or peri-implantation to postimplantation (GDs 4–8) (Li et al., 2006). On GD 9, animals were evaluated. The two highest TCDD doses (50 and 100 ng/kg-day) caused significant ($p < 0.05$) early embryo loss independent of gestational exposure time. At 100 ng/kg-day, however, the embryo loss was greater when administered during GDs 1–8 or GDs 1–3 compared to GDs 4–8 ($p < 0.01$). Uterine weight was significantly decreased in the pseudopregnant mice when administered 50 or 100 ng/kg-day TCDD during GDs 1–8 ($p < 0.001$) or 1–3 ($p < 0.01$), but was only decreased at 100- ng/kg-day in pseudopregnant mice when administered during GDs 4–8 ($p < 0.01$). Estradiol levels were increased at all TCDD treatment levels (100% at the lowest dose), but statistical significance was not indicated. All doses at all treatment times resulted in a significant reduction ($p < 0.01$) in serum progesterone levels, with a 45% decrease at the lowest dose. Because the hormone effects were observed

following 4 days of treatment, the nominal doses were averaged over the entire test period of 8 days prior to measurement. The resulting average daily doses of TCDD were 0, 1, 25, and 50 ng/kg-day.

A LOAEL of 2 ng/kg-day administered for 4 to 8 days is established in this study for a significant ($p < 0.01$) decrease in progesterone (45% above control) and an approximate 2-fold increase in estradiol levels (significance not indicated). A NOAEL cannot be determined.

2.4.2.2.9. *Markowski et al., 2001.*

Pregnant Holtzman rats (4–7 per treatment group) were administered a single gavage dose of 0, 20, 60, or 180 ng/kg TCDD (purity not specified) in olive oil on GD 18 (Markowski et al., 2001). One female rat from each liter (4–7 per treatment group) was assigned to training on a wheel apparatus to respond on a lever for brief opportunities to run. Once animals responded to an FR1 schedule of reinforcement, the requirement for lever pressing was increased to FR2, FR5, FR10, FR20, and FR30 schedules. After each training session, the estrous cycle stage was determined. Maternal body weight, length of gestation, number of pups per litter, and sex distribution within litters were unaffected by treatment. For each of the FR schedules, there was a significant dose-related ($p = 0.0001$) decrease in the number of earned run opportunities, lever response rate, and total number of revolutions in the wheel in the adult female offspring. There was no correlation between estrous cycle and responding for access to wheel running.

The developmental LOAEL for this study is a single dose of 20 ng/kg administered on GD 18 for neurobehavioral effects. A NOAEL cannot be determined for this study.

2.4.2.2.10. *Miettinen et al., 2006*

Miettinen et al. (2006) administered a single oral dose of 0, 30, 100, 300, or 1,000 ng/kg TCDD (purity >99%) in corn oil on GD 15 to pregnant Line C rats. The offspring (24–32 per treatment group) were assigned to a sugar-rich cariogenic diet (via feed and drinking water) and were orally inoculated three separate times with fresh cultures of *Streptococcus mutans*. Three control groups varied with regard to TCDD exposure and administration of a cariogenic diet. Two of the control groups received no TCDD, and the offspring were either maintained on a normal diet without inoculation with *S. mutans* (C1; $n = 48$) or were given the cariogenic diet with *S. mutans* inoculation (C2; $n = 42$). The final control group was maternally exposed to

1,000 ng/kg TCDD with offspring fed a normal diet without *S. mutans* inoculation (C3; $n = 12$). TCDD did not affect the maternal or offspring body weight. Survival of the offspring was reduced in the 1,000 ng/kg dose group (50–58% survival compared to 83–95% in C1 and C2, respectively). All offspring administered 1,000 ng/kg were missing all lower third molars. Two animals (8%) in the 100 ng/kg group were missing one of their lower third molars. All doses, except the 100 ng/kg dose, caused a significant ($p < 0.05$) increase in the number of caries lesions compared to group C2 (60, 79, 76, 83, and 91% in the C2, 30, 100, 300, and 1,000 ng/kg groups, respectively). Group C3 (1,000 ng/kg TCDD exposure, normal diet) animals also had increased caries lesions compared to C1 (8% versus 0%, respectively). There were no changes in tooth mineral composition that could explain the increase in caries susceptibility.

The developmental LOAEL from this study is a single dose of 30 ng/kg administered on GD 15 based on the significant ($p < 0.05$) increase in dental caries in pups (30% above control). A NOAEL cannot be determined from this study.

2.4.2.2.11. *Nohara et al., 2000.*

Pregnant Holtzman rats were administered 0, 12.5, 50, 200, or 800 ng/kg TCDD in corn oil by gavage on GD 15 (Nohara et al., 2000). On PND 2, five males were randomly selected from each litter and dose group. TCDD was detected in the thymus, spleen, and bone marrow of the male pups on PND 21 and PND 49. TCDD was still detected in the thymus and spleen on PND 120 but the levels decreased over time. The TCDD concentration was highest in the thymus at all time points. There were no changes in the body, thymus, or spleen weights of the male offspring on PND 5, PND 21, PND 49, or PND 120. On PND 5, there was a 200-fold increase in CYP1A1 in the thymus of the high-dose male pups. CYP1A1 was only slightly increased in the spleen. This induction decreased through PND 49. There was a slight (not statistically significant) dose-dependent decrease in thymus cellularity in the male offspring at PND 120. Spleen cellularity at PND 49 decreased in a dose-dependent manner (15–50% of the control), with a statistically significant ($p < 0.05$) decrease observed in the high-dose group. A slight but not significant reduction in spleen cellularity was noted in the high-dose group at PND 21. The same effect was not observed at PND 120, nor was there any change in the percent of B or T cells in the spleen. No changes in cytokine levels were observed in the 800-ng/kg group.

Although a change in spleen cellularity on PND 49 (puberty) was observed, this effect was transient and there were no coexisting changes in the percentage of splenic lymphocytes, spleen weight, and cytokine levels. Therefore, a developmental NOAEL of a single dose of 800 ng/kg administered on GD 15 is identified for this study. A LOAEL is not established.

2.4.2.2.12. *Ohsako et al., 2001.*

Pregnant Holtzman rats (6 per treatment group) were administered 0, 12.5, 50, 200, or 800 ng/kg TCDD (purity >99.5%) in corn oil by gavage on GD 15 (Ohsako et al., 2001). On PND 2, five males were randomly selected from each litter. Two male offspring from each litter were sacrificed on PND 49 and PND 120. Neither maternal nor male offspring body weight was affected by TCDD treatment. TCDD was detected in both fat and testes at all dose levels (including controls) with highest levels found in fat. There were no apparent treatment-related effects on testicular weight, epididymal weight, daily sperm production, cauda epididymal sperm reserves, luteinizing hormone, follicle stimulating hormone, or testosterone levels. There was, however, a clear dose-dependent decrease in urogenital complex weight and ventral prostate weight at both PND 49 and PND 120. For male offspring, statistically-significant ($p < 0.05$) decreases were noted in urogenital complex weight at PND 120 in the 200 and 800 ng/kg groups, in ventral prostate weight at PND 49 in 800 ng/kg group, and at PND 120 in the 200 and 800 ng/kg groups. There was also a dose-dependent decrease in anogenital distance (the length between the base of the genital tubercle and the anterior edge of the anus); the decrease was not statistically significant at PND 49. At PND 120, however, male offspring in all but the lowest dose group had significantly ($p < 0.05$) reduced anogenital distance compared to the control animals. There was also a dose-dependent increase in 5αR-II mRNA expression in the ventral prostate on PND 49 with significant increases ($p < 0.05$) in the 200 and 800 ng/kg animals. There was a significant ($p < 0.01$) decrease in the androgen receptor mRNA in the ventral prostate on PND 49 at all doses tested. Similar effects were not observed on PND 120 or in the caput epididymis on PND 49.

The developmental LOAEL for this study is a single dose of 50 ng/kg administered on GD 15 for significantly ($p < 0.01$) reduced anogenital distance in male offspring (approximately 14%). The NOAEL for this study is 12.5 ng/kg.

2.4.2.2.13. *Schantz et al., 1996.*

Schantz et al. (1996) studied the impact of in utero TCDD exposure on spatial learning in male and female pups. Groups of pregnant Harlan Sprague-Dawley rats (n = 108, divided into 4 cohorts; number of animals in each TCDD group approximately 4 per treatment group) were dosed via gavage with 0, 25, or 100 ng/kg-day TCDD (purity >98%) in corn oil on GDs 10−16. On the day of birth (post natal day [PND] 0), the pups were examined for gross abnormalities and the number of live pups, weight, and sex were recorded for each litter. On PND 2, litters were culled to eight animals and were balanced to include four males and four females whenever possible. To minimize litter-size effects, litters with fewer than five pups were excluded from the study. The exclusion of these litters resulted in 10−11 litters per treatment group. Pups were weaned on PND 21 and one male and one female pup from each litter were maintained for the learning tests. Pups were tested 5 days per week for spatial learning and memory in a radial arm maze and a T-maze. A radial arm maze working memory test and a T-maze DSA task were used a part of the testing process.

TCDD treatment did not affect dam gestational weight gain, dam liver weight, gestation length, litter size, percentage of live births, birth weight, or postnatal growth of the pups observed during the course of the study. Exposed pups, however, exhibited some signs of toxicity in all exposure groups. Thymus weight was decreased and liver weight was increased in the 100 ng/kg-day TCDD dose group. Also, liver microsomal 7-ethoxyresorufin-O-deethylase (EROD) activity was markedly induced in pups from both the 25 and 100 ng/kg-day dose groups. In the radial maze test, rats from all TCDD exposure groups displayed a significant ($p < 0.01$) learning behavior as shown by progressively fewer errors from the first block of sessions through the fourth session. The treatment by sex and treatment by session block interactions were not significant. Comparisons between the average number of errors per session block in the TCDD-exposed and control group indicated that both the 25 and the 100 ng/kg-day dose groups made significantly ($p < 0.05$ and $p < 0.001$, respectively) fewer errors compared to the control group. TCDD did not significantly affect adjacent arm selection behavior as measured by C statistic; hence the reduction in errors observed did not appear to be accounted for by an increased tendency to run into adjacent arms. Female pups had a significant ($p < 0.05$) shorter radial arm maze latency, however, compared to the male pups. In the T-maze test, TCDD did not significantly affect the percent of correct performance. All exposure groups

performed best at the shortest delay, which showed a decline as the length of the inter-trial delay interval was increased. Additionally, all treated groups improved their performance over a three-block session period. This finding indicated that animals in all groups could learn the task. These observations were confirmed by a highly significant main effect of delay ($p < 0.001$) and highly significant main effect of session blocks ($p < 0.001$). At the shortest 15-second delay, average percent correct performance increased from 75 to 92%, while at the longest 40-second delay, the average percent correct performance increased from 62 to 82%. A significant ($p < 0.05$) main effect of exposure was evident in latency to respond in the T-maze. Comparisons of the exposed group to control group, however, indicated that none of the individual exposure groups differed significantly from the controls. Because no clear pattern was observed in the various exposure groups, differences in latency to respond had no impact on learning of the task.

Based on these results, the study authors state that the fact TCDD seems to have a facilitatory effect on radial arm maze learning in rats should be interpreted with caution and needs further evaluation using different and more varied learning tasks. No toxicologically adverse endpoints were concurrently examined. Thus, a LOAEL and a NOAEL cannot be determined for this study.

2.4.2.2.14. *Seo et al., 1995.*

To study developmental effects of TCDD on thyroid hormone levels, time-mated female Sprague-Dawley rat dams (n = 10−14/treatment group) were administered 25 or 100 ng/kg-day of TCDD (>98% pure) in corn oil via gavage from GDs 10–16. Vehicle controls received equivalent amounts of corn oil. The study also investigated PCB treatment outcomes. At birth, pups were weighed and grossly examined for abnormalities. At 2 days of age, litters with fewer than 5 pups were excluded from the analysis and the remaining litters were culled to 4 males and 4 females. Each treatment group contained 10 or 11 litters. Pups remained with the dams until weaning. At weaning, 4−6 pups were retained for neurobehavioral tests (which were not reported as part of this study). The remaining offspring were sacrificed, which provided 5−9 litters per treatment group. Data were collected from one male and one female where possible. No signs of toxicity were evident in the dams; measurements on dams included gestational weight gain, liver weight, litter size, and live births. Pup birth weight and weaning

weight were unaffected by treatment. In pups sacrificed at weaning (21 days old), a significant ($p < 0.05$) decrease occurred in thymus weight for the high-dose group, but not in thyroid, liver, or brain weight. A significant ($p < 0.05$) decrease (20.4%) was observed in T4 in high-dose females. Thyroid stimulating hormone and T_3 were unaffected by treatment. Uridine diphosphate (UDP)-glucuronosyl transferase activity towards 4-nitrophenol significantly ($p < 0.05$) increased in both treatment groups over control values, and the increase in the high-dose group was significantly ($p < 0.05$) greater than in the low-dose group. Liver microsomal EROD activity was significantly ($p < 0.05$) increased in both treatment groups, but is considered to be an adaptive response and not adverse.

A LOAEL of 100 ng/kg-day for decreased thymus weights and decreased thyroxine is identified for this study. A NOAEL of 25 ng/kg-day is established.

2.4.2.2.15. *Simanainen et al, 2004.*

Simanainen et al. (2004) studied the impact of in utero and lactational TCDD exposure on the male reproductive system in three rat lines that are differentially sensitive to TCDD. Groups of 5 to 8 pregnant Line A, B, and C C57BL/6N CYP1A2 dams were given a single dose of 0, 30, 100, 300, or 1,000 ng/kg of TCDD (purity >99%) in corn oil on GD 15 via oral gavage. Control animals were similarly dosed with a corn oil vehicle. One day after birth, litters were randomly culled to include three males and three females to allow uniform postnatal exposure. Offspring were weaned on PND 28. Dam and pup viabilities were monitored throughout the study. Pup body weights were determined on PNDs 1, 4, 7, 14, and 28. Anogenital distance and crown-rump length were measured on PNDs 1 and 4. On day 70, pups were sacrificed and trunk blood was collected. Serum was collected for testosterone analysis. The testes, cauda of the right epididymis, ventral prostrate, seminal vesicles, and thymus was dissected and weighed. Absolute and relative organ weights were determined, and cauda epididymis and testes were also preserved for sperm count analysis.

TCDD caused no mortality or overt signs of toxicity to the dams. Pup survival from implantation to the day after birth also was not affected by TCDD exposure. Survival from the day of implantation to the day after birth, however, was uncharacteristically lower in control Line B rats (41%), resulting in a significant difference compared with the two lowest doses (30 and 100 ng/mg TCDD). The average survival percentage in the controls for Line A, B, and C

rats was 85% (range 80−86%); 64% (41−86%); and 74% (63−85%); respectively. Percentage of male pup survival in each line between PND 1 and PND 28 was 99% except for Line B males exposed to 30 ng/kg TCDD and Line C males exposed to 30 or 100 ng/kg, where male survival rate averaged 81% (range 81−83%). On PND 70, a significant ($p < 0.05$) reduction in body weight was observed only in Line B and C rats at 1,000 ng/kg. In pups exposed to 1,000 ng/kg TCDD, both absolute and relative weight of the ventral, anterior, and dorsolateral prostrate decreased in all three lines at most postnatal time points measured. The change was most consistent and significant ($p < 0.05$) in the ventral lobe. Animals exposed to 1,000 ng/kg TCDD had an average decrease in absolute weight of the anterior prostrate of 37, 32, and 34% in Lines A, B and C, respectively. Additionally, the average dorsolateral prostrate weight was also decreased by 34, 28, and 39% in Lines A, B, and C, respectively. The effect on the ventral prostrate was reversible with the only significant ($p < 0.05$) decrease in weight observed in Line B rats at PND 70 in the 1,000 ng/kg TCDD dose group. The authors reported that TCDD had no consistent effects on the weight of seminal vesicles. The absolute weights of the testis and epididymis showed a significant ($p < 0.05$) increase on PNDs 28−49, but the relative testis, epididymis, and cauda epididymis weights remained unchanged. In pups exposed to 1,000 ng/kg TCDD, severe malformation, including small caput and cauda and degeneration of corpus epididymis, was observed. Malformations in the epididymis were observed in 6 of 44 Line C male rat offspring and 3 of 47 Line A male rat offspring. In Line A, B, and C rats at PND 70in the 1,000 ng/kg TCDD dose group, daily sperm production was reduced by 9, 25, and 36% and cauda epididymal sperm reserves were reduced by 18, 42, and 49%, respectively. Daily sperm reduction (17%) was significant ($p < 0.05$) in Line C rats at a TCDD dose of 300 ng/kg and in Line B and C rats at 1,000 ng/kg. A reduction in cauda epididymal sperm reserves (25%) was significant ($p < 0.05$) in Line C rats at 300 and 1,000 ng/kg TCDD.

A LOAEL for TCDD of 300 ng/kg is identified for reduction in daily sperm production and cauda epididymal sperm reserves in Line C rats. A NOAEL of 100 ng/kg is identified for this study.

2.4.2.2.16. *Sugita-Konishi et al., 2003.*

Sugita-Konishi et al. (2003) examined the immunotoxic effects of lactational exposure to TCDD in newborn mice. Eight pregnant female C57BL/6NCji mice were administered 0, 1.8, or

18 ng/L of TCDD via drinking water from parturition to weaning of the offspring (for a total of 17 days). Based on an average water intake of 14–16 mL/day, the average daily intake of TCDD for the dams was 1.14 and 11.3 ng/kg-day in the low- and high-dose groups, respectively. In male offspring sacrificed at weaning (21 days after birth), there was a statistically-significant ($p < 0.05$) decrease in relative spleen weight and a statistically-significant ($p < 0.005$) increase in thymic CD4+ cells in the high-dose group. The changes in relative spleen weight and thymic CD4+ cells were dose related, but effects in the low-dose group did not achieve statistical significance. Changes in spleen weight and CD4+ cell numbers were not observed in the female offspring. In a separate experiment, offspring infected with *Listeria monocytogenes* following lactational TCDD exposure exhibited a statistically significant increase in serum tumor necrosis factor alpha (TNF-α) 2 days after infection in both sexes in the low- ($p < 0.05$) and high-dose ($p < 0.005$) groups. There was also a statistically significant increase in serum interferon gamma in *Listeria*-infected high-dose females ($p < 0.05$). The number of bacteria in the spleen was also significantly increased ($p < 0.05$) 2 days after infection in the high-dose females compared to the controls, but not in males. *Listeria* levels in the spleen returned to control levels by 4 days after infection in both sexes.

Based on these results, a LOAEL for TCDD of 11.3 ng/kg-day following a 17 day exposure to dams was identified for significantly ($p < 0.05$) decreased spleen weight (in male pups), a significant ($p < 0.005$) increase in thymic CD4+ cells (in male pups), and for increased susceptibility to *Listeria monocytogenes* (in male and female pups). The NOAEL for this study is 1.14 ng/kg-day.

2.4.2.3. *Acute Studies*

2.4.2.3.1. *Burleson et al., 1996.*

Burleson et al. (1996) studied the impact of TCDD exposure on mice that were challenged with the influenza virus 7 days after treatment with TCDD. Groups of 8-week-old female B6C3F1 mice ($n = 20$, 2 replicate groups) were treated one time with 0, 1, 5, 10, 50, 100, or 6,000 ng/kg TCDD (purity >99%, dissolved in corn oil) via oral gavage. In addition to the treated groups, randomly selected animals were assigned as a sentinel group and screened for numerous pathogens. Results of all tests performed on this sentinel group were negative. Seven days after TCDD treatment, all animals were lightly anesthetized and infected intranasally with a

highly lethal influenza A/Hong Kong/8/68 virus (H3N1; passage 14). The animals were infected with sufficient H3N1 virus to achieve a 30% mortality rate in the control animals. Animals were observed for mortality and morbidity for 21 days following viral infection. Six mice from each treatment group were sacrificed on days 3, 9, and 12 post-infection, and body, thymus, and wet lung weights were recorded. Influenza viral titers were examined by sacrificing eight mice each at 2 hours and at 1, 4, 6, 7, 8, 9, 10, and 11 days post infection.

Exposure to TCDD resulted in significantly ($p < 0.05$) increased mortality in the 10, 50, and 100 ng/kg dose groups. No statistically significant difference in the percentage alive was observed between these dose groups. TCDD doses of 1 and 5 ng/kg did not alter mortality in influenza infected animals. A time-related increase in the wet weights of the lungs in infected mice as a result of increased edema also was reflected in an increase in the lung weight-to-body weight ratio. The study authors stated that this ratio was not altered as a result of TCDD exposure. TCDD-only exposures at 1, 10, or 100 ng/kg did not affect thymus weight. Similarly, animals infected with the influenza virus following TCDD exposure also showed no loss in thymic weight. Enhanced mortality in TCDD-treated animals was not correlated with an increase in influenza virus titers. Additionally, animals treated with 1, 10, 100, or 1,000 ng/kg did not affect pulmonary viral titer assays on days 6, 7, and 8 post-infection. The authors also concluded that TCDD did not alter Hong Kong virus replication or clearance.

Although these results support immunotoxic effects induced by TCDD, the findings were not reproduced by Nohara et al. (2002) using the identical study design, and the translation of these findings to humans is dubious. Thus, no LOAEL/NOAEL was established. A LOEL for TCDD of 10 ng/kg for a single exposure is identified for significantly ($p < 0.05$) increased mortality in mice infected 7 days later with the influenza virus. The NOEL for this study is 5 ng/kg.

2.4.2.3.2. *Crofton et al., 2005.*

Crofton et al. (2005) studied the impact of TCDD exposure in addition to the impact of mixtures of thyroid disrupting chemicals and PCBs on serum total thyroxine (TT4) concentration. Groups of female Long-Evans rats were dosed via oral gavage with 0, 0.1, 3, 10, 30, 100, 300, 1,000, 3,000, or 10,000 ng/kg-day TCDD (purity >99%) in corn oil (n = 14, 6, 12, 6, 6, 6, 6, 6, 6, and 4, respectively) for 4 consecutive days. On the day following the last dose,

animals were sacrificed, trunk blood was collected, and serum obtained via centrifugation was assayed for TT4 concentration using standard radioimmunoassay methods.

No visible signs of toxicity or changes in animal body weight as a result of TCDD exposure were observed. Serum T4 levels showed a dose-dependent decrease, with the levels dropping sharply beginning at 100 ng/kg-day dose. Percent serum T4 levels were 96.3, 98.6, 99.8, 93.3, 70.9, 62.5, 52.7, 54.7, and 49.1% in the 0.1, 3, 10, 30, 100, 300, 1,000, 3,000, and 10,000 ng /kg-day groups, respectively.

A LOAEL for TCDD of 100 ng/kg-day for 4 consecutive days of exposure is identified in this study for a reduction in serum T4 levels (70.9% compared to 100% in controls). The NOAEL for this study is 30 ng/kg-day.

2.4.2.3.3. *Kitchin and Woods, 1979.*

Female Sprague-Dawley rats (nine per control and four per treatment group) were administered a single dose of 0, 0.6, 2, 4, 20, 60, 200, 600, 2,000, 5,000, or 20,000 ng/kg TCDD (purity >99%) in corn oil. Animals were sacrificed 3 days after treatment and CYP level and benzo(a)pyrene hydroxylase activity in the liver were measured. A significant ($p < 0.05$) increase in cytochrome P450 levels occurred with doses of 600 ng/kg or greater and in benzo(a)pyrene hydroxylase activity with doses of 2 ng/kg or greater. Cytochrome P450 was significantly ($p < 0.05$) higher 1 month after a single exposure of 2,000 ng/kg (the only dose measured), but not after 3 or 6 months. Aryl hydrocarbon hydralase (AHH; $p < 0.05$) and EROD ($p < 0.01$) were both significantly increased through 3 months after treatment, and although elevated at 6 months, the results were not significant.

CYP induction alone is not considered a significant toxicologically adverse effect given that CYPs are induced as a means of hepatic processing of xenobiotic agents. Thus, no LOAEL or NOAEL was established for this study because adverse endpoints (e.g., indicators of hepatotoxicity) were not measured. The acute LOEL, however, is 2 ng/kg based on a significant ($p < 0.05$) increase in benzo(a)pyrene hydroxylase activity (37% above control). The NOEL is 0.6 ng/kg.

2.4.2.3.4. *Li et al., 1997.*

Female Sprague-Dawley rats (22 days old; 10 per treatment) were administered a single oral dose of TCDD (>98% pure) in corn oil via gavage at doses of 3, 10, 30, 100, 300, 1,000, 3,000, 10,000, or 30,000 ng/kg. Vehicle controls received equivalent amounts of corn oil, while naïve controls were sham-treated only. In a preliminary time-course study, animals received a single dose of 10,000 ng/kg and were sacrificed at 1, 2, 4, 8, 16, 24, 48, and 72 hours. The time-course study showed two peaks in LH and FSH levels at 1 hour and 24 hours, with a decrease to control values by 48 hours. Thus, in the dose-response study, animals were sacrificed at 1 or 24 hours after treatment, blood was collected, and serum FSH and LH were measured. The dose-response study demonstrated that the peak at 1 hour was related to the vehicle as the peak also occurred in the vehicle controls, but did not occur in the naïve controls. At 24 hours, FSH was increased at 10 ng/kg and higher (>4-fold increase at 10 ng/kg). Doses of 10 to 1,000 ng/kg showed similar increases (not all reached statistical significance; $p < 0.05$). A dose-dependent increase occurred for doses ≥3000 ($p < 0.05$) with a maximum increase of 20-fold over the vehicle control. At 24 hours, the LH response significantly ($p < 0.05$) increased only for doses ≥300 ng/kg with a maximum increase of 15-fold above the vehicle control. The study authors calculated an ED_{50} of 500 ng/kg for gonadotropin increase. The dose-dependent release of LH was confirmed in in vitro studies, but did not occur with the same magnitude. The increase did not occur in calcium-free medium and was unrelated to gonadotropin releasing hormone.

Based on the increase in serum FSH, the LOAEL was 10 ng/kg and the NOAEL was 3 ng/kg.

2.4.2.3.5. *Lucier et al., 1986.*

Adult female Sprague-Dawley rats (six per treatment) were administered a single gavage dose of TCDD (purity not specified) in either corn oil or contaminated soil at doses of 15, 40, 100, 200, 500, 1,000, 2,000, 5,000 (corn oil), or 5,500 (contaminated soil) ng/kg. Animals were sacrificed 6 days later and livers were removed for analysis. No clinical signs of acute toxicity or changes in body weight were observed at any dose. AHH increased in a dose-dependent manner with significant ($p < 0.05$) increases observed at 15 ng/kg or greater in corn oil or 40 ng/kg or greater in contaminated soil. Cytochrome P450 was significantly ($p < 0.05$)

increased with doses of 1,000 ng/kg or greater in corn oil or 500 ng/kg or greater in contaminated soil. A dose-dependent increase was observed for UDP glucoronyltransferase (significance of individual doses not reported), with the results twice as high with corn oil than with contaminated soil. The authors state that the results indicate bioavailability from soils is 50%.

Because the association between AHH activity and TCDD-mediated hepatotoxicity is unknown and no adverse endpoints were measured, a LOAEL or NOAEL was not determined for this study. The acute LOEL for this study is 15 ng/kg, based on the significant ($p < 0.05$) increase (80% above control) in AHH. No NOEL is established.

2.4.2.3.6. *Nohara et al., 2002.*

Male and female B6C3F1 (C57BL/6 x C3H), BALB/c, C57BL/6N, and DBA2 mice (10−40 per treatment group) were administered a single dose of 0, 5, 20, 100, or 500 ng/kg TCDD in corn oil via gavage. Seven days following TCDD treatment, mice were infected with a mouse-adapted strain of influenza (A/PR/34/8; H1N1) at a plaque forming unit dose designed to target approximately 30% mortality in each strain. TCDD did not affect the body weight or survival in any of the infected mouse strains at any dose.

Therefore, no LOAEL is established in this study. The NOAEL is 500 ng/kg.

2.4.2.3.7. *Simanainen et al., 2003.*

Simanainen et al. (2003) studied the short-term effects of TCDD exposure to determine the efficacy and potency relationships among three differentially susceptible rat lines. The three rat lines used were A, B, and C, which were selectively bred from TCDD-resistant Han/Wistar and TCDD-sensitive Long-Evans rats. The study authors reported that Line A rats were most resistant to TCDD acute lethality followed by Line B and C. Groups of five or six randomly selected rats (sex not specified) were treated with a single oral dose of TCDD (purity >99%) in corn oil by oral gavage. The dose of TCDD was reported to range between 30 ng/kg and 3,000 µg/kg for Line A, 30 ng/kg and 1,000 µg/kg in Line B, and 30 ng/kg and 100 µg/kg for Line C. Control animals were similarly dosed with a corn oil vehicle. Rats were sacrificed on day 8 post-exposure, and trunk blood was collected and serum separated. Liver and thymus were removed and weighed, and liver samples were collected and preserved. Liver EROD activity,

serum aspartate aminotransferase (ASAT) activity, free fatty acid (FFA) concentration, and total bilirubin concentration were determined. Teeth were also examined.

Relative thymus weights were reduced 25% at 300 ng/kg relative to controls in Line B rats. Liver enzyme (CYP1A1) induction, as measured by EROD activity, was evident at all exposure levels; CYP induction is considered to be an adaptive effect and not adverse in itself. No other endpoints were affected below 1 µg/kg in any of the three rat lines.

A LOAEL for TCDD of 300 ng/kg is identified for decreased relative thymus weight in Line B rats. A NOAEL of 100 ng/kg is identified for this study.

2.4.2.3.8. *Simanainen et al., 2002.*

To study the short-term effects of TCDD on hormone levels, adult female Long-Evans (TCDD-sensitive) and Han/Wistar (TCDD-resistant) rats (n = 9−11/treatment) were administered a single dose of TCDD (>99% pure) in corn oil via gavage at doses ranging from 30 ng/kg to 100 µg/kg. Vehicle controls received an equivalent amount of corn oil. The study also examined other polychlorinated dibenzo-p-dioxins outcomes. Rats were sacrificed on day 8 post-exposure, and trunk blood was collected and serum separated. Liver and thymus were removed and weighed, and liver samples were collected and preserved. Liver EROD activity, serum ASAT activity, FFA concentration, and total bilirubin concentration were determined. Teeth were also examined.

Neither FFA or ASAT levels in Han/Wistar rats showed a dose-response relationship. In Long-Evans rats, however, a significant ($p < 0.05$) dose-dependent increase in FFA occurred at 300 ng/kg TCDD. Serum ASAT sharply increased in Long-Evans rats between 3,000 and 10,000 ng/kg. Body weight change and relative thymus weights were significantly decreased ($p < 0.05$) in Han/Wistar rats with doses ≥10,000 ng/kg and in Long-Evans rats with doses ≥1,000 ng/kg. Liver EROD activity was significantly ($p < 0.05$) increased with all doses in both strains. Serum T4 was significantly ($p < 0.05$) decreased in Long-Evans rats at concentrations ≥300 ng/kg, but were not significantly affected in Han/Wistar rats. Serum bilirubin was significantly ($p < 0.05$) increased with doses ≥10,000 ng/kg in Long-Evans rats and ≥30,000 ng/kg in Hans/Wistar rats. Both strains of rat showed a dose-dependent increase in mean severity of incisor tooth defects. The results indicate that TCDD was the most potent congener tested in both rat strains.

A LOAEL of 300 ng/kg for decreased T4 in the Long-Evans rat is identified for this study. A NOAEL of 100 ng/kg is is established.

2.4.2.3.9. *Smialowicz et al., 2004.*

Smialowicz et al. (2004) examined the impact of TCDD exposure on immunosuppression in mice. Groups of female (number not specified) C57BL/6N CYP1A2 (+/+) wild-type mice were administered a single dose of 0, 30, 100, 300, 1,000, 3,000, or 10,000 ng/kg TCDD (purity >99%) in corn oil via oral gavage. Control animals were similarly dosed with a corn oil vehicle. To assess immune function, 7 days after TCDD administration, all mice were immunized with sheep red blood cells (SRBCs) via injection into the lateral tail vein. Five days after immunization, mice were sacrificed, blood was collected, and enzyme-linked immunosorbant assays were performed. Additionally, spleen, thymus, and liver weights also were measured.

Body and spleen weights of the wild-type mice were unaffected by the TCDD exposure. A decrease in thymus weights of the mice appeared to be dose related. Only mice treated with 10,000 ng/kg TCDD, however, showed a statistically significant ($p < 0.05$) decrease in thymus weights compared to corresponding controls. Liver weights also showed a dose-related increase with only animals treated with 3,000 and 10,000 ng/kg TCDD showing statistical significance ($p < 0.05$) compared to the control group. The antibody response to SRBCs indicated a dose-related suppression in the wild-type mice, with animals treated with 1,000, 3,000, and 10,000 ng/kg TCDD showing statistically significant ($p < 0.05$) suppression compared to the controls.

A LOAEL for TCDD of 1,000 ng/kg is identified in female C57BL/6N CYP1A2 (+/+) wild-type mice for significant ($p < 0.05$) suppression of SRBCs. The NOAEL for this study is 300 ng/kg.

2.4.2.3.10. *Vanden Heuvel et al., 1994.*

Vanden Heuvel et al. (1994) examined the dose-response relationship between TCDD exposure and induction of hepatic mRNA. Groups of 10-week-old female Sprague-Dawley rats were administered TCDD (purity ~99%) in corn oil once at 0, 0.1, 0.05, 1, 10, 100, 1,000, or 10,000 ng/kg-BW. Four days after TCDD treatment, animals were sacrificed and livers were excised and preserved. Total hepatic RNA was extracted using guanidine thiocyanate and DNA

was removed using standard phenol-chloroform-isoamyl alcohol partitioning procedures. Quantitative competitive RNA-PCR method was used to analyze CYP1A1, UDP-glucuronosyltransferase I (UGT1), plasminogen activator inhibitor 2 (PAI2), β-actin, and transforming growth factor α (TGFα). In addition to hepatic mRNA levels, microsomal protein was assayed for EROD activity and livers were tested for TCDD concentration.

CYP1A1 mRNA induction levels in the TCDD-treated groups were low in the low-dose region and sharply increased to plateaus at higher doses. The lowest dose that showed a statistically significant ($p < 0.05$) difference compared to controls was the 1 ng/kg dose, which showed a three-fold increase in CYP1A1 mRNA levels. In contrast, a 130-fold increase occurred at 100 ng/kg and a 4,000- and 7,000-fold increase occurred at 1,000 and 10,000 ng/kg, respectively. A slight increase in the CYP1A1/β-actin levels was observed in the 0.1 ng/kg group, but this increase was not significant. EROD activity exhibited a pattern similar to CYP1A1 activity. EROD activity, however, was approximately 100-fold less sensitive compared to mRNA levels in TCDD-treated groups. Statistical significance (*p*-value not provided) in CYP1A1 level was observed at the 100 ng/kg dose compared to the 1 ng/kg dose. The study authors reported that, despite this difference in CYP1A1 and EROD activity, the correlation between CYP1A1 enzyme activity and mRNA levels was good. Dose-response relationships for the induction of UGT1, PAI2, and TGFα mRNA differed from what had been observed for CYP1A1 mRNA. UGT1 mRNA was induced, but at the much higher dose of 1,000 ng/kg. Additionally, the five-fold maximum induction of UGT1 mRNA was much less than the 7,000-fold induction observed for CYP1A1 mRNA at the 10,000 ng /kg dose. The authors state that this could be a result of the constitutive level of UGT1, which is much higher than CYP1A1, which makes detecting induction of UGT1 in the low dose regions more difficult. PAI2 and TGFα mRNA were not affected by TCDD in rat liver in the dose range tested. These results indicate that dioxin-inducible genes have a quite dissimilar dose-response relationship.

Induction of CYP1A1 expression is not considered an adverse effect, as the role of CYP1A1 in TCDD-mediated hepatotoxicity is unsettled. Therefore, in the absence of other indicators of hepatotoxicity, a NOAEL/LOAEL cannot be determined for this study. A LOEL for TCDD of 1 ng/kg for a single exposure was identified for statistically significant ($p < 0.05$) increase in CYP1A1 mRNA levels. The NOEL for this study is 0.1 ng/kg.

2.4.2.4. *Subchronic Studies*

2.4.2.4.1. *Chu et al., 2001.*

Adult female Sprague-Dawley rats (five per treatment group) were administered TCDD (purity >99%) in corn oil by gavage at doses of 0, 2.5, 25, 250, or 1,000 ng/kg-day for 28 days (Chu et al., 2001). The 1,000 ng/kg-day dose of TCDD caused a significant ($p \leq 0.05$) decrease in body weight gain (36% lower than the control), increase in relative liver weight (40% greater than the control), and decrease in relative thymus weight (50% lower than the control). There was a significant ($p \leq 0.05$) increase in EROD activity, methoxy resoufin-O-deethylase (MROD) activity, and UDP-glucuronosyl transferase (UDPGT) activity in the liver of female rats receiving 250 or 1,000 ng/kg-day TCDD. In addition, significant ($p \leq 0.05$) increases in serum cholesterol were observed in the 250 and 1,000 ng/kg-day dose groups, and liver ascorbic acid (AA) also was significantly increased in the 1,000 ng/kg-day dose group. There was ~1.5-fold increase in liver glutathione-*S*-transferase (GST), which was not statistically significant. Other significant ($p \leq 0.05$) findings for the 1,000 ng/kg-day group included a decrease in liver vitamin A (51% lower than the control), an increase in kidney vitamin A (15.5-fold increase above the control), an increase in liver benzyloxy resoufin-O-deethylase (BROD, 30-fold increase above control), a decrease in liver pentoxyresoufin-O-deethylase (PROD, 37% lower than the control), increase in serum albumin (18% above the control), and a decrease in mean corpuscular hemoglobin (MCH, 7% below the control) and mean corpuscular volume (MCV, 7% below the control).

Based on the numerous significant ($p \leq 0.05$) liver-related biochemical changes and significant ($p \leq 0.05$) increased relative liver weight, as well as significantly decreased body weight and relative thymus weight, the LOAEL for 28 days of exposure in this study is 1,000 ng/kg-day and the NOAEL is 250 ng/kg-day.

2.4.2.4.2. *Chu et al., 2007.*

Chu et al. (2007) examined the potential impact of TCDD on various organs and the toxicological impacts as a result of interactions between TCDD and PCBs in rats. Groups of female Sprague-Dawley rats ($n = 5$ per treatment group) were treated daily for 28 days via gavage with 0, 2.5, 25, 250, or 1,000 ng /kg-day TCDD (purity not specified) dissolved in corn oil. Body weights were determined three times per week, and clinical observations were made

daily. At study termination, all animals were sacrificed and blood was analyzed for various biochemical and hematological parameters. Liver, spleen, heart, thymus, brain, and kidneys were removed and weighed. A small portion of the liver was homogenized and assayed for BROD; EROD; MROD; and PROD. UDPGT, GST, and ascorbic acid levels also were measured. Vitamin A levels in the liver, kidney, and lungs were analyzed as free retinol (vitamin A), and histopathological analysis was conducted on various tissues.

Growth rate and thymic weights in rats treated with 1,000 ng/kg-day TCDD were significantly ($p \leq 0.05$) inhibited compared to the control group. Enzyme analysis indicated that measured levels of TCDD in the liver correlated with hepatic microsomal enzyme activity. The authors reported that liver microsomal EROD and MROD activities were significantly ($p < 0.05$ for EROD activity, significance level for MROD not reported) increased in the 250 and 1,000 ng/kg-day TCDD dose groups compared to the control group. UDPGT levels were significantly (significance level not reported) increased in the 250 and 1,000 ng/kg-day TCDD dose groups compared to the controls. Serum albumin levels were significantly ($p < 0.05$) increased in the 1,000 ng/kg-day TCDD dose group compared to the control group. Serum cholesterol levels were significantly (level not reported) increased compared to the control group at 250 ng/kg-day TCDD dose, while liver ascorbic acid concentrations were significantly (level not reported) increased in the 1,000 ng/kg-day dose group. Hematological analysis indicated that hemoglobin, packed cell volume, MCH, MCV, and platelet values were decreased in the 1,000 ng/kg-day TCDD dose group. Significant ($p \leq 0.05$) differences were observed only in MCH and MCV levels compared to the control. Vitamin A levels in the liver and kidney were significantly ($p < 0.05$) lower in the 1,000 ng/kg-day TCDD group compared to the control group. Histopathological evaluation of various tissues indicated that liver, thyroid, and thymus were the target organs. No TCDD-related affects were found in other tissues. A dose-dependent alteration in the thymus consisted of reduced thymic cortex and increased medullar volume with more animals exhibiting these changes at the 250 and 1,000 ng/kg-day dose level compared to the control group. Alterations in thyroid included reduced follicles, reduced colloid density, and increased epithelial height. A dose-dependent change in the thyroid was observed, with the highest impact evident in reduced follicles and reduced colloid density beginning at a dose of 25 ng/kg-day TCDD. Changes in liver were characterized by accentuated hepatic zones, anisokaryosis of hepatocytes, increased cytoplasmic density, and vacuolation. These changes

were also dose dependent, with more animals exhibiting these histopathological changes with increasing TCDD dose. Based on these results, the study authors concluded that exposure to TCDD resulted in a wide range of adverse effects with the thyroid proving to be most sensitive.

A LOAEL for TCDD of 25 ng/kg for a 28-day exposure is identified for alterations in thyroid, thymus, and liver histopathology. The NOAEL for this study is 2.5 ng/kg-day.

2.4.2.4.3. *DeCaprio et al., 1986.*

Hartley guinea pigs (10 per sex per dose) were administered TCDD (purity not specified) in the diet for 90 days at concentrations of 0, 2, 10, 76, or 430 ppt (equivalent to 0, 0.12, 0.61, 4.9, and 26 ng/kg-day in males and 0, 0.12, 0.68, 4.86, and 31 ng/kg-day in females calculated by the study authors using food consumption and body weights). Other animals were administered the high-dose diet (i.e., 430 ppt) for 11, 21, or 35 days and then administered the control diet (i.e., no exposure) for the remainder of the 90 days for recovery analysis. Four high-dose males died and two were sacrificed moribund by day 45; the remaining four animals were sacrificed on day 46 for necropsy. Four high-dose females also died and two were sacrificed moribund by day 55 with the remaining females sacrificed on day 60 for necropsy. Animals in the 76- and 430-ppt groups had significantly ($p < 0.05$) reduced body weights. Organ weights were not obtained in the 430-ppt group due to the early sacrifice, but in the 76-ppt group a significant decrease in relative thymus weight ($p < 0.05$) was observed, and relative liver ($p < 0.01$) and brain ($p < 0.05$) weights in males increased. Although a similar trend occurred in the females, the results were not statistically significant. Males administered 76 ppt in the diet also had a 53% increase in triglycerides ($p < 0.05$). The same increase was observed in females, but was not statistically significant. In the recovery groups, mortality during the recovery period after 11 or 21 days of treatment was 10% and after 35 days of treatment was 70%. Animals lost weight during the treatment period. Although the body weight increased during the recovery period, the body weight remained low compared to the control for the study duration.

The LOAEL from this study is 4.9 ng/kg-day for 90 days of exposure, based on decreased body weight (12−15%; $p < 0.05$) and changes in organ weights (10−30%, significant only in the males). The NOAEL is 0.61 ng/kg-day.

2.4.2.4.4. *Devito et al., 1994.*

Female B6C3F1 mice (5 per treatment) were administered 0, 1.5, 4.5, 15, 45, or 150 ng/kg TCDD (98% pure) in corn oil via gavage, 5 days a week for 13 weeks. This dose is equivalent to 0, 1.07, 3.21, 10.7, 32.1, 107 ng/kg-day (adjusted for continuous exposure, administered dose multiplied by 5 and divided by 7). Body weight was recorded weekly and animals were sacrificed 3 days after the last treatment. Examinations were performed on the lung, skin, uterus, and liver. No differences were observed in the liver or uterus weights or in the estrogen receptor levels in these two tissues. A dose-dependent increase in EROD activity (an indicator of CYP1A1 [CYP] induction) in the lung, skin, and liver was observed, with significant ($p < 0.05$) increases even at the lowest dose. The TCDD doses used did not achieve maximal EROD induction. A significant ($p < 0.05$) increase in liver acetanilide-4-hydroxylase (ACOH; an indicator of CYP1A2 induction) also was observed with all doses. A maximum induction of ACOH occurred with doses of 3.21 ng/kg-day and greater. A dose-dependent increase in specific phosphotyrosyl protein (pp) levels also was observed. Levels of pp34 and pp38 were significantly ($p < 0.05$) increased even at the lowest dose, while pp32 reached statistical significance ($p < 0.05$) with doses of 4.5 ng/kg-day and above.

The role of CYPs and phosphorylated pp32, pp34, and pp38 in TCDD-mediated toxicity is unknown, and changes in the activity or function of these proteins are not considered adverse Therefore, no LOAEL or NOAEL is established. The 13-week LOEL is 1.07 ng/kg-day, based on a significant ($p < 0.05$) increase in EROD, ACOH, pp34, and pp38 levels (all increased by at least 2-fold). No NOEL is established for this study.

2.4.2.4.5. *Fattore et al., 2000.*

Fattore et al. (2000) examined TCDD-induced reduction of hepatic vitamin A levels in a subchronic rat bioassay on Sprague-Dawley rats. Four experiments were conducted; Experiments 1, 2, and 3 were conducted in both male and female rats, while Experiment 4 was conducted only in female rats. The dosing regimens for each experiment were as follows

> ***Experiment 1***: Groups of six Iva:SIV 50 rats (male and female) were maintained on a diet consisting of 0, 200, 2,000, or 20,000 ng TCDD/kg diet and 3-µg vitamin A/kg diet for 13 weeks. Assuming food consumption of 10% of body weight per day, the average daily doses are 0, 20, 200, and 2,000 ng/kg-day TCDD.

Experiment 2: Groups of six male and female rats were treated with 0 or 200 ng TCDD/kg-day and 3 µg vitamin A/kg diet for 13 weeks.

Experiment 3: Groups of six male and female rats were fed 0, 200, or 1,000 ng TCDD/kg-day and 3 µg vitamin A/kg diet for 13 weeks.

Experiment 4: Groups of female rats (number not specified; IVA;SIV 50 Sprague-Dawley strain) were treated with TCDD for 26 and 39 weeks in addition to a 13-week dietary treatment with 0 or 100 ng TCDD/kg-day and 3 µg vitamin A/kg diet for 13 weeks.

For a 13-week exposure duration employed in all four experiments, male and female rats were treated at 0, 20, 100 (females only), 200, 1,000, or 2,000 ng/kg-day. In all four experiments, liver from control and treated animals was analyzed at termination for free retinol content to determine hepatic vitamin A levels.

***Results*:**

Experiment 1: Liver and body weights in both treated males and females were significantly affected at all but the lowest dose tested (20 ng/kg-day). Liver injury was severe, particularly in female rats treated with 2,000 ng TCDD/kg-day. Dietary intake of vitamin A in male rats was comparable to intake in controls, except in the 2,000 ng/kg-day group, which showed a reduction of 16% in the dietary intake of vitamin A compared to controls. There was no effect of TCDD on vitamin A intake in female rats. Hepatic vitamin A levels showed a dose-dependent reduction with levels dropping sharply in the 200 and 2,000 ng/kg-day dose groups, particularly in treated females. The reduction was significant at 200 ng/kg-day ($p < 0.05$) and 2,000 ng/kg-day ($p < 0.01$) in males, and at 200 ng/kg-day ($p < 0.5$) and 2,000 ng/kg-day ($p < 0.001$) in females. The reductions ranged from 68–99% in males and 72–99% in females when compared to corresponding controls.

Experiment 2: Changes in liver and body weights were not reported. Hepatic vitamin A level in males and females were reduced by 70% and 99%, respectively, compared to controls, in rats receiving 20 ng/kg-day (significance level in females: $p < 0.01$).

Experiment 3: Similar to the results of Experiments 1 and 2, a dose-related trend of significantly ($p < 0.001$) reduced hepatic vitamin A level was observed in both males and females, with males exhibiting a particularly sharp drop at the 1,000 ng/kg-day dose compared to controls.

Experiment 4: Females treated with 100 ng/kg-day showed significant reductions in hepatic vitamin A levels ($p < 0.05$–0.001) at all three treatment durations (13, 26, and 39 weeks).

A LOAEL for TCDD of 20 ng/kg-day for a 13-week subchronic exposure was identified in this study for decreased hepatic vitamin A levels (27 and 24 % lower than the corresponding

control in female and male rats, respectively). This LOAEL is determined using data from Experiment 1. A NOAEL was not identified in this study.

2.4.2.4.6. *Fox et al., 1993.*

Sprague-Dawley rats (6 per sex per dose) were gavaged with TCDD (purity not specified) in corn oil using a dose-loading regime to achieve and maintain steady-state levels of 0.03, 30, or 150 ng/g in the liver. The regime consisted of an initial loading dose of 5, 2,500, or 12,000 ng/kg followed every 4 days with a maintenance dose of 0.9, 600, or 3,500 ng/kg. Averaging the doses over the 14 days provides average daily doses of 0.55, 307, and 1,607 ng/kg-day (e.g., 5 ng/kg-day on day 1 and 0.9 ng/kg-day on days 5, 9, and 13 is 5 + 0.9 + 0.9 + 0.9/14 = 0.55 ng/kg-day). Body weight, liver weight, and liver gene expression were measured at 7 and 14 days. A significant ($p < 0.05$) decrease in body weight occurred in high-dose males (at 14 weeks only) and females (at 7 and 14 days). A significant ($p < 0.05$) increase in absolute and relative liver weights was observed in mid- and high-dose males and females at both 7 and 14 days. Although the liver of treated animals indicated moderate vacuolization and swelling, there was no indication of necrosis. An increase in gene expression (clone 1, CYP1A1, CYP1A2, and albumin) was observed in the mid- and high-dose groups. A significant ($p < 0.05$) decrease in labeling index (indication of cell proliferation) occurred in both females (all doses) and males (high-dose only) during week 1, but not during week 2.

The 14-day LOAEL is 307 ng/kg-day for significant ($p < 0.05$) increases in absolute and relative liver weights (25─34%). The NOAEL is 0.55 ng/kg-day.

2.4.2.4.7. *Hassoun et al., 1998.*

Female B6C3F1 mice (number not specified) received TCDD (>98% pure) in corn oil 5 days per week for 13 weeks via gavage at doses of 0, 0.45, 1.5, 15, or 150 ng/kg (equivalent to 0, 0.321, 1.07, 10.7, and 107 ng/kg-day adjusted for continuous exposure; administered dose multiplied by 5 and divided by 7). Three days after the final dose, animals were sacrificed and brains were removed for oxidative stress testing. Biomarkers for oxidative stress included production of superoxide anion, lipid peroxidation, and DNA single-strand breaks. A significant ($p < 0.05$) increase was observed in superoxide anion production, lipid peroxidation as measured

by thiobarbituric acid-reactive substances (TBARS), and DNA single-strand breaks with all doses tested.

No other indicators of brain pathology were assessed, and it is unfeasible to link the markers of oxidative stress to a TCDD-induced toxicological outcome in the brain. Thus, no LOAEL/NOAEL was established. The subchronic (13-week) LOEL is 0.32 ng/kg-day, based on significant ($p < 0.05$) increases in superoxide anion production (80% above control); lipid peroxide production (25% above the control); and DNA single-strand breaks (2-fold over the control). No NOEL is established.

2.4.2.4.8. *Hassoun et al., 2000.*

Hassoun et al. (2000) examined the effect of subchronic TCDD exposure on oxidative stress in hepatic and brain tissues. Groups of 8-week-old female Harlan Sprague-Dawley rats (6 rats/group) were administered TCDD (98% purity, dissolved in 1% acetone in corn oil) via gavage at 0, 3, 10, 22, 46, or 100 ng/kg-day, 5 days/week for 13 weeks (0, 2.14, 7.14, 15.7, 32.9, or 71.4 ng/kg-day adjusted for continuous exposure; administered doses were multiplied by 5 and divided by 7 days/week). Animals were sacrificed at the end of the study period, and brain and liver tissues were collected and used to determine the production of reactive oxygen species, lipid peroxidation, and DNA single-strand breaks (SSBs).

A dose-dependent effect was observed in both the liver and brain tissue as a result of TCDD treatment. Based on the maximal induction of superoxide anion by various doses, more production of superoxide anion was observed in the liver tissue when compared to the brain tissue with an observed increase of 3.1- and 2.2-fold respectively, when compared to the control group. A similar dose-dependent effect was observed in the induction of lipid peroxidation in TCDD-treated animals with an approximately 1.8-fold increase in lipid peroxidation in both tissues relative to the corresponding controls. A dose-dependent relationship was also observed for DNA SSBs in both the hepatic and brain tissues at all TCDD-treated doses compared to controls. Increases were statistically significant ($p \leq 0.05$) beginning at the lowest administered dose.

Similar to the statement above, because no adverse endpoints were measured, no LOAEL/NOAEL was established. However, a LOEL for TCDD of 2.14 ng/kg-day for a 13-week exposure duration was identified in this study for significant increases ($p \leq 0.05$) in

superoxide anion, lipid peroxidation, and DNA SSBs in the liver and brain tissues. A NOEL cannot be determined for this study.

2.4.2.4.9. *Hassoun et al., 2003.*

Hassoun et al. (2003) examined the role of antioxidant enzymes in TCDD-induced oxidative stress in various regions of the rat brain after subchronic exposure. Groups of 8-week-old female Harlan Sprague-Dawley rats (12 rats/group) were administered TCDD (98% purity, dissolved in 1% acetone in corn oil) via gavage at 0, 10, 22, or 46 ng/kg-day (0, 7.14, 15.7, or 32.9 ng/kg-day adjusted for continuous exposure; administered doses were multiplied by 5 and divided by 7) daily for 13 weeks. Animals were sacrificed at the end of the study period and the brain was immediately removed and dissected to the following regions: cerebral cortex (Cc), hippocampus (H), cerebellum (C), and brain stem including midbrain, pons, and medulla. Four pooled samples from each region per dose (i.e., 3 animals/pooled sample) were used in the study. Dissected regions were subsequently assayed for lipid peroxidation (thiobarbituric acid reactive substances, or TBARS), superoxide dismutase, catalase, and glutathione peroxidase. Because the cytochrome c reduction method was used to determine superoxide anion (SA) production in brain tissues, superoxide dismutase (SOD) was added to some of the brain tissue samples that had the highest SA production (tissue homogenates from Cc and H from rats treated with 46 ng/kg-day TCDD).

A dose-dependent increase in the production of SA was observed in the Cc and H, but significant changes in SA production were not observed in either the C or the mid-brain, pons, or medulla brain stem cells. Similar to SA production, there was a dose-dependent increase in the production of TBARS in the Cc and H regions of the brain, but no significant changes were observed in either the C or the B sections of the brain. The study authors also measured the activities of various enzymes as a result of TCDD treatment and reported a dose-dependent increase in SOD activity in the C and B sections, while there was dose-dependent suppression in SOD activity in Cc and H. In contrast, catalase activity was significantly ($p < 0.05$) increased in H and Cc at the 10 ng/kg-day TCDD dose level compared to controls and the mid- and high-dose animals. Catalase activity also was increased in a dose-dependent manner in the C section, but no significant changes in the activity of this enzyme were observed in the B section at any of the three TCDD tested doses. The effects of subchronic exposure to different doses of TCDD on

glutathione stimulating hormone peroxidase (GSH-Px) showed a different response compared to other enzymes. There was a dose-dependent increase in the activity of this enzyme in the C and B regions of the brain, while a significant increase in the activity of GSH-Px occurred in Cc and H only at the 10 ng/kg-day TCDD dose. In addition, the activity of this enzyme was suppressed in a dose-dependent manner in the Cc and H at 22 and 46 ng/kg-day TCDD doses. Based on these results, the study authors concluded that induction of oxidative stress by TCDD in the rat brain occurs mainly in the Cc and H regions.

Similar to the statement above, because no adverse endpoints were measured, no LOAEL/NOAEL was established. However, a LOEL for TCDD of 7.14 ng/kg-day for a 13-week exposure duration was identified for this study for increases in superoxide anion and lipid peroxidation production, as well as increased activity in SOD, catalase, and GSH-Px.

2.4.2.4.10. *Kociba et al., 1976.*

Adult Sprague-Dawley rats (12 per sex per treatment group) were administered TCDD (purity not reported) in corn oil via gavage 5 days per week at doses of 0, 1, 10, 100, or 1,000 ng/kg-day (equivalent to 0, 0.71, 7.14, 71.4, or 714 ng/kg-day averaged over 7 days; 5/7 of dose). Five animals per group were sacrificed at the end of treatment, and the remaining animals were observed over 13 weeks post treatment (only initial results for the post-treatment period were provided in the report). Body weights and food consumption were measured semi-weekly. Hematology and clinical chemistry were measured after 36−37 or 85−86 days of treatment and 59−60 days after termination of treatment. Forty-eight hour urine samples were collected from select rats from 85−89 days of treatment and 52−56 days after cessation of treatment. Gross and histopathological exams were conducted on the tissues.

Four high-dose females died during treatment. Two high-dose females and two high-dose males died during the post-treatment period. Animals treated with 714 ng/kg-day were less active during the treatment period, which became less evident during the post-treatment period. Yellow discoloration of the external pinnae also was noted in this group, both during treatment and during the post-treatment period. A significant ($p < 0.05$) reduction in body weight and food consumption was observed in the 71.4 and 714 ng/kg-day groups. The following significant ($p < 0.05$) hematology changes were observed in the high-dose (714 ng/kg-day) males at all measured time points: decreased packed cell volume, decreased red

blood cells, decreased hemoglobin, increased reticulocytes, and decreased thrombocytes. Significant ($p < 0.05$) changes also occurred in the high-dose females, but the only consistent observation was a decrease in thrombocytes and increased leukocytes. Significant changes in clinical chemistry ($p < 0.05$) and urinalysis ($p < 0.05$) were more consistent between the sexes in the high-dose group and included increases in total and direct serum bilirubin; increase in serum alkaline phosphatase; decreased urinary creatinine; and increased urinary coproporphyrin, uroporphyrin, and delta-amino-levulinic. The following significant ($p < 0.05$) changes were observed in the 71.4 ng/kg-day group: decreased packed cell volume (4−9%) in males; decreased red blood cells (2−10%) in males; decreased hemoglobin (2−13%) in males; increased urinary coproporyphin (2.2-fold increase during treatment) in females; increased urinary delta-amino-levulinic (47% increase during treatment) in females; increased total and direct serum bilirubin (48−61%) in females; and increased serum alkaline phosphatase (2-fold) in females. The following significant ($p < 0.05$) changes in relative organ weights were observed increased brain weight in 714 ng/kg-day males and females; increased liver weight in males (71.4 and 714 ng/kg-day) and females (7.14, 71.4, and 714 ng/kg-day); increased spleen weight in 714-ng/kg-day males and females; decreased thymus weight in 71.4 and 714 ng/kg males and females; and increased testes weight in 714 ng/kg-day males. Microscopic changes were observed in the thymus, and in other lymphoid tissues, and in the liver in rats treated with 71.4 ng/kg-day or greater.

The subchronic (13-week) LOAEL is 71.4 ng/kg-day, based on the numerous changes noted in body weight, hematology, clinical chemistry, urinalysis, and histopathology. The NOAEL is 7.14 ng/kg-day.

2.4.2.4.11. ***Mally and Chipman, 2002.***

Female F344 rats (3 per treatment group) were administered TCDD at concentrations of 0, 2.5, 25, or 250 ng/kg in corn oil via gavage for either 3 consecutive days or 2 days per week for 28 days (Mally and Chipman, 2002). The average daily doses for the 28-day study when adjusted for 7 days a week were 0, 0.71, 7.1, and 71 ng/kg-day (i.e., 2/7 of administered dose). No clinical signs of toxicity were observed. Histological examination of the liver revealed no abnormalities. All doses of TCDD reduced the number of connexin (Cx) 32 plaques and Cx32 plaque area in the liver, which was considered the target tissue. The reductions were not

statistically significant after the 3-day treatment, but were significant after the 28-day treatment ($p < 0.05$). TCDD also caused a reduction in the Cx32 plaque number and area in the thyroid after 28 days, but the results were not statistically significant. Although the reduction in Cx32 plaque number and plaque area in the liver and thyroid occurred at all dose levels, there was no relation to dose. TCDD did not induce hepatocyte proliferation.

In the absence of additional indicators of hepatotoxicity, changes in Cx32 plaques are not clearly linked to TCDD-mediated hepatotoxicity, nor are they considered an adverse effect. Additionally, no toxicologically-relevant endpoints were examined. Therefore, a NOAEL or LOAEL cannot be determined. A 28-day LOEL at the lowest dose of 0.71 ng/kg-day for significantly ($p < 0.05$) decreased Cx32 plaque area is evident (approximately 70% of the controls).

2.4.2.4.12. *Slezak et al., 2000.*

Slezak et al. (2000) studied the impact of subchronic TCDD exposure on oxidative stress in various organs of B6C3F1 female mice. Groups of 8- to 10-week-old female B6C3F1 mice (number not specified) were administered TCDD (purity >98%, dissolved in corn oil) via gavage at 0, 0.15, 0.45, 1.5, 15, or 150 ng/kg-day (0, 0.11, 0.32, 1.07, 10.7, or 107.14 ng/kg-day adjusted for continuous exposure) 5 days per week for 13 weeks. Three days after the last treatment, the animals were sacrificed and organs were removed for the measurement of oxidative stress indicators including SA, lipid peroxidation (TBARS), and GSH-Px. Tissue TCDD concentrations also were measured.

The study authors reported that TCDD dose range resulted in overlapping tissue concentrations for liver, lung, kidney and spleen. Liver had the highest TCDD concentration, with each tissue demonstrating a dose-dependent increase in TCDD concentration. Compared to controls, SA production was significantly ($p < 0.05$) lower at the 0.15 ng/kg-day TCDD dose, while it was significantly ($p < 0.05$) higher at 15 and 150 ng/kg-day. A dose-dependent increase in hepatic TBARS production was observed, although the rate of production was significant ($p < 0.05$) only at the highest TCDD administered dose (150 ng/kg-day) compared to controls. AA also followed the same pattern observed for SA and TBARS with AA production significantly ($p < 0.05$) increased at the 15 and 150 ng/kg-day TCDD doses. Contrary to the SA, TBARS, and AA responses, GSH levels were decreased at 0.15 ng/kg-day, were increased at

0.45 and 150 ng/kg-day, and did not change at 1.5 or 15 ng/kg-day when compared to the control group. Unlike the liver, there was no significant increase in SA production in the lung at any of the TCDD tested doses; a dose dependent reduction, however, was observed at 0.45, 15, and 150 ng/kg-day compared to controls. GSH and AA production was decreased at 0.15 ng/kg-day, while AA production was significantly ($p < 0.05$) increased at 15 and 150 ng/kg-day. Kidney SA production showed a statistically significant ($p < 0.05$) increase only at the 15 and 150 ng/kg-day doses. GSH, like the liver and the lung, exhibited a decrease in production following treatment at 0.15 ng/kg-day with this trend continuing at 0.45 and 1.5 ng/kg-day. AA levels were significantly ($p < 0.05$) lower at all subchronic doses, except at 1.5 ng/kg-day dose. SA levels in the spleen differed little from the control group at any of the TCDD doses. Total GSH was higher only at the 150 ng/kg-day dose level, while the AA levels were significantly ($p < 0.05$) decreased at 0.15, 1.5, and 150 ng/kg-day.

Similar to the statements regarding the Hassoun et al. studies above, because no adverse endpoints were measured, no LOAEL/NOAEL was established. Therefore, a NOAEL or LOAEL cannot be determined. However, a NOEL and LOEL of 1.07 and 10.7 ng/kg-day, respectively, are identified in this study for increases in superoxide anion in the liver.

2.4.2.4.13. *Smialowicz et al., 2008.*

Female B6C3F1 mice (8–15 per treatment group) were administered TCDD (purity >98%) in corn oil by gavage at doses of 0, 1.5, 15, 150, or 450 ng/kg-day, 5 days a week for 13 weeks (1.07, 10.7, 107, or 321 ng/kg-day, adjusted for continuous exposure; i.e., 5/7 of the dose) (Smialowicz et al., 2008). Mice were immunized 3 days after the final TCDD exposure with an intravenous injection of an optimal concentration of 4×10^7 SRBCs and sacrificed 4 days later. No TCDD-related effects on body weight were observed. There was a dose-related decrease in relative spleen weight (9–19% lower than control values) with statistically significant ($p < 0.05$) decreases at all but the lowest dose. Additionally, there was a statistically significant ($p < 0.05$) increase in relative liver weight (5–21%) in all treatment groups compared to controls. Statistically significant dose-dependent decreases were observed in the antibody response to SRBCs (24–89% lower than control values), as measured by both the number of plaque forming cells per 10^6 cells and plaque forming cells per spleen.

The 13-week LOAEL for this study is 1.07 ng/kg-day based on a significant ($p < 0.05$) increase in relative liver weight (10%) and a significant ($p < 0.05$) decrease in antibody response to SRBCs (24%). A NOAEL cannot be determined for this study.

2.4.2.4.14. *Van Birgelen et al. 1995a, b.*

Van Birgelen et al. (1995) studied the impact of TCDD exposure on various biochemical endpoints in rats. Groups of 7-week-old female Sprague-Dawley rats ($n = 8$ per treatment group) were treated with 0, 200, 400, 700, 5,000, or 20,000 ng/kg TCDD (purity >99%) in diet for 13 weeks. Daily TCDD intake based on food consumption, diet level, and mean weight was estimated to be 0, 14, 26, 47, 320, or 1,024 ng/kg-day. Blood samples were collected from treated animals and assayed for retinol (vitamin A), triiodothyronine, and total (TT4) and free (FT4) thyroxine. At study termination, the animals were sacrificed and the liver, thymus, spleen, and kidneys were removed and weighed. Parts of the liver were homogenized and assayed to determine EROD; CYP1A1; CYP1A2; and UDPGT activity. Liver samples also were analyzed for retinol content.

TCDD-treated animals showed a dose-related decrease in food consumption. Animals treated with 1,024 ng/kg-day TCDD consumed 32% less food compared to controls. Similarly, a dose-related decrease in body weight gain was observed in all animals treated with TCDD. Animals treated with ≥47 ng/kg-day of TCDD showed a statistically significant ($p < 0.05$) decrease in body weight gain. Relative liver weights were significantly ($p < 0.05$) increased in the 320 and 1,024 ng/kg-day TCDD dose groups compared to the controls. Absolute and relative thymus weights were significantly ($p < 0.05$) decreased at all TCDD dose groups compared to the control group. Relative kidney and spleen weights were significantly ($p < 0.05$) higher in animals dosed with ≥47 ng/kg-day of TCDD compared to the control group, with the greatest increase occurring in animals treated with 1,024 ng/kg-day TCDD (121 and 173% higher than controls for kidney and spleen, respectively). Cytochrome P450 enzymes, including EROD, CYP1A2, CYP1A1, and UDPGT, exhibited statistically significant ($p < 0.05$) increases in activity at all TCDD dose groups compared to the control group. TT4 and FT4 thyroid hormone concentrations were statistically significantly ($p < 0.05$) decreased only at TCDD doses ≥47 ng/kg-day. A dose-dependent increase was observed in the plasma retinol concentrations with significant ($p < 0.05$) increases occurring at ≥47 ng/kg-day TCDD after a 13-week

exposure. A dose-dependent reduction in liver retinoid levels also was observed after 13 weeks of TCDD exposure with the levels dropping significantly ($p < 0.05$) at all TCDD-treated doses compared to the control group.

A LOAEL for TCDD of 14 ng/kg for a 13-week exposure is identified for significantly ($p < 0.05$) decreased absolute and relative thymus weights and significantly ($p < 0.05$) decreased liver retinoid levels. A NOAEL cannot be determined for this study.

2.4.2.4.15. *Vos et al., 1973.*

Vos et al. (1973) conducted a study to examine the immune response in laboratory animals treated with TCDD. In one experiment, 10 female Hartley strain guinea pigs were orally treated with 8 weekly doses of 0, 8, 40, 200, and 1,000 ng/kg TCDD in corn oil (purity of TCDD not specified) (0, 1.14, 5.71, 28.6, and 143 ng/kg-day adjusted for continuous exposure; administered dose divided by 7). At study termination, the animals were sacrificed, and heart blood was used to determine total leukocyte and differential leukocyte counts. In another experiment, the effect of TCDD on humoral immunity was determined by injecting 0.1 mL of tetanus toxoid into the right hind-foot pad on day 28 (1 left foot tetanus toxoid, aluminum phosphate-adsorbed) and again on day 42 (1 left foot tetanus toxoid, unadsorbed). Blood was collected ($n = 10$) on days 35 and 49, and the serum tetanus-antitoxin concentrations were determined using a modified single radial immunodiffusion technique.

All guinea pigs receiving 1,000 ng/kg-day TCDD either died or were killed when moribund between 24 and 32 days. These animals showed severe weight loss, lymphopenia, and depletion of the lymphoid organs, especially the thymus. Microscopic observations revealed severe atrophy of the thymic cortex with substantial destruction of lymphocytes, with the nuclear debris being engulfed by macrophages. Large cystic Hassall bodies, filled with polymorphonuclear leukocytes were observed in the medulla. All animals treated with 0, 8, 40, or 200 ng/kg-day TCDD survived until study termination. Body weight gain was significantly ($p < 0.01$) lower in the 200 ng/kg-day group. Absolute thymus weight was significantly reduced in the 40 and 200 ng/kg-day treatment groups ($p < 0.01$ and $p < 0.05$, respectively). In contrast, relative thymus weight was significantly ($p < 0.01$) reduced only in the 200 ng/kg-day dose group. The absolute weight of the superficial cervical lymph nodes was significantly ($p < 0.05$) decreased in the 200 ng/kg-day group, while the relative adrenal weight was significantly

($p < 0.05$) increased in the 200 ng/kg-day dose group. Total leukocyte count was significantly ($p < 0.05$) decreased in the 40 ng/kg-day dose group and total lymphocyte count was significantly decreased at 8, 40, and 200 ng/kg-day ($p < 0.01$, $p < 0.05$, and $p < 0.05$, respectively). A significant (p-values not provided) monotonic dose-response relationship was determined for body weight (decrease), relative thymus weight (decrease), relative adrenal weight (increase), and total leukocyte and lymphocyte count (decrease). Microscopic examination of the lymphoid organs and adrenals showed no effects, while slight cortical atrophy of the thymus was observed at the 200 ng/kg-day dose.

Animals receiving the tetanus toxoid injection showed a small but significant increase in serum tetanus antitoxin concentrations at the 8 and 40 ng/kg-day dose ($p < 0.05$ and $p < 0.01$, respectively). Measurement at days 49 and 56 indicated that serum antitoxin levels had decreased sharply and the significant ($p < 0.05$ on day 49 and $p < 0.01$ on day 56) effect was seen only at the 200 ng/kg-day dose level.

A LOAEL for TCDD of 5.71 ng/kg-day for an 8-week exposure is identified in this study for significantly ($p < 0.01$) reduced absolute thymus weight, significantly ($p < 0.05$) reduced leukocyte and lymphocyte count, and significantly ($p < 0.01$) increased serum tetanus antitoxin concentration. The NOAEL for this study is 1.14 ng/kg-day.

2.4.2.4.16. *White et al., 1986.*

White et al. (1986) studied the impact of TCDD exposure on serum complement levels. Groups of female (C57BL/6 × C3H)F1(B6C3F1) mice were treated for 14 consecutive days with TCDD in corn oil (purity of TCDD not specified) at doses of 0, 10, 50, 100, 500, 1,000 or 2,000 ng/kg-day via gastric intubation ($n = 6-8$). At study termination, blood was collected from anesthetized animals and assayed for serum complement activity and complement component C3 levels.

Serum complement activity between the 10 and 100 ng/kg-day doses was between 69 and 59% compared to the vehicle control group, with all treatment groups being significantly ($p < 0.05$) low compared to the vehicle control. In contrast, C3 levels were comparable to the vehicle control with levels ranging between 98 and 94% of the control group. The higher doses of 500, 1,000, and 2,000 ng/kg-day, however, produced a marked decrease of the component

hemolytic activity (45, 35, and 19% of the vehicle control) and of C3 levels (91, 81, and 74 % of the vehicle control, respectively; significance level at $p < 0.05$).

A LOAEL for TCDD of 10 ng/kg-day for a 14-day exposure is identified in this study for significantly ($p < 0.05$) lower serum complement activity. A NOAEL cannot be determined for this study.

2.4.2.5. *Chronic Studies (Noncancer Endpoints)*

2.4.2.5.1. *Cantoni et al., 1981.*

CD-COBS rats (4 per treatment) were orally administered TCDD (purity not specified) dissolved in acetone:corn oil (1:6) at doses of 0 (vehicle alone), 10, 100, or 1,000 ng/kg per week (equivalent to 1.43, 14.3, and 143 ng/kg-day adjusted for continuous exposure, administered dose by dividing the dose by 7) for 45 weeks. Urine was collected several times during treatment and tested for porphyrin excretion. Twenty-four hours after the final dose, animals were sacrificed and their livers, spleens, and kidneys were removed for analysis of total porphyrins. All treatment groups had a significant ($p < 0.05$) increase in coproporphyrin excretion beginning at 6, 3, or 2 months, respectively. Uroporphyrin excretion was significantly ($p < 0.05$) increased in the 14.3 ng/kg-day group at 10 months and in the 143 ng/kg-day group beginning at 6 months. The high-dose group also had a significant ($p < 0.05$) increase in excretion of heptacarboxylic methyl ester beginning at 6 months. The high-dose group had a marked porphyric state beginning at 8 months as indicated by a 70-fold increase above controls in total urinary porphyrin excretion. This group also had a significant ($p < 0.05$) increase in total porphyrins in the liver, kidneys, and spleen.

The 45-week LOAEL for this study is 1.43 ng/kg-day, based on a 2- to 3-fold increase in urinary coproporphyrin excretion. No NOAEL was established for this study.

2.4.2.5.2. *Croutch et al., 2005.*

Croutch et al. (2005) examined the impact of TCDD exposure on body weight via insulin-like growth factor (IGF) signaling. Female Sprague-Dawley rats were randomly assigned in groups of five to initial loading doses of TCDD (purity >98.5%, dissolved in corn oil) at 0, 12.5, 50, 200, 800, or 3,200 ng/kg-day, followed by treatment with maintenance doses equivalent to 10% of the initial loading dose every third day to maintain a pharmacokinetic steady state

throughout the entire study (equivalent to: 14-day average = 0, 1.25, 5, 20, 80, or 320 ng/kg-day; 28-day average = 0, 0.85, 3.4, 13.6, 54.3, or 217 ng/kg-day; 63-day average = 0, 0.60, 2.4, 9.5, 38, or 152 ng/kg-day; and 128-day average dose = 0, 0.51, 2.0, 8.1, 32.5, or 130 ng/kg-day). Following 2, 4, 8, 16, 32, 64, or 128 days of initial dosing, the animals were sacrificed, livers were removed and weighed, and trunk blood was collected to analyze glucose content. Rat liver phosphoenolpyruvate carboxykinase (PEPCK) mRNA and protein levels also were analyzed, and PEPCK activity was measured.

Body weights of TCDD-treated animals decreased after the second week of the 3,200 ng/kg-day TCDD loading dose, with significant differences beginning at week 9. There was also a statistically significant ($p \leq 0.05$) difference in body weights at weeks 10, 11, 13, 18, and 19 at the highest loading dose (3,200 ng/kg-day). PEPCK activity in the liver was also decreased in a dose-dependent manner following TCDD administration at approximately 16 days. PEPCK inhibition was statistically significant ($p \leq 0.05$) on day 4 in rats treated with either 800 or 3,200 ng/kg-day TCDD when compared to animals treated with a loading dose of 200 ng/kg-day. A similar statistically significant change was observed in animals treated with 3,200 ng/kg-day on day 16 when compared to the 200 ng/kg-day treatment group. In contrast, differences in PEPCK activity at other doses or time points were not statistically significant. In TCDD-treated animals, there was also a dose-dependent decrease in PEPCK mRNA expression along with a decrease in PEPCK protein levels in the liver. In addition to body weight and PEPCK activity changes, animals treated with 3,200 ng/kg-day TCDD showed a sharp decline in circulating IGF-I levels on day 8 compared to the control group (corn oil) and TCDD-treated animals at lower doses. In the highest dose animals, IGF-I levels continued to decline to 42% of the control group by day 16 of the study. The IGF-I levels at the highest dose plateaued at an average decrease of 66% through day 128 when compared to controls. Beginning at day 8, the decrease in IGF-I was statistically significant at every time point through day 128 compared to the control group, as well as groups treated with either 12.5 or 50 ng/kg-day TCDD. Similar statistically significant decreases also were observed for the 800 ng/kg-day TCDD-treated groups with an initial decrease of 37% on day 16 followed by a further decline to approximately 45% thereafter compared to controls and the 12.5, 50, and 200 ng/kg-day dose groups. In contrast to these results, circulating levels of insulin and glucose were unaffected by TCDD treatment, while

the active or phosphorylated form of AMPK-α protein increased with dose as a result of TCDD treatment.

A LOAEL for TCDD of 217 ng/kg-day for a 28-day exposure duration (because this represented the most sensitive time for elicitation of effects) was identified in this study for decreased body weight, significant ($p \leq 0.05$) inhibition of PEPCK activity, and reduced IGF-I levels (42% lower than the control group). A NOAEL of 54.3 ng/kg-day was identified in this study.

2.4.2.5.3. *Hassoun et al., 2002.*

Hassoun et al. (2002) examined the potential of TCDD and other dioxin-like chemicals to induce oxidative stress in a chronic rat bioassay. Groups of six Harlan Sprague-Dawley female rats were treated with 0, 3, 10, 22, 46, or 100 ng/kg-day TCDD (98% purity), 5 days a week via gavage for 30 weeks. The administered doses adjusted for continuous exposure were 0, 2.14, 7.14, 15.7, 32.9, and 71.4 ng/kg-day, respectively (administered doses were multiplied by 5 and divided by 7). At study termination, hepatic and brain tissues from all treated rats were divided into two portions and examined for the production of reactive oxygen species and SSBs in DNA.

When compared to controls, there was a dose-dependent increase in the production of superoxide anion in TCDD-treated animals ranging from 21–998% and 66–257% in hepatic and brain tissues, respectively. Hepatic tissues had statistically significant ($p < 0.05$) increases in superoxide anion production at doses ≥7.14 ng/kg-day, while the brain tissue had a statistically significant ($p < 0.05$) increase over controls at all doses. Similarly, increases in lipid peroxidation were observed in hepatic and brain tissues with a 481% increase ($p < 0.05$) at 71.4 ng/kg-day in the hepatic tissue when compared to controls. The increase in lipid oxidation in brain tissue ranged from 33–188% ($p < 0.05$) in the 2.14–71.4 ng/kg-day dose groups. DNA SSBs were also observed in both hepatic and brain tissue in all treated groups. When compared to the control group, there was a dose-dependent statistically significant ($p < 0.05$) increase in DNA SSBs ranging from 58–322% and 29–137% in hepatic and brain tissues, respectively. Nonmonotonic dose-response relationships were observed for superoxide production and lipid peroxidation in liver tissues, with greater-than-linear increases in effect between the two highest dose levels.

As stated above, because no adverse endpoints were measured, no LOAEL/NOAEL was established. However, a LOEL for TCDD of 2.14 ng/kg-day for a 30-week exposure duration is identified in this study for significant ($p < 0.05$) increases in superoxide anion, lipid peroxidation production, and DNA SSBs in the liver and brain tissues. A NOEL cannot be determined for this study.

2.4.2.5.4. *Kociba et al., 1978.*

Sprague-Dawley rats (50 per sex per treatment group) were administered TCDD (purity >99%) in the diet at doses of 0, 1, 10, or 100 ng/kg-day for 2 years. Body weights and food consumption were routinely measured. Hematology, clinical chemistry, and urinalysis were measured after 3, 12, or 23 months of treatment. Animals were routinely palpitated for tumors. Gross and histopathological exams were conducted on the tissues of dead or dying animals or at terminal sacrifice. Specific organs also were weighed.

The high-dose females had a statistically significant ($p < 0.05$) increase in mortality compared to the controls during the second half of the study. Mortality changes in males were variable and of questionable toxicological significance. A significant ($p < 0.05$) reduction in body weight occurred in the 100 ng/kg-day males and females beginning at 6 months. Mid-dose females also had reduced body weight, but to a lesser degree during the same time frame. There were no consistent changes in food consumption. The following significant ($p < 0.05$) hematology changes were observed in the high-dose animals: decreased packed cell volume in males after 3 months and in females after 1 year, decreased red blood cells in females after 1 year and in males at terminal sacrifice, decreased hemoglobin in males after 3 months and in females after 1 year, and decreased total white blood cell count in females after 1 year. Changes in clinical chemistry ($p < 0.05$) occurred only in high-dose females and consisted of an increase in serum alkaline phosphatase and gamma glutamyl transferase. Significant changes in urinalysis occurred only in females and included increased urinary coproporphyrin in the mid- and high-dose groups, increased urinary uroporphyrin in the mid- and high-dose groups, and increased urinary delta-amino-levulinic acid in the high-dose group. Significant ($p < 0.05$) changes in relative organ weights were observed, including increased liver weight in mid- and high-dose females and decreased thymus weight in high-dose females. Mid- and high-dose rats showed hepatocellular degeneration and inflammatory and necrotic changes in the liver. Thymic

and splenic atrophy were noted in high-dose females. An increase in non-neoplastic lung lesions was noted in mid-dose females and high-dose males and females. High-dose females had an increase in uterine changes. High-dose males had a significant ($p < 0.05$) increase in the incidence of stratified squamous cell carcinomas of the tongue. High-dose males and females had a significant ($p < 0.05$) increase in the incidence of squamous cell carcinomas of the hard palate/turbinates.

The chronic (2-year) LOAEL is 10 ng/kg-day, based on the numerous significant ($p < 0.05$) changes noted in coproporphyrin excretion (67% increase above control) and an increase in liver and lung lesions in female rats. The NOAEL is 1 ng/kg-day.

2.4.2.5.5. *Maronpot et al., 1993.*

An initiation-promotion study was performed in female Sprague-Dawley rats (8−10 rats per group). Rats were initiated with saline or diethylnitrosamine (DEN), followed 2 weeks later by promotion with biweekly administration of TCDD (purity not specified) in corn oil via gavage for 30 weeks. The doses were stated to be equivalent to 3.5, 10.7, 35.7, or 125 ng/kg-day. Rats were sacrificed 7 days after the final treatment. A significant ($p < 0.05$) decrease in body weight occurred in the 125 ng/kg-day group. A significant ($p < 0.05$) increase in relative liver weight occurred in the 35.7 and 125 ng/kg-day groups. There was a significant ($p < 0.05$) increase in the labeling index in the 125 ng/kg-day group, but only with DEN initiation. In the TCDD-alone group, a 2-fold increase in labeling index occurred in the 125 ng/kg-day group that did not reach statistical significance. A significant ($p < 0.05$) trend for increased alkaline phosphatase levels was observed in TCDD-treated animals, but despite a 50% increase in the highest dose group the increase was not statistically significant. Total cholesterol and triglycerides were significantly ($p < 0.05$) higher in the 125 ng/kg-day TCDD-alone group. A significant ($p < 0.05$) increase in 5'-nucleotidase occurred in the 35.7 and 125 ng/kg-day TCDD-alone groups. A dose-dependent increase in the incidence and severity of liver toxicity as measured by microscopic lesions was observed.

The 30-week LOAEL is 35.7 ng/kg-day, based on a significant ($p < 0.05$) increase in relative liver weight (12%, accompanied by increases in incidence and severity of liver lesions). The 30-week NOAEL is 10.7 ng/kg-day.

2.4.2.5.6. *National Toxicology Program, 1982.*

National Toxicology Program (NTP, 1982) conducted a carcinogenic bioassay of TCDD on rats and mice. Fifty male and female Osborne-Mendel rats and male and female B6C3F1 mice were treated twice per week with TCDD (purity not specified) in corn oil via oral gavage at doses of 0, 5, 25, or 250 ng/kg for rats and male mice (1.4, 7.1, 71 ng/kg-day adjusted for continuous exposure; administered doses multiplied by 2 and divided by 7) and 0, 20, 100, or 1,000 ng/kg for female mice (5.7, 28.6, or 286 ng/kg-day adjusted for continuous dosing; administered doses multiplied by 2 and divided by 7) for 104 weeks. Seventy-five rats and mice of each sex served as vehicle controls. One untreated control group of 25 rats and mice of each sex was present in the TCDD treatment room and one untreated control group consisting of 25 rats and mice of each sex were present in the vehicle-control room. Animals surviving until study termination were sacrificed at 105 or 108 weeks. A complete histopathological evaluation was conducted on all animals.

Survival rates were not affected by TCDD exposure in rats or mice of either sex. Male rats exhibited a dose-related depression in mean body weight after week 55, while the females exhibited a dose-related body-weight depression after 45 weeks of TCDD exposure. However, the magnitude of the body weight response is not indicated. Mean body weights in male and female mice were comparable to the vehicle control group throughout the bioassay. Noncancer histopathologic findings included increased incidences of liver lesions (termed toxic hepatitis) from TCDD exposure, and were detected in the high-dose rats and high-dose mice of each sex.

A LOAEL for TCDD of 1.4 ng/kg-day for a 104-week exposure duration is identified for increased incidences of liver lesions in mice of both sexes. A NOAEL cannot be determined for this study.

2.4.2.5.7. *National Toxicology Program, 2006.*

Female Sprague-Dawley rats (81 control; 82 treatment group) were administered TCDD (purity >98%) in corn oil:acetone (99:1) via gavage at doses of 0, 3, 10, 22, 46, or 100 ng/kg-day, 5 days per week for 105 weeks (0, 2.14, 7.14, 15.7, 32.9, or 71.4 ng/kg-day, adjusted for continuous exposure) (NTP, 2006). In addition to this primary group, a stop group of 50 animals was administered 100 ng/kg-day TCDD in corn oil:acetone (99:1) via gavage for 30 weeks and then just the vehicle for the remainder of the study. Up to 10 rats per dose group

from the primary study were sacrificed and evaluated at 14, 31, or 53 (n = 8) weeks for biologically noteworthy changes in the incidences of neoplasms or non-neoplastic lesions in the liver, lung, oral mucosa, uterus, pancreas, thymus, adrenal cortex, heart, clitoral gland, ovary, kidney, forestomach, bone marrow, mesentery gland, and pituitary gland. All animals also received a complete necropsy and microscopic examination, and the following organs were weighed: the left kidney, liver, lung, left ovary, spleen, thymus (14 weeks only), and thyroid gland. At study termination, the number of surviving animals had declined to 25 in the control group and to 21, 23, 19, 22, and 21 in five treatment groups, respectively, due to accidental deaths, moribund animals, or death due to natural causes.

Survival rate was not affected by TCDD treatment. Mean body weights in the high dose primary study group and the 100 ng/kg stop group were less than the vehicle control group after week 13 of the study. The mean body weights of animals in the 46 ng/kg-day group were less than in the vehicle control at study termination (2 years), whereas animals in the 22 ng/kg-day had lower mean body weights compared to controls during the last 10 weeks of study. In addition to body weight changes, liver weights were also impacted as a result of TCDD exposure. Absolute and relative liver weights were significantly (either $p \leq 0.01$ or $p \leq 0.05$) higher in all dose groups compared to controls at the 14- and 31-week evaluation period, whereas the relative liver weights were significantly (either $p \leq 0.01$ or $p \leq 0.05$) higher only at ≥10 ng/kg-day at 53 weeks.

No clinical findings associated with TCDD treatment were observed. TCDD caused changes in thyroid hormone levels at 14, 31, and 53 weeks. The following changes were statistically significant ($p \leq 0.05$) compared to the vehicle control: decrease in TT4 at doses ≥22 ng/kg-day at 14 and 31 weeks and at doses ≥46 ng/kg-day at 53 weeks; decrease in FT4 at doses ≥22 ng/kg-day at 14 and 31 weeks; increase in total T_3 at doses ≥46 ng/kg-day at 14 and 31 weeks and at doses ≥10 ng/kg-day at 53 weeks; and increase in TSH at doses ≥46 ng/kg-day at 14 weeks. There was a statistically-significant ($p \leq 0.05$) increase in hepatocyte proliferation at 14 weeks (22 ng/kg-day group only); 31 weeks (all doses); and 53 weeks (≥46 ng/kg-day). There were statistically significant ($p \leq 0.01$) dose-dependent increases in liver (includes EROD [CYP1A1-associated] activity; 7-pentoxyresorufin-O-deethylase [PROD; CYP2B-associated] activity; and acetanilide-4-hydroxylase [CYP1A2-associated] activity) and lung (EROD) cytochrome P450 enzyme activities in all treatment groups at all three evaluation periods

compared to the vehicle control group. The largest effect was an 82-fold induction of hepatic EROD activity in the 46 ng/kg-day group at 31 weeks.

TCDD was detected at the greatest concentration in the liver, followed by fat tissue, with tissue concentration increasing in both of these tissues in a dose-dependent manner. TCDD tissue levels generally remained constant after the first measurement at week 14. Pathological examination at week 14 revealed increased incidences of hepatocellular hypertrophy in animals administered ≥10 ng/kg-day TCDD. Examinations at weeks 31 and 53 indicated that incidence and or severity of hepatocellular hypertrophy was increased at all treatment doses although incidences were statistically significant ($p \leq 0.05$) only at ≥10 ng/kg-day doses. The incidence of non-neoplastic hepatic lesions (including inflammation, necrosis, multiple eosinophilic focus, diffuse fatty change, pigmentation, toxic hepatopathy) in the liver increased at doses ≥22 ng/kg-day beginning at 14 weeks. Severity of the lesions increased at 14 weeks at doses ≥46 ng/kg-day and were also observed at lower dose levels during later evaluation periods (31 and 53 weeks). By terminal sacrifice, numerous non-neoplastic changes were noted in TCDD treated rats, even at the lowest dose tested.

Noncancer cardiovascular and pulmonary effects were evident after 2 years of TCDD exposure. Significantly increased incidences of minimal to mild cardiomyopathy were seen in male and female rats at ≥10 ng/kg-day. In the lung, there was a significant ($p \leq 0.01$) dose-dependent increase, when compared to the vehicle control, in the incidence of bronchiolar metaplasia of the alveolar epithelium at all dose groups in the primary study.

A LOAEL for TCDD of 2.14 ng/kg-day adjusted dose for a 105-week exposure duration is identified in this study for significantly (either $p \leq 0.01$ or $p \leq 0.05$) increased absolute and relative liver weights, increased incidence of hepatocellular hypertrophy, and increased incidence of alveolar to bronchiolar epithelial metaplasia. A NOAEL cannot be determined for this study.

2.4.2.5.8. *Rier et al., 2001a, b.*

Female rhesus monkeys (8 per treatment group) were administered 0, 5, or 25 ppt TCDD (purity not specified) in the diet for 4 years. Previously, Bowman et al. (1989b) determined that these dietary concentrations were equivalent to 0, 0.15, and 0.67 ng/kg-day, respectively. Thirteen years after termination of TCDD treatment, serum concentrations of TCDD and dioxin-like polyhalogenated aromatic hydrocarbons (PHAH) were measured in six control

monkeys, six monkeys treated with 0.15 ng/kg-day, and three monkeys treated with 0.67 ng/kg-day (Rier et al., 2001a). Even after 13 years without treatment, there was significantly ($p < 0.05$) elevated serum levels of TCDD and other dioxin-like compounds in treated monkeys. There was a significant increase in triglycerides and total lipids in the serum of monkeys treated with either 0.15 or 0.67 ng/kg-day, but not in cholesterol or phospholipids. In addition to these 15 animals, 8 other female monkeys (4 treated with 0.67 ng/kg-day TCDD that died 7 to 11 years after treatment and 4 lead-treated animals with no history of PHAH exposure) were evaluated for endometriosis. Elevated serum concentrations of TCDD were not correlated with endometriosis. Increased serum levels of 3,3',4,4'-tetrachlorobiphenyl (TCB), however, were associated with the presence and severity of endometriosis ($p < 0.05$). TCB was found in none of the animals without endometriosis, including TCDD-treated animals, nor was it found in control animals with endometriosis. Animals with elevated serum levels of TCB, pentachlorobiphenyl, and total serum analyte TCDD equivalents (TEQ) had an increased incidence of endometriosis, but severity was associated only with increased levels of TCB. EPA did not develop a LOAEL for TCDD for this study, because of DLC contamination.

In a separate study that evaluated the same 15 monkeys 13 years after exposure, Rier et al. (2001b) examined effects on systemic immunity. Peripheral blood mononuclear cells (PBMC) obtained from untreated monkeys secreted no detectable levels of TNF-α in response to T-cell mitogen exposure. There was, however, a significant ($p < 0.05$) dose-dependent increase in TNF-α production in PBMC from the TCDD-treated monkeys. Although PBMC from treated monkeys with endometriosis produced more TNF-α than cells from unexposed controls without the disease (median 128 pg/mL compared to not detected; $p < 0.01$), PBMC from TCDD-treated animals without endometriosis also produced more TNF-α than controls (median 425 pg/mL, $p < 0.067$). TNF-α production from the animals without endometriosis, however, was much more variable and was not statistically significant compared to controls. In addition, there was a dose-related but statistically insignificant decrease in PBMC cytotoxicity against natural killer-sensitive RAJI cells in TCDD-treated animals compared to the unexposed controls. The results were again related to TCDD exposure and not the presence of endometriosis. TCDD alone was not associated with changes in PBMC surface antigen expression, but increased serum levels of TCDD. 1,2,3,6,7,8-Hexachlorodibenzofuran and 3,3',4,4',5-pentachlorobiphenyl were correlated with increased numbers of CD3+/CD25- and CD3-/CD25+ leukocytes, as well as

increased secretion of TNF-α in response to T-cell mitogen exposure. Although TNF-α production is considered to be a general indicator of inflammation, relative adversity of increased TNF-α secreted by PBMCs in and of itself cannot be substantiated in the absence of concurrent physiological measurements of an inflammatory response. Therefore, neither a LOAEL nor NOAEL can be determined for this study.

2.4.2.5.9. *Sewall et al., 1993.*

Sewall et al. (1993) examined the impact of TCDD exposure on the hepatic epidermal growth factor receptor (EGFR) as a critical effect in hepatocarcinogenicity. In two separate experiments, groups of 6- to 8-week-old female Sprague-Dawley rats were randomly assigned to the following groups: control group, receiving saline and corn oil; a promoted group that received four different doses of TCDD along with saline; a DEN-only initiated control group; and a DEN and TCDD initiated and promoted group that received four different doses of TCDD. DEN was administered via intraperitoneal injection at a dose of 175 mg/kg [saline (S) vehicle] as the initiating agent to animals that were 70 days old. The control animals received saline only. In the first experiment, each treatment group (S/TCDD and DEN/TCDD) that included sham-operated or ovariectomized and intact animals were treated with TCDD (purity >98%) at 125 ng/kg-day. In the second dose-response experiment, DEN-initiated and saline control treatment groups (intact animals, 84 days old) were administered TCDD (purity >98%) in corn oil via oral gavage once every 2 weeks for 30 weeks at doses equivalent to 0, 3.5, 10.7, 35.7, or 125 ng/kg-day ($n = 9$). A week after the last treatment, all animals were sacrificed and livers were harvested and fixed for immunohistochemistry. Sections of the fixed liver were tested for EGFR binding, EGFR autophosphorylation, immunolocalization of EGFR, and hepatic cell proliferation.

In the first experiment, intact animals treated with 125 ng/kg-day TCDD exhibited a 65% reduction in EGFR binding capacity. In contrast, the EGFR equilibrium maximum binding capacity (B_{max}) of the ovariectomized rats was not statistically different from the ovariectomized control rats, and no changes in the K_d were detected in any treatment group. In the dose-response experiment with intact animals, a significant ($p < 0.05$) TCDD dose-dependent decrease in the B_{max} of EGFR was shown. A two-factor, five-level ANOVA indicated that the effect of TCDD exposure on EGFR B_{max} was significant ($p = 0.0001$), whereas, the effect of

DEN treatment on EGFR B_{max} was not significant. Comparative analysis using Fisher's protected least significant difference indicated that the lowest TCDD dose resulting in a statistically significant ($p < 0.05$) decrease in the EGFR B_{max} was 10.7 ng/kg-day S/TCDD group. At the highest TCDD dose of 125 ng/kg-day, the EGFR B_{max} was reduced by 38% compared to controls in both the DEN initiated and noninitiated groups. A two-factor, five-level ANOVA showed no significant effect on EGFR K_d in either the DEN- or the TCDD-treated groups. The EGFR autophosphorylation assay indicated that, with increasing TCDD dose, the amount of EGFR autophosphorylation in DEN/TCDD-treated animals decreased. The study authors state that this decrease is similar to the dose-response alterations observed for the EGFR B_{max}. Additionally, EGFR autophosphorylation in control and 125 ng/kg-day noninitiated animals was similar to the corresponding dose levels for the DEN-treated animals, suggesting that DEN treatment did not affect the EGFR or the EGFR response to TCDD under the experimental conditions. The immunolocalization assay indicated that staining was more apparent in the centrilobular and midzonal regions of the liver in the DEN initiated control animals, whereas, the amount of hepatocyte plasma membrane staining in DEN/TCDD treated animals substantially decreased. The cell proliferation assay showed a decrease in the cell labeling index in the 3.5 ng/kg-day DEN/TCDD dose group that was statistically less ($p \leq 0.05$) than the labeling index for the control group. In contrast, the labeling index for the 125 ng/kg-day DEN/TCDD treatment group was significantly ($p \leq 0.05$) higher compared to controls. Except for the low-dose (3.5 ng/kg-day) group, a clear dose-response trend (two mid-level doses were not statistically significant) was observed in the other three TCDD treated groups.

The role of EGFR in TCDD-mediated hepatotoxicity is unknown, and as such, this endpoint cannot be unequivocally linked to TCDD-induced hepatotoxicity nor labeled as adverse. Thus, no LOAEL/NOAEL was established. A LOEL for TCDD of 3.5 ng/kg-day for a 30-week exposure duration was identified in this study for a significant ($p = 0.0001$ using ANOVA) decrease in EGFR B_{max} levels. A NOEL cannot be determined for this study.

2.4.2.5.10. *Sewall et al., 1995.*

Sewall et al. (1995) studied the dose-response relationship for thyroid function alterations in female rats as a result of TCDD exposure. Groups of female Sprague-Dawley rats were

initiated with DEN at 70 days of age at a dose of 175 mg/kg in a saline vehicle via an i.p. injection. DEN was administered as a liver-initiating agent for a concurrent study to determine TCDD promotion of hepatic preneoplastic foci. Saline-treated animals served as controls. At 84 days of age, both the DEN-initiated and the saline-noninitiated groups of animals were administered TCDD (purity >98%) or corn oil vehicle via oral gavage once every 2 weeks for 30 weeks at dose levels equivalent to 0, 0.1, 0.35, 1.0, 3.5, 10.7, 35.7, or 125 ng/kg-day (n = 9 per group). One week after the last TCDD treatment, the animals were sacrificed and the thyroid was removed and fixed for further analysis. Blood was drawn from the abdominal aortic vein, and the serum was isolated and preserved for hormone analysis. Liver was also removed and prepped for further analysis. Thyroid hormone analysis was performed to determine serum TSH, T3, and T4 levels using radioimmunoassay kits. Histological examination was conducted on eosin-stained sections of the thyroid tissue. RNA level in the hepatic tissue was determined using a reverse transcription polymerase chain reaction (RT-PCR) technique.

TCDD treatment did not affect thyroid weight. A dose-dependent decrease in serum T4 levels was observed in both noninitiated and DEN-initiated animals with T4 levels dropping significantly ($p < 0.05$) at the 35 and 125 ng/kg-day TCDD doses in the noninitiated group. Compared to the noninitiated control group, DEN alone did not significantly affect T4 levels. Serum T3 level in the 125 ng/kg-day treatment group was slightly elevated but was not significantly different from levels in the control group. TSH levels in DEN initiated rats were increased at a dose of 3.5 ng/kg-day. In the noninitiated group, TSH level in the 125 ng TCDD/kg-day group was 3.27 ± 0.34 ng/mL (n = 9) compared to 1.3 ± 0.18 ng/mL in the corn oil control group (n = 7). This result, in conjunction with the T4 data, demonstrates that TCDD had a similar effect on thyroid hormone levels in both the noninitiated and DEN initiated groups. Histological sections examined for nodular lesions or neoplasms exhibited thyroid follicular adenoma in one DEN/corn oil control animal. The DEN/TCDD-treated animals exhibited diffuse follicular hyperplasia, with the size of colloidal follicles decreasing with TCDD treatment. Other qualitative DEN/TCDD-related changes included increased frequency of abnormally shaped follicles. The study authors reported that image analysis demonstrated a significant (p = 0.013) TCDD dose-related decrease in mean follicle size along with a significant (p = 0.001) TCDD dose-related increase in parenchymal area. Additionally, like T4 and TSH

levels, DEN treatment alone or in combination with TCDD did not influence thyroid follicular or C-cell morphology.

RT-PCR results for UGT1 and CYP1A1 mRNA levels indicated that the amount of UGT1 mRNA at the 125 ng/kg-day dose was approximately 2.5-fold higher compared to the concurrent controls. The study authors also stated that the maximal response for the UGT1 mRNA levels was reached at a dose between 1.0 and 3.5 ng TCDD/kg-day. In contrast, the maximum induction of CYP1A1 mRNA was 260-fold higher at the 125 ng/kg-day compared to the concurrent controls.

A LOAEL for TCDD of 35 ng/kg-day for a 30-week exposure duration was identified in this study for a significant ($p < 0.05$) decrease in T4 levels. The NOAEL for this study is 10.7 ng/kg-day.

2.4.2.5.11. *Toth et al., 1979.*

Toth et al. (1979) examined the impact of TCDD exposure on the formation of liver tumors in male mice. Ten-week-old, outbred Swiss/H/Riop male mice were administered sunflower oil or TCDD (purity not specified; in sunflower oil) at 0, 7, 700 or 7,000 ng/kg (0, 1, 100, or 1,000 ng/kg-day adjusted for continuous dosing; administered dose divided by 7; $n = 38$, 44, 44, and 43, respectively) once per week via gastric tube for 1 year. Once exposure had ceased, animals were followed for the rest of their lives. After spontaneous death or when mice were moribund, autopsies were performed and all organs were examined histologically.

Average life span in the 1,000 ng/kg-day dose group decreased considerably (72%) when compared to the control group. TCDD also caused dose-dependent, severe chronic and ulcerous skin lesions (12, 30, and 58% in the 1, 100, and 1,000 ng/kg-day dose groups, respectively) that was followed by generalized lethal amyloidosis (12, 23, and 40% in the 1, 100, and 1,000 ng/kg-day dose groups, respectively).

A LOAEL for TCDD of 1 ng/kg-day for 1-year exposure duration was identified in this study for severe chronic and ulcerous skin lesions (12% higher than controls), and generalized lethal amyloidosis (12% higher than controls). A NOAEL cannot be determined for this study.

2.4.2.6. *Chronic Studies (Cancer Endpoints)*

2.4.2.6.1. *Kociba et al., 1978.*

As discussed above, Kociba et al. (1978) conducted a lifetime (2-year) feeding study of male and female Sprague-Dawley rats using doses of 0, 1, 10, and 100 ng/kg-day. There were 50 males and 50 females in each group.

With respect to the cancer endpoints examined, the most significant finding was an increase in hepatocellular hyperplastic nodules and hepatocellular carcinomas in female rats. The incidence of hepatocellular carcinomas was significantly elevated above the control incidence at the 100 ng/kg-day dose, whereas increased incidence of hyperplastic nodules was evident in the 10 ng/kg-day dose group.

There have been two reevaluations of slides of liver sections from the Kociba et al. study (Squire, 1980; Sauer, 1990; Goodman and Sauer, 1992). The Squire Review was requested by EPA as an independent review of the slides. The Sauer Review was carried out using refined criteria for the diagnosis of proliferative hepatocellular lesions (Maronpot et al., 1986, 1989). Liver tumor incidences for the three evaluations are compared in Appendix F. Although there are some quantitative differences between the evaluations, the lowest detectable effect for liver tumor incidence is consistently observed at 10 ng/kg-day.

In the 10 ng/kg-day dose group, significant increases in the incidence of hyperplastic nodules of the liver were observed in female rats (18/50 in the Kociba evaluation, 27/50 in the Squire evaluation). Two females (2/50) had hepatocellular carcinomas. In the 1990 reevaluation (Sauer, 1990; Goodman and Sauer, 1992), nine females (9/50) were identified with hepatocellular adenomas and none with carcinomas; thus only one-third of the previously observed "tumors" were identified when using the refined diagnostic criteria. As discussed below, the tumor reclassification of Goodman and Sauer (1992) was used in the dose-response modeling for the Kociba et al. (1978) data set.

In addition to nodules in the liver, increased incidence of stratified squamous cell carcinoma of the tongue and nasal turbinates/hard palate, and keratinizing squamous cell carcinoma of the lung were also observed in female rats in the 100 ng/kg-day dose group. One possible cause for the induction of lung tumors in the Kociba feeding study may have been the aspiration of dosed feed into the lungs. However the promotion of lung tumors has been observed in mice treated systemically by intraperitoneal (i.p.) injections of TCDD (Beebe et al.,

1995). In addition the induction of hyperplastic and metaplastic lesions in rats has been observed following chronic oral gavage treatment with TCDD (Tritscher et al., 2000). More recently, chronic oral exposure to HCDD resulted in the induction of lung tumors in treated female rats (Rozman, 2000). These data indicate that the induction of lung tumors in the Kociba was most likely primarily the result of systemic chronic dietary exposure to TCDD rather than due to a localized exposure to aspired dosed feed.

There was no detectable increase in liver tumor incidences in male rats in any of the dose groups. The mechanism responsible for dioxin-mediated sex specificity for hepatocarcinogenesis in rats is not clear, but may involve ovarian hormones (Lucier et al., 1991).

Although there was no increase in liver tumors in male rats in this study, in the 100 ng/kg-day group, there was an increased incidence of stratified squamous cell carcinoma of the hard palate/nasal turbinate, stratified squamous cell carcinoma of the tongue, and adenoma of the adrenal cortex.

Kociba et al. (1978) had reported that chemically related increases in preneoplastic or neoplastic lesions were not found in the 1 ng/kg-day dose group. However, Squire identified two male rats in the 1 ng/kg-day dose group with squamous cell carcinoma of the nasal turbinates/hard palate, and one of these male rats had a squamous cell carcinoma of the tongue. These are both rare tumors in Sprague-Dawley rats, and these sites are targets for TCDD, implying that 1 ng/kg-day may not represent a NOEL. However, no dose-response relationships were evident for tumors at these sites (Huff et al., 1991)

There is considerable controversy concerning the possibility that TCDD-induced liver tumors are a consequence of cytotoxicity. Goodman and Sauer (1992) have extended the reevaluation of the Kociba slides to include liver toxicity data and have reported a correlation between the presence of overt hepatotoxicity and the development of hepatocellular neoplasms in female rats. With the exception of two tumors in controls and one each in the low- and mid-dose groups, all liver tumors occurred in livers showing clear signs of toxicity. However, male rat livers exhibit cytotoxicity in response to high TCDD doses, yet they do not develop liver tumors. Moreover, both intact and ovariectomized female rats exhibit liver toxicity in response to TCDD, yet TCDD is a more potent promoter in intact but not ovariectomized rats (Lucier et al., 1991). Therefore, if cytotoxicity is playing a role in liver tumorigenesis, other factors must also be

involved. Also, there is little information on the role of cytotoxicity in TCDD-mediated cancer at other sites such as the lung and thyroid.

2.4.2.6.2. *Toth et al., 1979.*

In a study of 10-week-old outbred male Swiss/H/Riop mice, Toth et al. (1979) administered oral gavage TCDD doses of 0, 7, 700, and 7,000 ng/kg-day in sunflower oil weekly for 1 year (0, 1, 100, or 1,000 ng/kg-day adjusted for continuous dosing; see details above). All mice (100/group) were followed for their entire lives. The study authors identified the effective number of mice in each group to be the number of surviving animals when the first tumor-bearing animal was identified. The average lifespan of the control, low, mid and high dose groups was 588, 649, 633, and 424 days, respectively.

In the 100 ng/kg-day dose group, liver tumor incidence was twice that of the control group and was statistically significant ($p < 0.01\%$). A dose-related increase in liver tumor incidence was observed (18, 29, 48, and 30% in the control and three TCDD-treated groups, respectively) in all treated mice. Increases were not statistically significant, however, at 1 and 1,000 ng/kg-day. The study authors also stated that spontaneous and induced liver tumors were not histologically different. Additionally, the ratio of benign hepatomas to hepatocellular carcinomas in the control group was not affected by treatment and an increase was observed only in the absolute number of liver tumors. Cirrhosis was not observed with the tumors.

2.4.2.6.3. *NTP, 1982.*

As discussed above, the NTP (1982) study was conducted using Osborne-Mendel rats and B6C3F1 mice (NTP, 1982). Groups of 50 male rats, 50 female rats, and 50 male mice received TCDD as a suspension in corn oil:acteone (9:1) by gavage twice each week at doses of 0, 5, 25, or 250 ng/kg-day (daily averaged doses of 0, 1.4, 7.1, or 71 ng/kg-day for rats and male mice and doses of 0, 5.7, 28.6, or 286 ng/kg-day for female mice.

There were no statistically significant dose-related decreases in survival in any sex-species group. TCDD-induced malignant liver tumors occurred in the high-dose female rats and in male and female mice. These can be considered to result from TCDD exposure because they are relatively uncommon lesions in control Osborne-Mendel rats (male, 1/208; female, 3/208), are seen in female rats and mice of both sexes, and their increasing incidence with

increasing dose is statistically significant (Cochran-Armitage trend test, $p = 0.004$). Because liver tumors were increased in both sexes of mice, this effect is not female-specific as was observed in rats. Interestingly, liver tumor incidences were decreased in female rats in both the NTP and Kociba low doses (not statistically significant compared with controls). For example, the combined control incidence data were 11/161 (7%) compared with 4/99 (4%) in the low-dose group.

The incidences of thyroid gland (follicular cell) tumors were increased in all three dose groups in male rats. Because the responses in the two highest dose groups are highly significant, the statistically significant elevation of incidence in the lowest dose group (Fisher exact p-value = 0.042) is considered to be caused by exposure to TCDD, suggesting that thyroid tumor incidence may be the most sensitive site for TCDD-mediated carcinogenesis. Because 71 ng/kg-day is above the maximum tolerated dose (MTD) (Huff et al., 1991), thyroid tumors occur at doses more than 50 times lower than the MTD.

TCDD-induced neoplasms of the adrenal gland were observed in the 7.1 ng/kg-day/dose group in male rats and in high-dose female rats. Fibrosarcomas of the subcutaneous tissue were significantly elevated in high-dose female mice and female rats. One additional tumor type, lymphoma, was seen in high-dose female mice. Lung tumors were elevated in high-dose female mice; the increase was not statistically significant when compared with concurrent controls, but the increase was dose related (Cochran-Armitage trend test, $p = 0.004$).

Huff (1992) concluded, based on the NTP bioassay results, that TCDD was a complete carcinogen and induced neoplasms in rats and mice of both sexes. As was observed in the Kociba study (Kociba et al., 1978), liver tumors were observed with greater frequency in treated female rats, but in male rats the thyroid appears to be the most sensitive (increased tumor incidence at doses as low as 1.4 ng/kg-day).

2.4.2.6.4. *NTP, 2006.*

As discussed above, female Sprague-Dawley rats (81 control; 82 treatment group) were administered TCDD (purity >98%) in corn oil:acetone (99:1) via gavage at doses of 0, 3, 10, 22, 46, or 100 ng/kg-day, 5 days per week for 105 weeks (0, 2.14, 7.14, 15.7, 32.9, or 71.4 ng/kg-day, adjusted for continuous exposure) (NTP, 2006). In addition to this primary group, a stop-dose group of 50 animals was administered 100 ng/kg-day TCDD in corn

oil:acetone (99:1) via gavage for 30 weeks and then just the vehicle for the remainder of the study. At study termination, the number of surviving animals had declined to 25 in the control group and to 21, 23, 19, 22, and 21 in five treatment groups, respectively, due to accidental deaths, moribund animals, or death due to natural causes.

Incidence of hepatocellular adenomas was significantly ($p < 0.001$) increased in the 100 ng/kg-day dose group in the primary study and exceeded incidences seen in historical vehicle control range at study termination. A dose-related increase in the incidence of cholangiosarcoma was seen in the primary study group in animals receiving 22 ng/kg-day or higher doses of TCDD. The high dose group of 100 ng/kg-day had the highest incidence of cholangiosarcoma with a significant ($p < 0.001$) number of animals exhibiting multiple cholangiosarcomas. Such an incidence was not seen in historical vehicle controls. In contrast, only two cholangiosarcomas and hepatocellular adenomas were seen in the 100 ng/kg-day group in the stop-exposure study.

In the lung, at 2 years, there was a significantly ($p = 0.002$) increased incidence of cystic keratinizing epithelioma in the 100 ng/kg-day dose group of the primary study, while there were no epitheliomas in the 100 ng/kg-day group of the stop-exposure study. There was also a significant ($p \leq 0.01$) dose-dependent increase, when compared to the vehicle control, in the incidence of bronchiolar metaplasia of the alveolar epithelium at all dose groups in the primary study. Squamous metaplasia was also present in the 46 and 100 ng/kg-day dose groups in the primary study, and was also observed in the 100 ng/kg-day dose group in the stop-exposure study.

A positive trend in the incidence of gingival squamous cell carcinoma of the oral cavity was seen at all doses (except 22 ng/kg-day), with the incidence significantly ($p = 0.007$) high in the 100 ng/kg-day dose group. In addition, the occurrence of this lesion in the 46 and 100 ng/kg-day group of the primary study and 100 ng/kg-day group of the stop-exposure study exceeded the historical control range. The incidence of gingival squamous hyperplasia was significantly (either $p \leq 0.01$ or $p \leq 0.05$) increased in all dose groups of the primary study as well as the 100 ng/kg-day group of the stop-exposure study.

In the uterus, at 2 years, there was a significantly ($p = 0.032$) higher rate of squamous cell carcinoma in the 46 ng/kg-day group compared to vehicle controls. In addition there were two

squamous cell carcinomas in the 100 ng/kg-day group of the stop-exposure study. No squamous cell carcinomas have been reported in historical vehicle controls.

These results indicate that TCDD is carcinogenic to female Sprague-Dawley rats and causes tumors at multiple sites.

2.4.3. Summary of Key Data Set Selection for TCDD Dose-Response Modeling

To meet the NAS concerns regarding transparency and clarity in the identification of TCDD studies for dose-response assessment, EPA has, in this section, developed and applied two sets of criteria, for animal bioassays and epidemiologic studies. EPA has collected and evaluated these studies, including studies from the 2003 Reassessment and newer studies found via literature searches and through public submissions. Tables 2-4 and 2-5 contain the final lists of key cancer and noncancer studies, respectively, that have met EPA's inclusion criteria for epidemiologic data. Tables 2-6 and 2-7 provide the final lists of key studies that have met EPA's inclusion criteria for animal bioassay data for cancer and noncancer studies, respectively. Collectively, these four tables contain the final set of key studies that EPA has used to develop noncancer and cancer dose-response assessments for TCDD in Sections 4 and 5 of this document. In Sections 4 and 5, additional evaluations are made to determine which study/endpoint data sets are the most appropriate for development of the RfD and OSF for TCDD, using statistical criteria, dose-response modeling results and decisions regarding toxicological relevance of the endpoints. The approaches taken to select the final candidate study/endpoint data sets are discussed in Sections 4 and 5 and are illustrated in Figures 4-1 and 5-1 of those sections.

Table 2-1. Summary of epidemiologic cancer studies (key characteristics)

Publication	Length of follow-up	Latency period	Half-life for TCDD	Fraction of TEQs accounted for by TCDD
NIOSH cohort studies				
Fingerhut et al., 1991	1942−1987	0, 20 years	N/A	N/A
Steenland et al., 1999	1942−1993	0, 15 years	N/A	N/A
Steenland et al., 2001	1942−1993	0, 15 years	8.7 years (Michalek et al., 1996)	TCDD accounted for all occupational TEQ; 10% of background
Cheng et al., 2006	1942−1993	0, 10, 15 years	8.7 years (Michalek et al., 1996), and CADM (Aylward et al., 2005a)	N/A
Collins et al., 2009	1942−2003	None	7.2 years (Flesch-Janys et al., 1996)	N/A
BASF cohort studies				
Thiess et al., 1982	1953−1980	None	N/A	N/A
Zober et al., 1990	1953−1987	Years since first exposure: 0−9, 10−19, and 20+	N/A	N/A
Ott and Zober, 1996	1953−1991	None	5.8 years	N/A
Hamburg cohort studies				
Manz et al., 1991	1952−1989	None, used duration of employment (<20, >20 years)	N/A	N/A
Flesch-Janys et al., 1995	1952−1992	None	7.2 years (Flesch-Janys et al., 1994)	Mean TEQ without TCDD was 155 ng/kg; mean TEQ with TCDD was 296.5 ng/kg
Flesch-Janys et al., 1998	1952−1992	None	7.2 years (Flesch-Janys et al., 1996), also used decay rates that were function of age and fat composition	Mean concentration of TCDD was 101.3 ng/kg; for TEQ (without TCDD) mean exposure was 89.3 ng/kg
Becher et al., 1998	1952−1992	0, 5, 10, 15 and 20 years	7.2 years (Flesch-Janys et al., 1996) took into account age and fat composition	Not described

1

Table 2-1. Summary of epidemiologic cancer studies (key characteristics) (continued)

Publication	Length of follow-up	Latency period	Half-life for TCDD	Fraction of TEQs accounted for byTCDD
Seveso cohort studies				
Bertazzi et al., 2001	1976−1996	Periods postexposure: 0, 0−4, 5−9, 10−14, 15−19	N/A	N/A
Consonni et al., 2008	1976−2001	Periods postexposure: 0, 0−4, 5−9, 10−14, 15−19, 20−24	N/A	N/A
Warner et al., 2002	1976−1998	None	8 years (Pirkle et al., 1989)	N/A
Chapaevsk cohort studies				
Revich et al., 2001	Cross-sectional study (1995−1998)	N/A	N/A	N/A
Ranch Hand cohort studies				
Akhtar et al., 2004	1962−1999	None	N/A	N/A
Michalek and Pavuk, 2008	1962−2004	None, but stratified by period of service	7.6 years	N/A
New Zealand cohort studies				
McBride et al., 2009a	1969−2004	None	N/A	N/A
McBride et al., 2009b	1969−2004	None	7 years	N/A

Table 2-2. Epidemiology cancer study selection considerations and criteria

	Methods use to ascertain health outcomes were unbiased, highly sensitive and specific.	Risk estimates are not susceptible to biases from confounding exposures or from study design or statistical analysis.	Association between TCDD and adverse health effect, with exposure-response relationship.	Exposure assessment methodology clear and adequately characterizes individual-level exposures. Limitations and uncertainties in exposure assessment considered.	Study size and follow-up large enough to yield precise estimates of risk and ensure adequate statistical power.	Published in peer-reviewed literature with appropriate discussion of strengths, limitations.	Exposure primarily TCDD and quantified so that dose-response relationship can be assessed.	Effective dose & oral exposure estimable & consistent w/ current biological understanding. Latency and appropriate window(s) of exposure examined.	Pass for dose-response analyses?
Cancer	**Considerations**					**Criteria**			**Y/N**
NIOSH Cohort Studies									
Fingerhut et al., 1991 all cancer sites, site-specific analyses	√	X	X	X	√	√	X	√	N
Steenland et al., 1999 all cancer sites combined, site-specific analyses	√	√	√	√	√	√	√	√	[a]
Steenland et al., 2001									
all cancer sites combined	√	√	√	√	√	√	√	√	Y
Cheng et al., 2006								N	
all cancer sites combined	√	√	√	√	√	√	√	√	Y
Collins et al., 2009 all cancer sites combined, site-specific analyses	√	√	X	√	√	√	√	X	N
BASF Cohort Studies									
Zober et al., 1990 all cancer sites combined, site-specific analyses	√	√	X	X	X	√	X	X	N
Ott and Zober, 1996									
all cancer sites combined	√	√	√	√	√	√	√	√	Y

Table 2-2. Epidemiology cancer study selection considerations and criteria (continued)

	Considerations					Criteria			
Cancer	Methods use to ascertain health outcomes were unbiased, highly sensitive and specific.	Risk estimates are not susceptible to biases from confounding exposures or from study design or statistical analysis.	Association between TCDD and adverse health effect, with exposure-response relationship.	Exposure assessment methodology clear and adequately characterizes individual-level exposures. Limitations and uncertainties in exposure assessment considered.	Study size and follow-up large enough to yield precise estimates of risk and ensure adequate statistical power.	Published in peer-reviewed literature with appropriate discussion of strengths, limitations.	Exposure primarily TCDD and quantified so that dose-response relationship can be assessed.	Effective dose & oral exposure estimable & consistent w/ current biological understanding. Latency and appropriate window(s) of exposure examined.	Pass for dose-response analyses? **Y/N**
Hamburg Cohort									
Manz et al., 1991 all cancer sites combines, site-specific analyses	√	√	√	√	√	√	X	√	N
Flesh-Janys et al., 1995 all cancer sites combined	√	√	√	√	√	√	√	X	N
Flesh-Janys et al., 1998 all cancer sites combined, site-specific analyses	√	√	√	√	√	√	√	√	[b]
Becher et al., 1998 all cancer sites combined	√	√	√	√	√	√	√	√	Y
Seveso Cohort								N	
Bertazzi et al., 2001 all cancer sites combined, site-specific analyses	√	√	√	X	√	√	X	X	N
Pesatori et al., 2003 all cancer sites combined, site-specific analyses	√	√	X	X	√	√	X	X	N
Consonni et al., 2008 all cancer sites combined, site-specific analyses	√	√	√	X	√	√	X	X	N

Table 2-2. Epidemiology cancer study selection considerations and criteria (continued)

	Methods use to ascertain health outcomes were unbiased, highly sensitive and specific.	Risk estimates are not susceptible to biases from confounding exposures or from study design or statistical analysis.	Association between TCDD and adverse health effect, with exposure-response relationship.	Exposure assessment methodology clear and adequately characterizes individual-level exposures. Limitations and uncertainties in exposure assessment considered.	Study size and follow-up large enough to yield precise estimates of risk and ensure adequate statistical power.	Published in peer-reviewed literature with appropriate discussion of strengths, limitations.	Exposure primarily TCDD and quantified so that dose-response relationship can be assessed.	Effective dose & oral exposure estimable & consistent w/ current biological understanding. Latency and appropriate window(s) of exposure examined.	Pass for dose-response analyses?
Cancer	**Considerations**					**Criteria**			**Y/N**
Baccarelli et al., 2006 site specific analysis	√	√	X	√	√	√	√	√	[c]
Warner et al., 2002 breast cancer incidence	√	√	√	√	√	√	√	√	Y
Chapaevsk Study								N	
Revich et al., 2001 all cancer sites combined, site-specific analyses	X	X	X	X	√	X	X	X	N
Ranch Hands Cohort									
Akhtar et al., 2004 all cancer sites combined, site-specific analyses	√	X	√	√	√	√	X	√	N
Michalek and Pavuk, 2008 all cancer sites combined	√	X	√	√	√	√	X	√	N
Others									
tMannetje et al., 2005 all cancer sites combined, site-specific analyses	√	X	√	√	√	X	X	X	N
McBride et al., 2009b all cancer sites combined, site-specific analyses	√	X	X	√	X	√	X	X	N

Table 2-2. Epidemiology cancer study selection considerations and criteria (continued)

	Considerations					Criteria			
Cancer	Methods use to ascertain health outcomes were unbiased, highly sensitive and specific.	Risk estimates are not susceptible to biases from confounding exposures or from study design or statistical analysis.	Association between TCDD and adverse health effect, with exposure-response relationship.	Exposure assessment methodology clear and adequately characterizes individual-level exposures. Limitations and uncertainties in exposure assessment considered.	Study size and follow-up large enough to yield precise estimates of risk and ensure adequate statistical power.	Published in peer-reviewed literature with appropriate discussion of strengths, limitations.	Exposure primarily TCDD and quantified so that dose-response relationship can be assessed.	Effective dose & oral exposure estimable & consistent w/ current biological understanding. Latency and appropriate window(s) of exposure examined.	Pass for dose-response analyses? **Y/N**
McBride et al., 2009a									
all cancer sites combined, site-specific analyses	√	√	X	√	X	√	√	√	[d]
Hooiveld et al., 1998									
all cancer sites combined, site-specific analyses	√	√	√	√	X	√	√	X	N

N

[a]This study has been superseded and updated by Steenland et al. (2001).
[b]Becher et al. (1998) assessed this same cohort taking cancer latency into account, thereby superseding this study.
[c]It is unknown whether the frequency of t(14;18)translocations in lymphocytes relates specifically to an increased risk of non-Hodgkin's lymphoma. Given this lack of obvious adverse effect, dose-response analyses for this outcome were not conducted.
[d]No dose-response associations were noted.

√ = Consideration/criteria satisfied; X= Consideration/criteria not satisfied.

Table 2-3. Epidemiology noncancer study selection considerations and criteria

	Methods use to ascertain health outcomes were unbiased, highly sensitive and specific.	Risk estimates are not susceptible to biases from confounding exposures or from study design or statistical analysis.	Association between TCDD and adverse health effect, with exposure-response relationship.	Exposure assessment methodology clear and adequately characterizes individual-level exposures. Limitations and uncertainties in exposure assessment considered.	Study size and follow-up large enough to yield precise estimates of risk and ensure adequate statistical power.	Published in peer-reviewed literature with appropriate discussion of strengths, limitations.	Exposure primarily TCDD and quantified so that dose-response relationships can be assessed.	Effective dose & oral exposure estimable & consistent w/ current biological understanding. Latency and appropriate window(s) of exposure examined for a Non-Fatal endpoint.	Pass for dose-response analyses?
Noncancer	**Considerations**					**Criteria**			**Y/N**
NIOSH Cohort									
Steenland et al., 1999 mortality (noncancer) -ischemic heart disease	√	X	√	√	√	√	X	X	N
Collins et al., 2009 mortality (noncancer)	√	√	X	√	√	√	√	X	N
BASF Cohort									
Ott and Zober, 1996 mortality (noncancer)	√	√	X	√	√	√	√	X	N
Hamburg Cohort									
Flesch-Janys et al., 1995 mortality (noncancer)	√	√	√	√	√	√	√	X	N
Seveso Cohort									
Eskenazi et al., 2002a menstrual cycle characteristics	√	√	√	√	√	√	√	√	Y
Eskenazi et al., 2002b endometriosis	X	X	X	√	X	√	√	X	N

Table 2-3. Epidemiology noncancer study selection considerations and criteria (continued)

	Considerations					Criteria			
Noncancer	Methods use to ascertain health outcomes were unbiased, highly sensitive and specific.	Risk estimates are not susceptible to biases from confounding exposures or from study design or statistical analysis.	Association between TCDD and adverse health effect, with exposure-response relationship.	Exposure assessment methodology clear and adequately characterizes individual-level exposures. Limitations and uncertainties in exposure assessment considered.	Study size and follow-up large enough to yield precise estimates of risk and ensure adequate statistical power.	Published in peer-reviewed literature with appropriate discussion of strengths, limitations.	Exposure primarily TCDD and quantified so that dose-response relationships can be assessed.	Effective dose & oral exposure estimable & consistent w/ current biological understanding. Latency and appropriate window(s) of exposure examined for a Non-Fatal endpoint.	Pass for dose-response analyses? **Y/N**
Eskenazi et al., 2003 birth outcomes	X	X	X	√	√	√	√	X	N
Warner et al., 2004 age at menarch	√	√	X	√	√	√	√	X	N
Eskenazi et al., 2005 age at menopause	√	√	X	√	√	√	√	X	N
Warner et al., 2007 ovarian function	√	√	X	√	√	√	√	X	N
Eskenazi et al., 2007 uterine leiomyoma	√	√	√	√	√	√	√	X	N[a]
Mocarelli et al., 2008 semen quality	√	√	√	√	√	√	√	√	Y
Mocarelli et al., 2000 sex ratio	√	√	√	√	√	X	√	X	N[b]
Baccarelli et al., 2008 neonatal thyroid function	√	√	√	X	√	√	√	√	Y

Table 2-3. Epidemiology noncancer study selection considerations and criteria (continued)

	Methods use to ascertain health outcomes were unbiased, highly sensitive and specific.	Risk estimates are not susceptible to biases from confounding exposures or from study design or statistical analysis.	Association between TCDD and adverse health effect, with exposure-response relationship.	Exposure assessment methodology clear and adequately characterizes individual-level exposures. Limitations and uncertainties in exposure assessment considered.	Study size and follow-up large enough to yield precise estimates of risk and ensure adequate statistical power.	Published in peer-reviewed literature with appropriate discussion of strengths, limitations.	Exposure primarily TCDD and quantified so that dose-response relationships can be assessed.	Effective dose & oral exposure estimable & consistent w/ current biological understanding. Latency and appropriate window(s) of exposure examined for a Non-Fatal endpoint.	Pass for dose-response analyses?
Alaluusua et al., 2004									
oral hygiene	√	√	√	√	√	√	√	√	Y
Bertazzi et al., 2001									
mortality (noncancer)	√	√	X	X	√	√	X	X	N
Consonni et al., 2008									
mortality (noncancer)	√	√	X	X	√	√	X	X	N
Baccarelli et al., 2005									
chloracne	√	√	√	√	√	√	√	√	[c]
Baccarelli et al., 2002, 2004									
immunological effects	√	√	√	√	√	√	√	X	N
Chapaevsk Study								N	
Revich et al., 2001									
mortality (noncancer) and reproductive health	√	X	X	X	√	√	X	X	N
Ranch Hands Cohort									
Michalek and Pavuk, 2008									
diabetes	√	X	√	√	√	√	X	√	N
Other									
McBride et al., 2009a									

Table 2-3. Epidemiology noncancer study selection considerations and criteria (continued)

	Methods use to ascertain health outcomes were unbiased, highly sensitive and specific.	Risk estimates are not susceptible to biases from confounding exposures or from study design or statistical analysis.	Association between TCDD and adverse health effect, with exposure-response relationship.	Exposure assessment methodology clear and adequately characterizes individual-level exposures. Limitations and uncertainties in exposure assessment considered.	Study size and follow-up large enough to yield precise estimates of risk and ensure adequate statistical power.	Published in peer-reviewed literature with appropriate discussion of strengths, limitations.	Exposure primarily TCDD and quantified so that dose-response relationships can be assessed.	Effective dose & oral exposure estimable & consistent w/ current biological understanding. Latency and appropriate window(s) of exposure examined for a Non-Fatal endpoint.	Pass for dose-response analyses?
mortality (noncancer)	X	X	X	√	X	√	√	X	N
McBride et al., 2009b mortality (noncancer)	X	√	X	√	X	√	X	X	N
Ryan et al., 2002 sex ratio	X	X	X	X	√	√	X	X	N

[a]Categorical measures of TCDD suggest an inverse association between TCDD exposure and uterine fibroids. The observed direction of the reported associations precluded quantitative dose-response modeling.

[b]The somewhat arbitrary cut off age of 19 for statistically significant exposure associations results in a highly uncertain critical exposure window. It is difficult to determine whether effects are a consequence of the initial high exposure during childhood or a function of the cumulative exposure for this entire exposure window. The differences between these two dose estimates are quite large.

[c]Chloracne is recognized to occur following high TCDD exposure levels. This study provides limited relevance to TCDD RfD development, as exposure levels observed in the general population are much lower.

√ = Consideration/criteria satisfied. X= Consideration/criteria not satisfied.

Table 2-4. Epidemiology studies selected for TCDD cancer dose-response modeling

Health outcome	Location, time period	Cohort description	Exposure assessment	Exposure measures	No. of cases/ deaths	Effect Measure/ RR (95% CI)	Risk factors	Comments	Reference
Mortality from all cancers	USA, 1942–1993	NIOSH cohort including 3,538 occupationally exposed male workers at 8 plants in the United States; 256 cancer deaths	Cumulative serum lipid TCDD concentrations (CSLC) based on work histories, job-exposure matrix, and concentration and age-dependent two-compartment model of elimination kinetics	No exposure categories provided	256 cancer deaths	The slope (β) was 3.3×10^{-6} for lag of 15 years excluding upper 5% of TCDD exposures. The slopes ranged two orders of magnitude depending on modeling assumption	Available: age, year of birth, and race Risks adjusted for: year of birth, age, and race	Confounding by smoking was considered indirectly by analysis of smoking-related and smoking-unrelated cancers. Other occupational exposures were considered indirectly by repeated analyses removing one plant at a time. Based on indirect evaluation, there was no clear evidence of confounding.	Cheng et al., 2006
Mortality from all cancers	USA, 1942–1993	NIOSH cohort including 3,538 male workers, 256 cancer deaths	CSLC based on work histories, job-exposure matrix, and a simple one-compartment first-order pharmacokinetic elimination model with 8.7-year half-life	CSLC (ppt-years) <335 335−520 520−1,212 1,212−2,896 2,896−7,568 7,568−20,455 ≥20,455	 64 29 22 30 31 32 48	 1.00 1.26 (0.79−2.00) 1.02 (0.62−1.65) 1.43 (0.91−2.25) 1.46 (0.93−2.30) 1.82 (1.18−2,82) 1.62 (1.03−2,56)	Available: date of birth and age Adjusted for: date of birth, and age was used as time scale in Cox model	Included in U.S. EPA (2003)	Steenland et al., 2001

Table 2-4. Epidemiology studies selected for TCDD cancer dose-response modeling (continued)

Health outcome	Location, time period	Cohort description	Exposure assessment	Exposure measures	No. of cases/ deaths	Effect Measure/ RR (95% CI)	Risk factors	Comments	Reference
Mortality from all cancers combined	Hamburg, Germany, production period was 1950−1984 and mortality follow-up extended through 1992	Boehringer cohort including approximately 1,189 workers employed in the production of herbicides	Cumulative TCDD serum lipid concentrations based on area under curve (in µg/kg years); back-extrapolation to date of last employment took into account age and percent body fat; half-life value was 7.2 years	Categorical exposures (Cox model) 0− <1 1− <4 4− <8 8− <16 16− <64 64+	124	 1.0 1.12 (0.70−1.80) 1.42 (0.70−2.85) 1.77 (0.81−3.86) 1.63 (0.73−3.64) 2.19 (0.76−6.29)	 Available: year of entry, age of entry, duration of employment, birth cohort, β-HCH; TEQ other than TCDD	Included in U.S. EPA (2003) A large number of models were fitted. These included models for 5 different latency intervals (0, 5, 10, 15, and 20 years), as well as multiplicative, additive and power models, and different offset variables (person years and expected deaths)	Becher et al., 1998
				Continuous exposure TCDD (µg/kg years)	124	β = 0.0089, p = 0.0047	Available: year of entry, age of entry, duration of employment, birth cohort, β-HCH; TEQ other than TCDD		

Table 2-4. Epidemiology studies selected for TCDD cancer dose-response modeling (continued)

Health outcome	Location, time period	Cohort description	Exposure assessment	Exposure measures	No. of cases/ deaths	Effect Measure/ RR (95% CI)	Risk factors	Comments	Reference
Mortality and incidence for all cancers combined, as well as for specific cancer sites	Ludwig-shafen, Germany, 1954−1992	BASF cohort, 243 men exposed from accidental release that occurred in 1953 during production of trichlorophenol, or who were involved in clean-up activities	Cumulative TCDD serum lipid concentrations expressed in µg/kg based on TCDD half-life of 5.1-8.9 years, Cox regression model	Internal comparisons based on continuous measure of TCDD. External comparisons exposure categories: <0.1, 0.1−0.99, 1.0−1.99 >2 µg/kg	*Internal cohort analysis* 31 cancer deaths 47 incident cancers *External cohort analyses* Deaths 8 8 8 7	Date of 1st TCDD exposure 1.22 (95% CI: 1.00−1.50) 1.11 (95% CI: 0.91−1.35) SMRs 0.8 (0.4−1.6) 1.2 (0.5−2.3) 1.4 (0.6−2.7) 2.0 (0.8−4.0)	Available: age, BMI, smoking status and history of occupational exposure to aromatic amines and asbestos	Included in U.S. EPA (2003) Positive associations noted for digestive cancer, but not for respiratory cancer Associated between TCDD and increased SMRs found only among current smokers Last published account of this cohort	Ott and Zober, 1996

Table 2-4. Epidemiology studies selected for TCDD cancer dose-response modeling (continued)

Health outcome	Location, time period	Cohort description	Exposure assessment	Exposure measures	No. of cases/ deaths	Effect Measure/ RR (95% CI)	Risk factors	Comments	Reference
Breast cancer incidence	Italy 1976–1998	981 women from zones A and B with available archive serum samples, 15 breast cancer cases	TCDD serum lipid concentrations (ppt) collected between 1976 and 1981. For most samples collected after 1977, serum TCDD levels were back-extrapolated using a first-order kinetic model with a 9-year half-life.	<20 ppt 20.1−44 ppt 44.1−100 ppt >100 ppt Log_{10}TCDD also modeled as continuous variable	Cases 1 2 7 5 15	 1.0 1.0 (0.1−10.8) 4.5 (0.6−36.8) 3.3 (0.4−28.0) 2.1 (1.0−4.6)	Available: gravidity, parity, age at first pregnancy, age at last pregnancy, lactation, family history of breast cancer, age at menarche, current body mass index, oral contraceptive use, menarcheal status at explosion, menopause status at diagnosis, height, smoking, alcohol consumption. Adjusted for age, which was used as time scale in Cox model; other covariates were evaluated but were not identified as confounders.	Included in U.S. EPA (2003)	Warner et al., 2002

Table 2-5. Epidemiology studies selected for TCDD noncancer dose-response modeling

Health outcome	Location, time period	Cohort description	Exposure assessment	Exposure measures	No. of cases/ deaths	Effect Measure/ RR (95% CI)	Risk factors	Comments	Reference
b-TSH measured 72 hours after birth from a heel pick (routine screening for all newborns in the region	Italy, 1976; children, 1994–2005	*Population-based study:* 1,041 singletons (56 from zone A, 425 from zone B and 533 from reference) born between Jan. 1, 1994-June 30, 2005. *Plasma dioxin study:* 51 children born to 38 women of fertile age who were part of the Seveso Chloracne Study.	Based on zone of residence, estimated mean values from a previous study. Maternal plasma TCDD levels estimated at the date of delivery using a first-order pharmacokinetic model and elimination rate estimated in Seveso women (half-life =9.8 years).	*Population-based study:* Reference Zone B Zone A *Plasma dioxin study:* Continuous maternal plasma TCDD	 533 births 425 births 56 births	*Population-based study* Mean b-TSH Reference: 0.98 (95% CI: 0.90−1.08) Zone B: 1.66 (95% CI: 1.19−2.31) Zone A: 1.35 (95% CI: 1.22−1.49) Association between neonatal b-TSH with plasma TCDD: adjusted $\beta = 0.75$ ($p < 0.001$)	Available: gender, birth weight, birth order, maternal age at delivery, hospital, type of delivery. There was limited evidence of confounding, so mean TSH results presented here are unadjusted.	An association with serum TCDD levels of mothers was found with b-TSH among the 51 births in the plasma dioxin study.	Baccarelli et al., 2008

Table 2-5. Epidemiology studies selected for TCDD noncancer dose-response modeling (continued)

Health outcome	Location, time period	Cohort description	Exposure assessment	Exposure measures	No. of cases/ deaths	Effect Measure/ RR (95% CI)	Risk factors	Comments	Reference
Sperm conc. (million/ mL) Progressive motility (%) Serum E_2 (pmol/L)	Italy, 1976, 1998	135 exposed (from zone A) and 184 non-exposed men aged 1−26 in 1976 were included. These subjects were selected from the cohort of 257 exposed and 372 unexposed people.	Serum TCDD (in ppt) from 1976-1977 samples (for exposed men); background values were assumed for unexposed men based on serum analysis of residents in uncontaminated areas.	TCDD quartiles		Mean values were compared between the exposed and comparison groups for sperm concentration, volume, motility and count, FSH, E_2, LH, and Inhibin B.	Available: age, abstinence time, smoking status, education, alcohol use, maternal smoking during pregnancy, employment status, BMI, chronic exposure to solvents and other toxic substances. Adjusted for smoking status, organic solvents, age at time of tests, BMI, alcohol use, education, employment status and abstinence (days) for sperm data. Hormone data not adjusted for education level, employment status, and abstinence time.	Results stratified by timing of exposure (1−9 yrs old vs. 10−17 yrs old in 1976).	Mocarelli et al., 2008

Table 2-5. Epidemiology studies selected for TCDD noncancer dose-response modeling (continued)

Health outcome	Location, time period	Cohort description	Exposure assessment	Exposure measures	No. of cases/ deaths	Effect Measure/ RR (95% CI)	Risk factors	Comments	Reference
Dental defects	Seveso, Italy, Dental exams administered in 2001 among those exposed to TCDD in 1976	65 subjects <9.5 years old at time of Seveso explosion and residing in zones ABR; 130 subjects recruited from the non-ABR region (unexposed)	Serum TCDD (ng/kg) from 1976 samples for those who resided in Zone ABR; no serum levels for non-ABR residents (unexposed). TCDD exposure represent levels as of 1976 (after accident)	Non-ABR Zone 31−226 ng/kg 238−592 ng/kg 700−26000 ng/kg <5 years of age at time of accident Non-ABR Zone or 31−226 ng/kg serum TCDD 238-26,000 ng/kg serum TCDD	10/39 1/10 5/11 9/15 25/75	Dental defect % 26% 10% 45% 60% Odds Ratios (among those <5 years of age at time of accident) 1.0 2.4 (1.3−4.5)	Available: medical history, age, sex, education, smoking	Dose-response pattern observed with dental defects in the ABR zone; however, the control population had a much higher prevalence of dental defects (26%) than those in the lowest exposure group (10%). Also assessed hypodontia and other dental and oral aberrations, but these were too rare to allow modeling by ABR zone.	Alaluusua et al., 2004

Table 2-5. Epidemiology studies selected for TCDD noncancer dose-response modeling (continued)

Health outcome	Location, time period	Cohort description	Exposure assessment	Exposure measures	No. of cases/ deaths	Effect Measure/ RR (95% CI)	Risk factors	Comments	Reference
Menstrual cycle characteristics: menstrual cycle length.	Seveso, Italy, follow-up interview conducted in 1996-1997 of women exposed to TCDD in the 1976 accident	Women who were <40 years from zones A or B in 1976, A positive association found among women who were pre-menarcheal at the time of accident (n=134)	Serum TCDD (ng/kg) from 1976 samples. TCDD exposure level was back-extrapolated to 1976 using the Filser or the first-order kinetic models.	Interquartile range was 64-322 ppt TCDD examined as continuous measure (per 10-fold increase in serum levels).		Lengthening of the menstrual cycle by 0.93 days (95% CI: -0.01, 1.86)	Interview data: medical history, personal habits, work history, reproductive history, age, smoking, body mass index, alcohol and coffee consumption, exercise, illness, abdominal surgeries.		Eskenazi et al., 2002a

Table 2-6. Animal bioassays selected for cancer dose-response modeling

Species/strain	Sex exposure route/duration	*n*	Average daily dose levels (ng/kg-day)	Cancer types	Statistical significant tumors (pairwise with controls or trend tests)	Reference
Rat/Sprague-Dawley	Male/female Oral-lifetime feeding; 2 years	50 each (86 each in vehicle control group)	0, 1, 10, or 100	Females: liver, lung, oral cavity Males: adrenal, oral cavity, tongue	Adrenal cortex: adenoma Liver: hepatocellular adenoma(s) or carcinoma(s); hyperplastic nodules Lung: keratinizing squamous cell carcinoma Oral cavity: stratified squamous cell carcinoma of hard palate or nasal turbinates Tongue: stratified squamous cell carcinoma	Kociba et al., 1978; Female liver tumors analysis updated in Goodman and Sauer, 1992
Mouse/ B6C3F1	Male/female Oral-gavage twice per week; 104 weeks	50 each (75 each in vehicle control group)	0, 1.4, 7.1, or 71 for males; 0, 5.7, 28.6, or 286 for females	Females: hematopoietic system, liver, subcutaneous tissue, thyroid Males: liver, lung	Hematopoietic system: lymphoma or leukemia Liver: hepatocellular adenoma or carcinoma Lung: alveolar/bronchiolar adenoma or carcinoma Subcutaneous tissue: fibrosarcoma Thyroid: follicular-cell adenoma	NTP, 1982a
Rat/Osborne-Mendel	Male/female Oral-gavage twice per week; 104 weeks	50 each (75 each in vehicle control group)	0, 1.4, 7.1, or 71	Females: adrenal, liver, subcutaneous tissue, thyroid Males: adrenal, liver, thyroid	Adrenal: cortical adenoma, or carcinoma or adenoma, NOS Liver: neoplastic nodule or hepatocellular carcinoma Subcutaneous tissue: fibrosarcoma Liver: neoplastic nodule or hepatocellular carcinoma Thyroid: follicular-cell adenoma or carcinoma	NTP, 1982a
Rat/Harlan Sprague-Dawley	Female Oral-gavage 5 days per week; 2 years	53 or 54	0, 2.14, 7.14, 15.7, 32.9, or 71.4	Liver Lung Oral mucosa Pancreas	Liver: hepatocellular adenoma Liver: cholangiocarcinoma Lung: cystic keratinizing epithelioma Oral mucosa: squamous cell carcinoma Pancreas: adenoma or carcinoma	NTP, 2006
Mouse/Outbred Swiss/H/Riop	Male Gastric intubation once per week; 1 year	43 or 44 (vehicle control group = 38)	0, 1, 100, or 1,000	Liver	Liver: tumors	Toth et al., 1979

Table 2-7. Animal bioassay studies selected for noncancer dose-response modeling

Species/ strain	Exposure protocol	Sex (exposure group)	*n*	Average daily dose levels (ng/kg-day)	NOAEL (ng/kg-day)	LOAEL (ng/kg-day)	Endpoint(s) examined	LOAEL/NOAEL Endpoint(s)	Reference
Reproductive toxicity studies									
Monkey/ Rhesus	Daily dietary exposure in female monkeys (3.5−4 years)	F (F0, F1, F2, F3)	3 to 7 (F1)	0, 0.15, or 0.67	0.15	0.67	Reproductive and developmental effects	Neurobehavioral effects (e.g., discrimination-reversal learning affected)	Bowman et al., 1989a, b (and related Schantz and Bowman, 1989; Schantz et al., 1986)
Mink	Daily dietary exposure (132 days)	F	12	0.03 (control), 0.8, 2.65, 9, or 70	None	2.65	Reproductive effects	Reduced kit survival	Hochstein et al., 2001
Rat/Holtzman	Corn oil gavage (initial loading dose followed by weekly dose during mating, pregnancy, and lactation– about 10 weeks)	F (F0) F and M (F1 and F2)	12 (F0) Not specified (F1 and F2)	0 or 16.5	None	16.5 (maternal exposure)	Reproductive and developmental effects	Decreased development of the ventral prostrate (F1), decreased sex ratio (percentage of males) (F2)	Ikeda et al., 2005a
Mouse/ICR	Sesame oil gavage (initial loading dose followed by weekly doses for 5 weeks)	M (F0)	42 or 43	0, 0.095, or 950	0.1	100	Reproductive effects	Decreased male/female sex ratio (percentage of males) (F1)	Ishihara et al., 2007

Table 2-7. Animal bioassay studies selected for noncancer dose-response modeling (continued)

Species/ strain	Exposure protocol	Sex (exposure group)	*n*	Average daily dose levels (ng/kg-day)	NOAEL (ng/kg-day)	LOAEL (ng/kg-day)	Endpoint(s) examined	LOAEL/NOAEL Endpoint(s)	Reference
Reproductive toxicity studies (continued)									
Rat/Wistar albino	Olive oil gavage (daily for 45 days)	M	6	0, 1, 10, or 100	None	1	Reproductive effects	Reduced sperm production, decreased reproductive organ weights	Latchoumy-candane and Mathur, 2002 (and related Latchoumy-candane et al., 2002a, b, 2003)
Rat/Sprague-Dawley	Daily dietary exposure (3 generations)	F and M, (F0) F and M, (F1 and F2)	10–32 (F0) 22 (F1) 28 (F2)	0, 1, 10, or 100	1	10	Reproductive and developmental effects	Decrease in fertility, decrease in the number of live pups, decrease in gestational survival; decrease in postnatal survival, decreased postnatal body weight in one or more generations	Murray et al., 1979
Monkey/ Rhesus	Daily dietary exposure (4 years)	F	8	0, 0.15, or 0.67	None	0.15	Reproductive effects	Increased incidence of endometriosis (disease ranged from moderate to severe)	Rier et al., 1993, 1995
Rat/Sprague-Dawley	Maternal corn oil gavage (weekly on GD 14 and 21; PND 7 and 14) Offspring corn oil gavage (weekly for 11 months)	F (F0) F (F1)	3 (F0) 10 (F1)	0, 0.14, 0.71, 7.14, or 28.6	0.14	0.71	Reproductive effects	Decrease serum estradiol levels (F1)	Shi et al., 2007

Table 2-7. Animal bioassay studies selected for noncancer dose-response modeling (continued)

Species/ strain	Exposure protocol	Sex (exposure group)	*n*	Average daily dose levels (ng/kg-day)	NOAEL (ng/kg-day)	LOAEL (ng/kg-day)	Endpoint(s) examined	LOAEL/NOAEL Endpoint(s)	Reference
Reproductive toxicity studies (continued)									
Rhesus monkey/ Cynomolgus	Fed gelatin capsules (5 days/week for 12 months)	F	6 (treatment) 5 (controls)	0, 0.71, 3.57, or 17.86	17.86	None	Endometriosis effects	Increased endometrial implant survival, increased maximum and minimum implant diameters, growth regulatory cytokine dysregulation	Yang et al., 2000
Developmental toxicity studies									
Rat/Harlan Sprague-Dawley	Corn oil gavage (GD 10−16	F (F0)	80−88 (F1)	0, 25, or 100	None	25	Developmental effects	Decreased preference in the consumption of 0.25% saccharin solution (F1)	Amin et al., 2000
Rat/CRL:WI (Han)	Maternal daily dietary exposure for an estimated 20 weeks (12 weeks prior to mating through parturition)	F (F0) M (F1)	65 (F0 treatments) 75 (F0 controls) at study initiation; following interim sacrifice ~30 animals were allowed to litter; F1 on PND 21 was ~7	0, 2.4, 8, or 46	None	2.4 (maternal exposure)	Reproductive and developmental effects	Delayed BPS (F1)	Bell et al., 2007a

Table 2-7. Animal bioassay studies selected for noncancer dose-response modeling (continued)

Species/ strain	Exposure protocol	Sex (exposure group)	*n*	Average daily dose levels (ng/kg-day)	NOAEL (ng/kg-day)	LOAEL (ng/kg-day)	Endpoint(s) examined	LOAEL/NOAEL Endpoint(s)	Reference
Developmental toxicity studies (continued)									
Rat/Sprague-Dawley	Maternal corn oil gavage (GD 14 and 21; PND 7 and 14) Offspring corn oil gavage (weekly for 8 months)	F (F0 and F1)	2 or 3 (F0) 7 (F1)	0, 7.14, or 28.6	None	7.14	Developmental effects	Decreased serum estradiol levels (F1)	Franczak et al., 2006
Rat/Sprague-Dawley	Maternal single corn oil gavage (GD 8) Offspring exposed during gestation and lactation (35 days)	F (F0) F and M (F1)	12 (F0) 50 or 60 (F1)	0, 20, 60, or 180	None	20 (maternal exposure)	Developmental effects	Abrogation of sexually dimorphic neuro-behavioral responses (F1)	Hojo et al., 2002 (and related Zareba et al., 2002)
Rat/ Han/Wistar and Long-Evans	Maternal single corn oil gavage (GD 15)	F (F0) F and M (F1)	4 to 8 (F0) 3F/3M per treatment group (F1)	0, 30, 100, 300, or 1,000	None	30 (maternal exposure)	Developmental effects	Reduced mesiodistal length of the lower third molar (F1)	Kattainen et al., 2001

Table 2-7. Animal bioassay studies selected for noncancer dose-response modeling (continued)

Species/ strain	Exposure protocol	Sex (exposure group)	*n*	Average daily dose levels (ng/kg-day)	NOAEL (ng/kg-day)	LOAEL (ng/kg-day)	Endpoint(s) examined	LOAEL/NOAEL Endpoint(s)	Reference
Developmental toxicity studies (continued)									
Mouse/ C57BL/6J, BALB/cByJ, A/J, CBA/J, C3H/HeJ, and C57BL/10J	Maternal single corn oil gavage (GD 13)	F (F0) F and M (F1a, b, c)	Dams not specified (F0); 23−36 (F1a); 4−5 (F1b); 107−110 (F1c)	0, 10, 100, or 1,000	None	10 (maternal exposure)	Developmental effects	Variation in M1 morphology in C57BL/10J males and females (F1a); decreased mandible shape and size in C3H/HeJ males (F1b); variation in molar shape in C3H/HeJ males (F1c)	Keller et al., 2007, 2008a, b
Mouse/ddY	Maternal olive oil gavage (weekly for 8 weeks prior to mating)	F (F0) M (F1)	7 (F0) 3 (F1 immuno-cytochemical analysis) 6 (F1 cell number count)	0, 0.7, or 70	None	0.7 (LOEL) (maternal exposure)	Neurotoxicity	Decreased serotonin-immunoreactive neurons in raphe nuclei of male offspring (F1)	Kuchiiwa et al., 2002
Mouse/NIH (pregnant and pseudo-pregnant)	Maternal sesame oil gavage daily for 8 days (GD 1−8)	F	10	0, 2, 50, or 100	None	2	Developmental effects	Decreased progesterone and increased serum estradiol levels	Li et al., 2006
Rat/Holtzman	Maternal single olive oil gavage (GD 18)	F (F0 and F1)	4−7 (F0 and F1)	0, 20, 60, or 180	None	20 (maternal exposure)	Behavioral effects	Decreased training responses (F1)	Markowski et al., 2001
Rat/Line C	Maternal single corn oil gavage (GD 15)	F (F0) F and M (F1)	24−32 (treatment) 12−48 (controls)	0, 30, 100, 300, or 1,000	None	30 (maternal exposure)	Developmental effects	Increase in dental caries (F1)	Miettinen et al., 2006

Table 2-7. Animal bioassay studies selected for noncancer dose-response modeling (continued)

Species/ strain	Exposure protocol	Sex (exposure group)	*n*	Average daily dose levels (ng/kg-day)	NOAEL (ng/kg-day)	LOAEL (ng/kg-day)	Endpoint(s) examined	LOAEL/NOAEL Endpoint(s)	Reference
Developmental toxicity studies (continued)									
Rat/Holtzman	Maternal single corn oil gavage (GD 15)	F (F0) M (F1)	Not specified (F0) 5 males and 3 females (F1)	0, 12.5, 50, 200, or 800	800 (maternal exposure)	None	Immunotoxicity	Decreased spleen cellularity (F1)	Nohara et al., 2000
Rat/Holtzman	Maternal single corn oil gavage (GD 15)	F (F0) M (F1)	6 (F0) 5 males and 3 females (F1)	0, 12.5, 50, 200, or 800	12.5 (maternal exposure)	50 (maternal exposure)	Developmental effects	Decreased anogenital distance (F1)	Ohsako et al., 2001
Rat/Harlan Sprague-Dawley	Maternal corn oil gavage (GD 10−16	F(F0)	~4 (F0); 80−88 (F1)	0, 25, or 100	Not determined	Not determined	Developmental effects	Facilitatory effect on radial arm maze learning (F1)	Schantz et al., 1996
Rat/Sprague-Dawley	Maternal corn oil gavage (GD 10−16)	F and M (F1)	~15 (F0); 5−9 (F1)	0, 25, or 100	25	100	Developmental effects	Decreased thymus weight	Seo et al., 1995
Rat/TCDD-resistant Han/Wistar bred with TCDD-sensitive Long-Evans	Maternal corn oil gavage (GD 15)	F (F0) M (F1)	5−8 (F0)	0, 30, 100, 300, or 1,000	100	300	Reproductive effects	Reduction in daily sperm production and cauda epididymal sperm reserves	Simanainen et al., 2004
Mouse/C57/6 NCji	Maternal drinking water exposure (daily for 17-day lactational period)	F (F0) F and M (F1)	8 (F0) Not specified (F1)	0, 1.14, or 11.3	1.14 (NOEL) (maternal exposure)	11.3 (LOEL) (maternal exposure)	Immunotoxicity	Increased susceptibility to *Listeria* (F1 males and females); increase in thymic CD4+ cells (F1 males); decreased spleen weight (F1 males)	Sugita-Konishi et al., 2003

Table 2-7. Animal bioassay studies selected for noncancer dose-response modeling (continued)

Species/ strain	Exposure protocol	Sex (exposure group)	*n*	Average daily dose levels (ng/kg-day)	NOAEL (ng/kg-day)	LOAEL (ng/kg-day)	Endpoint(s) examined	LOAEL/NOAEL Endpoint(s)	Reference
Acute toxicity studies									
Mouse/B6C3F1	Corn oil gavage (single exposure)	F	20	0, 1, 5, 10, 50, 100, or 6,000	5	10	Immunotoxicity	Increased mortality from influenza infection 7 days after a single TCDD exposure	Burleson et al., 1996
Rat/Long-Evans	Corn oil gavage (4 consecutive days)	F	14, 6, 12, 6, 6, 6, 6, 6, 6, and 4, respectively in control and treated groups	0, 0.1, 3, 10, 30, 100, 300, 1,000, 3,000, or 10,000	30	100	Thyroid effects	Reduction in serum T4 levels	Crofton et al., 2005
Rat/Sprague-Dawley	Corn oil gavage (single dose)	F	4 (treated); 9 (control)	0, 0.6, 2, 4, 20, 60, 200, 600, 2,000, 5,000, or 20,000	0.6 (NOEL)	2 (LOEL)	Enzyme induction	Increased benzo(a)pyrene hydroxylase (BPH)	Kitchin and Woods, 1979
Rat/Sprague-Dawley	Corn oil dose via oral gastric intubation (single dose)	F	10	0, 3, 10, 30, 100, 300, 1,000, 3,000, 10,000, or 30,000	3	10	Hormonal effects	Increased serum FSH	Li et al., 1997

Table 2-7. Animal bioassay studies selected for noncancer dose-response modeling (continued)

Species/ strain	Exposure protocol	Sex (exposure group)	*n*	Average daily dose levels (ng/kg-day)	NOAEL (ng/kg-day)	LOAEL (ng/kg-day)	Endpoint(s) examined	LOAEL/NOAEL Endpoint(s)	Reference
Acute toxicity studies (continued)									
Rat/Sprague-Dawley	Corn oil gavage or TCDD-contaminated soil (single dose)	F	6	0, 15, 40, 100, 200, 500, 1,000, 2,000, or 5,000 in corn oil 0, 15, 44, 100, 220, 500, 1,100, 2,000, or 5,500 in contaminated soil	None	15 (LOEL)	Enzyme induction	Induction of arylhyrocarbon hydroxylase (at low dose in both treatment protocols)	Lucier et al., 1986
Mouse/ B6C3F1 (BALB/c (C57BL/6N (and DBA2	Corn oil gavage (single dose)	M, F	10−40	0, 5, 20, 100, or 500	500	None	Mortality and body weight changes	No increased mortality of virus-infected mice or treatment-related changes in body weight	Nohara et al., 2002
Rat/TCDD-resistant Han/Wistar bred; TCDD-sensitive Long-Evans	Corn oil gavage (single dose)	M, F	9−11	30–100,000	100	300	General toxicological endpoints, organ weights, dental defects	Reduction in serum T4 levels	Simanainen et al., 2002
Rat/TCDD-resistant Han/Wistar bred with TCDD-sensitive Long-Evans	Corn oil gavage (single dose)	M, F	5−6	Line A: 30−3,000,000 Line B: 30−1,000,000 Line C: 30−100,000	100	300	General toxicological endpoints, organ weights, dental defects	Decreased thymus weight	Simanainen et al., 2003

Table 2-7. Animal bioassay studies selected for noncancer dose-response modeling (continued)

Species/ strain	Exposure protocol	Sex (exposure group)	*n*	Average daily dose levels (ng/kg-day)	NOAEL (ng/kg-day)	LOAEL (ng/kg-day)	Endpoint(s) examined	LOAEL/NOAEL Endpoint(s)	Reference
Acute toxicity studies (continued)									
Mouse/ C57BL/6N CYP1A2 (+/+) wild-type	Corn oil gavage (single dose)	F	Not specified	0, 30, 100, 300, 1000, 3000, or 10,000	300	1,000	Immunotoxicity	Decreased antibody response to SRBCs	Smialowicz et al., 2004
Rat/Sprague-Dawley	Corn oil gavage (single dose)	F	5−15	0, 0.05, 0.1, 1, 10, 100, 1,000, or 10,000	0.1 (NOEL)	1 (LOEL)	Liver effects	Increase in hepatic EROD activity and CYP1A1 mRNA levels	Vanden Heuvel et al., 1994
Subchronic toxicity studies									
Rat/Sprague-Dawley	Corn oil gavage (daily for 28 days)	F	5	0, 2.5, 25, 250, or 1,000	250	1,000	Body and organ weight changes	Decreased body weight, increased relative liver weight and related biochemical changes, decreased relative thymus weight	Chu et al., 2001
Rat/Sprague-Dawley	Corn oil gavage (daily for 28 days)	F	5	0, 2.5, 25, 250, or 1,000	2.5	25	Liver effects	Alterations in thyroid, thymus, and liver histopathology	Chu et al., 2007
Guinea pig/ Hartley	Daily dietary exposure (90 days)	M, F	10/sex	0, 0.12, 0.61, 4.9, or 26 (males); 0, 0.12, 0.68, 4.86, or 31 (females)	0.61	4.9	Body and organ weight changes	Decreased body weight (male and females); increased relative liver weights (males); decreased relative thymus weight (males)	DeCaprio et al., 1986
Mice/B6C3F1	Corn oil gavage (5 days/week for 13 weeks)	F	5	0, 1.07, 3.21, 10.7, 32.1, or 107	None	1.07 (LOEL)	Body and organ weight changes; enzyme induction	Increased EROD, ACOH and phosphotyrosyl proteins at all doses	Devito et al., 1994

Table 2-7. Animal bioassay studies selected for noncancer dose-response modeling (continued)

Species/ strain	Exposure protocol	Sex (exposure group)	*n*	Average daily dose levels (ng/kg-day)	NOAEL (ng/kg-day)	LOAEL (ng/kg-day)	Endpoint(s) examined	LOAEL/NOAEL Endpoint(s)	Reference
Subchronic toxicity studies (continued)									
Rat/Iva:SIV 50-Sprague-Dawley	Daily dietary exposure (13 weeks)	M, F	6	0, 20, 200, or 2,000	None	20	Liver effects	Reduced hepatic vitamin A levels	Fattore et al., 2000
	Daily dietary exposure (13 weeks)	M, F	6	0 or 200					
	Daily dietary exposure (13 weeks)	M, F	6	0, 200, or 1,000					
	Daily dietary exposure (13 weeks, 26, and 39 weeks)	F	6	0 or 100					
Rat/Sprague-Dawley	Gavage loading/ maintenance doses (every 4 days for 14 days)	M, F	6	0, 0.55, 307, or 1,607	0.57	327	Body and liver weight changes; hepatic cell proliferation	Increased absolute and relative liver weight	Fox et al., 1993
Mouse/ B6C3F1	Corn oil gavage (5 days/week for 13 weeks)	F	Not specified	0, 0.32, 1.07, 10.7, or 107	None	0.32 (LOEL)	Brain effects	Induction of biomarkers of oxidative stress at all doses	Hassoun et al., 1998
Rat/Harlan Sprague-Dawley	Corn oil gavage (5 days/week for 13 weeks)	F	6	0, 2.14, 7.14, 15.7, 32.9, or 71.4	None	2.14 (LOEL)	Liver and brain effects	Induction of biomarkers of oxidative stress at all doses in liver and brain	Hassoun et al., 2000

Table 2-7. Animal bioassay studies selected for noncancer dose-response modeling (continued)

Species/ strain	Exposure protocol	Sex (exposure group)	*n*	Average daily dose levels (ng/kg-day)	NOAEL (ng/kg-day)	LOAEL (ng/kg-day)	Endpoint(s) examined	LOAEL/NOAEL Endpoint(s)	Reference
Subchronic toxicity studies (continued)									
Rat/Harlan Sprague-Dawley	Corn oil gavage (5 days/week for 13 weeks)	F	12	0, 7.14, 15.7, or 32.9	None	7.14 (LOEL)	Brain effects	Induction of biomarkers of oxidative stress at all doses	Hassoun et al., 2003
Rat/Sprague-Dawley	Corn oil gavage (5 days/week for 13 weeks)	M, F	12	0, 0.71, 7.14, 71.4, or 714	7.14	71.4	Liver effects, body weight changes, and hematologic and clinical effects	Reduced body weight and food consumption, slight liver degeneration, lymphoid depletion, increased urinary porphyrins and delta aminolevulinic acid, increased serum alkaline phosphatase and bilirubin	Kociba et al., 1976
Rat/F344	Corn oil gavage (2 days/week for 28 days)	F	3	0, 0.71, 7.14, or 71.4	None	0.71 (LOEL)	Clinical signs and histopathology	Decreased Cx32 plaque number and area in the liver	Mally and Chipman, 2002
Mouse/ B6C3F1	Corn oil gavage (5 days/week for 13 weeks)	F	Not specified	0, 0.11, 0.32, 1.07, 10.7, or 107.14	1.07 (NOEL)	10.7 (LOEL)	Liver, lung, kidney, and spleen effects	Increased hepatic superoxide anion	Slezak et al., 2000
Mouse/ B6C3F1	Corn oil gavage (5 days/week for 13 weeks)	F	8−15	0, 1.07, 10.7, 107, or 321	None	1.07	Immunotoxicity and organ weight	Reduced antibody response to SRBC, increased relative liver weight	Smialowicz et al., 2008
Rat/Sprague-Dawley	TCDD in diet (13 weeks)	F	8	0, 14, 26, 47, 320, or 1,024	None	14	Multiple end-points	Decreased absolute and relative thymus weights, decreased liver retinoid levels	Van Birgelen 1995a, b

Table 2-7. Animal bioassay studies selected for noncancer dose-response modeling (continued)

Species/ strain	Exposure protocol	Sex (exposure group)	*n*	Average daily dose levels (ng/kg-day)	NOAEL (ng/kg-day)	LOAEL (ng/kg-day)	Endpoint(s) examined	LOAEL/NOAEL Endpoint(s)	Reference
Subchronic toxicity studies (continued)									
Guinea pig/ Hartley	Corn oil gavage (weekly for 8 weeks)	F	10	0, 1.14, 5.71, 28.6, or 143	1.14	5.71	Immunotoxicity	Decreased total leukocytes and lymphocyte count, decreased absolute thymus and weight, increase in primary serum tetanus antitoxin	Vos et al., 1973
Mouse/ B6C3F1	Corn oil gavage (daily for 14 days)	F	6−8	0, 10, 50, 100, 500, 1,000, or 2,000	None	10	Immunotoxicity	Reduction of serum complement activity	White et al., 1986
Chronic toxicity studies									
Rat/CD-COBS	Corn oil gavage (weekly for 45 weeks)	F	4	0, 1.43, 14.3, or 143	None	1.43	Hepatic porphyria	Increased urinary porphyrin excretion	Cantoni et al., 1981
Rat/Sprague-Dawley	Loading/ maintenance dose (every 3 days for different durations up to 128 days)	F	5	0, 0.85, 3.4, 13.6, 54.3, or 217 (28-day duration)	54.3 (28-day duration)	217 (28-day duration)	Body weight changes and changes in PEPCK activity and IGF-I levels	Decreased body weight, decreased PEPCK activity, and reduced IGF-I levels	Croutch et al., 2005
Rat/Sprague-Dawley	Corn oil gavage (5 days/week for 30 weeks)	F	6	0, 2.14, 7.14, 15.7, 32.9, or 71.4	None	2.14 (LOEL)	Brain effects	Induction of biomarkers of oxidative stress at all doses	Hassoun et al., 2002

Table 2-7. Animal bioassay studies selected for noncancer dose-response modeling (continued)

Species/ strain	Exposure protocol	Sex (exposure group)	*n*	Average daily dose levels (ng/kg-day)	NOAEL (ng/kg-day)	LOAEL (ng/kg-day)	Endpoint(s) examined	LOAEL/NOAEL Endpoint(s)	Reference
Chronic toxicity studies (continued)									
Rat/Sprague-Dawley	Daily dietary exposure (2 years)	M, F	50	0, 1, 10, or 100	1	10	Multiple endpoints measured	Increased urinary porphyrins, hepatocellular nodules, and focal alveolar hyperplasia	Kociba et al., 1978
Rat/Sprague-Dawley	Biweekly gavage (30 weeks)	F	9	0, 3.5, 10.7, 35, or 125	10.7	35	Body and organ weight changes, clinical chemistry, hepatocellular proliferation	Increased relative liver weight	Maronpot et al., 1993
Mouse/ B6C3F1; Rat/Osborne Mendel	Corn oil gavage (2 days/week for 104 weeks)	M, F	50	0, 1.4, 7.1, or 71 for rats and male mice; 0, 5.7, 28.6, or 286 for female mice	None	1.4	Liver and body weight changes	Increased incidences of liver lesions in mice (males and females)	NTP, 1982
Rat/Sprague-Dawley	Corn oil gavage (5 days/week for 105 weeks)	F	53	0, 2.14, 7.14, 15.7, 32.9, or 71.4	None	2.14	Liver and lung effects	Increased absolute and relative liver weights, increased incidence of hepatocellular hypertrophy, increased incidence of alveolar to bronchiolar epithelial metaplasia	NTP, 2006
Monkey/ Rhesus	Daily dietary exposure (4 years)	F	8	0, 0.15, or 0.67	None	0.15	General toxicological endpoints and reproductive effects	Elevated serum triglycerides and total lipids	Rier et al., 2001a, b

Table 2-7. Animal bioassay studies selected for noncancer dose-response modeling (continued)

Species/ strain	Exposure protocol	Sex (exposure group)	*n*	Average daily dose levels (ng/kg-day)	NOAEL (ng/kg-day)	LOAEL (ng/kg-day)	Endpoint(s) examined	LOAEL/NOAEL Endpoint(s)	Reference
Chronic toxicity studies (continued)									
Rat/Sprague-Dawley	Biweekly gavage (30 weeks)	F	9	0, 3.5, 10.7, 35, or 125	None	3.5 (LOEL)	EGFR kinetics and auto-phosphorylation, hepatocellular proliferation	Decrease in EGFR maximum binding capacity	Sewall et al., 1993
Rat/Sprague-Dawley	Biweekly gavage (30 weeks)	F	9	0, 0.1, 0.35, 1, 3.5, 10.7, 35, or 125	10.7	35	Thyroid function	Decreased serum T^4 levels	Sewall et al., 1995
Mouse/Swiss/ H/Riop	Sunflower oil gavage (weekly for 1 year)	M	38−44	0, 1, 100, or 1,000	None	1	Skin effects	Dermal amyloidosis and skin lesions	Toth et al., 1979

ND = not determined.

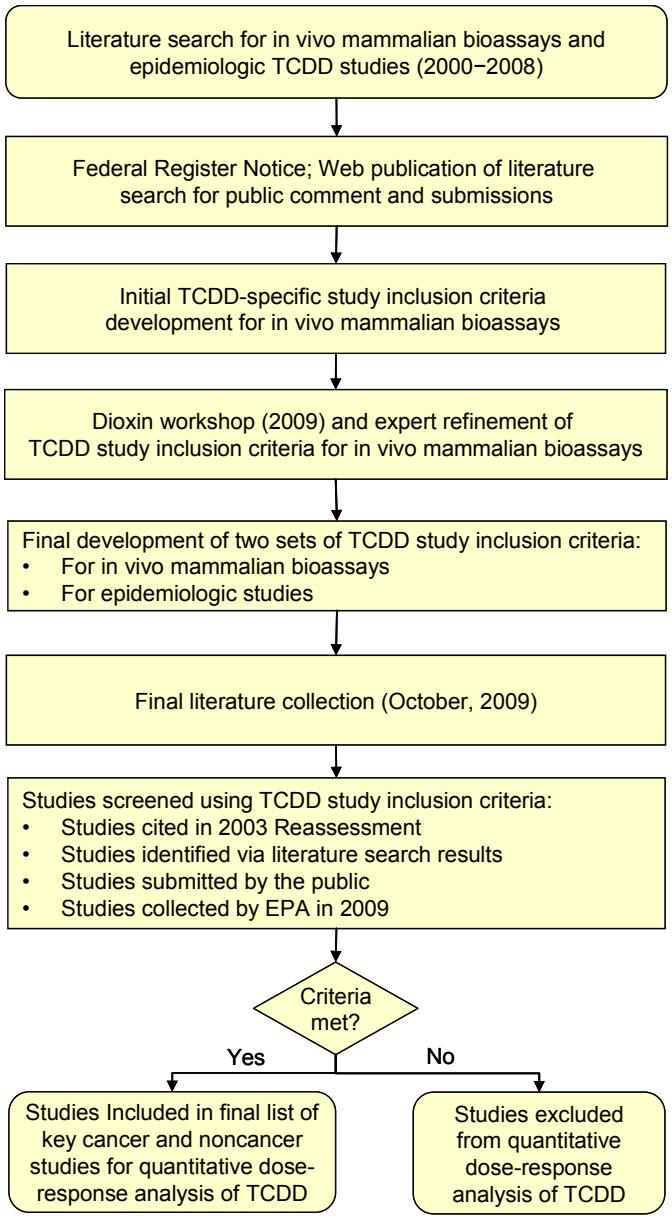


Figure 2-1. EPA's process to select and identify in vivo mammalian and epidemiologic studies for use in the dose-response analysis of TCDD. EPA first conducted a literature search to identify studies published since the 2003 Reassessment. Results were published and additional study submissions were accepted from the public. Next EPA developed TCDD-specific study inclusion criteria for in vivo mammalian studies and held a Dioxin Workshop where these criteria were discussed and refined. Third, EPA developed two final sets of study inclusion criteria, one for in vivo mammalian studies and another for epidemiologic studies. Finally, EPA applied these two sets of criteria to all studies from the literature search, public submissions, 2003 Reassessment, and additional studies identified by EPA after the Dioxin Workshop through October 2009. The studies that met these criteria formed a list of key studies for EPA's consideration in TCDD dose-response assessment.

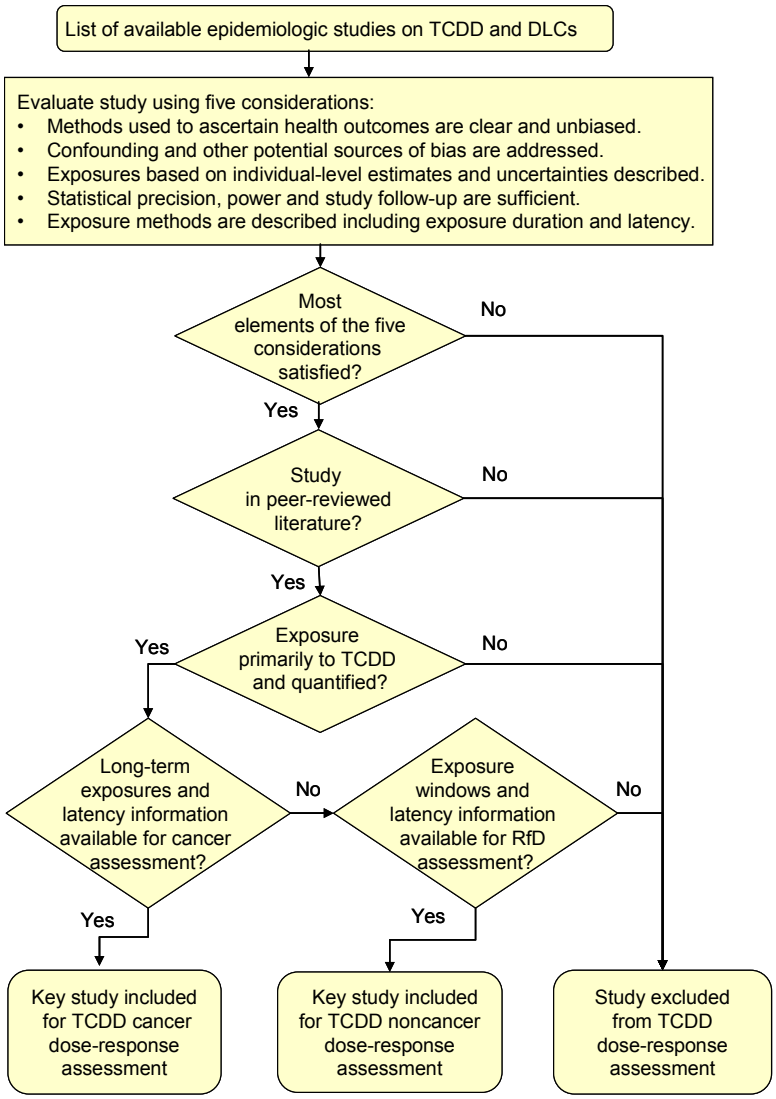


Figure 2-2. EPA's process to evaluate available epidemiologic studies using study inclusion criteria for use in the dose-response analysis of TCDD. EPA applied its TCDD-specific epidemiologic study inclusion criteria to all studies published on TCDD and DLCs. The studies were initially evaluated using five considerations regarded as providing the most relevant kind of information needed for quantitative human health risk analyses. For each study that satisfied most of these considerations and was published in the peer-reviewed literature, EPA then examined whether the exposures were primarily to TCDD and if the TCDD exposures could be quantified so that dose-response analyses could be conducted. Finally, EPA required that the effective dose and oral exposure be estimable: (1) for cancer, information is required on long-term exposures, (2) for noncancer, information is required regarding the appropriate time window of exposure that is relevant for a specific, nonfatal health endpoint, and (3) for all endpoints, the latency period between TCDD exposure and the onset of the effect is needed. Only studies meeting these criteria were included in EPA's TCDD dose-response analysis.

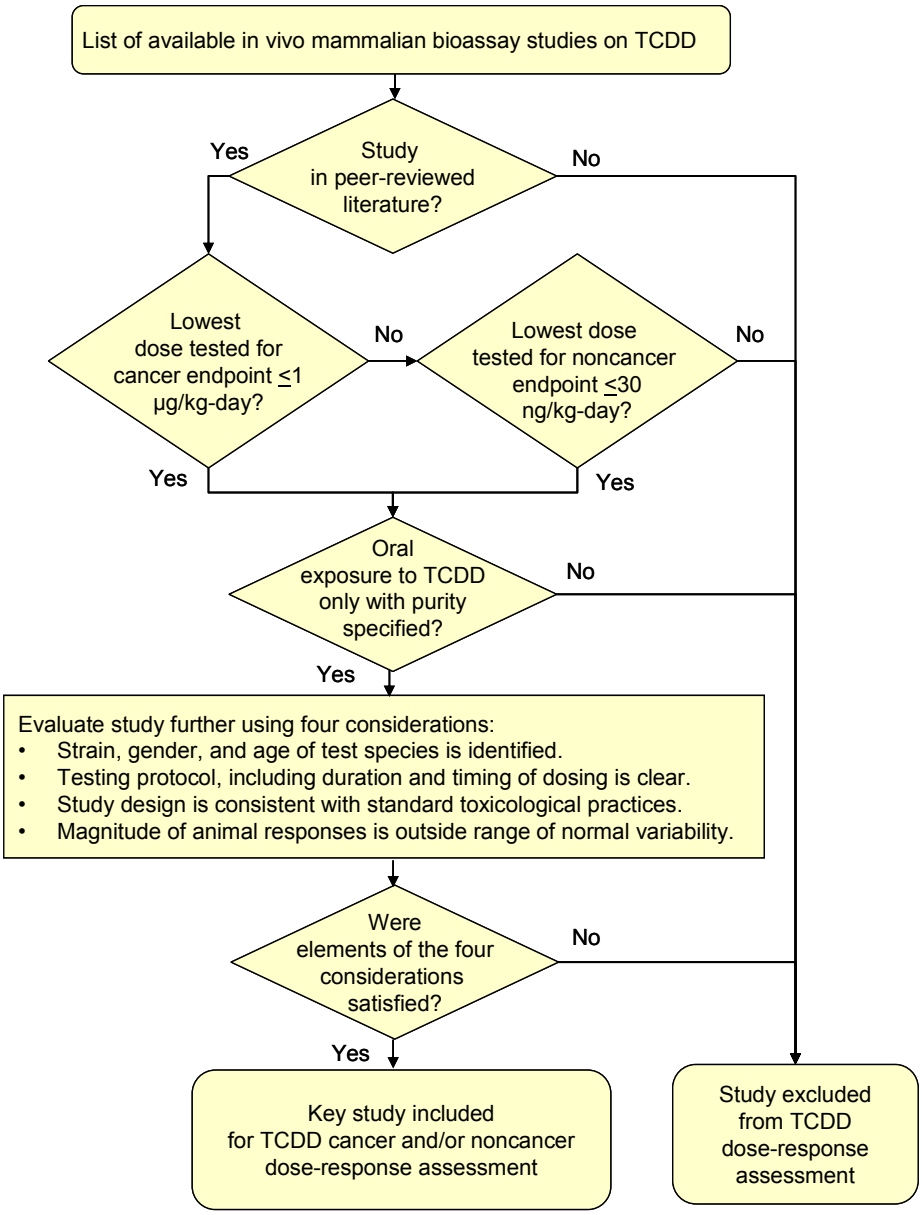


Figure 2-3. EPA's process to evaluate available animal bioassay studies using study inclusion criteria for use in the dose-response analysis of TCDD. EPA evaluated all available in vivo mammalian bioassay studies on TCDD. Studies had to be published in the peer-reviewed literature. Next, to ensure working in the low-dose range for TCDD dose-response analysis, EPA applied dose requirements to the lowest tested average daily doses in each study, with specific requirements for cancer (≤1 µg/kg-day) and noncancer (≤30 ng/kg-day) studies. Third, EPA required that the animals were exposed via the oral route to only TCDD and that the purity of the TCDD was specified. Finally, the studies were evaluated using four considerations regarded as providing the most relevant kind of information needed for quantitative human health risk analyses from animal bioassay data. Only studies meeting all of these criteria and considerations were included in EPA's TCDD dose-response analysis.

3. THE USE OF TOXICOKINETICS IN THE DOSE-RESPONSE MODELING FOR CANCER AND NONCANCER ENDPOINTS

A key recommendation from the National Academy of Sciences (NAS) for improving the 2003 Reassessment was that U.S. Environmental Protection Agency (EPA) should justify its approaches to dose-response modeling for cancer and noncancer endpoints. Further, the NAS suggested that EPA incorporate the most up-to-date and relevant state of the science for 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) dose-response assessment.

While EPA believes that at the time of its release, the 2003 Reassessment offered a substantial improvement over the general state-of-the-science regarding dose-response modeling, EPA agrees that the justification of the approaches to dose-response modeling can be improved and the methodologies updated to reflect the most current EPA practices and science. In Section 3, EPA describes the use of toxicokinetic (TK) information in the dose-response modeling assessment of TCDD. Section 3.1 summarizes the NAS comments regarding the use of TK in the dose-response approaches for TCDD. Section 3.2 overviews EPA's responses to the NAS comments. Section 3.3 discusses TCDD kinetics, including TK models developed to simulate disposition of this compound in rodents and humans (see Section 3.3.4), alternative measures of dose that could be used in a TCDD dose-response analysis and uncertainties in the TCDD dose estimates (see Section 3.3.5). Sections 4 and 5 of this document incorporate the TK information into noncancer and cancer dose response modeling, respectively.

3.1. SUMMARY OF NAS COMMENTS ON THE USE OF TOXICOKINETICS IN DOSE-RESPONSE MODELING APPROACHES FOR TCDD

The NAS commented on the appropriate use of TK models in dose-response modeling for TCDD. Specifically, the committee requested that EPA consider using such models to provide refined estimates of dose, for example, as the underlying science and predictive capabilities of these models improved.

> [Discussing Kinetic models]…the committee encourages further development and use of these models as data become available to validate and further develop them (NAS, 2006a, p. 59).

Although the NAS basically agreed with EPA's use of body burden as a dose metric in the 2003 Reassessment (e.g., see NAS, 2006a, p. 7), the NAS was concerned about the limitations of first order kinetic models, such as the one used in the 2003 Reassessment, to estimate TCDD body burdens.

> TCDD, other dioxins, and DLCs act as potent inducers of CYP, a property that can affect both the hepatic sequestration of these compounds and their half-lives. Hepatic sequestration of dioxin may influence the quantitative extrapolation of the rodent liver tumor results because the body-burden distribution pattern in highly dosed rats would differ from the corresponding distribution in humans subject to background levels of exposure. EPA should consider the possible quantitative influence of dose-dependent toxicokinetics on the interpretation of animal toxicological data (NAS, 2006a, p. 129).

The NAS also asked EPA to evaluate the impact of kinetic uncertainty and variability on dose-response assessment. The NAS committee asked EPA to use TK models to examine both interspecies and human interindividual differences in the disposition of TCDD, which would better justify EPA dose-response modeling choices.

> The Reassessment does not adequately consider the use of a PBPK model to define species differences in tissue distribution in relation to total body burden for either cancer or noncancer end points (NAS, 2006a, p. 62).
>
> EPA …should consider physiologically based pharmacokinetic modeling as a means to adjust for differences in body fat composition and for other differences between rodents and humans (NAS, 2006a, p. 10).
>
> The Reassessment does not provide details about the magnitudes of the various uncertainties surrounding the decisions EPA makes in relation to dose metrics (e.g., the impact of species differences in percentage of body fat on the steady-state concentrations present in nonadipose tissues). The committee recommends that EPA use simple PBPK models to define the magnitude of any differences between humans and rodents in the relationship between total body burden at steady-state concentrations (as calculated from the intake, half-life, bioavailability) and tissue concentrations. The same model could be used to explore human variability in kinetics in relation to elimination half-life. EPA should modify the estimated human equivalent intakes when necessary (NAS, 2006a, p. 73).

Finally, the NAS asked EPA to use TK considerations to better justify its choice of dose metric.

> EPA makes a number of assumptions about the appropriate dose metric and mathematical functions to use in the Reassessment's dose-response analysis … but does not adequately comment on the extent to which each of these assumptions could affect the resulting risk estimates…EPA did not quantitatively describe how this particular selection affected its estimates of exposure and therefore provided no overall quantitative perspective on the relative importance of the selection (NAS, 2006a, p. 51).

3.2. OVERVIEW OF EPA'S RESPONSE TO THE NAS COMMENTS ON THE USE OF TOXICOKENTICS IN DOSE-RESPONSE MODELING APPROACHES FOR TCDD

In response to the NAS recommendations regarding TCDD kinetics and choice of dose metrics, this document presents an in depth evaluation of TCDD TK models, exploring their differences and commonalities and their possible application for the derivation of dose metrics relevant to TCDD. Initially, EPA discusses the application of first order kinetics to estimate body burden as a dose metric for TCDD. This first order kinetic model is used to predict TCDD body burden for all of the studies identified as Key Studies (see Section 2.4); this model uses a constant half-life to simulate the elimination of TCDD from the body. However, given the observed data indicating early influence of cytochrome P450 1A2 (CYP1A2) induction and binding to TCDD in the liver and later redistribution of TCDD to fat tissue, the use of a constant half-life for TCDD clearance following long term or chronic TCDD exposure is not biologically supported. Therefore, using half-life estimates based on observed terminal steady state levels of TCDD will not account for the possibility of an accelerated dose-dependent clearance of the chemical during early stages following elevated TCDD exposures. The biological processes leading to dose-dependent TCDD excretion are better described using physiologically based pharmacokinetic (PBPK) models than by simple first order kinetic models. Additionally, as part of its preparation for developing this document, EPA evaluated recent TCDD kinetic studies as NAS advocated. Although the NAS agreed with continued use of body burden metric as the dose metric of choice, EPA believes that the state-of-the-practice has advanced sufficiently to justify the consideration of alternative dose metrics (other than administered dose) based on an application of a physiologically-based TK model.

EPA identified a number of advances in the overall scientific understanding of TCDD disposition; many of these are documented in a summary discussion introducing the section on TCDD kinetics (see Section 3.3). The increased understanding warranted an evaluation of current kinetic modeling of TCDD to determine if the use of such models would improve the dose-response assessment for TCDD. Justification of the final PBPK model choice is detailed in Section 3.3. Through the choice of a published PBPK model to estimate dose metrics for dioxin, EPA has addressed several of the NAS concerns. The PBPK model can be applied to estimate dose metrics other than body burden that may be more directly related to response, e.g., tissue levels, serum levels, blood concentrations, or dose metrics related to TCDD-protein receptor binding. The selected PBPK model included explicit description of physiological and biochemical parameters, therefore, it can also provide an excellent tool for investigating differences in species uptake and disposition of TCDD. One of the criteria used to select a PBPK model for TCDD kinetics was the availability of both human and animal models so that differences in species uptake and disposition of TCDD can be investigated. Additionally, the PBPK model includes quantitative information that is suitable for addressing the impact of physiological (e.g., body weight [BW] or fat tissue volume), or biochemical (e.g., induction of CYP1A2) variability on overall risk of TCDD between species, in response to another area of concern in the NAS report. The sensitivity analysis and uncertainty in dose metrics derived for the risk assessment of TCDD are also presented in Section 3.3. Detailed discussion on the uncertainty in choice of PBPK model-driven dose metrics is also provided in Section 3.3.

3.3. PHARMACOKINETICS (PK) AND PK MODELING

3.3.1. PK Data and Models in TCDD Dose-Response Modeling: Overview and Scope

In general, the use of measures of internal dose in dose-response modeling is considered to be superior to that of administered dose (or uptake) because the former is more closely related to the response. The evaluation of internal dose, or dose metric, in exposed humans and other animals is facilitated by an understanding of pharmacokinetics (i.e., absorption, distribution, metabolism, and excretion). When measurements of internal dose (e.g., blood concentration, tissue concentration) are not available in animals and humans, pharmacokinetic models can be used to estimate them. The available data on the pharmacokinetics of TCDD in animals and humans have been reviewed (Van Birgelen and Van den Berg, 2000; U.S. EPA, 2003; NRC,

2006). It is evident based on these reviews and other analyses that three distinctive features of TCDD play important roles in determining its pharmacokinetic behavior, as discussed below:

TCDD is very highly lipophilic and thus is more soluble in fat or other relatively nonpolar organic media than in water. The *n*-octanol/water partition coefficient is a commonly-used measure of lipophilicity equal to the equilibrium ratio of a substance's concentration in *n*−octanol (a surrogate for biotic lipid) to the substance's concentration in water (Leo et al., 1971). For TCDD, this coefficient is on the order of 10,000,000 or more (ATSDR, 1998). It follows that the solubility of TCDD in the body's lipid fraction, i.e., the fatty portions of various tissues, including adipose, organs, and blood, is extremely high.

TCDD is very slowly metabolized compared to many other organic compounds, with an elimination half life in humans on the order of years following an initial period of distribution in the body (Carrier et al., 1995a; Michalek et al., 2002). Most laboratory animals used for toxicologic testing tend to eliminate TCDD much more quickly than people, although even in animals TCDD is eliminated much more slowly than most other chemicals.

- **TCDD induces binding proteins in the liver** that have the effect of sequestering some of the TCDD. The ability of TCDD to alter gene expression and the demonstration that the induction of CYP1A2 is responsible for hepatic TCDD sequestration suggest that both pharmacokinetic and pharmacodynamic events must be incorporated for a quantitative description of TCDD disposition (Santostefano et al., 1998). The induction of these proteins implies that TCDD tends to be eliminated more rapidly in the early years following short-term, high-level exposures than it is after those initial levels have declined. Recent efforts of pharmacokinetic modeling have supported the concentration-dependent elimination of TCDD in animals and humans (Aylward et al., 2005b; Emond et al., 2006).

Sections 3.3.2 and 3.3.3 of this section present the salient features of TCDD pharmacokinetics in animals and humans, with particular focus on mechanisms and data of relevance to interspecies and intraspecies variability. Section 3.3.4 describes the various dose metrics for the dose-response modeling of TCDD and the characteristics of pharmacokinetic models potentially useful for estimating these metrics. Finally, Sections 3.3.5 and 3.3.6 summarize the results of application of pharmacokinetic models to derive dose metrics as well as the uncertainty associated with the predictions of dose metrics used in dose-response modeling. Dose metrics derived via PBPK modeling approaches are utilized in Sections 4 and 5 of this document for noncancer and cancer TCDD dose-response modeling, respectively.

3.3.2. PK of TCDD in Animals and Humans

3.3.2.1. *Absorption and Bioavailability*

When administered via the oral route in the dissolved form, TCDD appears to be well absorbed. Animal studies indicate that oral exposure to TCDD in the diet or in an oil vehicle results in the absorption of >50% of the administered dose (Nolan et al., 1979; Olson et al., 1980). Human data from Poiger and Schlatter (1986) indicate that >87% of the oral dose (after ingestion of 105 ng [^{3}H]−2,3,7,8−TCDD [1.14 ng/kg BW] in 6 mL corn oil) was absorbed from the gastrointestinal tract. Lakshmanan et al. (1986), investigating the oral absorption of TCDD, suggested that it is absorbed primarily by the lymphatic route and transported predominantly by chylomicrons.

Oral absorption is generally less efficient when TCDD is more tightly bound in soil matrices. Based on experiments in miniature swine, Wittsiepe et al. (2007) reported an approximately 70% reduction in bioavailability when TCDD was administered in the form of contaminated soil, relative to TCDD after extraction from the same soil matrix with solvents. Working with soil from the prominent contamination site at Times Beach, Missouri, Shu et al. (1988) reported an oral bioavailability of approximately 43% based on experiments in rats. Percent dose absorbed by the dermal route is reported to be less than the oral route, whereas absorption of TCDD by the transpulmonary route appears to be efficient (see, for example, Banks et al., 1990; Banks and Birnbaum, 1991; Nessel et al., 1992; Diliberto et al., 1996; U.S. EPA, 2003; Roy et al., 2008).

3.3.2.2. *Distribution*

TCDD in systemic circulation equilibrates and moves into the tissues where it is then accumulated, bound, or eliminated. Whereas the bulk of the body tissues are expected to equilibrate in a matter of hours, the adipose tissue will approach equilibrium concentrations with blood much more slowly. Consistent with these assertions, a number of experimental and modeling studies in rats and humans have shown that TCDD has a large volume of distribution (Vd), i.e., the apparent volume in which it is distributed. The Vd corresponds to the volume of blood plus the product of internal tissue volumes and the corresponding tissue:blood partition coefficients. This parameter is a key determinant of the elimination rate of TCDD in exposed organisms. The tissue:blood partition coefficients of TCDD, in turn, are determined by the

relative solubility of TCDD in tissue and blood components (including neutral lipids, phospholipids, and water).

Column 1 in Table 3-1 presents the tissue:blood partition coefficients for TCDD (Wang et al., 1997; Emond et al., 2005). Column 3 of this table lists the physical volume of each tissue, scaled to a person weighing 60 kg. The last column shows the implications of the tissue volumes and tissue:blood partition coefficients for the effective volumes of distribution for each tissue and for the body as a whole. It can be seen that, purely on the basis of solubility space, the fat should be expected to contain about 94% of the TCDD in the body, and that the body as a whole behaves as if it is about 1,200 liters in terms of blood-equivalents (i.e., approximately 22−fold larger than its physical volume).

Maruyama et al. (2002) have published another set of tissue/blood partition coefficients for TCDD and other dioxin congeners based in part on observations of tissue concentrations measured in autopsy specimens from eight Japanese people without known unusual exposures to TCDD. Their estimates of TCDD partition coefficients seem to be rather large and variable, with a fat:blood value of 247 ± 78 (standard deviation [SD]), a liver:blood value of 9.8 ± 5.7 and a muscle:blood value of 18 ± 10.6. Depending on time of autopsy, tissue samples may not be an accurate source of information on observed, in vivo partition coefficients because weight loss is likely to occur pre and post mortem. In particular, a decline in fat stores volume could lead to an increased concentration of dioxin in fat in autopsy specimens relative to what would be observed in vivo.

The calculations shown in Table 3-1 do not include the additional amount that will be bound to induced proteins in the liver. That induction and binding will tend to increase the contribution of the liver on the effective volume of distribution (Birnbaum, 1986).

It is also of interest to point out some basic implications of the data in Table 3-1 for the expected rates of perfusion-mediated transfer of TCDD between blood and each of the organ/tissues. The rate of loss from a tissue (occurring primarily via blood flow) and the corresponding half-life can be calculated using the following equations:

$$\text{Rate constant for loss (hour}^{-1}) = \frac{\text{Blood flow (liters / hour)}}{\text{Tissue volume (liters)} \times \text{Tissue / Blood Partition Coefficent}} \quad \text{(Eq. 3-1)}$$

$$t_{1/2} \text{ for tissue perfusion loss} = \frac{\ln(2)}{\text{Rate constant for loss}}$$

$$= \frac{\ln(2) \times \text{Tissue volume (liters)} \times \text{Tissue/Blood Partition Coefficent}}{\text{Blood flow (liters/hour)}} \qquad \text{(Eq. 3-2)}$$

Because TCDD is highly lipophilic, its concentration in the aqueous portion of the blood is very small, and TCDD tends to partition from blood components into cellular membranes and tissues, probably in large part via diffusion. As a result, full equilibrium concentrations of TCDD are not attained by the end of the transit time through organs from the arterial to venous blood. For organs in which this occurs, diffusion coefficients or "permeability factors" have been estimated to assess the fractional attainment of equilibrium concentration that occurs by the time the blood leaving each organ reaches the venous circulation. Table 3-2 presents the permeability factors and implications for perfusion half-lives for TCDD, per Emond et al. (2005, 2006).

Despite the high lipid bioconcentration potential of TCDD, it does not always occur at the highest concentration in the adipose tissue (Poiger and Schlatter, 1986; Geyer et al., 1986; Abraham et al., 1988). Further, the ratios of tissue:tissue concentrations of TCDD and related compounds (e.g., the liver:adipose ratio) may not remain constant during nonsteady-state conditions. TCDD concentrations have been observed to decrease more rapidly in the liver than in adipose tissue. For example, Abraham et al. (1988) found that the liver:adipose tissue concentration ratio in female Wistar rats exposed to a subcutaneous TCDD dose of 300 ng/kg decreased from 10.3 at 1 day postexposure to 0.5 at 91 days postexposure. It should be noted that even at a ratio of 0.5, the amount of TCDD in the liver is greater than that based on lipid content of the tissue alone, consistent with the presence of hepatic TCDD binding proteins. The liver/adipose tissue ratio also was dose-dependent, such that the liver TCDD burden increased from ~11% of the administered dose at low doses (i.e., 1−10 ng/kg) to ~37% of the dose at an exposure level of 300 ng/kg. The increase in TCDD levels in liver, accompanied by a decrease in concentration in the adipose tissue, is a particular behavior to be considered in high dose to low dose extrapolations. This behavior is essentially a result of dose-dependent hepatic processes, as described below.

3.3.2.3. *Metabolism and Protein Binding*

The metabolism of TCDD is slow, particularly in humans, and it is thought to be mediated by the CYP1A2 enzyme that is inducible by TCDD (Ramsey et al., 1982; Wendling et al., 1990; Olson et al., 1994; Weber et al., 1997). The low rate of metabolism in combination with sequestration appear to account for the retention of TCDD in liver, and these processes collectively contribute to the long half-life for elimination of TCDD from the body.

Dynamic changes in TCDD binding in liver and partitioning to fat have been studied extensively in rats and mice (Diliberto et al., 1995, 2001). Figure 3-1 shows observations by Diliberto et al. (1995) of the ratio of liver concentrations to adipose tissue concentrations for mice given doses spread over a 100−fold range and studied at four different times following exposure. It can be seen that even for the lowest dose studied the liver:fat concentration ratio is higher than would be expected based on the lipid contents of the tissues (i.e., 0.06:1, corresponding to the ratio of human liver:blood and fat:blood partition coefficients; see Table 3-1). Moreover, the relative concentration in the liver consistently rises with dose, with the steepest rise observed during the first two weeks after dosing. If the distribution of TCDD were governed solely by passive partitioning into fat, there should be no such change in relative concentrations with dose. However, data presented in Figure 3-1 illustrate that at longer time points, the ratio of TCDD in the liver to TCDD in fat decreases, indicating that a redistribution of the chemical occurs as time goes on for each applied dose. The redistribution of TCDD tissue levels from liver to fat with increasing time suggests that binding of the chemical in the liver (including via induction of CYP1A2) is an important kinetic consideration at early exposure points with relatively high applied doses. At steady state levels (longer than 35 days, and low applied doses), there seems to be a tendency for TCDD to redistribute to fat tissue.

Experiments with CYP1A2 “knock-out” mice (i.e., congenic strains differing in only a single gene that is “knocked out” in one of the strains) indicate that the inducible binding of TCDD is attributable to CYP1A2 (Diliberto et al., 1997, 1999). As noted previously, this enzyme is believed to make an important contribution to metabolism of TCDD. Given the critical role of CYP1A2 induction in the kinetics of TCDD, dose-and time-dependent induction of this protein in rats has been examined and modeled (Wang et al., 1997; Santostefano et al., 1998; Emond et al., 2004, 2006). Accordingly, the amount of CYP1A2 in the liver can be

computed as the time-integrated product of inducible production and a simple first-order loss process (Wang et al., 1997):

$$\frac{dCYP_{2A1}}{dt} = S(t)K_0 - K_2 C_{A2t} \quad \text{(Eq. 3-3)}$$

where CYP_{2A1} is the concentration of the enzyme, K_2 is the rate constant for the first order loss, C_{A2t} is the concentration of CYP1A2 in the liver, K_0 is the basal rate of production of CYP1A2 in the liver, and $S(t)$ is a multiplicative stimulation factor for CYP1A2 production in the form of a Hill-type function:

$$S(t) = 1 + \frac{In_{A2}(C_{Ah\text{-}TCDD})^h}{(\mathrm{IC}_{\mathrm{A2}})^{\mathrm{h}} + (C_{Ah\text{-}TCDD})^h} \quad \text{(Eq. 3-4)}$$

where IC_{A2} corresponds to the concentration of the aryl hydrocarbon (Ah)-TCDD complex at which half of the maximum fold stimulation of CYP2A production is reached, and h, the Hill exponent, determines the curvature of the stimulation in relation to concentration of the Ah-TCDD complex at relatively low doses. A value of 0.6 as the Hill exponent has been used by Wang et al. (1997, 2000) and Emond et al. (2004, 2005, 2006), indicative of a negative cooperation, i.e., the curve is convex-upward (supralinear), depicting a faster increase in the low-dose region compared to a straight line. Additional parameters in this expression include In_{A2}, the maximum fold increase in the CYP1A2 synthesis rate over the basal rate that can occur at high levels of TCDD, and ($C_{Ah\text{-}TCDD}$), the concentration of TCDD bound to the aryl hydrocarbon receptor (AhR). This concentration in turn depends on the concentration of TCDD in the liver (C_{Lif}), the concentration of the AhR (Ah_{Li}) in liver, and the dissociation constant for the Ah-TCDD receptor complex, K_{DAh}:

$$C_{Ah\text{-}TCDD} = \frac{Ah_{Li} \times C_{Lif}}{K_{DAh} + C_{Lif}} \quad \text{(Eq. 3-5)}$$

3.3.2.4. *Elimination*

Elimination half-lives (i.e., the time taken for the concentration to be reduced to one-half of its initial level) of TCDD range from 11 days in the hamster to 2,120 days in humans (U.S. EPA, 2003). Hepatic metabolism and binding processes, fecal excretion, and accumulation in adipose tissue collectively determine the dose-dependent elimination half-lives in various species. Aylward et al. (2005a) depicted the relationship between the elimination rate versus initial level of lipid-corrected TCDD in serum for 36 people (see Figure 3-2). Even though this analysis was done using the initial TCDD level, rather than the geometric mean or midpoint level in the decline for each person, it indicated a concentration-dependency of the half-life and elimination of TCDD in exposed individuals.

3.3.2.5. *Interspecies Differences and Similarities*

Among the pharmacokinetic determinants of TCDD, some are known to vary markedly between species whereas others are not characterized sufficiently in this regard. Overall, the qualitative determinants of the body burden and elimination half-lives appear to be similar across species. Based on empirical observations for TCDD as well as with other PCDFs, Carrier et al. (1995a,b) argued that in rats, monkeys, and humans, the dose-dependent changes in the fraction contained in liver and adipose tissue follow a similar pattern across species. The authors suggested that the half-saturation body burden is around 100 ng/kg and the plateau of liver dose (as fraction of body burden) appears to occur around 1,000 ng/kg. Literature also indicates that AhR is conserved phylogenetically (Nebert et al., 1991; Fujii-Kuriyama et al., 1995; Harper et al., 2002) and is present in mammalian species, including experimental animals and humans (Roberts et al., 1985, 1986; Manchester et al., 1987; Lorenzen and Okey, 1991; Okey et al., 1994). These qualitative similarities in pharmacokinetic determinants and outcome support the use of animal data to infer general patterns of the pharmacokinetic behavior of TCDD in humans. However, quantitative differences in determinants, including physiological, physicochemical, and biochemical, need to be taken into account. Even though species-specific physiological parameters can be obtained from the literature, key data on species-specific biochemical parameters (particularly binding constants, maximal capacity, induction rates, and other parameters) are not available for humans at this time. However, these can be inferred by using a pharmacokinetic model fit to in vivo data on the rate of TCDD elimination from specific

compartments in humans (Carrier et al., 1995a, b; Emond et al., 2004, 2005, 2006; Aylward et al., 2005b).

3.3.3. PK of TCDD in Humans: Interindividual Variability

The pharmacokinetics and tissue dose of TCDD in humans can exhibit variability within a population as a function of the interindividual variability in the key determinants. If a chronic, lifetime exposure is considered to be the relevant scenario for developing guidance for human exposure standards, the key determinants of concern are clearance, binding, and temporal changes in volume of distribution. When considering the interindividual variability in pharmacokinetics and dose metrics of TCDD, it is important to recognize that higher lipid-corrected serum concentrations in highly exposed persons are associated with greater elimination rates, probably due to greater degrees of induction of CYP1A2 in the liver and possibly other related metabolic enzymes (Grassman et al., 2000; Abraham et al., 2002; Aylward et al., 2005b; Emond et al., 2006).

The interindividual variability in fat content is a critical parameter given the pharmacokinetic characteristics of TCDD (see Section 3.3.2). Both metabolic elimination and elimination via the GI tract depend on the fraction of TCDD in the body that is available outside of adipose tissue. As body fat content rises, a smaller portion of the total body TCDD will be contained in the relatively available fraction outside of fat. Because elimination of TCDD by both metabolism and fecal excretion depends on the small proportion of TCDD that exists outside of fat tissue, people with larger proportions of body fat—including many older people—will tend to require longer times to reduce TCDD levels by a given proportion than leaner people (Van der Molen et al., 1996, 1998; Rohde et al., 1999; Emond et al., 2006).

The sections that follow highlight key aspects of interindividual variability in TCDD pharmacokinetics, with an emphasis on the available data related to elimination half-lives and volume of distribution.

3.3.3.1. *Life Stage and Gender*

The influence of the variability of fat content in human population on the distribution and clearance of TCDD has been evaluated by several investigators. Figure 3-3 shows the results of

a rare experiment in which TCDD elimination via feces was measured in six people in relation to their body fat content (Rohde et al., 1999).

There are data which suggestive an inverse dependency of half life on percent body fat. Observations of TCDD elimination rates in a small number of men and women in the Seveso cohort provide a modest opportunity to compare TCDD elimination rates on this basis. Based on the partition coefficients reported by Emond et al. (2006), the elimination rates for the men in the sampled group are expected to be greater than the elimination rates in the women. Based on calculations similar to those shown in Table 3-2, and fat proportions inferred from body mass indices using the equations of Lean et al. (1996), the Seveso men studied are expected to have an overall average of about 3.92% of their TCDD body burden outside of fat, whereas the women are expected to have an average of only 2.36% outside of fat. On this basis, the TCDD elimination rates in the men are expected to be 3.92/2.36 = 1.66 times faster than the elimination rates in the women. By comparison, Michalek et al. (2002) reported observed elimination rates in men and women that result in a slightly lower ratio:

$$\frac{\text{men: } 0.111\ \text{year}^{-1} \pm 0.010\ \text{(std.error)}}{\text{women: } 0.071\ \text{year}^{-1} \pm 0.010\ \text{(std.error)}} = 1.56 \qquad \text{(Eq. 3-6)}$$

The central estimates for the elimination rates correspond to half lives of 6.5 and 9.6 years for men and women, respectively.

A further point of comparison can be derived using the observed body mass index (BMI)[8] and TCDD elimination rate of each of the male Ranch Hand military veterans, whose TCDD elimination rates were observed between 9 and 33 years after their time in Vietnam. The average BMI over that time was 29.44 (based on 287 measurements for the 97 veterans, tabulated in three periods by Michalek et al., 2002), and their average age was about 44.5 for the measurements. Based on these data, the corresponding average estimated percent body fat is 29.7% using the Lean et al. (1996) formula for men. The observed average TCDD elimination rate constant for these men for the period was 0.092 year^{-1} ± 0.004 (standard error), corresponding to a half life of 7.5 years. This half life is slightly longer than the central estimate of the half life of 6.2 years

[8]The body mass index, or BMI, is calculated as the body weight in kilograms divided by the square of the height in meters.

(i.e., ln(2)/0.111) for the smaller group of Seveso males with their slightly smaller estimated percent body fat. Figure 3-4 shows a simple plot of these data and a fitted unweighted regression line characterizing the relationship between estimated fat content and TCDD elimination rates. Variation in metabolic enzyme activities and other routes of loss is also likely to be important, but there is little human quantitative information available on these issues.

More recently, Kerger et al. (2006) estimated the slope of the relationship between half-life and age to be 0.12 years (95% confidence interval, 0.10−0.14), which corresponds to the rate of increase in TCDD half-life for each year of age. The authors speculated that although age explained most of the variance in the individual half-life trends, it was also correlated with TCDD concentration, BMI, and body fat mass. The regression model developed by these authors discriminated between the high and low TCDD exposures or concentrations. Thus, after accounting for the TCDD (concentration × age) term's effect on the slope of age, the final model for TCDD concentration ≤ 700 ppt was

$$t_{1/2} = 0.35 + 0.12 \times Age \quad \text{(Eq. 3-7)}$$

For TCDD concentration >700 ppt, the final model was:

$$t_{1/2} = 0.35 + 0.088 \times Age \quad \text{(Eq. 3-8)}$$

where $t_{1/2}$ is the half-life and Age is the age at time of subsequent sampling. Pharmacokinetic information relevant to specific age groups is presented in the sections that follow.

3.3.3.1.1. *Prenatal period.*

Data to estimate TCDD elimination rates for fetuses are not available. Levels of TCDD in fetal tissues for rats were experimentally estimated at different gestational periods and utilized in a developmental model by Emond et al. (2004). There is information on body composition that is relevant to prediction of TCDD dose to fetus. These data, summarized as part of the radiation dosimetry model of the International Commission on Radiological Protection, are consistent with the idea that early fetuses are nearly all water and less than 1% lipid, and lipid levels rise toward parity with protein near the time of normal delivery.

Bell et al. (2007b) reported that the disposition of TCDD into the fetus shows dose dependency, with a greater proportion of the dose reaching the fetus at lower doses of TCDD. Further, both CYP1A1 and CYP1A2 are highly inducible (~103−fold) in fetal liver, whereas CYP1A2 shows much lower induction (10−fold) in maternal liver. It has been speculated that this is due to the lower basal levels of CYP1A2 in fetal liver, as compared to maternal liver (Bell et al., 2007b).

3.3.3.1.2. *Infancy and childhood.*

Hattis et al. (2003) describe the general pattern of change of body fat content with age in children. Central tendency values for percent body fat begin at about 12% at birth and rise steeply to reach about 26% near the middle of the first year of life. Fat content then falls to reach a minimum of approximately 15% at 5−8 years of age, followed by a sex-dependent "adiposity rebound" that takes females to about 26% body fat while the males remain near 16−17% on average by age 20. The interindividual variability distributions about these central values are complex, as some children experience the "adiposity rebound" earlier than others, and this creates patterns that are not simply interpretable as unimodal normal distributions. Hattis et al. (2003) did find it possible to fit distributions of body fat content inferred from NHANES skin fold measures to mixtures of two normal distributions for children between age 5 and 18.

At least two groups of authors have published PBPK modeling results indicating generally more rapid clearance of TCDD in children than in adults, a trend that is consistent with the generally lower fat content of children (Kreuzer et al., 1997; Van der Molen et al., 2000; Leung et al., 2006). The rapid expansion of the adipose tissue compartment can contribute, in part, to the reduced apparent half-life in children (Clewell et al., 2004). This reduction may also be due to varying rates of metabolism and/or fecal lipid excretion (Abraham et al., 1996; Kerger et al., 2007).

Furthermore, very young children have different modes and quantities of exposure compared to adults. Lakind et al. (2000) characterize distributions of milk intake for nursing infants to characterize distributions of TCDD exposure. This is also a corresponding route of loss of TCDD stores for lactating women, as described in Section 3.3.3.2 below.

3.3.3.1.3. *Adulthood and old age.*

The fraction of fat in relation to body weight in adulthood and old age can be computed as a function of the BMI and age (e.g., Lean et al., 1996):

$$\%\ Body\ Fat\ (\text{males}) = 1.33 \times BMI + 0.236 \times Age - 20.2 \quad \text{(Eq. 3-9)}$$

$$\%\ Body\ Fat\ (\text{females}) = 1.21 \times BMI + 0.262 \times Age - 6.7 \quad \text{(Eq. 3-10)}$$

The above equations are the result of analysis of data based on underwater weighing of 63 men and 84 women (age range 16.8−65.4). The salient observation with respect to TCDD for these data is that age and BMI-dependent variability in fat content have implications for the variability in TCDD elimination rates and internal dose among adults.

3.3.3.2. *Physiological States: Pregnancy and Lactation*

Data on body fat content in pregnant women at various stages of gestation (Pipe et al., 1979) have potential implications for TCDD elimination rates during pregnancy, even though the relationship between these parameters has not been formally analyzed.

Lactation is viewed as an additional route of elimination for some chemicals such as TCDD. According to a recent study, a breast-feeding woman expels through lactation an estimated 8.76 kg fat per year [q_f (kg/day), 0.8 kg milk/day with an average 3% lipid], and the partition coefficient between blood lipid and milk fat (K_{BM}) for TCDD is 0.92 (Wittsiepe et al., 2007; Milbrath et al., 2009). The estimated rate of elimination of TCDD due to breast-feeding (k_{bfed}) can then be computed as follows (Milbrath et al., 2009):

$$k_{bfed} = \frac{q_f \times \Delta t_{bfed}}{K_{BM} \times \frac{pbf_i}{100} \times BW_i} \quad \text{(Eq. 3-11)}$$

where

Δt_{bfed} (unitless) = the fraction of the year during which the woman was actively breast-feeding;

pbf_i = woman's percent body fat; and

BW = woman's body weight in kg.

Assuming no interaction between breast-feeding and other half-life determinants Milbrath et al. (2009), the authors predicted a half-life of 4.3 years for TCDD in a 30-year-old, nonsmoking woman with 30% body fat if she did not breast-feed that year, and a half-life of 1.8 years if she breast-fed for 6 months.

3.3.3.3. *Lifestyle and Habits*

One of the factors related to lifestyle and habits that could influence TCDD kinetics is smoking. Smoking has been reported to enhance the elimination of dioxin and dioxin-like compounds (Flesch-Janys et al., 1996; Ferriby et al., 2007). Milbrath et al. (2009) accounted for interindividual variation in body composition as well as smoking habits in an empirical model. The predicted half-life (years) for an individual *i* as a function of age, smoking status, and percent body fat *i* was as follows

$$t_{1/2}(age, smoke, pbf)_i = [\beta_{(0age)} + \beta_{(age)} \times age_i] \times SF_i \times \frac{pbf_i}{pbf_{ref(age_i)}} \qquad \text{(Eq. 3-12)}$$

where

$\beta_{(0age)}$ = intercept constant derived from regressed data;
$\beta_{(age)}$ = slope constant derived from regressed data;
age_i = specific age *i* (years);
pbf_i = individual percent body fat;
$pbf_{ref(age_i)}$ = reference percent body fat; and
SF_i = the unitless, multiplicative smoking factor.

3.3.3.4. *Genetic Traits and Polymorphism*

One particular genetic locus that is potentially related to TCDD pharmacokinetics and tissue dose is the gene for the AhR. Eight candidate AhR polymorphisms have been identified to date (Harper et al., 2002; Connor and Aylward, 2006). Given the role of AhR in regulating the induction of CYP1 isozymes (Baron et al., 1998; Toide et al., 2003; Connor and Aylward, 2006), the polymorphism might lead to interindividual differences in metabolic clearance, the significance of which would depend upon the dose, fat content, and exposure scenario. In this regard, it should be noted that the inducibility of aromatic hydrocarbon hydroxylase in human

tissues has been reported to be highly variable, up to 100−fold (Wong et al., 1986; Smart and Daly, 2000; Connor and Aylward, 2006).

Finally, the scientific literature contains values of K_d (the dissociation constant of the TCDD−AhR complex) ranging from about 1 to much higher values (corresponding to lower binding affinity) (reviewed in Connor and Aylward, 2006). This provides suggestive evidence for a heterogeneous human AhR, with functionally important polymorphisms (Roberts et al., 1986; Micka et al., 1997), even though some of the range may be attributed to experimental procedural differences and to other factors (Manchester et al., 1987; Lorenzen and Okey, 1991; Harper et al., 2002; Connor and Aylmard, 2006).

The various pharmacokinetic processes and determinants (see Sections 3.3.2 and 3.3.3), individually or together, might influence the dose metrics of relevance to the dose-response modeling of TCDD.

3.3.4. Dose Metrics and Pharmacokinetic Models for TCDD

3.3.4.1. *Alternative Dose Metrics for Dose-Response Modeling*

The **dose metric** related to a toxicologic endpoint can range from the maximal concentration, the area under a time-course curve (AUC), or the time-averaged concentration of the toxic moiety in the body, blood, or target tissue, to an appropriate measure of the resulting interactions in the target tissue (e.g., receptor occupancy or functional biomarkers related to specific effects). A single dose metric, however, is unlikely to be sufficient for all endpoints and exposure durations. Further, the ideal dose metric chosen on the basis of the mode of action (MOA) may not be the dose metric for which model predictions can be obtained with a high level of confidence. Consideration of these issues is critical to the selection of the dose metrics of relevance to dose-response modeling of TCDD.

Figure 3-5 lists a range of alternative dose metrics for TCDD in terms of their relevance based on considerations of pharmacokinetic mechanisms and MOA. The **administered dose** or daily intake (ng/kg-day) is the least relevant dose metric for dose-response modeling of TCDD. This dose adjusts only for body weight differences between species. When used with an additional uncertainty factor for kinetics (or $BW^{3/4}$) and for dynamics, it can also account for allometrically-predicted pharmacokinetic (clearance) and pharmacodynamic differences between

species in deriving the human equivalent dose (HED). In effect, it facilitates the computation of administered dose associated with the same steady-state blood concentration of parent chemical in humans and rats by accounting for differences in metabolic clearance (assumed to be related to body surface area, with no corresponding temporal changes in the volume of distribution; see, for example, Krishnan and Andersen, 1991). Such calculations of HED for TCDD may not be appropriate given that (1) steady-state was not attained in all critical toxicological studies chosen for the assessment, (2) the clearance is mainly due to enzyme(s) and processes whose levels/rates do not necessarily vary across species or life stages as a function of body surface differences, and (3) there is a likelihood of change in volume of distribution over time. Furthermore, the use of administered dose does not explicitly account for the dose-dependent elimination of TCDD from tissues as demonstrated in multiple studies (reviewed in Sections 3.3.2 and 3.3.4). The use of administered dose in TCDD dose-response modeling is unlikely to facilitate the characterization of the true relationship between the response and the relevant measures of internal dose that are influenced by dose-dependent elimination and binding processes. Additionally, the use of administered dose to extrapolate across species or life stages would not effectively take into account the differences in fat content or the demonstrated dose-dependent and species-dependent differences in elimination half-life of TCDD.

The other alternative dose metrics for TCDD include absorbed dose, body burden, serum or whole blood concentration, tissue concentration, and functional-related metrics of relevance to the MOA (e.g., receptor occupancy, change in protein levels). These measures can be calculated as a current (terminal), average (over a defined period), or integral quantity. The applicability of the integral measures, such as the AUC (i.e., the area under the curve of a plot of blood or plasma concentration vs. time), traditionally used for analyzing chronic toxicity data, is questionable in the case of TCDD because of differences in lifespan and uncertainties regarding the appropriateness of the duration to be specified for averaging the AUC in experimental animals and humans for certain critical effects (NRC, 2006).

Among the alternative dose metrics, the **absorbed dose** accounts for differences in body weight as well as species-specific differences in bioavailability. However, this dose measure, while more closely related to the endpoint than administered dose, does not account for the physiological and biochemical mechanisms responsible for interspecies differences in internal dose.

Body burden, or more appropriately the body concentration, represents the amount of TCDD per kg body weight. TCDD body burdens, like other dose measures, can be determined as the peak, the average over the period of the bioassays, or the level at the end of the experiments. Thus, the terminal or average body burdens can be obtained either using data or pharmacokinetic models and used in dose-response modeling. The body burden is a measure of TCDD dose that reflects the net impact of bioavailability, uptake, distribution, and elimination processes in the organism. It is essentially a function of the volume of distribution and clearance processes, and as such it does take into account the temporal changes in volume of distribution as well as the concentration-dependent clearance. These are phenomena that are critical to the understanding of TCDD dose to the target. However, the body burden may not accurately reflect the tissue dose (NRC, 2006), and as such does not allow for analysis of species-specific differences in target organ sensitivity to TCDD. In essence, the body burden represents only an "overall average" of TCDD concentration in the body, without regard to the differential partitioning and accumulation in specific tissues, including the target tissue(s).

Serum (or blood) concentration of TCDD is a dose metric that reflects both the body burden and the dose to target tissues. Serum or blood concentration, at steady-state, would be reflective of the impact of clearance processes, and expected to be directly proportional to the tissue concentrations of TCDD (NRC, 2006). This dose metric for lipophilic chemicals such as TCDD is often expressed as a lipid-normalized value, to adjust for varying serum lipid content (e.g., DeKoning and Karamus, 2000; Niskar et al., 2009), particularly in human biomonitoring studies, thus of relevance to dose-response modeling; however, the serum lipid-normalized concentrations of TCDD are not routinely collected and reported in animal toxicologic studies. The lipid-adjusted serum concentration, however, would be reflective of the lipid-adjusted concentration of TCDD in other organs (reviewed in Aylward et al., 2008) depending upon the extent of steady-state attained and the similarity of lipid composition across tissues in each species. In essence, the serum lipid-normalized measure is representative of the amount of TCDD per specified volume of total lipids, whereas the whole blood measure will be reflective of the ensemble of free, lipid-bound and protein-bound TCDD in plasma and erythrocytes, which may be species-specific. Even though these dose metrics are thought to be more closely and directly related to the tissue concentrations associated with an effect, a less direct association

might occur at increasing doses when nonlinear processes dominate the kinetics and distribution of TCDD into organs such as the liver.

Tissue concentration of TCDD, as free, bound, or total TCDD, is a more relevant pharmacokinetic measure of dose, given that it provides a measure of exposure of the target cells to the chemical. In this regard, the CYP1A2-bound fraction may be considered as a relevant dose metric for certain toxic effects; however, the available data contain mixed results regarding the mechanistic linkage of this dose metric to toxicity and carcinogenicity (reviewed in Budinsky et al., 2006). In such cases, the use of alternative dose metrics (e.g., bound concentration as well as the serum concentration) in dose-response modeling could be considered. Other function-related biomarkers and dose metrics could facilitate the additional consideration of pharmacodynamic aspects reflecting tissue- and species-specific sensitivity. These metrics represent the most relevant measures of tissue exposure and sensitivity to TCDD.

Empirical time-course data on the alternative dose metrics of TCDD associated with epidemiologic and experimental (animal) studies are not available, requiring the use of pharmacokinetic models to obtain estimates of these dose metrics. These models may be simple, based on first order kinetics (see Section 3.3.4.2), or more complex based on physiochemical, biochemical, and physiological parameters for simulating uptake, distribution (including sequestration to proteins), and clearance of TCDD (see Section 3.3.4.3).

3.3.4.2. *First-Order Kinetic Modeling*

Figure 3-6 illustrates the process of estimating a human-equivalent TCDD oral exposure from an experimental animal-administered dose, based on the assumption that body burden is the effective dose metric for TK equivalence across species. The primary assumption is that the time-weighted average (TWA) TCDD body burden over some critical time period is the proximate toxicokinetically-effective dose eliciting a toxicologic effect.[9] The process consists of estimating the effective average body burden in the experimental animal over some time t_A (generally the experimental duration) using a TK model, then "back-calculating" a daily human exposure level that would result in that average body burden over some time t_H (the human equivalent to t_A).

[9]The conversion depicted in Figure 3-6 does not account for toxicodynamic differences between species.

The following closed-form equation is the general formula used to calculate a TCDD terminal body burden in an experimental animal or human at time (t).

$$BB(t) = BB(0) + \frac{d(1 - e^{-kt})\,fa}{k} \qquad \text{(Eq. 3-13)}$$

where

$BB(t)$ = the body burden at time t (ng/kg);
$BB(0)$ = the initial body burden (ng/kg);
d = the daily dose (ng/kg-day);
k = the whole-body elimination rate (days^{-1});
t = the time at which the body burden is determined (days); and
fa = the fraction of oral dose absorbed (unitless).

For the experimental animal, *BB(t)* is $BB_A(t) = BB_A(0)e^{-k_A t_A} + \frac{d_A(1 - e^{-k_A t_A})\,fa_A}{k_A}$, and for humans, this parameter is $BB_H(t) = BB_H(0)e^{-k_H t_H} + \frac{d_H(1 - e^{-k_H t_H})\,fa_H}{k_H}$.

Setting $BB_H(t) = BB_A(t)$ obtains the following expression:

$$BB_H(0)e^{-k_H t_H} + \frac{d_H(1 - e^{-k_H t_H})\,fa_H}{k_H} = BB_A(0)e^{-k_A t_A} + \frac{d_A(1 - e^{-k_A t_A})\,fa_A}{k_A} \qquad \text{(Eq. 3-14)}$$

Rearranging yields the general solution for d_H.

$$d_H = d_A \frac{k_H}{k_A} \frac{fa_A}{fa_H} \frac{(1 - e^{-k_A t_A})}{(1 - e^{-k_H t_H})} + BB_A(0)e^{-k_A t_A} - BB_H(0)e^{-k_H t_H} \qquad \text{(Eq. 3-15)}$$

Assuming that initial body burdens are very small compared to BB(t) and that the fraction of TCDD absorbed is the same for humans and experimental animals, and using the relationship $k = \frac{\ln(2)}{t_{1/2}}$, where $t_{½}$ is the whole-body half-life, a simplified solution for d_H is obtained.

$$d_H = d_A \frac{t_{1/2A}}{t_{1/2H}} \frac{(1 - e^{-k_A t_A})}{(1 - e^{-k_H t_H})} \quad \text{(Eq. 3-16)}$$

The term $1-e^{-kt}$ is the daily fraction eliminated. Therefore, d_H can be seen to be the average daily administered dose to the experimental animal times the ratio of the animal:human half-life times the ratio of the animal:human daily fraction eliminated over the respective times, t_A and t_H. For both species at (theoretical) steady state ($t \rightarrow \infty$; daily fraction eliminated $\rightarrow$ 1), the latter ratio approaches unity, reducing the animal:human conversion factor to the ratio of the half-lives. The latter approach was used in the 2003 Reassessment for conversion of animal cancer slope factors to the human equivalent, where only lifetime exposures are relevant.[10]

However, for less-than-lifetime exposures eliciting noncancer effects, specific values for t_A and t_H must be considered. Furthermore, Eq. 3-16 computes d_H on the basis of *terminal* body burdens at times t_A and t_H. The more representative metric for toxicokinetic equivalence based on average body burden over the respective time periods is given in Eq. 3-17.

$$BB(t) = BB(0)\frac{1}{t}\int_0^t e^{-k\tau} d\tau + d\frac{fa}{k}\frac{1}{t}\int_0^t (1 - e^{-k\tau}) d\tau = BB(0)\frac{\left(1 - e^{-kt}\right)}{kt} + d\frac{fa}{k}\left[1 - \frac{\left(1 - e^{-kt}\right)}{kt}\right] \quad \text{(Eq. 3-17)}$$

On the basis of average body burden as given in Eq. 3-17, is transformed again assuming minimal initial body burden (BB(0) ~ 0), as follows:

$$d_H = d_A \frac{t_{1/2A}}{t_{1/2H}} \frac{\left[1 - \frac{\left(1 - e^{-k_A t_A}\right)}{k_A t_A}\right]}{\left[1 - \frac{t_{H0}}{t_H} - \frac{\left(e^{-k_H t_{H0}} - e^{-k_H t_H}\right)}{k_H t_H}\right]} \quad \text{(Eq. 3-18)}$$

where t_{H0} is the initial human exposure time.

The value of t_A is the duration of the experimental exposure period. For some gestational exposures, if a critical exposure window is defined, t_A will be the duration of the critical

[10]No conversions to human-equivalent exposures were attempted for other effects in the 2003 Reassessment.

exposure window. The value of t_H is the human-equivalent duration corresponding to t_A. However, for less than lifetime exposure in humans, t_H does not begin at 0, but must end at 25,550 days (70 years) to include the terminal (pseudo) steady-state level, at which the $BB_H(t)$:d_H ratio is highest. The average is determined from the terminal end of the human exposure period because the daily exposure achieving the target blood concentration is smaller than for the same exposure period beginning at birth and is health protective for effects occurring after shorter-term exposure.[11] Figure 3-7 depicts the relationship of daily dose to TWA body burden graphically for several exposure duration scenarios. For shorter durations occurring later in life, the average body burden over the exposure period does not differ substantially from the steady-state value. Even for half-lifetime exposures, the deviation of the average from steady state is minimal. Only for lifetime exposures does the difference become more marked, but only by about 15%. Note that in the 2003 Reassessment, a constant value of 3,000 was used for $BB_H(t)$:d_H, based on the relationship of continuous exposure to theoretical steady-state body burden (t = lifetime, $t_{½}$ = 2,593 days); this approach, while conservative, does not account for exposure scenarios of different durations and does not strictly reflect the average body burden dose metric.

The simulation in Figure 3-7 is based on a unit daily exposure to humans, such that the target body burden represents $BB_H(t_H)$:d_H as a general scalar for calculating d_H from any given d_A. Table 3-3 shows the resulting TK conversion factors for the rodent species and strains comprising the bulk of the experimental animals in TCDD studies. Monkey and mink values are not shown in this table because, for the former, only chronic exposures were evaluated and, for the latter, no TCDD half-life information is available. Monkey (Rhesus) half-life estimates range from about 200−500 days. A representative value of 365 days is used for this TCDD assessment. The d_A to d_H conversion factor for the chronic monkey exposures (3.5−4 years) in TCDD studies is 9.2−9.7 (BB_A:d_A = 279−263).

Application of first order kinetics for the risk assessment of TCDD can only be used to estimate total body burdens or back-calculate administered dose from experimental data. Body burden calculations using first order kinetics is based on the assumption of a first order decrease in the levels of administered dose as function of time. In that sense, any loss of TCDD from the body is described by using a rate constant that is not specific to any biological process. This

[11] See the following section (3.3.4.3) for a more detailed discussion of this concept.

constant is usually estimated from estimates of half-life of TCDD. Assuming a constant half-life value for the clearance for long-term or chronic TCDD exposure is not biologically supported given the observed data indicating early influence of CYP1A2 induction and binding to TCDD and later redistribution of TCDD to fat tissue. Abraham et al. (1988) found that the liver:adipose tissue concentration ratio in female Wistar rats exposed to a subcutaneous TCDD dose of 300 ng/kg decreased from 10.3 at 1 day postexposure to 0.5 at 91 days postexposure. Consequently, using half-life estimates based on observed steady-state levels of TCDD will not account for the possibility of accelerated dose-dependent clearance of the chemical at the early stages and thus would result in estimation of lower administered levels of the chemical. The dynamic change in half-life due to dose-dependent elimination at the early stages of TCDD exposure and its later redistribution to fat tissues for steady-state levels is better described using biologically-based models, such as the PBPK models and concentration- and age-dependent elimination (CADM) models (Carrier et al., 1995a, b; Emond et al., 2004, 2005, 2006; Aylward et al., 2005b). Additionally, these models provide estimates for other dose metrics (e.g., serum or tissue levels) that are more biologically relevant to response than administered dose or total body burden (see Section 3.3.4.3).

3.3.4.3. *Biologically-Based Kinetic Models*

The development and evolution of biologically-based kinetic models for TCDD have been reviewed by EPA (2003) and Reddy et al. (2005). The initial PBPK model of Leung et al. (1988) was developed with the consideration of TCDD binding to CYP1A2 in the liver. The next level of PBPK models by Andersen et al. (1993) and Wang et al. (1997) used diffusion-limited uptake and described protein induction by interaction of DNA binding sites. The models of Kohn et al. (1993) and Andersen et al. (1997) further incorporated extensive hepatic biochemistry and described zonal induction of CYP by TCDD. TCDD PBPK models have evolved to include detailed descriptions of gastrointestinal uptake, lipoprotein transport, and mobilization of fat, as well as biochemical interactions of relevance to organ-level effects (Kohn et al., 1996; Roth et al., 1994). Subsequently, developed PBPK models either used constant hepatic clearance rate (Wang et al., 1997, 2000; Maruyama et al., 2002) or implemented varying elimination rates as an empirical function of body composition or dose (Andersen et al., 1993, 1997; Kohn et al., 1996; Van der Molen et al., 1998, 2000). The more recent pharmacokinetic

models explicitly characterize the concentration-dependent elimination of TCDD (Carrier et al., 1995a, b; Emond et al., 2004, 2005, 2006; Aylward et al., 2005b). The biologically-based pharmacokinetic models describing the concentration-dependent elimination (i.e., the pharmacokinetic models of Aylward et al. [2005b] and Emond et al. [2005, 2006]) are relevant for application to simulate the TCDD dose metrics in humans and animals exposed via the oral route. The rationale for considering the application of Aylward et al. (2005b) and Emond et al. (2004, 2005, 2006) models for estimating dose metrics for possible application to TCDD risk assessment is based on the following considerations.

- Both models represent research results from the more recent peer-reviewed publications.
- Both models are relatively simple and less parameterized than earlier kinetic models for TCDD. The Aylward et al. (2005b) model is based on two-time scale TCDD kinetics described by Carrier et al. (1995a), and the Emond et al. (2004, 2005, 2006) PBPK models are reduced versions of earlier complex PBPK models. Although simple, both the Aylward et al. (2005b) and Emond et al. (2004, 2005, 2006) models are still inclusive of important kinetic determinants of TCDD disposition.
- Both models are uniquely formulated with dose-dependent hepatic elimination consistent with the physiological interpretations commonly accepted by the scientific community.
- Both models and extrapolated human versions were tested against human data collected in a variety of human exposure scenarios (Alyward et al., 2005b; Emond et al., 2005).
- Both models are capable of deriving one or more of the candidate dose-metrics that are of interest to EPA's dose-response assessment of TCDD.

3.3.4.3.1. *CADM model.*

3.3.4.3.1.1. ***Model structure.***

The pharmacokinetic model of Aylward et al. (2005b), referred to as the CADM model in this report, is based on an earlier model developed by Carrier et al. (1995a,b) that describes the dose-dependent elimination and half-lives of polychlorinated dibenzo-*p*-dioxins and furans. This model describes the TCDD levels in blood (body), liver, and adipose tissue. Blood itself is not characterized physically as a separate compartment within the model, and the distribution of TCDD to tissues other than adipose tissue and liver (usually less than 4%) is not accounted for by the model. The original structure of the Carrier et al. (1995a, b) model was modified by Aylward et al. (2005b) to include TCDD elimination through partitioning from circulating lipids

across the lumen of the large intestine into the fecal content (see Figure 3-8). The most recent version of the Carrier model (Aylward et al., 2005b, 2008) includes fecal excretion of TCDD from two routes: (1) elimination from circulating blood lipid through partitioning into the intestinal lumen; and (2) elimination of unabsorbed TCDD from dietary intake.

A basic assumption of this model is that metabolic elimination of TCDD is a function of its current concentration in the liver. The current concentration of TCDD in the liver increases with increasing body burden in a nonlinear fashion as a result of the induction of (and binding of TCDD to) specific proteins (i.e., CYP1A2). Consequently, the fraction of TCDD body burden contained in the liver increases nonlinearly (with a corresponding decrease in the fraction contained in adipose tissues) with increasing body burden of TCDD (Carrier et al., 2005a).

Of particular note is that the adipose tissue compartment of the model is considered to represent the lipid contained throughout the body. It then assumes that the concentrations of TCDD in lipids of plasma and various organs is essentially equivalent to that of adipose tissue, and as such these concentrations are included in the adipose compartment of the model. Even though this approximation is fairly reasonable given the available data, there is some concern that the adipose compartment of this model also includes the lipid content of the liver to some unknown extent. Removal of lipid volume from the liver would mathematically alter total hepatic concentration and therefore would affect the estimated levels of the chemical available for binding to proteins.

Distribution in the body is modeled to occur between hepatic and adipose/lipid compartments, with the fraction of body burden in liver increasing according to a function that parallels the induction of the binding protein CYP1A2. Elimination is modeled to occur through hepatic metabolism (represented as a first-order process with rate constant *K* that decreases with age) and through lipid-based partitioning of unmetabolized TCDD across the intestinal lumen into the gut, which is also modeled as a first-order process. As the body burden increases, the amount of TCDD in the liver increases nonlinearly, resulting in an increased overall elimination rate.

3.3.4.3.1.2. ***Mathematical representation.***

The CADM model describes the distribution to tissues (including liver and adipose tissue) based on exchange from blood at time intervals of one month. The model is based on

quasi-steady-state-approximation, and thus it is also based on the consideration that the intertissue processes reach their equilibrium values "quasi-instantaneously." In this regard, absorption and internal distribution reflective of kinetics at the cellular level (e.g., diffusion, receptor binding, and enzyme induction) likely occur on a relatively fast time scale (a few hours to a few days). However, the overall body concentration (i.e., body burden) varies slowly with time such that it remains virtually unchanged during short time intervals.

The CADM model does not differentiate between binding to AhR and CYP1A2, and it lacks explicit descriptions of CYP1A2 induction, a key determinant of TCDD kinetics. However, the empirical equation in the CADM model is based on five parameters (i.e., f_{min}, f_{max}, K, W_a, and W_l; see Tables 3-4 and 3-5) that allow the successful description of the behavior of TCDD in liver and adipose tissue (i.e., TCDD half-lives in each compartment increase with decreasing body burden). This observation implies that the model adequately accounts for the ensemble of the processes. Essentially, the CADM model describes the rate of change in tissue concentrations of TCDD as a function of total body burden such that the global elimination rate decreases with decreasing body burden or administered dose.

3.3.4.3.1.3. ***Parameter estimation.***

The CADM model is characterized by its simplicity and fewer parameters compared to physiologically-based models. Reflecting this simplicity, hepatic extraction is computed with a unified empirical equation that accounts for all relevant processes (i.e., protein induction and binding).

The key parameters (f_{min}, f_{max}, K, and k_e) were all obtained by fitting to species-specific pharmacokinetic data. The physiological parameters (such as tissue weights) used in the model are within ranges documented in the literature. The fat content is described to vary as a function of age, sex, and BMI. However, the BMI of the model is not allowed to change during an individual simulation (which can range from 20 years to 70+ years) when in reality the percentage of fat in humans changes over time. None of the TCDD-specific parameters were estimated *a priori* or independent of the data set simulated by the model.

3.3.4.3.1.4. ***Model performance and degree of evaluation.***

The CADM model was not evaluated for its capabilities in predicting data sets not used in its parameterization. In other words, one or more of the key input parameters (f_{hmin}, f_{hmax}, k_e, K) was or were obtained essentially by fitting to the species-specific pharmacokinetic data, such that there was no "external" validation data set to which the model was applied. Despite the lack of emphasis on the "external" validation aspect, the authors have demonstrated the ability of the model to describe multiple data sets covering a range of doses and species.

The visual comparison of the simulated data to experimental values suggests that the model could, to an approximate degree, correctly reproduce the whole set of data (e.g., pharmacokinetic [PK] profile over a range of dose and time) and not just part of the PK curve, essentially with the use of a single set of equations and parameters.

The pharmacokinetic data sets for TCDD that were used to calibrate/evaluate the CADM model included the following:

- Adipose tissue and liver concentrations of TCDD following a single oral dose of 1 μg/kg in monkeys (McNulty et al., 1982);
- Percent dose retained in liver for a total dose of 14 ng in hamsters (Van den Berg et al., 1986);
- Elimination kinetics of TCDD in female Wistar rats following a single subcutaneous dose of 300 ng/kg (data from Abraham et al., 1988);
- Liver and adipose tissue concentrations (terminal measurements) in Sprague−Dawley rats given 1, 10 or 100 ng TCDD/kg bw during 2 years (Kociba et al., 1978); and
- Serum lipid concentrations of TCDD over a period of several years in 54 adults (29 men and 25 women) from Seveso and in three Austrian patients (Aylward et al., 2005a).

For illustration purposes, Figure 3-9 shows model simulations of rat data from Carrier et al. (1995a). Figure 3-2 (see Section 3.3.2.4) depicts the human data that were used by the authors to support the concentration-dependent elimination concept; the model was parameterized to fit approximately to these data (Aylward et al., 2005a).

The authors did not report any specialized analyses that quantitatively evaluated the uncertainty, sensitivity, and/or variability of CADM model parameters and structure.

3.3.4.3.1.5. ***Confidence in CADM model predictions of dose metrics.***

The level of confidence associated with the predictability and reliability of absorbed dose and body burden for oral exposures in humans (as well as several animal species) by this model can be ranked as high (see Table 3-6). This model, however, does not account for the differential solubility of TCDD in serum lipids and adipose tissue lipids, nor does it account for the diffusion-limited uptake by adipose tissue. Due to these limitations, the confidence associated with the predictions of the serum lipid concentration of TCDD is considered medium, particularly when it is not documented that steady-state is reached during the critical toxicologic studies and human exposures. Furthermore, the CADM model does not facilitate the computation of TCDD concentrations in specific internal organs (other than liver and adipose tissue). The reliability of this model for simulating the liver concentration (free, bound, or total) of TCDD at low doses is considered to be low. This low confidence level is a result of the uncertainty associated with the key parameter f_{hmin}. This parameter needs to be re-calibrated for each study/species/population to effectively represent the free fraction of TCDD in liver and the amount of TCDD contained in the hepatic lipids and bound to the liver proteins (whose levels might be reflective of background exposures of various sources; see Carrier et al., 1995a). The uncertainty related to the numerical value of this parameter in animals and humans—particularly at very low exposures—raises concern regarding the use of this model to predict TCDD concentration (free, bound, or total) in liver as the dose metric for dose-response modeling. Although the use of the parameter f_{hmax} permits the prediction of the dose to liver at high doses, it does not specifically facilitate the simulation of the amount bound to the protein or level of induction in liver. Because the CADM model is not capable of simulating enzyme induction based on biologically-relevant parameters, its reliability for predicting the concentration of TCDD bound specifically to the AhR is not known. Finally, due to the lack of parameterization or verification with kinetic data in pregnant, lactating, or developing animals or humans, the CADM model is unlikely to be reliable in the current form for use in *predicting* potential dose metrics in these subpopulations or study groups that might form the basis of points of departure (PODs) for the assessment.

3.3.4.3.2. *PBPK model.*

3.3.4.3.2.1. ***Model structure.***

Emond et al. (2004, 2006) simplified the eight-compartment rat model of Wang et al. (1997) to a four-compartmental model (liver, fat, rest of body and placenta with fetal transfer) (Emond et al., 2004), and later to a three-compartment adult model (liver, fat, rest of the body) (Emond et al., 2006) (see Figures 3-10 and 3-11). Their rationale for simplification of the model was based on evaluating, critiquing, and improving all earlier PBPK models by Wang et al. (1997). In general, the main reason for the simplification was that extrapolation of a PBPK model to humans with these many (i.e., eight compartments) compartments would be problematic due to the limited availability of relevant human data for validation (Emond et al., 2004). One major difference from earlier models, repeatedly emphasized by Emond et al. (2005, 2006), was their description (included in their simplified PBPK models) of the dose-dependent, inducible elimination of TCDD. The rationale for including TCDD binding and induction of CYP1A2 into the model was earlier described by Santostefano et al. (1998).

The most recent version of the rat and human PBPK models developed by Emond et al. (2006) describes the organism as a set of three compartments corresponding to real physical locations—liver, fat, and rest of the body—interconnected by systemic circulation (see Figure 3-10). The liver compartment includes descriptions of CYP1A2 induction, which is critical for simulating TCDD sequestration in liver and dose-dependent elimination of TCDD. In this model, the oral absorption of TCDD from the GI tract accounts for both the lymphatic (70%) and portal (30%) systems.

The biological relationship between TCDD "sequestration" by liver protein and its "elimination" by the liver is not entirely clear. TCDD is metabolized slowly by unidentified enzymes. CYP1A2 is known to metabolize TCDD based on studies in CYP1A2 KO mice (Diliberto et al., 1997, 1999), in which the metabolic profile is different compared to wild-type mice. However, since several metabolites appear in the feces of CYP1A2 knock out mice, it is assumed that there are other enzymes involved in TCDD metabolism. TCDD binds to the AhR and induces not only CYP1A2, but also CYP1A1, CYP1B1, and several UGTs and transporters (Gasiwicz et al., 2008). Both hydroxylated and glucuronidated hydroxyl metabolites are found in the feces of animals treated with TCDD (Hakk et al., 2009). Because the exact enzymes involved with TCDD are unknown and yet the metabolism is induced by TCDD, an assumption

of increased the elimination rate of TCDD in proportion to the induction of CYP1A2 is made. In the PBPK model, CYP1A2 is needed because TCDD binds to rat, mouse, and human CYP1A2 (Staskal et al., 2005; Diliberto et al., 1999). Thus CYP1A2 induction is necessary to describe TCDD pharmacokinetics due to TCDD binding. Hence, CYP1A2 can be used as a marker of Ah-receptor induction of "TCDD metabolizing enzymes." Other models use AhR occupancy as a marker of induction of "TCDD metabolizing enzymes" (Andersen et al., 1997; Kohn et al., 2001).

Figure 3-11 depicts the structure of the PBPK model developing rat (Emond et al., 2004).

3.3.4.3.2.2. *Mathematical representation.*

The key equations of the PBPK model of Emond et al. (2004) are reproduced in Text Boxes 3-1 and 3-2, whereas those from Emond et al. (2005, 2006) are listed in Table 3-7. The rate of change of TCDD in the various tissue compartments is modeled on the basis of diffusion limitation considerations. Accordingly, mass balance equations are used to compute the rate of change in the tissue (i.e., intracellular compartment) and tissue blood (i.e., extracellular compartment). The membrane transfer of TCDD is computed using a permeation coefficient-surface area cross product (PA) for each tissue. Metabolism and binding of TCDD to the AhR and inducible hepatic protein (CYP1A2) are described in the liver. The total mass in the liver was then apportioned between free dioxin (C_{lf}) and bound forms of TCDD (see Figure 3-12). The dose- and time-dependent induction of hepatic CYP1A2 in the liver is described per Wang et al. (1997) and Santostefano et al. (1998). Accordingly, the amount of CYP1A2 in the liver was computed as the time-integrated product of inducible production and a simple first-order loss process (Wang et al., 1997):

$$\frac{dCYP_{1A2}}{dt} = S(t)K_0 - K_2C_{A2t} \qquad \text{(Eq. 3-19)}$$

In this expression, CYP_{1A2} is the concentration of the enzyme (nmol/g), K_2 is the rate constant for the first order loss (liter/hour), C_{A2t} is the concentration of CYP1A2 in the liver (nM/hour), K_0 is the basal rate of production of CYP1A2 in the liver, and $S(t)$ (unitless) is a multiplicative

stimulation factor for CYP1A2 production in the form of a Hill-type function (see Section 3.3.2.3):

$$S(t) = 1 + \frac{In_{A2}(C_{Ah\text{-}TCDD})^h}{(IC_{A2})^h + (C_{Ah\text{-}TCDD})^h} \quad \text{(Eq. 3-20)}$$

where, $S(t)$ is the stimulation function, In_{A2} is the maximum fold of CYP1A2 synthesis rate over the basal rate, $C_{Ah\text{-}TCDD}$ is the amount of AhR occupied by TCDD, and IC_{A2} is the Michaelis-Menten constant of CYP1A2 induction (nM). The dose-dependent or variable elimination of TCDD was described using the relationship:

$$KBILE\ LI = \left[\frac{CYP1A2_{induced} - CYP1A2_{basal}}{CYP1A2_{basal}}\right] \times Kelv \quad \text{(Eq. 3-21)}$$

where $CYP1A2_{induced}$ is the concentration of induced CYP1A2 (nmol/mL), $CYP1A2_{basal}$ is the basal concentration of CYP1A2 (nmol/mL), and *Kelv* is the interspecies constant adjustment for the elimination rate ($hour^{-1}$).

There are various ways of formulating the dose-dependent elimination as a function of the level of CYP1A2, and the above equation (used by the authors) can be viewed as one means of describing this behavior quantitatively. The numerator in the equation above will always be greater than zero when there is TCDD in the system (including TCDD derived from either background exposures or defined external sources). Consequently, the rate of elimination will correspond to a nonzero value for situations involving TCDD exposures. The above equation, however, does not describe changes in elimination rate in direct proportionality with the CYP1A2 levels; also, the *Kelv* value by itself does not reflect a scalable basal metabolic rate. Rather, these two terms collectively describe the outcome related to the TCDD elimination processes, based on fitting to observations in rats (Santostefano et al., 1998). The impact of CYP1A2 induction and sequestration on binding and elimination of TCDD is simulated using the Emond et al. (2004) model.

The gestational model consisted of a fetal compartment, and the transfer of TCDD between the placental and fetal compartments was described as a diffusion-limited (rather than a perfusion-limited) process (see Text Boxes 3-1 and 3-2).

Text Box 3-1.

Variation of Body Weight with Age: $$BW_{Time}(g) = BW initial \times \left(\frac{0.41 \times Time}{1402.5 + Time} \right)$$

Cardiac Output: $$Qc(mL/h) = Qcc \times 60 \left(\frac{BWmother}{1,000} \right)^{0.75}$$

A factor of 60 corresponds to the conversion of minutes to hours, and 1,000 is the conversion of body weight from g to kg.

Blood Compartment:

$$Cb(nmol/mL) = \frac{((Qf \times Cfb) + (Qre \times Creb) + (Qli \times Clib) + (Qpla \times Cplab) + Lymph)) - (Cb \times Clru)}{Qc}$$

Text Box 3-2.

Placenta Tissue Compartment

(a) Tissue-blood subcompartment

$$\frac{dAplab}{dt}(nmol/h) = Qpla(Ca - Cplab) + PApla(Cplab - Cplafree)$$

$$Cplab = \frac{Aplab}{Wplab}$$

(b) Tissue cellular matrices

$$\frac{dApla}{dt}(nmol/h) = PApla(Cplab - Cplafree) - \frac{dApla\_fet}{dt} + \frac{dAfet\_pla}{dt}$$

$$Cpla(nmol/mL) = \frac{Apla}{Wpla}$$

Free TCDD Concentration in Placenta

$$Cplafree(nmol/mL) = Clpla - \left[(Cplafree \times Ppla + \left(\frac{Plab\max \times Cplafree}{Kdpla + Cplafree}\right)\right]$$

Dioxin Transfer from Placenta to Fetuses

$$\frac{dAPla\_fet}{dt}(nmol/h) = Cl_{Pla\_fet} \times Cpla$$

Dioxin Transfer from Fetuses to Placenta

$$\frac{dAfet\_Pla}{dt}(nmol/h) = Cl_{Pla\_fet} \times CfetV$$

Fetal Dioxin Concentration (Fetuses 5 10 Per Litter)

$$\frac{dAfet}{dt}(nmol/h) = \frac{dAPla\_fet}{dt} - \frac{dAfet\_Pla}{dt}$$

$$Cfet(nmol/h) = \frac{Afet}{Wfet}$$

$$CfetV(nmol/mL) = \frac{Cfet}{Pfet}$$

1

2

3.3.4.3.2.3. *Parameter estimation.*

Table 3-8 lists the numerical values of the adult rat and human PBPK models of Emond et al. (2005, 2006). The values for key input parameters of the rat gestational model are summarized in Table 3-8 as well as Figure 3-13.

The parameters for the rat model were obtained primarily from Wang et al. (1997) except that the value of affinity constant for CYP1A2 was changed from 0.03 to 0.04 nmol/mL to get better fit to experimental data (Emond et al., 2004) and the variable elimination parameter (*Kelv*) was obtained by optimization of model fit to kinetic data from Santostefano et al. (1998) and Wang et al. (1997) (Emond et al., 2005, 2006). Wang et al. (1997) used measured tissue weights whereas the tissue blood flows and tissue blood weights were obtained from International Life Sciences Institute (ILSI, 1994). The partition coefficients (which were similar to those of Leung et al. [1988, 1990]), the permeability x area (PA) value for tissues, the dissociation constant for binding to CYP1A2 (IC_{A2}) and the Hill coefficient (*h*) were estimated using a two-stage process of fitting to dose-response and time-course data on TCDD tissue distribution (Wang et al., 1997). In the initial stage, the experimental data of arterial blood concentrations were used as input to the individual compartment to estimate the parameters; then, with the values obtained during stage one as initial estimates, those unknown parameters were re-estimated by solving the entire model at once using an optimization route (Wang et al., 1997). The receptor concentrations and dissociation constant of TCDD bound to AhR were obtained by fitting the model to TCDD tissue concentration combining with enzyme data reported by Santostefano et al. (1998) whereas the basal CYP1A2 in liver was based on literature data (Wang et al., 1997).

The parameters for the human PBPK model were primarily based on the rat model (Wang et al., 1997; Emond et al., 2005, 2006). Specifically, the blood fraction in the tissues, the tissue:blood partition coefficients, tissue permeability coefficient, the binding affinity of TCDD to Ah and CYP, and the maximum binding capacity in the liver for AhR were all set equal to the values used in the rat model. The species-specific *Kelv* was estimated by fitting to human data (Emond et al., 2005).

For the gestational rat model, the parameters describing the growth of the placental and fetal compartments as well as temporal change in blood flow during gestation were incorporated based on existing data. Exponential equations for the growing compartments were used (see Figure 3-13), except for adipose tissue for which a linear increment based on literature data was

specified. While physiological parameters for the pregnant rat were obtained from the literature, all other input parameters were set equal to that of nonpregnant rat (obtained from Wang et al. [1997]), see Tables 3-7 and 3-8. The current version of the rat gestational model contains parameters for variable elimination from Emond et al. (2006; Table 3-8), and still provides the essentially the same predictions as the original publication (Emond et al., 2004).

3.3.4.3.2.4. *Model performance and degree of evaluation*.

The PBPK model of Emond et al. (2004, 2005, 2006) had parameters estimated by fitting to kinetic data, such that the resulting model consistently reproduced the kinetic data. The same model structure with a single set of species-specific parameters could reproduce the kinetics of TCDD following various doses and exposure scenarios not only in the rat but also in humans. The simulations of the PBPK model of Emond et al. (2006) have been compared with two sets of previously published rat data: blood pharmacokinetics following a single dose of 10 μg/kg (the dose corresponding to the mean effective dose for induction of CYP1A2) (Santostefano et al., 1998) (see Figure 3-14); and hepatic TCDD concentrations during chronic exposure to 50, 100, 500, or 1,750 ng/kg (Walker et al., 1999) (see Figure 3-15). It is relevant to note that the PBPK model of Emond et al. (2004, 2006) is essentially a reduced version of the Wang et al. (1997) model, and it therefore provides simulations of liver and fat concentrations of TCDD that deviated by not more than 10−15% of those of Wang et al. (1997). The nongestational model of Emond et al. (2004) simulated the kinetic data in liver, fat, blood and rest of body of female Sprague-Dawley rats given a single dose of 10 μg TCDD/kg (data from Santostefano et al., 1996) and in liver and fat of male Wistar rats treated with a loading dose of 25 ng/kg followed by a weekly maintenance dose of 5 ng TCDD/kg by gavage (data from Krowke et al., 1989).

The gestational rat PBPK model simulated the following PK data sets (Emond et al., 2004):

- TCDD concentration in blood, fat, liver, placenta, and fetus of female Long−Evans rats given 1, 10, or 30 ng/kg, 5 days/week, for 13 weeks prior to mating followed by daily exposure through parturition (Hurst et al., 2000a);
- TCDD concentration in tissues (liver, fat), blood, placenta and fetus determined on gestation day (GD) 16 and GD 21 following a single dose of 0.05, 0.8, or 1 μg/kg given on GD 15 to pregnant Long Evans rat (Hurst et al., 2000b);

- Maternal and fetal tissue concentrations on GD 9, GD 16 and GD 21 after a single dose of 1.15 μg TCDD/kg given to Long−Evans rats on GD 9 or GD 15 (Hurst et al., 1998); and
- Fetal TCDD concentrations determined on GD 19 and GD 21 in rats exposed to 5.6 μg TCDD/kg on GD 18 (Li et al., 2006)

Furthermore, the scaled rat model was shown to be capable of simulating human data from the Austrian and Seveso subjects (see Figures 3-16 and 3-17). In this regard, it is useful to note that the computational version of the PBPK model of Emond et al. (2005, 2006) also contained the necessary equation to transform the model output of blood concentration into serum lipid adjusted concentration of TCDD.

The human model of Emond et al. (2005) has advantages for improving the TCDD dosimetry used in existing human epidemiological studies because its PBPK structure naturally develops expectations for the redistribution of TCDD within the body (to stores in fat and liver) relative to metabolism. However, because the dose-dependency of metabolic elimination in the Emond et al. (2005) human model was not calibrated to human data, it is important to review the expectations of this model using a database of human observations that is as extensive as possible and a spread of internal TCDD concentrations that is as wide as possible. Thus, presented below is a juxtaposition of expectations of elimination rates for the Emond model with observations for two highly exposed Austrian patients and nine of 10 Ranch Hand veterans[12] used for the original "validation" comparisons in the Emond et al. (2005).

Figure 3-18 shows the time course of the declines in TCDD serum concentrations in two highly-exposed Austrian subjects compared with expectations from the Emond et al. (2005) model. The comparison in Figures 3-17 and 3-18 indicates that the Emond model adequately describes the rate of TCDD elimination for the more highly exposed Austrian patients, but predicts a somewhat faster rate of decline than that observed for the less heavily exposed patient.

Figure 3-19 shows the results of combining the simulated and observed rates of loss for a group of Austrian and Ranch Hand subjects evaluated by Emond et al. (2005), counting only one data point per person. The X-axis in this figure is the TCDD serum concentration at the

[12]In preliminary comparisons, the simulation run for the 10th Ranch Hand veteran appeared anomalous and was therefore excluded from this summary.

midpoint of the observations for each subject. The error bars in the figure represent ±1 standard error.

Table 3-9 presents the results of regression analyses of the observed rates of decline in relation to the estimated TCDD serum levels at the midpoint of the observations for each subject in the Ranch Hand study (see Figure 3-19). These results indicate that some appreciable dose dependency of TCDD elimination is unequivocally supported. However, the central estimate of the slope of the relationship between the log of the TCDD elimination rate and the log of the TCDD level is only about 75% of that expected under the Emond et al. PBPK model (i.e., 0.092 ÷ 0.123 = 0.748).

Based on Table 3-9 and Figure 3-19, the expected slope from the Emond model (slope = 0.123) is about 2.8 standard errors above the regression line (slope = 0.092 with standard errors value of 0.011) indicated by the observations, suggesting that the departure is statistically significant. It can be seen that some appreciable dose dependency of TCDD elimination is unequivocally supported, but it is indicated that the dependency is somewhat less than presently incorporated into the Emond model. The difference might be slightly enhanced if the modest effects of the special measures taken to accelerate removal of TCDD from the Austrian subjects were subtracted from the observed elimination rates (i.e., oral administration of Olestra for both patients and LDL-apheresis for the more highly exposed patient).

It would be desirable in the future to extend this juxtaposition to data for other intensively studied subjects. Figure 4 in Aylward et al. (2005a) (reproduced here as Figure 3-2) shows a plot of elimination rate versus initial level of lipid-corrected TCDD in serum for 36 people. This is not the most desirable comparison for characterizing the relationship, however, because the rate constant for loss should be related to the geometric mean or midpoint level in the decline for each person (rather than to the initial level) in order to be most accurate in relating current TCDD levels to elimination rates and to avoid possible "regression to the mean" type statistical errors due to measurement imperfections. Overall, the conclusion from this analysis is that it is not unreasonable to use the Emond model as it stands, but the model might be modestly improved by adapting it to (1) include the two nondose-dependent pathways of elimination documented in the Geusau papers (GI elimination via the feces and loss via the sloughing of skin cells), and (2) reducing the extent of loss via the dose-dependent metabolism pathway from the

liver (Geusau et al., 2002; Harrad et al., 2003) so that overall loss rates for the average elimination rates from the Ranch Hand veterans is maintained.

A sensitivity analysis of inputs used to estimate inducible elimination rate for a single oral dose of 0.001 to 10 μg/kg in the rat indicated that the number of key parameters ranged from seven at the low dose region to 12 at the high dose (see Figure 3-20) (Emond et al., 2006). The sensitive parameters identified included the oral absorption parameters (KABS), volumes of liver and adipose tissue (WLIO, WFO), adipose tissue:blood partition coefficient (PF), and the basal CYP1A2 level (CYP1A2 1A2). At high doses, the most sensitive parameters also included those related to the maximal induction of CYP1A2 and AhR binding capacity (see Figure 3-20) (Emond et al., 2006).

The gestational rat model described in Emond et al. (2004), upon reparameterization, could simulate the kinetics of TCDD in mice. The initial changes to the rat model parameters included: rest of the body:blood partition coefficient (PRE), basal concentration (CYP1A2_1A2), delay in induction time (CYP1A2_1TAU) and adipose tissue permeability coefficient (PAFF), in accordance with Wang et al. (2000) (see Table 3-8). Subsequently, four parameters (adipose tissue:blood partition coefficient, CYP1A2 affinity parameter, GI tract elimination transit constant ($hour^{-1}$) and the interspecies metabolic parameter *Kelv* ($hour^{-1}$) were re-estimated based on visually fit of model simulations to the PK data from Deliberto et al. (2001), following an oral dose 150 ng TCDD/kg/day, 5 days/week for 17 weeks (see Table 3-7). The resulting mouse model is capable of reproducing the kinetics of TCDD in the adult (see Figures 3-21 through 3-27), as well as, to a very limited extent, the kinetics during gestation (see Figure 3-28).

3.3.4.3.2.5. *Confidence in PBPK model predictions of dose metrics*.

The PBPK model facilitates prediction of absorbed dose, body burden, and blood concentration of TCDD for oral exposures in adult humans and rats (adult and developing) with high confidence (see Table 3-10). The model output of blood concentration can be normalized to lipid content representative of the study group (species, sex, age, lifestage, and diet). However, the PBPK model of Emond et al. (2004, 2005, 2006) does not simulate plasma and erythrocyte TCDD concentrations separately, and it predicts tissue concentrations on the basis of tissue:whole blood partition coefficients and not on the basis of serum lipid-normalized values.

The reliability of this model for simulating the liver concentration of TCDD in rats is considered to be high but it is considered to be medium for humans. Although empirical data on bound or free concentrations were not used to evaluate model performance in humans, the biological phenomena (consistent with available data) related to the hepatic sequestration, enzyme induction, and dose-dependent elimination are described in the model. This is one of the situations where PBPK models are uniquely useful; that is, they permit the prediction of system behavior based on understanding of the mechanistic determinants, even though the required data cannot be directly obtained in the system (e.g., bound concentrations in the liver of exposed humans). For these dose measures (i.e., bound concentration and total liver concentration), the level of confidence can be further improved or diminished by the outcome of sensitivity analysis. In this regard, the results of a focused sensitivity analysis indicate that the most sensitive parameters of the human model are among the most uncertain (i.e., those parameters for which estimates were not obtained in humans) with respect to prediction of liver TCDD concentration, contrary to the animal model (see Section 3.3.6).

With respect to the mouse model, however, the level of confidence is medium to low, given that it has not been verified extensively with blood, body burden, or tissue concentration time-course or dose-response data. However, the mouse PBPK model, based on the rat model that has been evaluated with several PK data sets, has been shown to reproduce limited liver kinetic data (see Figures 3-21 through 3-28; Boverhoff et al., 2005). The same model structure has been used for simulating kinetics of TCDD in humans successfully. Overall, the adult mouse model, given its biological basis combined with its ability to simulate TCDD kinetics in multiple species, is considered to exhibit a medium level of confidence for simulating dose metrics for use in high to low dose extrapolation and interspecies (mouse to human) extrapolation. Even though similar considerations are applicable to gestational model in mice, the confidence level is considered to be low since very limited comparison with empirical data has been conducted (see Figure 3-28). Despite the uncertainty in these predictions, the scaled rat gestational model, given its biological and mechanistic basis, might be of use in predicting dose metrics in these groups that might form the basis of PODs in certain key studies.

3.3.4.4. *Applicability of PK Models to Derive Dose Metrics for Dose-Response Modeling of TCDD: Confidence and Limitations*

Both the CADM and PBPK models describe the kinetics of TCDD following oral exposure to adult animals and humans by accounting for the key processes affecting kinetics, including hepatic sequestration phenomena, induction, and nonlinearity in elimination, and distribution in adipose tissue and liver. Both models can be used for estimating body burdens and serum lipid adjusted concentrations of TCDD. However, there are several differences between these two models. The PBPK model calculates the free and bound concentrations of TCDD in the intracellular subcompartment of tissues. The total or receptor-bound concentrations in liver are unambiguous and more easily interpretable with the PBPK model than with the CADM model. In addition, the PBPK model computes bound and total concentrations as a function of the free concentration in the intracellular compartment of the tissue. By contrast, the CADM model simulates the total concentration based on empirical consideration of hepatic processes. Consequently, the amount of TCDD bound to AhR or CYP1A2 cannot be simulated with the CADM model. The CADM model computes only the total TCDD concentration in liver, and describes TCDD elimination through partitioning from circulating lipids across the lumen of the large intestine into the feces, while the PBPK model accounts for this process empirically within its hepatic elimination constant. Elimination of TCDD via skin, a minor process, is not described by either model. Thus, dose-response modeling based on body burden of TCDD in adult animals and humans can be conducted with either of the models, provided the duration of the experiment is at least one month, due to limitations in the CADM model. As shown in Figure 3-29, the predicted slope and body burden over a large dose range are quite comparable (generally within a factor of two).

Results of simulations of serum lipid concentrations or liver concentrations vary for the two models to a larger extent (up to a factor of 7), particularly for simulations of short duration. These differences reflect two characteristics of the PBPK model: first, quasi-steady-state is not assumed in the PBPK model; second, the serum lipid composition used in the model is not the same as the adipose tissue lipids. The CADM model does not account for differential solubility of TCDD in serum lipids and adipose tissue lipids, nor does it account for the diffusion-limited uptake by adipose tissue. Therefore, the PBPK model would appear to be superior to the CADM

model with respect to the ability to simulate serum lipid and tissue concentrations during exposures that do not lead to the onset of steady-state condition in the exposed organism.

The CADM model is simple and based on fewer parameters than the PBPK model. Because the CADM model is constructed by fitting to data, its performance is likely to be reliable for the range of exposure doses, species, and life stages from which the parameter estimates were obtained. On the other hand, the PBPK model structure and parameters are biologically-based and can be adopted for each species and life stage. Accordingly, the PBPK model has been adopted to simulate the kinetics of TCDD in the fetus and in pregnant rats, as well as in adult humans and rats (Emond et al., 2004, 2005, 2006). The time step for calculation and dosing in the CADM model corresponds to 1 month. This requirement represents a constraint in terms of the use of this model to simulate a variety of dosing protocols used in animal toxicity studies. This requirement, however, is not a constraint with the PBPK models. So, simulating the body burden and serum lipid concentrations for a longer duration of exposure, either model would appear to be useful; but the PBPK model would be the tool of choice for simulating alternative dose metrics of TCDD (e.g., blood concentration, total tissue concentration, bound concentration) for various exposure scenarios (including single dose studies), routes and life stages in the species of relevance, to TCDD dose-response assessment, particularly, mice, rats, and humans.

Two minor modifications, to enhance the biological basis, were made to the PBPK model of Emond et al. (2006), before its use in the computation of dose metrics for TCDD. The first one involved the recalculation of the volume of the rest of the body as follows:

$$WRE0 = (0.91 - (WLIB0 \times WLI0 + WFB0 \times WF0 + WLI0 + WF0)/(1 + WREB0)) \quad (3\text{-}22)$$

where

$WRE0$ = weight of cellular component of rest of body compartment (as fraction of body weight);

$WLI0$ = weight of cellular component of liver compartment (as fraction of body weight);

$WF0$ = weight of cellular component of fat compartment (as fraction of body weight);

*WREB*0 = weight of the tissue blood component of the rest of body compartment (as fraction of body weight);

*WLIB*0 = weight of the tissue blood component of the liver compartment (as fraction of body weight); and

*WFB*0 = weight of the tissue blood component of the fat compartment (as fraction of body weight).

In the original code, the weight of the rest of body compartment was calculated as the difference between 91% of body weight and the sum total of the fractional volumes of blood, liver tissue (intracellular component), and adipose tissue (intracellular component). The blood compartment in the PBPK model is not explicitly characterized with a volume; as a result, the total volume of the compartments is less than 91%. The recalculations shown above were used to address this problem. Given the very low affinity of TCDD for blood and rest of the body, reparameterizing the model resulted in less than a 1% change in output compared to the published version of the PBPK model for chronic exposure scenarios (Emond et al., 2006).

The second minor modification related to the calculation of the rate of TCDD excreted via urine. The original model code computed the rate of excretion by multiplying the urinary clearance parameter with the concentration in the rest of the body compartment. Instead, the code was modified to use the blood concentration in this equation. This resulted in the re-estimation of the urinary clearance value in the rat and human models but it did not result in any significant change in the fit and performance of the original model.

The revised parameter estimates of the rat, mouse, and human models are captured in Table 3-8 with a footnote.

3.3.4.5. *Recommended Dose Metrics for Key Studies*

The selection of dose metrics for the dose-response modeling of key studies is largely the result of (1) the relevance of a dose metric on the basis of current knowledge of TCDD's mechanism of action for critical endpoints and (2) the feasibility and reliability of obtaining the dose metric with available PK models. Secondarily, the goodness-of-fit of the dose-response models (which reflects the relationship of the selected internal dose measures to the response) can be used to inform selection of the most appropriate dose metric for use in deriving TCDD toxicity values.

Body burden—even though this metric is based on mechanistic considerations—is a somewhat distant measure of dose with respect to target tissue dose, and this metric represents the "overall" average concentration of TCDD in the body. However, a benefit of body burden is that this metric represents a dose measure for which the available PK models can provide highly certain estimates. Thus, the overall confidence associated with the use of body burden in TCDD assessment is categorized as medium.

The confidence in the ability of PK models to simulate blood concentration as a dose metric is high, given that the models have been shown to consistently reproduce whole blood (or serum lipid-normalized) TCDD concentration profiles in both humans and rats. Considering the facts that the PBPK models simulate whole blood rather than the serum lipid-normalized concentrations of TCDD and that the study-specific values of serum lipid content are not known with certainty, it is preferable to rely on TCDD blood concentrations as the dose metric. The blood concentrations, if intended, can be normalized on the basis of appropriate total lipid levels. However, based on mechanistic considerations, the confidence in their use would be somewhat lower for hepatic effects. This conclusion reflects the concern regarding the inconsistent relationship between the two variables with increasing dose levels and the fraction of steady-state attained at the time of observation. For other systemic effects related to tissue concentrations, the confidence in the use of TCDD serum or blood concentration is high, particularly for chronic exposures, given the absence of organ-specific nonlinear mechanisms. The tissue concentration typically cannot be calculated as a reliable dose metric with either PK model. One exception to this conclusion is the use of PBPK models to estimate levels in liver, a metric that is highly relevant based on MOA considerations. Finally, the bound concentration may be evaluated for receptor-mediated effects. This dose metric, of medium-low confidence, can be obtained with PBPK models for high dose-low dose and interspecies extrapolations. The alternative dose metrics for dose-response modeling of TCDD selected on the basis of MOA and PK modeling considerations are summarized in Tables 3-11 and 3-12.

These measures of internal dose can be obtained as peak, average, integral (AUC), or terminal values. For chronic exposures (e.g., exposures reflected by the results of a cancer bioassay) in rodents, the terminal and average values would be fairly comparable. For less-than lifetime exposures, however, the terminal and average values will differ, and therefore an overall average or integrated value (AUC) would be more appropriate. Similarly, for developmental

exposures, these alternative dose metrics can be obtained with reference to the known or hypothesized exposure window of susceptibility.

3.3.5. Uncertainty in Dose Estimates

3.3.5.1. *Sources of Uncertainty in Dose Metric Predictions*

3.3.5.1.1. *Limitations of available PK data.*

3.3.5.1.1.1. *Animal data.*

The available animal data relate to blood, liver, and adipose tissue concentrations for certain exposure doses and scenarios. Although these data are informative regarding the dose- and time-dependency of TCDD kinetics for the range covered by the specific studies (see Section 3.3.2), they do not provide the peak, average, terminal, or lipid-normalized values of dose metrics associated with the key studies selected for this assessment. The limited available animal PK data are useful, however, in the evaluation of the pharmacokinetic models (see Section 3.3.4).

3.3.5.1.1.2. *Human data.*

The human data on potential dose metrics are restricted to the serum lipid-adjusted TCDD concentrations associated with mostly uncharacterized exposures (see Sections 3.3.2 and 3.3.3). While these data are useful in estimating half-lives in exposed human individuals, they do not provide estimates of hepatic clearance or reflect target organ exposure. Some autopsy data have been used to infer the partition coefficients; however, these data were collected without quantification of the temporal nature of TCDD uptake (see Section 3.2). Despite the limitations associated with the available human data, there has been some success in using these data to infer the half-lives and elimination rates in humans using pharmacokinetic models (Aylward et al., 2005b; Carrier et al., 2005a; Emond et al., 2006).

3.3.5.1.2. *Uncertainties associated with model specification.*

Uncertainty associated with model specification should be viewed as a function of the specific application, such as interspecies extrapolation, intraspecies variability, or high dose to low dose extrapolation. Because the use of pharmacokinetic models in this assessment is limited to interspecies extrapolation and high dose to low dose extrapolation, it is essential to evaluate

the confidence in predicted dose metrics for these specific purposes. For interspecies extrapolation, the PBPK and CADM models calculate differences in dose metric between an average adult animal and an average adult human. Both models have a biologically and mechanistically-relevant structure along with a set of parameters with reasonable biological basis, and reproduce a variety of pharmacokinetic data on TCDD in both rodents and humans. These models possess low uncertainty with respect to body burden, blood, and TCDD/serum (lipid) concentration for the purpose of conducting rat to human extrapolation. However, for other dose metrics, such as free, total, or bound hepatic concentrations, the uncertainty is higher in the CADM model compared to the PBPK model due to model specification differences related to the mechanisms of sequestration and induction in the liver (see Section 3.3.3).

For the purpose of high dose to low dose extrapolation in experimental animals, confidence in both models is high with respect to a variety of dose metrics (see previous discussion). The high confidence results from the use of the PBPK models to reproduce a number of data sets covering a wide range of dose levels in rodents (rats, mice). This dose range likely covers that of most of the key toxicological studies. Given that the TCDD levels during and at the end of exposures were not measured in most of the key studies, use of the PBPK models is preferred because these models account for dose-dependent elimination, induction, and sequestration. Despite the empirical nature of the specification of these key processes in PBPK models, they essentially reproduce the dose-dependent behavior in rodents, supporting their use in deriving dose metrics for dose-response modeling of TCDD. Overall, the uncertainty associated with the use of the dose metrics (identified in Table 3-10) is less than the uncertainty associated with the use of administered dose of TCDD, for relating to the concentration within tissues to produce an effect. The administered dose does not take into account interspecies differences in the volume of distribution and clearance or the complex nonlinear processes determining the internal dose.

The PBPK model of Emond et al. (2006) could benefit from further refinement and validation, including a more explicit consideration of nondose-dependent elimination pathways. As indicated in Section 4, there is some uncertainty associated with the way the elimination of TCDD is described in the existing human PBPK model. The current model essentially treats all TCDD elimination as related to dose dependent metabolism in the liver. In this regard, the classical and more recent PK data on TCDD may be useful in further improving the confidence

in their predictions. However, it is likely that there is nondose-dependent elimination of TCDD via feces and, to a lesser extent skin; juxtaposition of available elimination rate data with the PBPK model predictions suggests that the current PBPK model modestly overestimates the dose dependency of overall TCDD elimination. (The central estimate of the slope of the relationship between the log of the TCDD elimination rate and the log of the TCDD level is only about three−fourths of that expected using the unmodified PBPK model). Emond et al. (2005) acknowledge that the model did not describe the elimination of TCDD from the blood into the intestines, but it indirectly accounted for this phenomenon with the use of the optimized elimination rate.

3.3.5.1.3. *Impact of human interindividual variability.*

The sources and extent of human variability suggested by the available data are presented in Section 3.3.3, although there is some discussion of the impact of individual differences in body fat content. The CADM model facilitates the simulation of body burden and serum lipid concentrations on the basis of BMI and tissue weights of people, and the PBPK model simulates alternative dose metrics in the fetus and in pregnant animals in addition to adult animals and humans. However, neither of these models has been parameterized for simulation of population kinetics and distribution of TCDD dose metrics. Therefore, at the present time, a quantitative evaluation of the impact of human variability on the dose metrics of TCDD is not feasible, and dose metric-based replacement of the default interindividual factor has not been attempted.

3.3.5.2. *Potential Magnitude and Sources of Uncertainty in Dose Metrics*

3.3.5.2.1. *Magnitude of uncertainty.*

The usefulness of the CADM and PBPK models for conducting high dose to low dose and interspecies extrapolations is determined by their reliability in predicting the desired dose metrics. The confidence, or conversely the magnitude of uncertainty, in the model predictions of dose metrics is dictated by the extent to which the model has been verified with empirical data relevant to the dose metric, supplemented by sensitivity and uncertainty analyses. Analysis of sensitivity or uncertainty has not been conducted with the CADM model. For the PBPK model, Emond et al. (2006) published the initial results from sensitivity analyses of acute exposure modeling (see Section 3.3.3). One of the objectives of a sensitivity analysis that is of highest

relevance to this assessment is to identify the most critical model parameters with respect to the model output (i.e., dose metric).

If the model simulations have only been compared to entities that do not correspond to the moiety representing the dose metric, or if the comparisons have only been done for some but not all relevant dose levels, routes, and species, then the reliability in the predictions of dose metric can be an issue. The extent to which model results are uncertain will depend largely upon the extent to which the dose metric is measurable (e.g., serum concentrations of TCDD) or inferred (e.g., receptor-bound TCDD concentration).

With respect to body burden and blood concentration, extrapolation uncertainty is low, and therefore the need for sensitivity and uncertainty analysis is less critical. For serum lipid-based metric, the lipid content used for normalizing the animal and human blood concentrations will have a direct impact on the outcome. Because the PBPK models have directly evaluated blood or tissue concentration data of relevance to dose metric calculations in the species and life stage of interest to the assessment, extrapolation uncertainty is low and therefore confidence in these simulated dose metrics is high. For those dose metrics that are not directly measurable or are less easily determined by available calibration methods (e.g., free liver concentration, receptor bound concentration), sensitivity and uncertainty analyses are crucial to inform the reliability of the PBPK model predictions

Sensitivity analysis for the PBPK model predictions of liver concentration of TCDD indicated hepatic CYP1A2 concentration is the most sensitive parameter in the rat model (Emond et al., 2006). In addition, the absorption parameters, basal concentration of CYP1A2, and adipose tissue:blood partition coefficients were identified as highly sensitive model parameters for simulations of human kinetics. These results indicate that the confidence in the use of the rat PBPK model for high dose to low dose extrapolation is high. Confidence is low for the purpose of rat to human extrapolation given that the values of these key human model parameters are uncertain. A similar inference can be made with respect to the concentration of bound TCDD in liver, suggesting that the use of bound concentration as a dose metric in the rat is reliable for high dose to low dose extrapolation, and less reliable for extrapolating from rat to human due to model sensitivity to parameters with high uncertainty.

With regard to the predictability of body burden, the absorption and excretion parameters were among the sensitive parameters in the rat. Several other parameters were also identified as

being sensitive in humans. Despite the sensitivity to these parameters and the uncertainty associated with individual parameter estimates, the overall confidence in the model predictions of body burden appears to be high given the reproducibility of empirical data on tissue burdens and blood concentrations of TCDD in various experiments by both models. Similar conclusions can be drawn for serum lipid concentration of TCDD predicted by the PBPK model, except that the assigned value of lipid content will have additional, direct impact on this dose metric. Therefore, knowledge of the statistical representativeness of the lipid levels for rodents and humans, intended for use in normalization, will be crucial. In this regard, variability of the total lipid levels and the variability of the contribution of phospholipids and neutral lipids to the total lipid pool across species, lifestage and study groups is to be expected (Poulin and Theil, 2001; Bernert et al., 2007).

Both conceptual uncertainty and prediction uncertainty are relevant to dose metrics used in the dose-response modeling. In the current context, conceptual uncertainty arises from the assumed relevance of the dose metrics to the MOA and target organ toxicity of TCDD. Prediction uncertainty reflects the lack of confidence in the model predictions of dose metrics. Tables 3-13 and 3-14 provide a qualitative ranking of the importance and magnitude of each dose metric with respect to these two sources of uncertainty. Overall, uncertainty is high for the use of administered dose and absorbed dose at one end of the spectrum and receptor occupancy at the other end of the spectrum. Based on the information presented in Table 3-13, for instances involving the use of rat and human PBPK models, the overall uncertainty associated with the use of body burden and serum concentrations are medium, whereas for hepatic effects, the use of liver concentration would be considered to be medium/low. While using mouse PBPK model along with the human model (see Table 3-14), the contribution of the prediction uncertainty to the overall uncertainty increases due to the limited comparison of the mouse model simulations with empirical data.

3.3.6. Use of the Emond PBPK Models for Dose Extrapolation from Rodents to Humans

EPA has selected the Emond et al. (2004, 2005, 2006) PBPK models, as modified by EPA for this assessment, for establishing toxicokinetically-equivalent exposures in rodents and

humans.[13] The 2003 Reassessment (U.S. EPA, 2003) presented a strong argument for using the relevant tissue concentration as the effective dose metric. However, no models exist for estimation of all relevant tissue concentrations. Therefore, EPA has decided to use the concentration of TCDD in blood as a surrogate for tissue concentrations, assuming that tissue concentrations are proportional to blood concentrations. Furthermore, because the RfD and cancer slope factor are necessarily expressed in terms of average daily exposure, the blood concentrations are expressed as averages over the relevant period of exposure for each endpoint. Specifically, blood concentrations in the model simulations are averaged from the administration of the first dose to the administration of the last dose plus one dosing interval unit in order to capture the peaks and valleys for each administered dose. That is, for daily dosing, 24 hours of TCDD elimination following the last dose is included in the average (the modeling time interval is one hour); for a weekly dosing protocol, a full week is included. In addition, because of the accumulation of TCDD in fat and the large differences in elimination kinetics between rodent species and humans, exposure duration plays a much larger role in TK extrapolation across species than for rapidly-eliminated compounds. Because of these factors, EPA is using discrete exposure scenarios that relate human and rodent exposure durations. The use of discrete exposure scenarios was introduced previously in Section 3.4.4.2 describing first-order kinetic modeling and is further described in the following paragraphs. This section concludes with a quantitative evaluation of the impact of exposure duration on the rodent-to-human TK extrapolation from both the human and rodent "ends" of the process.

Figure 3-30 shows the TCDD blood concentration-time profile for continuous exposure at 0.01 ng/kg-day, as predicted by the Emond human PBPK model, and the target TCDD concentrations corresponding to the three discrete exposure scenarios used by EPA in this document. The target concentrations are those that would be identified in the animal bioassay studies that correspond to a particular POD (no-observed-adverse-effect level, lowest-observed-adverse-effect level, or benchmark dose lower confidence bound) established for that bioassay. That is, the target concentrations represent the toxicokinetically-equivalent internal exposure to be translated into an equivalent human intake (or HED).

[13]The models will be referred to hereafter as the "Emond human PBPK model" and the "Emond rodent PBPK model," with variations when referring to individual species or components (e.g., gestational).

For the lifetime exposure scenario, the HED is "matched" to the lifetime average TCDD blood concentration from a lifetime animal bioassay result by determining the continuous daily intake that would result in that average blood concentration for humans over 70 years.

For the gestational exposure scenario, the effective TCDD blood concentration (usually the peak) determined for the particular POD in a particular developmental study is matched to the average TCDD blood concentration over the gestational portion of the human gestational exposure scenario. The HED is determined as the continuous daily intake, starting from birth that would result in that average blood concentration over the 9-month gestational period for a pregnancy beginning at 45 years of age. The choice of 45 years as the beginning age of pregnancy is health protective of the population in that the daily exposure achieving the target blood concentration is smaller than for earlier pregnancies.

For a less-than-lifetime exposure, the average TCDD blood concentration over the exposure period in the animal bioassay associated with the POD is matched to the average over the 5-year period that includes the peak concentration (58 years for an intake of 0.01 ng/kg-day). The HED is determined as the continuous daily intake that would result in the target concentration over peak 5-year period. The use of the peak is analogous to the approach in the 2003 Reassessment, where the terminal steady-state body burden played the same role. The 5-year average over the peak is taken to smooth out sharp peaks and more closely approximate a plateau. The choice of peak is health protective because humans of any age must be protected for short-term exposures, and the daily intake achieving a given TCDD blood concentration is smallest when matched to the peak exposure as opposed to an average over shorter durations. Thus, target concentrations for any exposure duration of less-than-lifetime must be averaged backwards from the end of the lifetime scenario, rather than from the beginning. The only exception would be if the short-term endpoints evaluated in the animal bioassay were associated with a specific life stage (such as for the gestational scenario). Note that this scenario lumps all exposures from 1 day to over 1 year in rodents into the same less-than-lifetime category. Conceptually, duration-specific scenarios could be constructed by defining equivalent rodent and human exposure durations. However, for the most part, defining duration equivalents across species is a somewhat arbitrary exercise, not generally based on physiologic or toxicologic processes, but relying primarily on fraction-of-lifetime conversions. EPA defines "lifetime" exposure as 2 years and 70 years for rodents and humans, respectively. So, a half-lifetime

equivalence of 1 year in rodents and 35 years in humans is defined easily. Also, considering a subchronic exposure to be 10–15% of lifetime, leads to an equivalence of 90 days in rodents and 7–10 years in humans. However, in the practical sense with respect to the Emond human PBPK model predictions, the difference in the dose-to-target-concentration ratios are not significantly different from the peak 5-year average scenario, differing by less than 5%.

The net effect of using three different scenarios for estimating the HED from rodent exposures is that, for the same target concentration, the ratio of administered dose (to the rodent) to HED will be larger for short-term exposures than for chronic exposures. Figure 3-31 is similar to Figure 3-30, except that it shows the relationship of daily intake to a fixed target TCDD blood concentration level. Figure 3-31 shows that, for human intakes of approximately 0.01 ng/kg-day, the difference in the defined scenarios is 40% or less, with a lifetime-scenario daily intake of 0.014 ng/kg-day required to reach the same target concentration for a shorter-term exposure of 0.01 ng/kg-day. The corresponding daily intake for the gestational scenario is 0.011 ng/kg-day. Because of the nonlinearities in the Emond human PBPK model, the magnitude of the difference between the lifetime and less-than-lifetime exposure scenarios increases at lower intake levels, but not to a substantial degree.

The differential effect of short- and long-term exposures is much more accentuated at the rodent end of the exposure kinetic modeling. Analogous to the processes described in the previous section for first-order body burden (see Section 3.4.2.2), the TCDD blood concentration for single exposures is essentially the immediate absorbed fraction of the administered dose, which will be somewhat lower than the administered dose, while for chronic exposure, the TCDD blood concentration will reflect the long-term accumulation from daily exposure, which will be very much larger than the administered dose (expressed as a daily intake). Table 3-15 shows the overall impact of TK modeling on the extrapolation of administered dose to HED, comparing the Emond PBPK and first-order body burden models. For comparison purposes, the administered dose is fixed at 1 ng/kg-day for all model runs. Large animal-to-human TK extrapolation factors (TK_{EF}) are evident for short-term mouse studies, decreasing in magnitude with increasing exposure duration. The only exception is the slightly lower extrapolation factor for the mouse 1-day exposure, which is the result of the relatively short TCDD half-life (10 days) in mice and the use of the peak TCDD blood concentration as representative of single exposures, compared to the average TCDD blood concentration over the exposure period used for multiple

exposures. The TK_{EF}s are lower for rats because of the slower elimination of TCDD in rats compared to mice. Also, because of the nonlinear kinetics inherent in the Emond PBPK model, the span of the HED (13-fold for mice) across these exposure durations is greater than the span of the lipid-adjusted serum concentration (LASC; 4-fold for mice). Because of the dose-dependence of TCDD elimination in the Emond model, the TK_{EF} becomes smaller with decreasing intake. The result of this nonlinearity is that, although Table 3-15 shows much lower TK_{EF}s for the Emond PBPK model than for the first-order body burden metric, at much lower HED levels the two models give much closer predictions.

Table 3-1. Partition coefficients, tissue volumes, and volume of distribution for TCDD in humans

Tissue	Tissue/blood partition coefficient	Tissue volume (liters, for a 60 kg person)	Effective volume of distribution (Vd—liters of blood equivalent)	Percent total Vd
Blood	1	3	3	0.25
Fat	100	11.4	1.140	94.19
Liver	6	1.56	9	0.77
Rest of the body	1.5	38.64	58	4.79
	Total	**54.6***	**1.210**	**100.00**

*The total tissue volume presented here represents only 91% of body weight because some of the weight and volume of the body is occupied by bone and other structures where TCDD uptake and accumulation do not occur to a significant extent.

Source: Wang et al. (1997), Emond et al. (2005, 2006).

Table 3-2. Blood flows, permeability factors and resulting half lives (t½) for perfusion losses for humans as represented by the TCDD PBPK model of Emond et al. (2005, 2006)

Tissue	Permeability (fraction of compartment blood flow)	Rate constant for compartmental elimination ($hour^{-1}$)	$t_{½}$ (hrs)
Fat	0.12	0.0049	143
Liver	0.03	0.77	0.90
Rest of the body	0.35	3.84	0.18

Table 3-3. Toxicokinetic conversion factors for calculating human equivalent doses from rodent bioassays

Half-life (days)[a]	Mouse	Rat (Wistar)	Rat (other)	Guinea pig
	10	20	25	40
Exposure duration (days)	Conversion factor (CF)[b] $BB_A(t_A)$:d_A given in parentheses			
1	3882 (0.77)	3815 (0.79)	3802 (0.79)	3783 (0.79)
7	1107 (2.71)	1020 (2.94)	1004 (2.99)	979 (3.07)
14	681 (4.41)	587 (5.11)	569 (5.27)	543 (5.53)
28	453 (6.62)	350 (8.56)	331 (9.06)	303 (9.90)
90	307 (9.76)	186 (16.1)	163 (18.4)	130 (23.0)
180	282 (10.6)	154 (19.5)	129(23.2)	93 (32.1)
365	270 (11.1)	141 (21.3)	115(26.0)	77 (38.9)
730	226 (11.3)	115 (22.2)	93 (27.4)	60 (42.5)

[a]Half-life for humans = 2,593 days (7.1 years).
[b]$d_H = d_A$/CF; $BB_H(t_H)$:d_H = 2,185 (1−180 days), 2,202 (365 days), 2,555 (730 days).

Table 3-4. Equations used in the CADM PBPK model*

Parameter	Equation
Hepatic Concentration (ng/kg)	$C_{hepatic} = \frac{Q_{body}}{W_l} * (f_{\min} + \frac{(f_{\max} - f_{\min}) * C_{body}}{K + C_{body}}$
Fat Concentration (ng/kg)	$C_{adipose} = \frac{Q_{body}}{Wa} * (1 - (f_{\min} + \frac{(f_{\max} - f_{\min}) * C_{body}}{K + C_{body}}))$
Hepatic Elimination	$Exr\_hepatic = k_e * Q_{body} * (1 - (f_{\min} + \frac{(f_{\max} - f_{\min}) * C_{body}}{K + C_{body}}))$
Excretion via gut of Unchanged TCDD (Exsorption)	$Exr\_gut = k_a * Q_a$
Change of TCDD due to bodyweight change	$ChangeTCDD\_BW = Q_{body} * \frac{(BW(t+dt) - BW(t))}{BW(t)}$
Amount in body as a function of time	$Q_{body}(t+dt) - Q_{body}(t) = Exr\_hepatic + Exr\_gut + ChangeTCDD\_BW$
Adipose tissue growth	$W_a = \frac{1.2 * BMI + (0.23 * Age) - 10.8 * sex}{100}$
Change of hepatic elimination constant with age	$k_e = k_{e0} - k_{eslope} * Age$

*For abbreviations and parameters, see Table 3-5.
Source: Aylward et al. (2005b).

Table 3-5. Parameters of the CADM model

Parameter	Value	Units	Comments/sources
f_{hmin}	0.01	unitless	Minimum body burden fraction in liver
f_{hmax}	0.7	unitless	Maximum body burden fraction in liver
K	100	ng/kg	Body burden at half-maximum of fraction liver
k_e	Calculated	per year	$k_e = k_{e0} - k_{e\_slope}$*(age) with enforced minimum of k_{e_min}
k_{e0}	0.85	per year	CADM-mean hepatic elimination base rate at age 0
k_{e_slope}	0.011	per year	Change in k_e per year of age
k_{e_min}	0.2	per year	Minimum hepatic elimination rate
w_a (adipose weight fraction)	Calculated	unitless	w_a = [(1.2*BMI)+0.23*Age-10.8*sex]/100
w_h (liver body weight fraction)	0.03	unitless	Assumed constant
k_a (adipose clearance factor)	0.0025	per month	Passive elimination rate from intestinal tract
Monthly dose	0.15507069	ng	per month
Estimated absorption fraction	0.97	unitless	From Moser and McLaghlan (2001)
Body weight	70	kg	Standard male weight
Sex	1	unitless	1 = male; 0 = female
Time of administration	840	months	
Initial Cbody	0.2	ng/kg	Estimated background young adults UMDES sampling
Absorbed monthly dose 1	0.150418569	ng	per month

Note: The values of fhmin, fhmax, and K were obtained by best fit of the model simulations to the experimental data with the method of least squares (Carrier et al., 2005a).

Source: Aylward et al. (2005b).

Table 3-6. Confidence in the CADM model simulations of TCDD dose metrics

Dose metric	Level of confidence
Administered dose	N/A
Absorbed dose	H
Body burden	H
Serum lipid concentration	M
Total tissue (liver) concentration	L
Receptor occupancy (bound concentration)	N/A

H = high, M = medium, L = low, NA = not applicable.

Table 3-7. Equations used in the TCDD PBPK model of Emond et al. (2006)

Aspect	Equation
Body weight growth with age	$BW_{time}(g) = BW\_T0 \times \left(\frac{0.41 \times time}{1402.5 + time}\right)$
Cardiac output	$Qc(mL/hr) = QCCAR \times 60\left(\frac{BW}{1000}\right)^{0.75}$ A factor of 60 corresponds to the conversion of minutes to hours, and 1,000 is conversion of BW from grams to kilograms.
Blood compartment	$Cb(nmol/mL) = \frac{\left[\left(Qf \times Cfb\right) + \left(Qre \times Creb\right) + \left(Qli \times Clib\right) + lymph\right]}{Qc} - \frac{\left(Cb \times CLURI\right)}{Qc}$
Tissue compartment (fat, rest of the body)	
Tissue blood subcompartment	$\frac{dAtb}{dt}(nmol/mL) = Qt(Ca - Ctb) - PAt\left(Ctb - \frac{Ct}{Pt}\right)$ $Ctb(nmol/mL) = \frac{Atb}{Wtb}$
Tissue cellular matrices	$\frac{dAt}{dt}(nmol/mL) = PAt\left(Ctb - \frac{Ct}{Pt}\right)$ $Ct(nmol/mL) = \frac{At}{Wt}$
Liver tissue compartment	
Tissue blood subcompartment	$\frac{dAlib}{dt}(nmol/mL) = Qli(Ca - Clib) - PALI(Clib - Clifree) + input_{oral}$ $Clib(nmol/mL) = \frac{Alib}{WLIB}$
Tissue cellular matrices	$\frac{dAli}{dt}(nmol/mL) = PALI(Clib - Clifree) - (KBILE\_LI \times Clifree \times WLI)$ $Cli(nmol/mL) = \frac{Ali}{Wli}$
Free TCDD concentration in liver	$Clifree(nmol/mL) = Cli - \left[Clifree \times PLI + \left(\frac{LIBMAX \times Clifree}{KDLI + Clifree}\right) + \left(\frac{CYP1A2 \times Clifree}{KDLI1A2 + Clifree}\right)\right]$
Concentration bound to AhR in hepatic tissue	$Ct_{AhRbound}(nmol/mL) = \frac{LIBMAX \times Clifree}{KDLI + Clifree}$ All other induction processes and equations have been described and presented by Wang et al. (1997).

1

Table 3-7. Equations used in the TCDD PBPK model of Emond et al. (2006) (continued)

Aspect	Equation
Gastrointestinal absorption and distribution of TCDD to the portal lymphatic circulation	
Amount of TCDD remaining in lumen cavity	$\frac{dLumen}{dt}(nmol/hr) = \left[\left(KST + KABS\right) \times lumen\right] + intake$ Lumen in the amount of TCDD remaining in the GI tract (nmol); intake is the rate of intake of TCDD during a subchronic exposure (nmol/hr).
Amount of TCDD eliminated in the feces	$\frac{dFeces}{dt}(nmol/hr) = KST \times lumen$
Absorption rate of TCDD to the blood via the lymphatic circulation	$\frac{dLymph}{dt}(nmol/hr) = KABS \times lumen \times 0.7$
Absorption rate of TCDD by the liver via portal circulation	$\frac{dPortal}{dt}(nmol/hr) = KABS \times lumen \times 0.3$

2
3 Note: Key parameters and abbreviations are defined in Table 3-10.

Table 3-8. Parameters of the PBPK model for TCDD

Parameter Description	Symbol	Parameter values					
		Human nongestational[a]	Human gestational[a]	Mouse nongestational	Mouse gestational	Rat nongestational	Rat gestational
Body weight (g)	BW	Calculated	Calculated	23-28[b]	23-28	125-250[b]	85-190[b]
Cardiac output (mL/hour/kg)	QCCAR	15.36[c,d]	Calculated	275[c]	275[c]	311.4[e]	311.4[e]
Tissue (intracellular) volumes (fraction of BW)							
Liver	WLI0	Calculated	Calculated	0.0549[f]	0.0549[f]	0.036[e]	0.036[e]
Fat	WF0	Calculated	Calculated	0.069[e]	Calculated	0.069[e]	Calculated
Tissue blood volumes							
Liver (fraction of WLI0)	WLIB0	0.266[e]	0.266[e]	0.266[e]	0.266[e]	0.266[e]	0.266[e]
Fat (fraction of WF0)	WFB0	0.05[e]	0.05[e]	0.05[e]	0.05 [e]	0.05[e]	0.05[e]
Rest of body (fraction of WRE0)	WREB0	0.03[e]	0.03[e]	0.03[e]	0.03 [e]	0.03[e]	0.03[e]
Placenta tissue fraction of tissue blood weight (unitless)	WPLAB0	N/A	0.5[g]	N/A	0.5[e]	N/A	0.5[e]
Tissue blood flow (fraction of cardiac output)							
Liver	QLIF	0.26[c]	0.26[c]	0.161[f]	0.161[f]	0.183[e]	0.183[e]
Fat	QFF	0.05[c]	0.05[c]	0.07[h]	0.07[h]	0.069[e]	0.069[e]
Placenta	QPLAF	N/A	Calculated	N/A	Calculated	N/A	Calculated
Tissue permeability (fraction of tissue blood flow)							
Liver	PALIF	0.35[e]	0.35[e]	0.35[e]	0.35[e]	0.35[e]	0.35[e]
Fat	PAFF	0.12[i]	0.12[i]	0.12[i]	0.12[i]	0.091[e]	0.091[e]
Placenta diffusional permeability fraction (unitless)	PAPLAF	N/A	0.3[g]	N/A	0.03[g]	N/A	0.3[g]
Rest of body	PAREF	0.03[e]	0.03[e]	0.03[e]	0.03[e]	0.0298[e]	0.0298[e]

Table 3-8. Parameters of the PBPK model for TCDD (continued)

Parameter Description	Symbol	Parameter values					
		Human nongestational[a]	Human gestational[a]	Mouse nongestational	Mouse gestational	Rat nongestational	Rat gestational
Partition coefficient							
Liver	PLI	6[e]	[e]	[e]	[e]	[e]	[e]
Fetus/blood partition coefficient (unitless)	PFETUS	N/A	4 [j]	N/A	4[j]	N/A	4[j]
Placenta/blood partition coefficient (unitless)	PPLA	N/A	1.5[j]	N/A	3[g]	N/A	1.5[j]
Fat	PF	100[e]	100[e]	400[e]	400[e]	100[e]	100[e]
Rest of body	PRE	1.5[e]	1.5[e]	[k]	[k]	1.5[e]	1.5[e]
Metabolism constants							
Urinary clearance elimination (mL/hour)	CLURI	4.17E-08[l]	4.17E-08[l]	0.09[i]	0.09[i]	0.01[j]	0.01[j]
Clearance - transfer from mother to fetus (mL/hour)	CLPLA_FET	N/A	16[e] [3]	N/A[3]	0.17[i]	N/A	0.17[i]
Liver (biliary elimination and metabolism; $hour^{-1}$)	KBILE_LI	Inducible	Inducible	Inducible	Inducible	Inducible	Inducible
Interspecies constant ($hour^{-1}$)	Kelv	0.0011[i]	0.0011[i]	0.4[i]	0.4[i]	0.15[e]	0.15[e]
AhR							
Affinity constant in liver (nmol/mL)	KDLI	0.1[e]	0.1[e]	0.0001[e]	0.0001[e]	0.0001[e]	0.0001[e]
Binding capacity in liver (nmol/mL)	LIBMAX	0.35[e]	0.35[e]	0.00035[e]	0.00035[e]	0.00035[e]	0.00035[e]
Placenta binding capacity (nmol/mL)	PLABMAX	N/A	0.2[j]	N/A	0.0002[j]	N/A	0.0002[j]
Affinity constant protein (AhR) in placenta (nmol/mL)	KDPLA	N/A	0.1[j]	N/A	0.0001[j]	N/A	0.0001[j]

Table 3-8. Parameters of the PBPK model for TCDD (continued)

Parameter Description	Symbol	Parameter values					
		Human nongestational[a]	Human gestational[a]	Mouse nongestational	Mouse gestational	Rat nongestational	Rat gestational
CYP1A2 induction parameters							
Dissociation constant CYP1A2 (nmol/mL)	KDLI2	40[j]	40[j]	0.02[i]	0.02[i]	0.04[j]	0.04[j]
Degradation process CYP1A2 (nmol/mL)	CYP1A2_1OUTZ	1,600[e]	1,600[e]	1.6[e]	1.6[e]	1.6[e]	1.6[e]
Dissociation constant during induction (nmol/mL)	CYP1A2_1EC50	130[e]	130[e]	0.13[e]	0.13[e]	0.13[e]	0.13[e]
Basal concentration of CYP1A2 (nmol/mL)	CYP1A2_1A2	1,600[e]	1,600[e]	1.5[k]	1.5[k]	1.6[e]	1.6[e]
First-order rate of degradation ($hour^{-1}$)	CYP1A2_1KOUT	0.1[e]	0.1[e]	0.1[e]	0.1[e]	0.1[e]	0.1[e]
Time delay before induction process (hour)	CYP1A2_1TAU	0.25[e]	0.25[e]	1.5[k]	1.5[k]	0.25[e]	0.25[e]
Maximal induction of CYP1A2 (unitless)	CYP1A2_1EMAX	9,300[i]	9,300[i]	600[e]	600[e]	600[e]	600[e]
Other constants							
Oral absorption constant ($hour^{-1}$)	KABS	0.06[i]	0.06[i]	0.48[i]	0.48[i]	0.48[e]	0.48[e]
Gastric nonabsorption constant ($hour^{-1}$)	KST	0.01[m]	0.01[m]	0.30[i]	0.30[i]	0.36[e]	0.36[e]

[a]Units for human nongestational parameters are L rather than mL and kg rather than g where applicable.
[b]Body weight varies by study (Emond et al., 2004).
[c]Krishnan (2008).
[d]Units are L/kg/hr.
[e]Wang et al. (1997).
[f]ILSI (1994).
[g]Fixed.
[h]Leung et al. (1990).
[i]Optimized.
[j]Emond et al. (2004).
[k]Wang et al. (2000).
[l]Lawrence and Gobas (1997).
[m]Calculated to estimate 87% bioavailability of TCDD in humans (Poiger and Schlatter, 1986).

Table 3-9. Regression analysis results for the relationship between log_{10} serum TCDD at the midpoint of observations and the log_{10} of the rate constant for decline of TCDD levels using Ranch Hand data

Item	Aspect	Value
Summary of fit	RSquare	0.894
	RsquareAdj	0.871
	Root mean square error	0.044
	Mean responses	0.130
	Observations (or sum weights)	11
Parameter estimates	Intercept	
	Estimate	−0.054
	Standard deviation	0.026
	t ratio	−2.07
	Prob>|t|	0.0679
	Log (TCDDpg/g)	
	Estimate	0.092
	Standard error	0.011
	t ratio	8.28
	Prob>|t|	<0.0001

Table 3-10. Confidence in the PBPK model simulations of TCDD dose metrics

Dose metric	Human model	Rat model	Mouse model
Administered dose	N/A	N/A	N/A
Absorbed dose	H	H	M
Body burden	H	H	M
Serum (blood)concentration	H	H	M
Total liver concentration	M/L	H	M
Receptor occupancy (bound concentration)	L	L	L

H = high, M = medium, L = low.

Table 3-11. Overall confidence associated with alternative dose metrics for cancer and noncancer dose-response modeling for TCDD using rat PBPK model

End point	Body burden	Blood or serum concentration	Liver concentration	Bound concentration in liver
Liver effects	M		H	M/L
Nonhepatic effects	M	H		M/L

H = high, M = medium, L = low.

Table 3-12. Overall confidence associated with alternative dose metrics for cancer and noncancer dose-response modeling for TCDD using mouse PBPK model

End point	Body burden	Blood or serum concentration	Liver concentration	Bound concentration in liver
Liver effects	M		M	L
Nonhepatic effects	M	M		L

H = high, M = medium, L = low.

Table 3-13. Contributors to the overall uncertainty in the selection and use of dose metrics in the dose-response modeling of TCDD based on rat and human PBPK models

Dose metric	Conceptual uncertainty	Prediction uncertainty
Administered dose	H	NA
Absorbed dose	H	L
Body burden	M	L
Blood or serum concentration	M	L
Tissue concentration	L	M
Receptor occupancy	L(?)	H

H = high, M = medium, L = low, NA = not applicable, ? = if relevant to MOA of response.

Table 3-14. Contributors to the overall uncertainty in the selection and use of dose metrics in the dose-response modeling of TCDD based on mouse and human PBPK models

Dose metric	Conceptual uncertainty	Prediction uncertainty
Administered dose	H	NA
Absorbed dose	H	L
Body burden	M	M
Blood or serum concentration	M	M
Tissue concentration	L	M/H
Receptor occupancy	L(?)	H

H = high, M = medium, L = low, NA = not applicable, ? = if relevant to MOA of response.

Table 3-15. Impact of toxicokinetic modeling on the extrapolation of administered dose to HED, comparing the Emond PBPK and first-order body burden models

Exposure duration (days)	1st-order BB		Emond PBPK		
	HED (ng/kg−day)	TK_{EF}	LASC (ng/kg)	HED (ng/kg−day)	TK_{EF}
Mouse					
1	2.57E-4	3,882	75.5	9.49E-4	1,054
14	1.47E-3	681	64.4	8.17E-4	1,224
90	3.25E-3	307	173	3.83E-3	261
365	3.70E-3	270	248	6.66E-3	150
730	4.43E-3	226	263	1.08E-2	93
Rat					
1	2.63E-4	3,802	110	1.87E-3	535
14	1.76E-3	569	208	5.22E-3	192
90	6.13E-3	163	599	2.81E-2	36
365	8.68E-3	115	811	4.52E-2	22
730	1.07E-2	93	853	6.47E-2	15

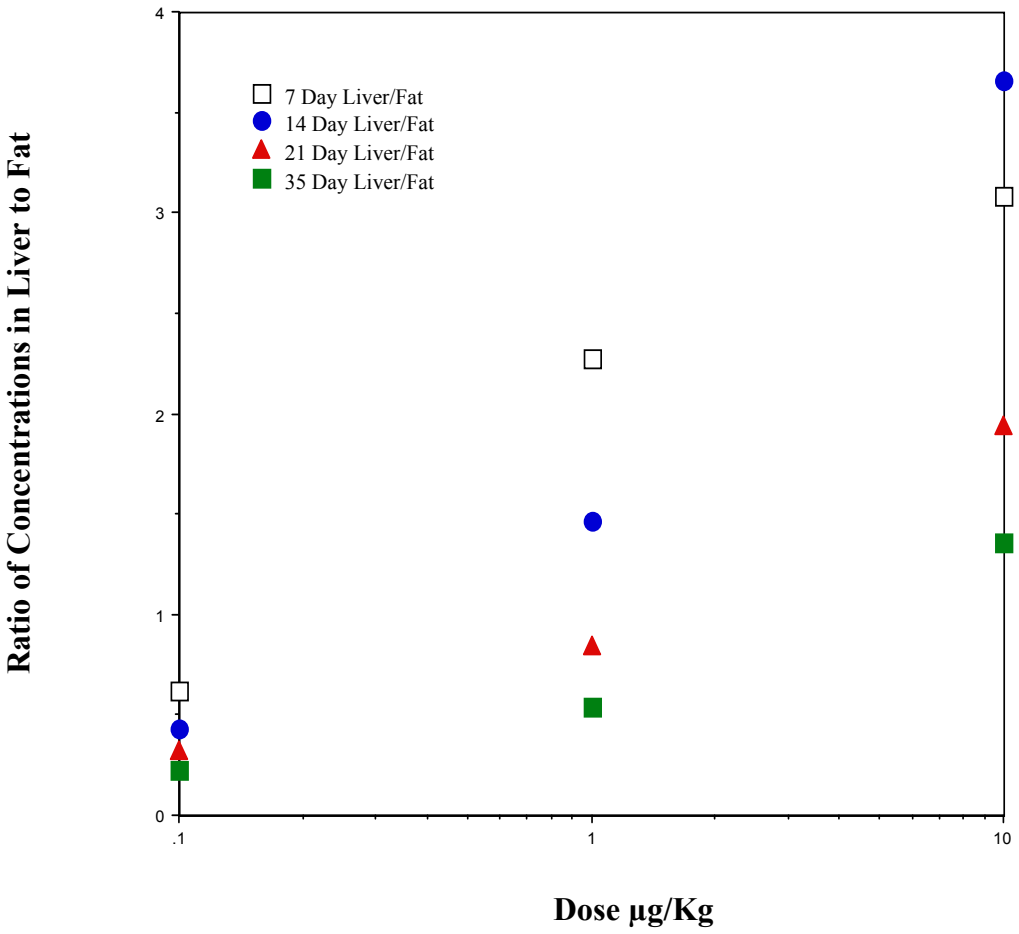


Figure 3-1. Liver/fat concentration ratios in relation to TCDD dose at various times after oral administration of TCDD to mice.

Source: Dilberto et al. (1995).

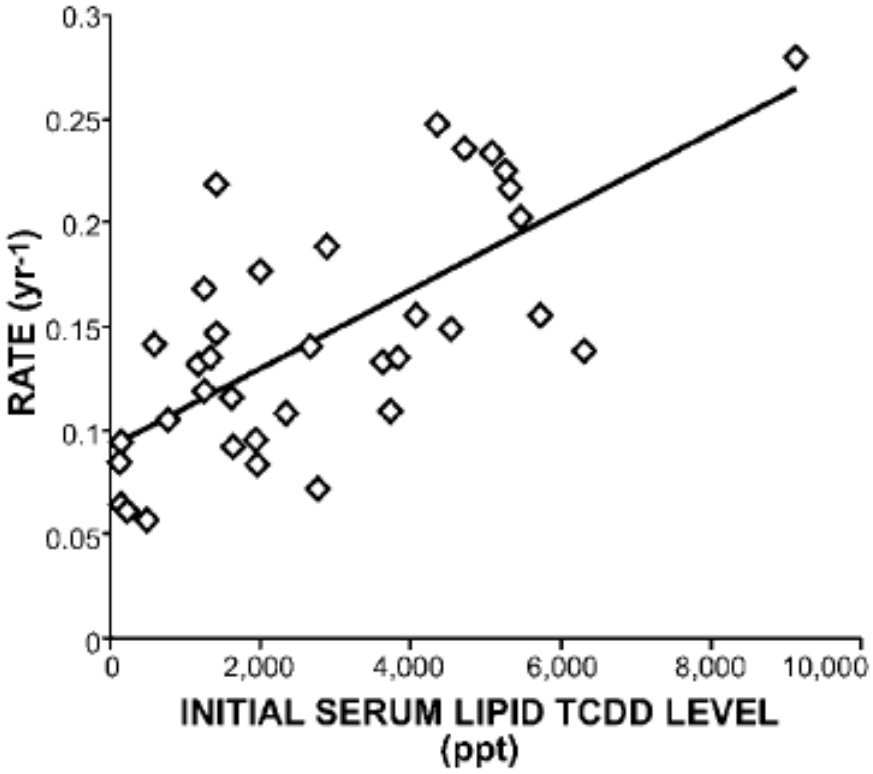


Figure 3-2. First-order elimination rate fits to 36 sets of serial TCDD sampling data from Seveso patients as function of initial serum lipid TCDD.

Source: Aylward et al. (2005).

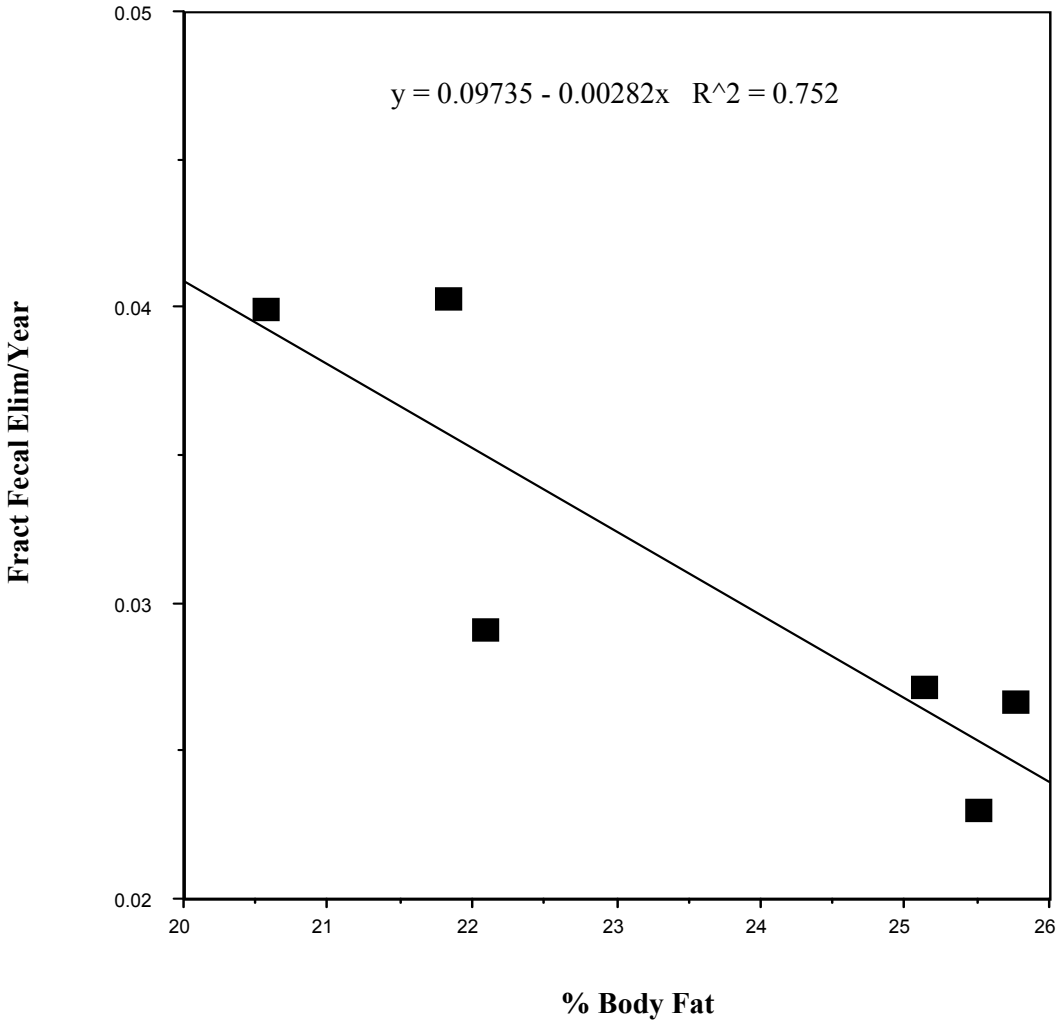


Figure 3-3. Observed relationship of fecal 2,3,7,8-TCDD clearance and estimated percent body fat.

Source: Rohde et al. (1999).

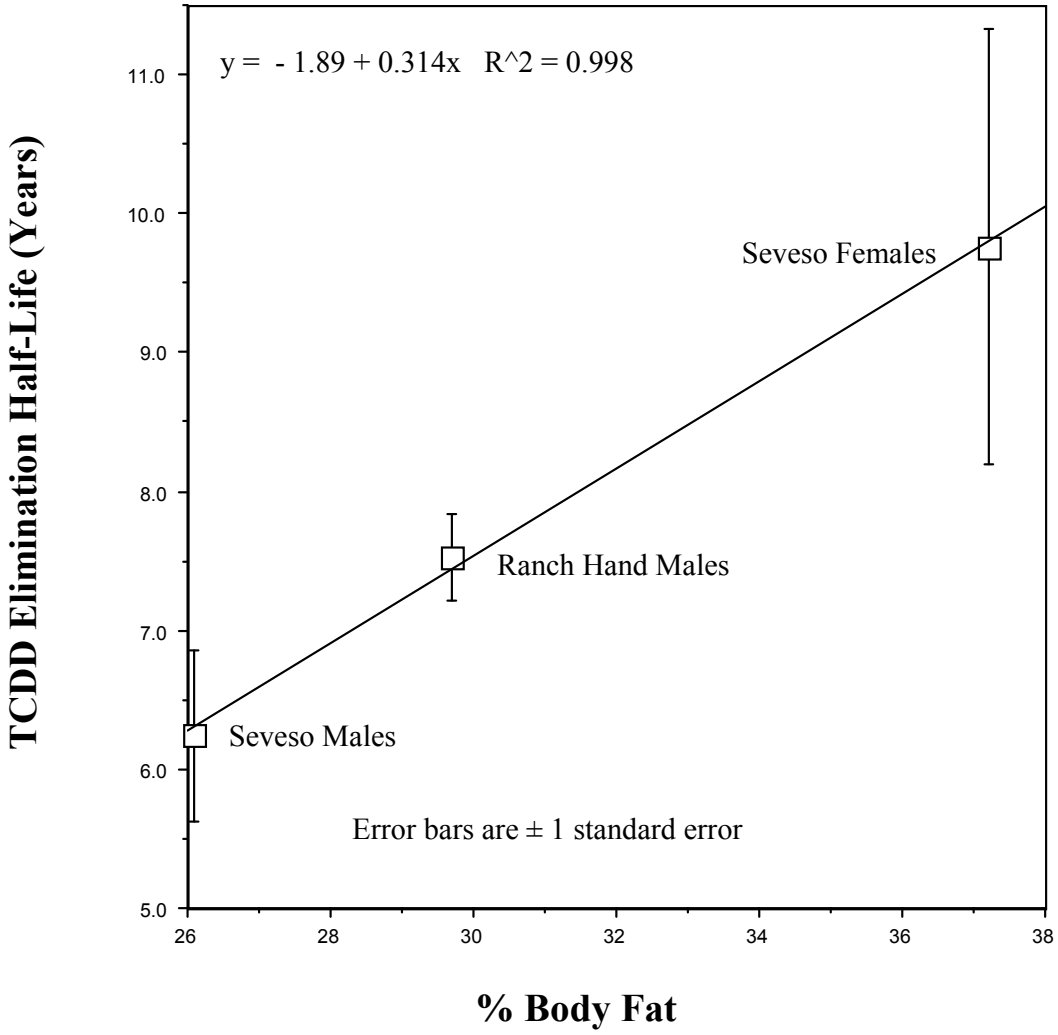


Figure 3-4. Unweighted empirical relationship between percent body fat estimated from body mass index and TCDD elimination half-life—combined Ranch Hand and Seveso observation.

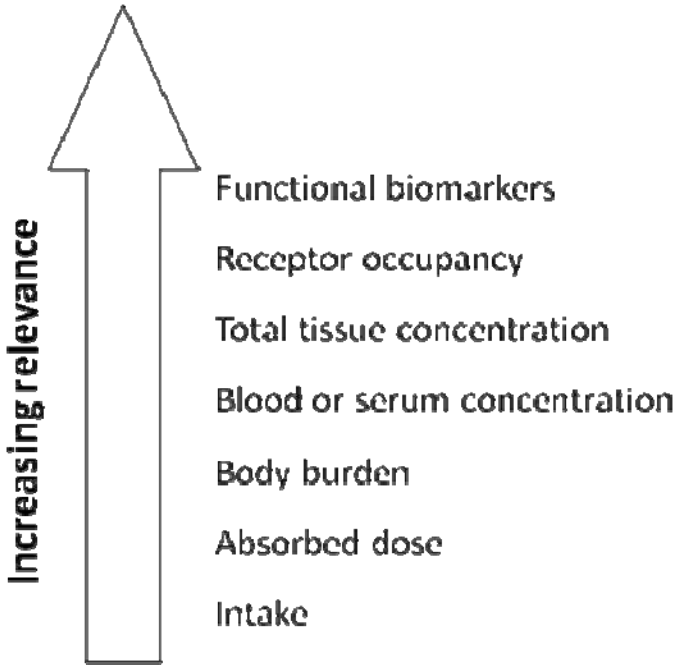


Figure 3-5. Relevance of candidate dose metrics for dose-response modeling, based on mode of action and target organ toxicity of TCDD.

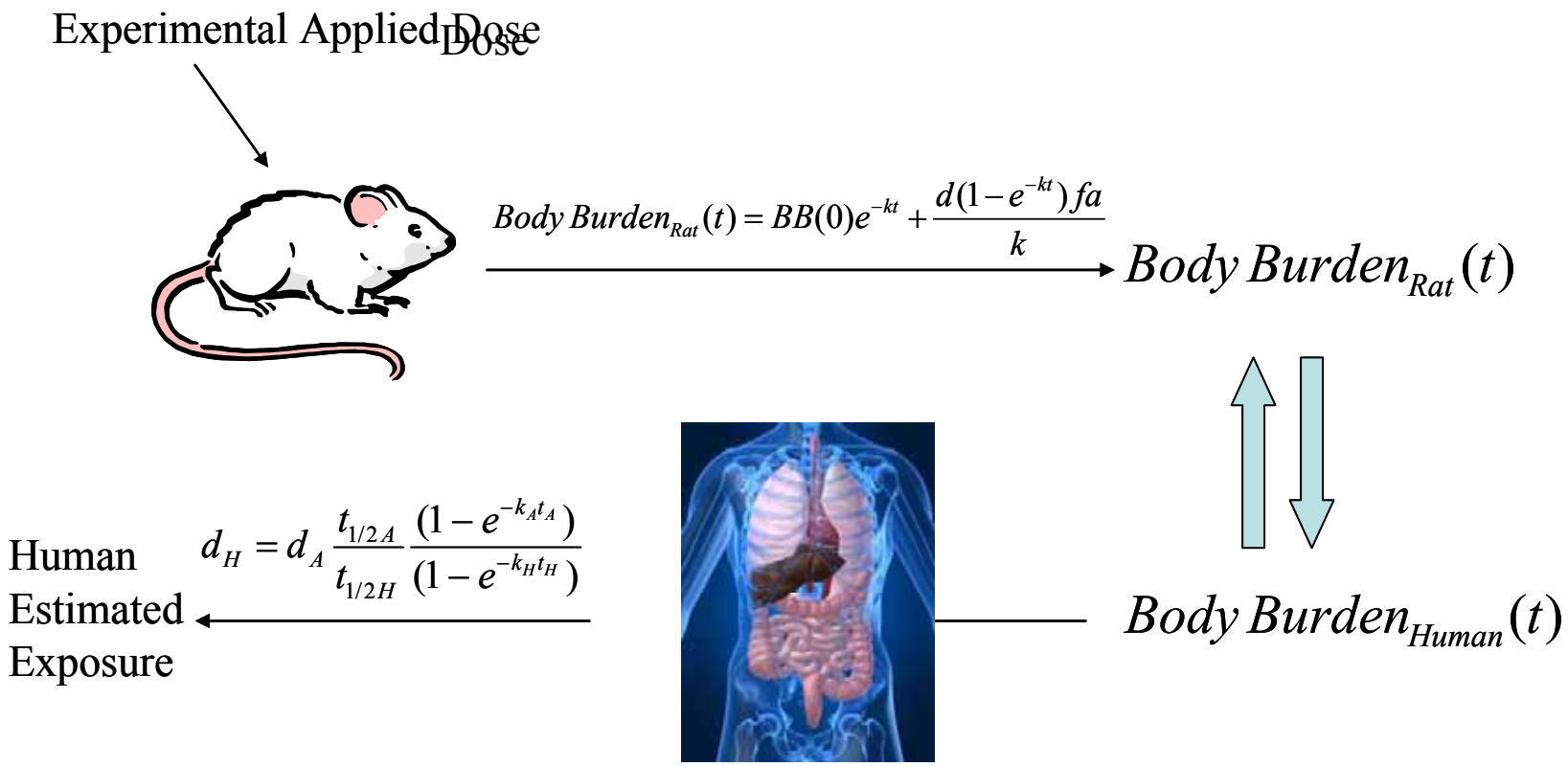


Figure 3-6. Process of estimating a human-equivalent TCDD lifetime average daily oral exposure (d_H) from an experimental animal average daily oral exposure (d_A) based on the body-burden dose metric. The arrows represent mathematical conversions based on toxicokinetic modeling. BB_A (TWA animal body burden) and BB_H (TWA human body burden) are assumed to be toxicokinetically equivalent. See text for further explanation.

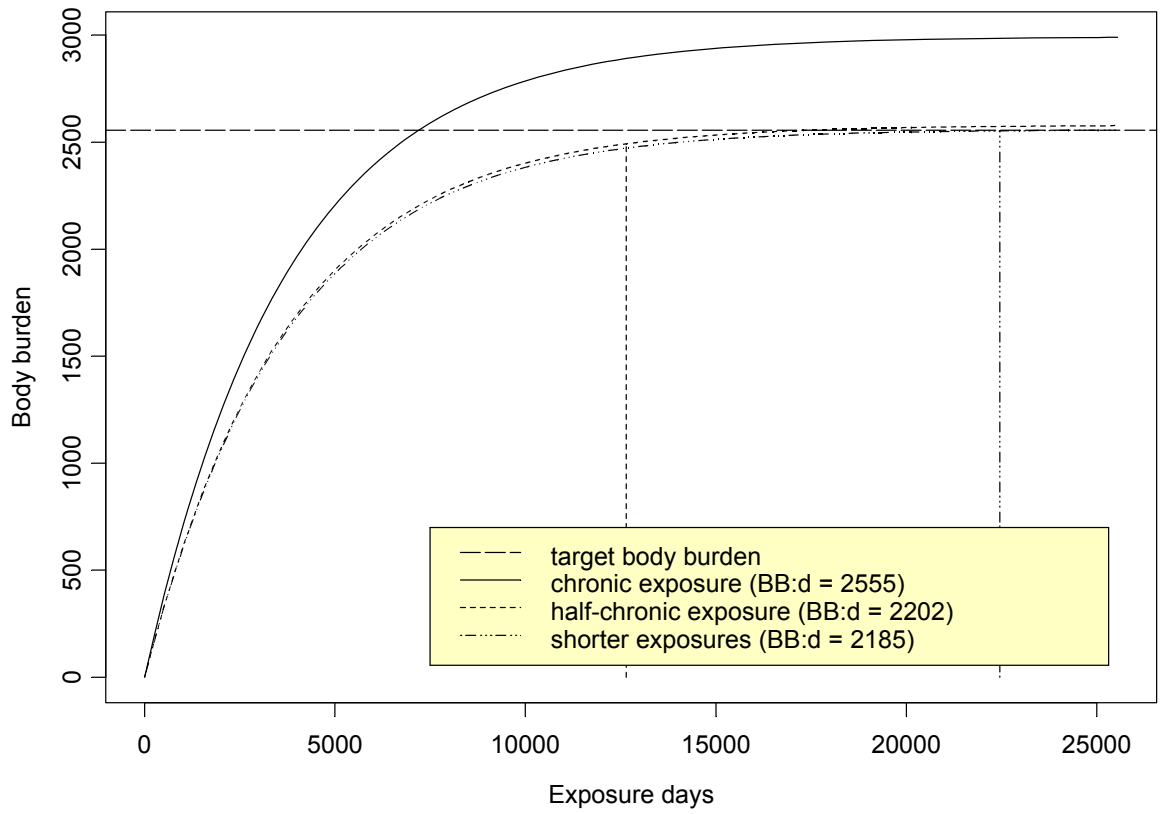


Figure 3-7. Human body burden time profiles for achieving a target body burden for different exposure duration scenarios. BB:d is $BB_H(t_H)$:d_H in Figure 3-6. The curve depicted using the solid line illustrates the increase in the human body burden over time for a hypothetical human administered a daily TCDD dose where the time-weighted average human body burden estimate over the lifetime is equal to the target body burden attained in a rodent bioassay. When compared to shorter durations (dashed lines), a higher average daily TCDD dose is required to yield a time-weighted average human body burden over a lifetime that is equal to the target body burden attained in a rodent bioassay. The half-chronic exposure scenario (depicted using a dashed line) is equivalent to a 1-year exposure in rodents. When compared to a chronic BB_H, a lower value of d_H is needed to attain the target body burden in a rodent bioassay when the time-weighted average is over the last 35 years of life; the dose to plateau ratio is also smaller (i.e., $d_{H,C} < d_{H,SC}$ to attain the target body burden in a rodent bioassay). The shorter exposure scenario is equivalent to most other shorter rodent exposure durations, from 1 day to subchronic, which are indistinguishable with respect to the BB:d ratio (subchronic shown).

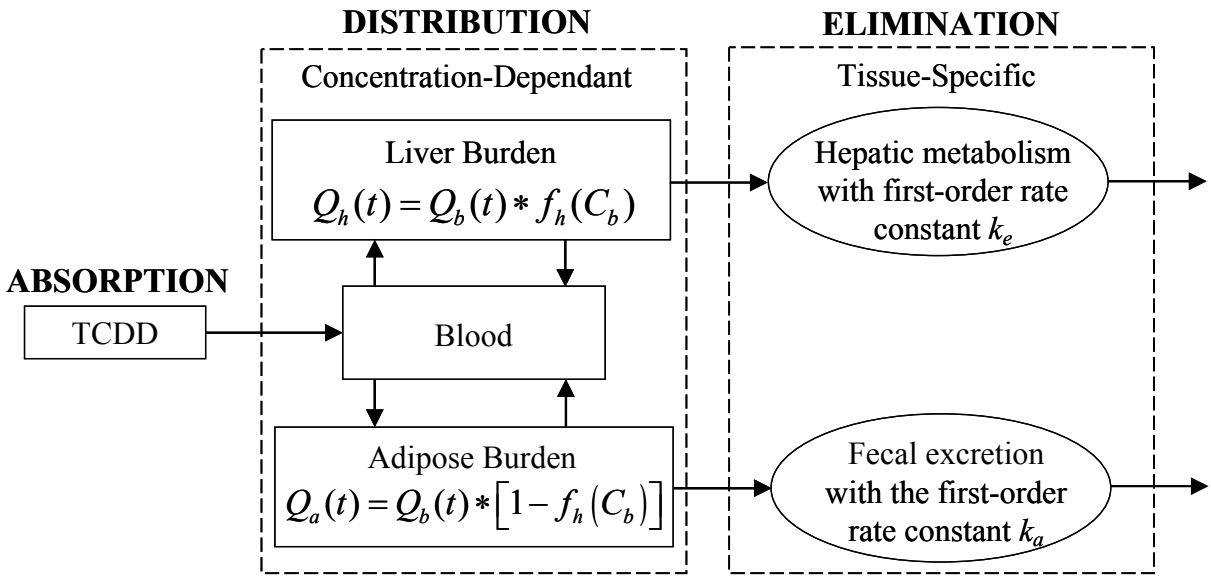


Figure 3-8. Schematic of the CADM structure.

Source: Aylward et al. (2005).

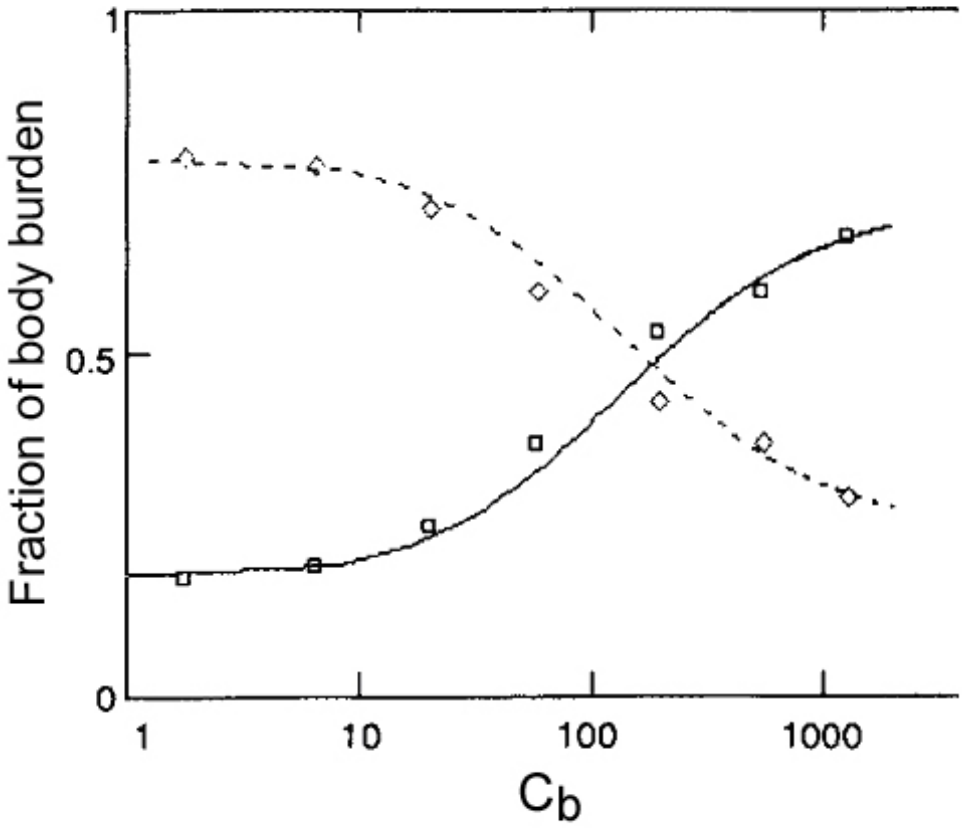


Figure 3-9. Comparison of observed and simulated fractions of the body burden contained in the liver and adipose tissues in rats. f_h, fraction contained in liver (observation) (□); $f_{h\text{-}sim}$, fraction contained in liver (simulation) (—); f_{at}, fraction contained in the adipose tissue (observation) (◊); $f_{at\text{-}sim}$, fraction contained in the adipose tissue (simulation) (---); and $C_{b,}$ body concentration in ng TCDD/kg body wt.

Source: Carrier et al. (1995a); data from Abraham et al. (1988) measured 7 days after dosing.

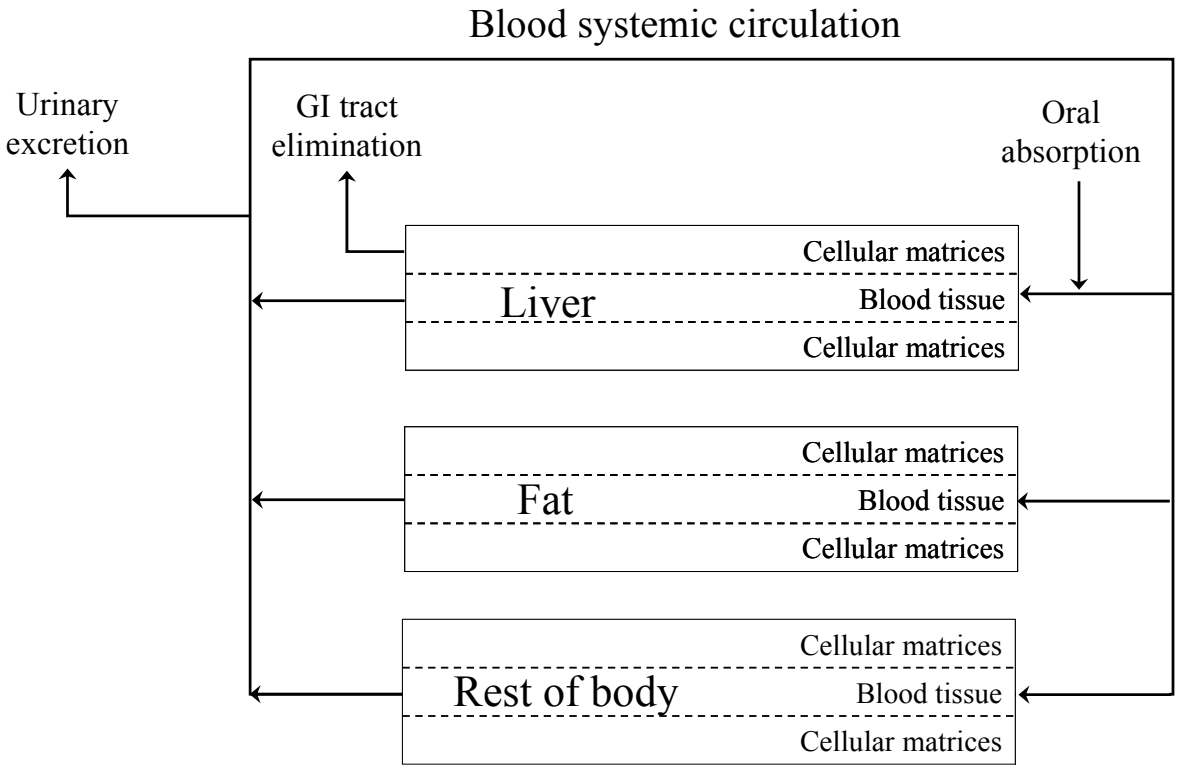


Figure 3-10. Conceptual representation of PBPK model for rat exposed to TCDD.

Source: Emond et al. (2006).

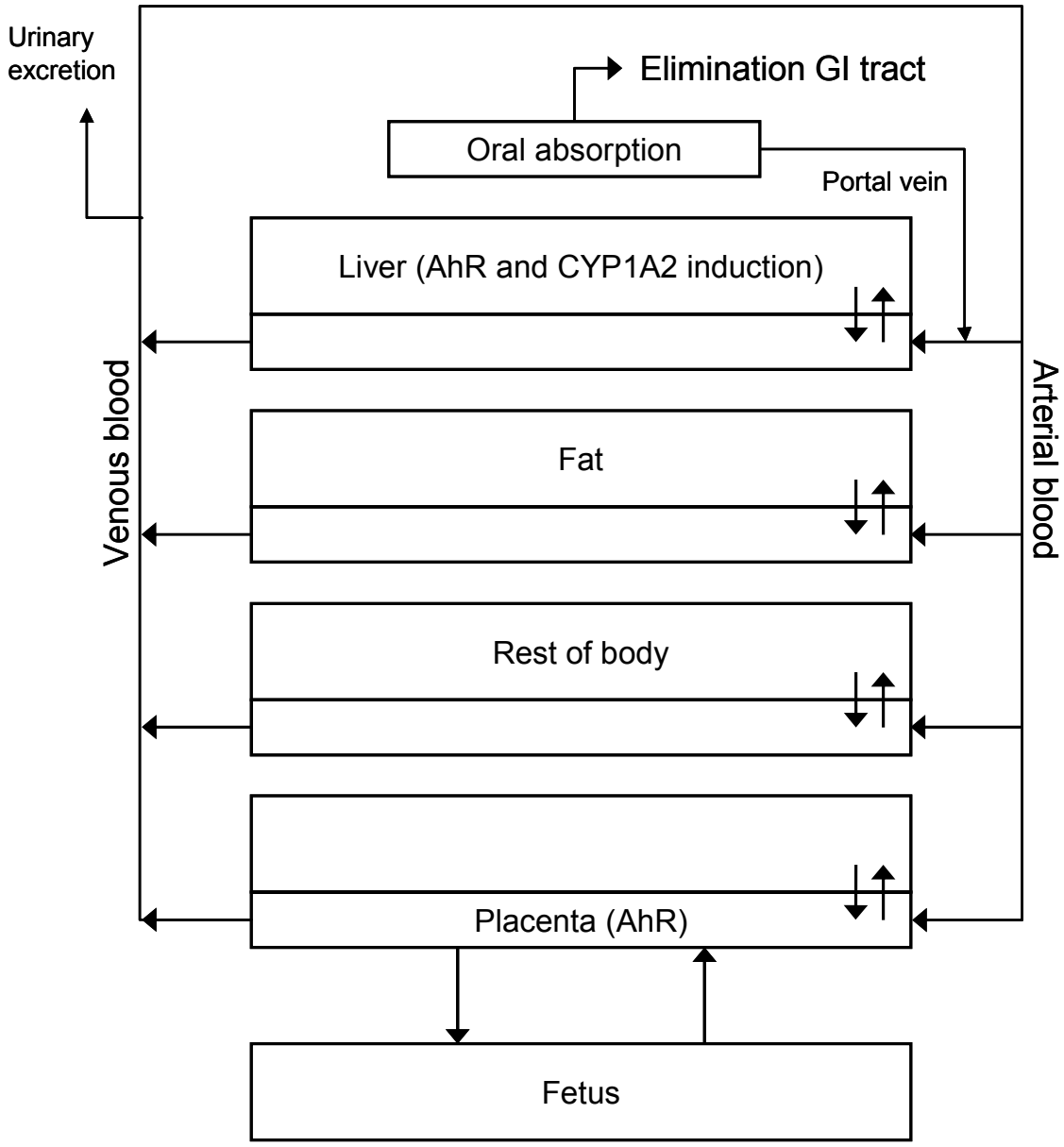


Figure 3-11. Conceptual representation of PBPK model for rat developmental exposure to TCDD.

Source: Emond et al. (2004).

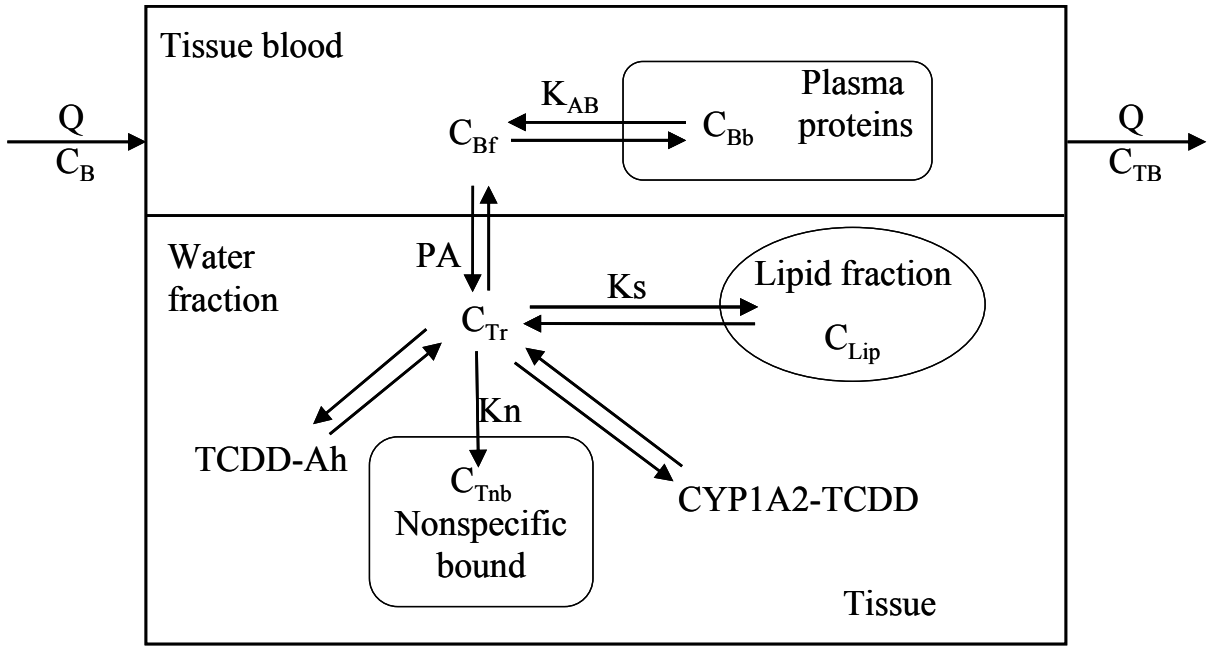


Figure 3-12. TCDD distribution in the liver tissue.

Source: Wang et al. (1997).

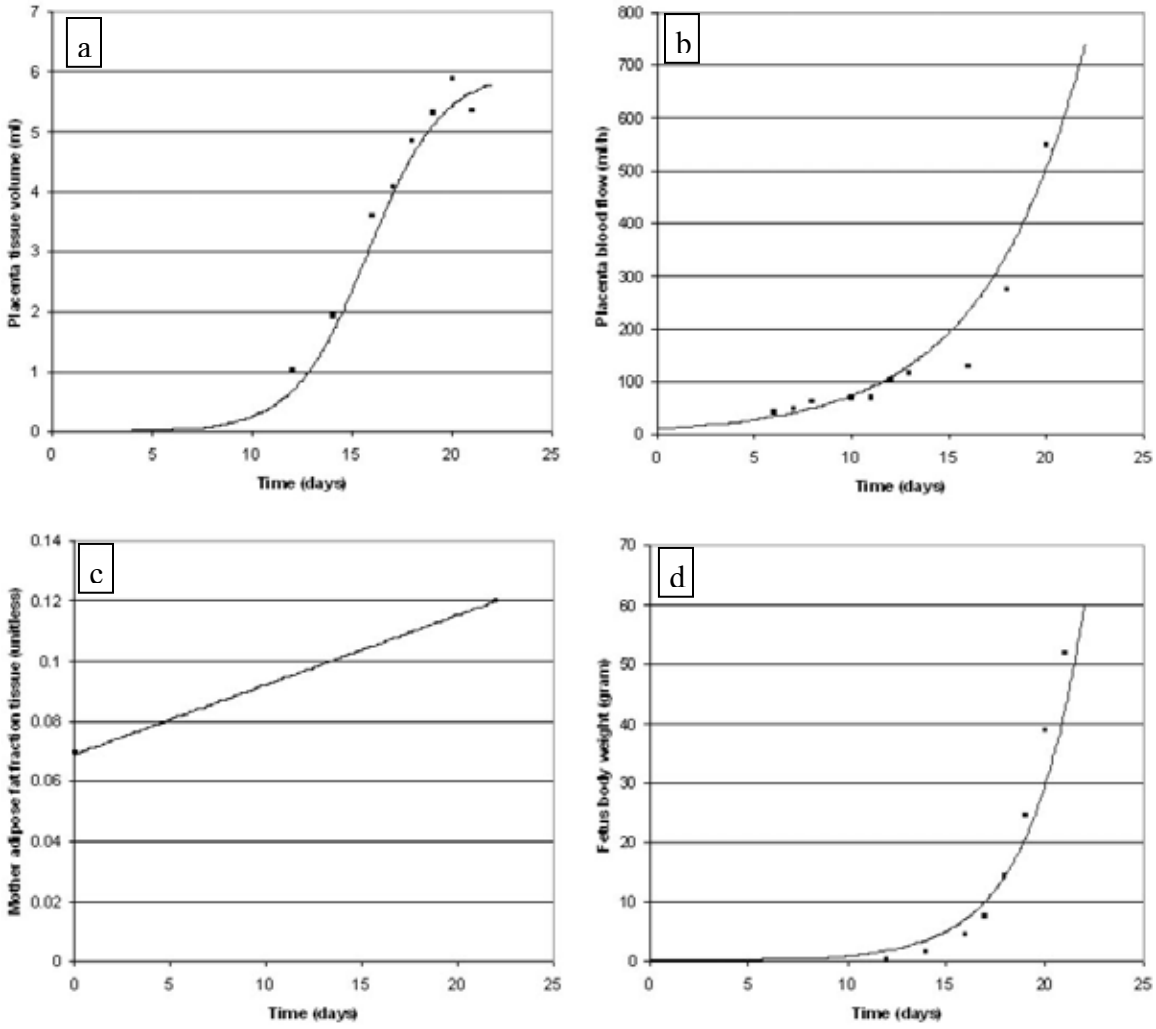


Figure 3-13. Growth rates for physiological changes occurring during gestation. (a) Placental growth during gestation (calculated for $n = 10$ placenta). Experimental data from Sikov (1970). (b) Blood flow rate in Placental compartment during gestation. Experimental data from Buelke-Sam et al. (1982a, b). (c) Fat fraction of body weight during gestation. Experimental data came from Fisher et al. (1989), and (d) Fetal growth during gestation. Experimental data obtained from Sikov (1970).

A

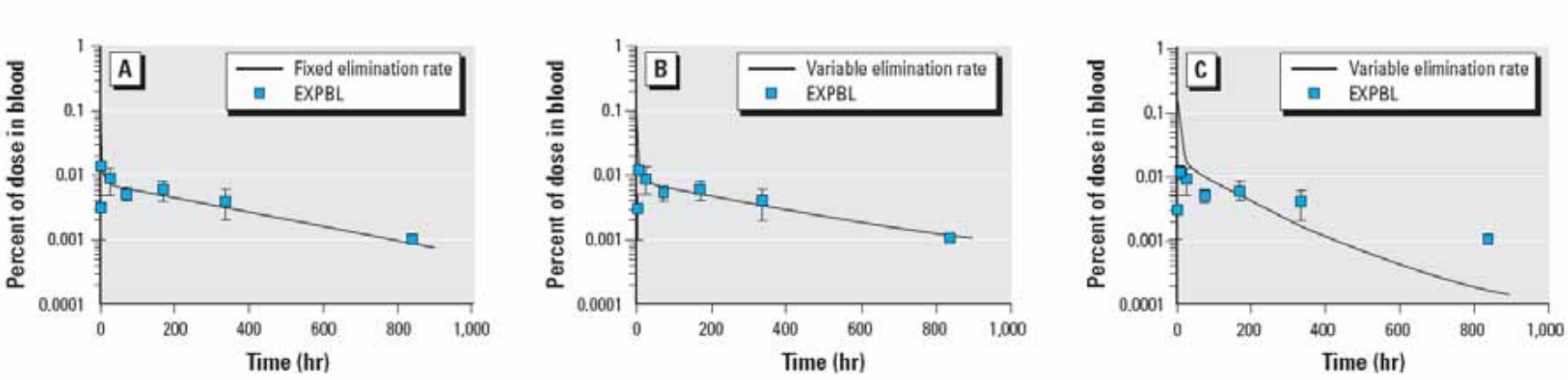


Figure 3-14. Comparisons of model predictions to experimental data using a fixed elimination rate model with hepatic sequestration (A) and an inducible elimination rate model with (B) and without (C) hepatic sequestration. EXBL, experimental blood levels. Model predictions were compared with the data of Santostefano et al. (1998), where female rats were exposed to a single oral dose of 10 μg of TCDD/kg BW. Error bars are ± SD.

Source: Edmond et al. (2006).

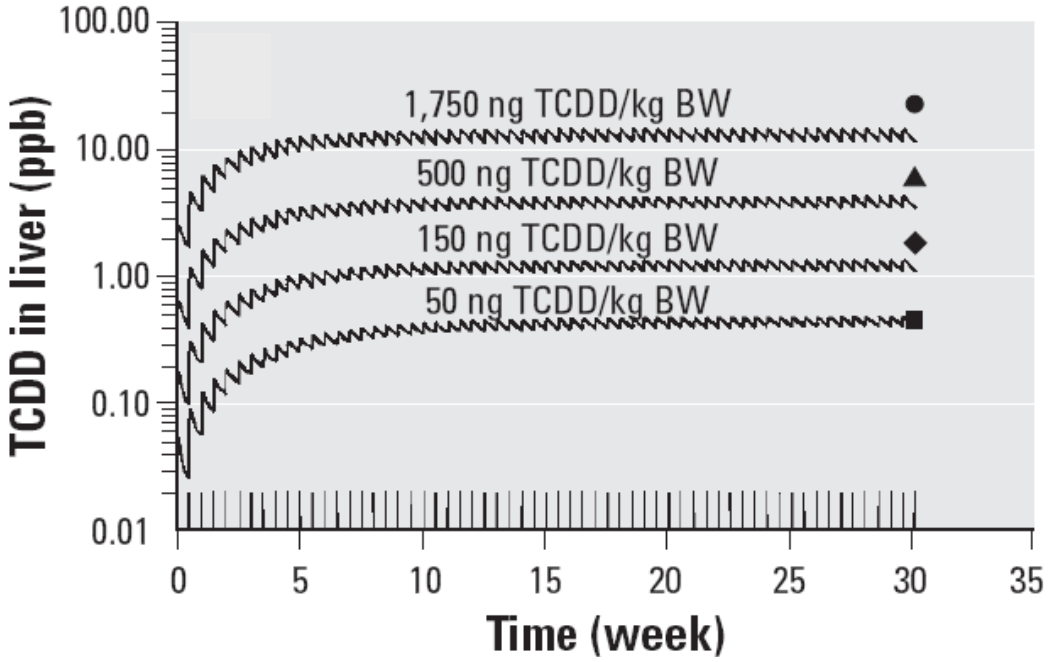


Figure 3-15. PBPK model simulation of hepatic TCDD concentration (ppb) during chronic exposure to TCDD at 50, 150, 500, 1,750 ng TCDD/BW using the inducible elimination rate model compared with the experimental data measured at the end of exposure.

Source: Emond et al. (2006).

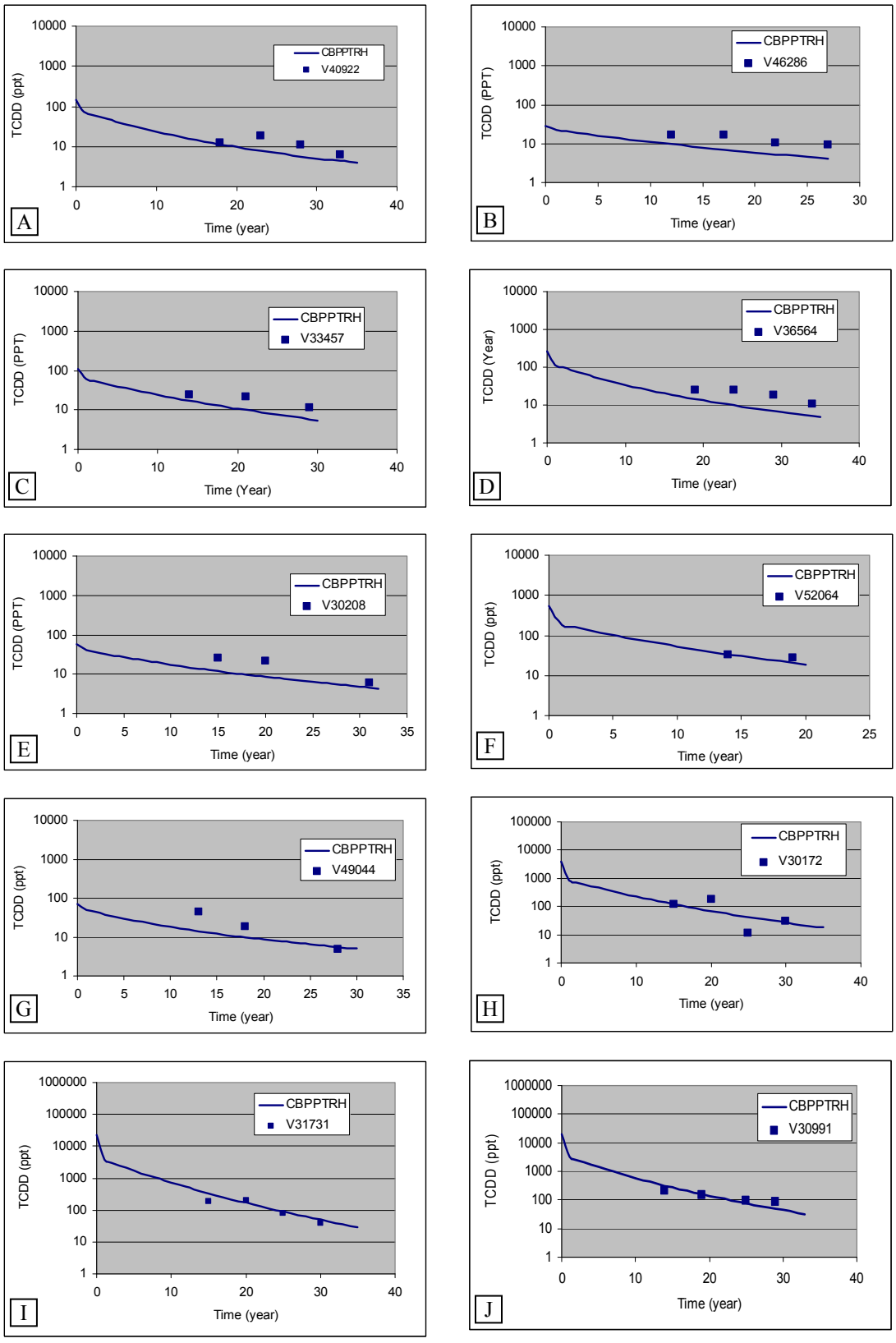
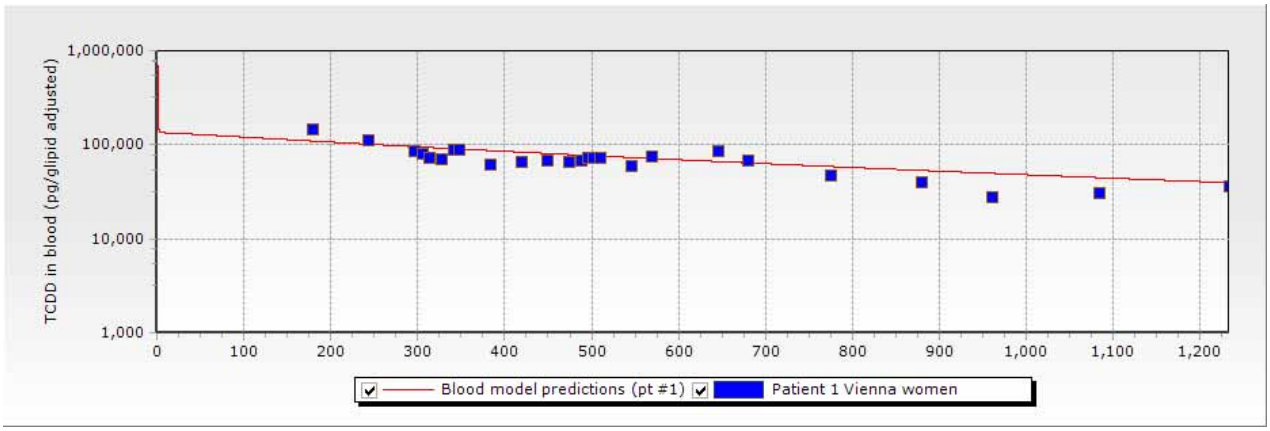


Figure 3-16. Model predictions of TCDD blood concentration in 10 veterans (A-J) from Ranch Hand Cohort.

Source: Emond et al. (2005).



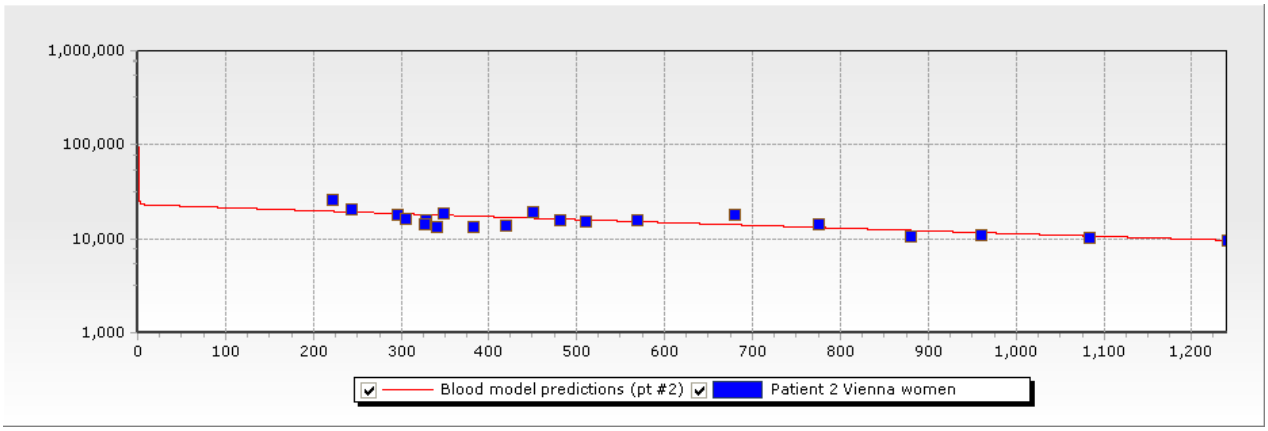


Figure 3-17. Time course of TCDD in blood (pg/g lipid adjusted) for two highly exposed Austrian women (patients 1 and 2). Symbols represent measured concentrations, and lines represent model predictions. These data were used as part of the model evaluation (Geusau et al., 2002).

Source: Emond et al. (2005).

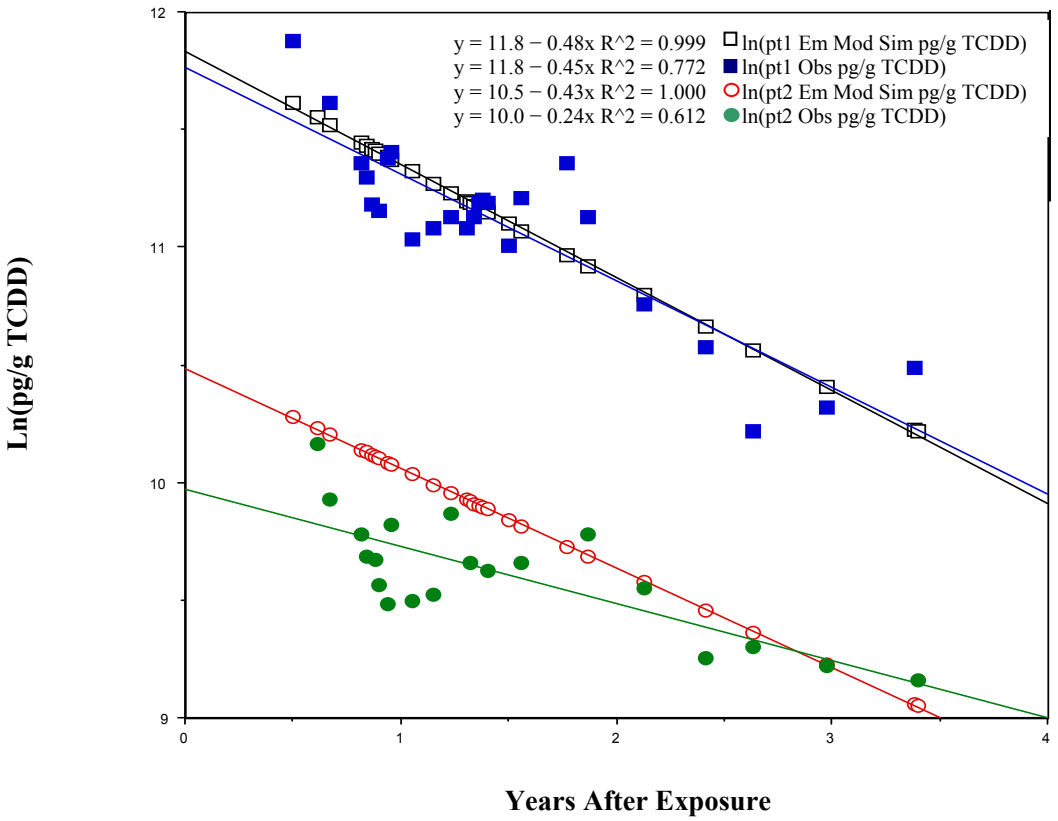


Figure 3-18. Observed vs. Emond et al. (2005) model simulated serum TCDD concentrations (pg/g lipid) over time (ln = natural log) in two Austrian women. Data from Geusau et al. (2002).

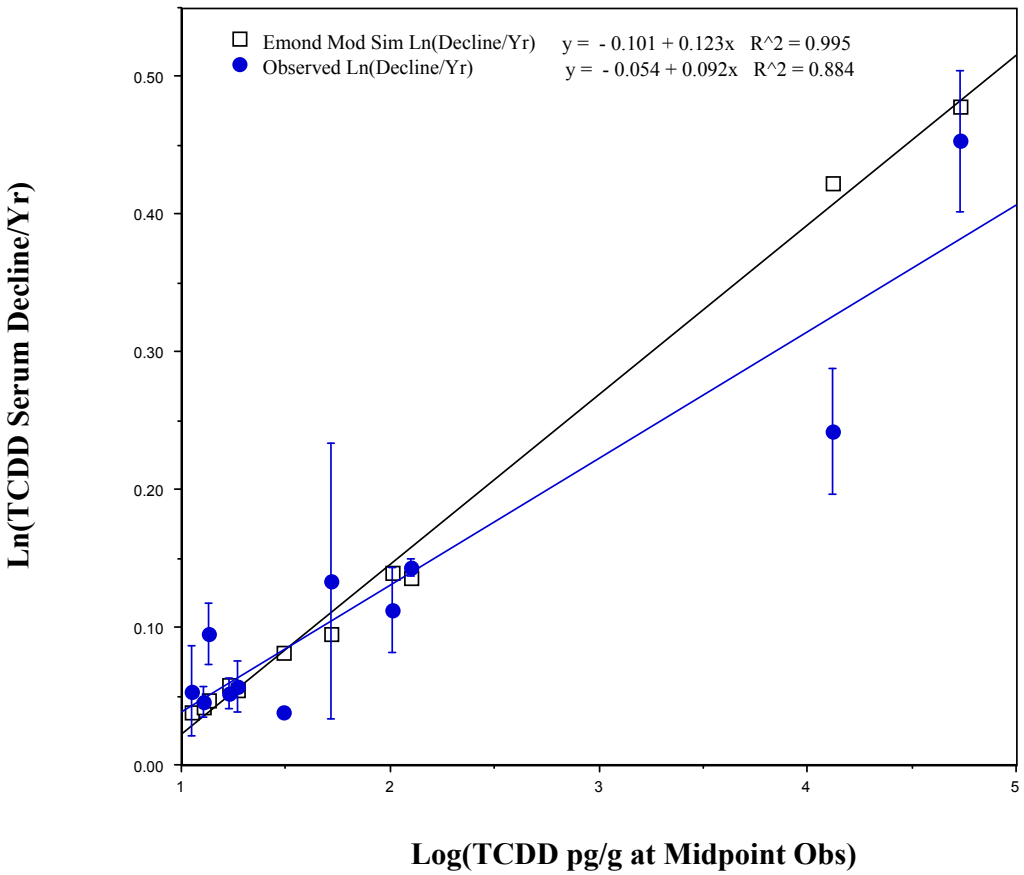


Figure 3-19. Comparison of the dose dependency of TCDD elimination in the Emond model vs. observations of nine Ranch Hand veterans and two highly exposed Austrian patients. Circles are observed data.

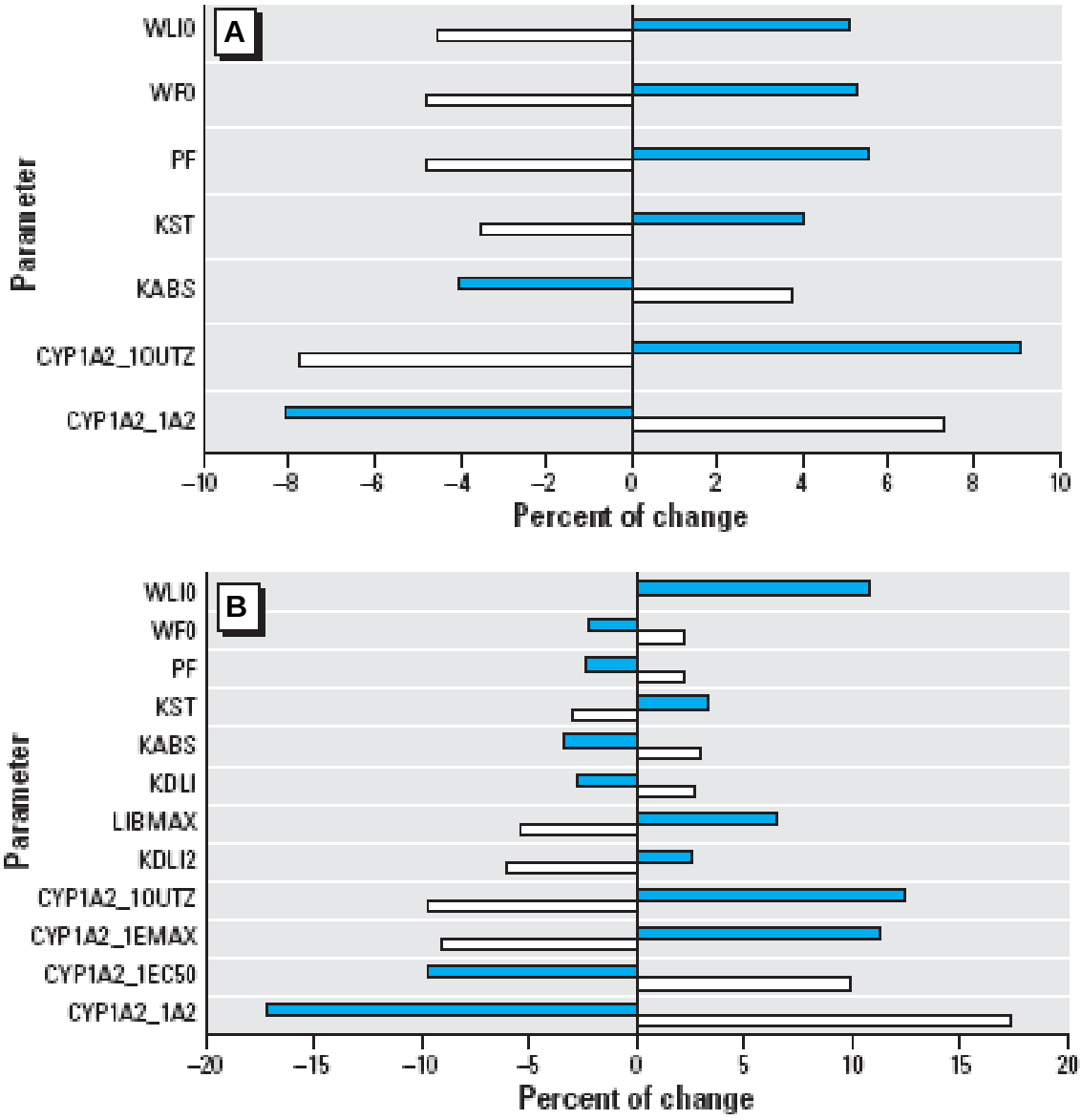
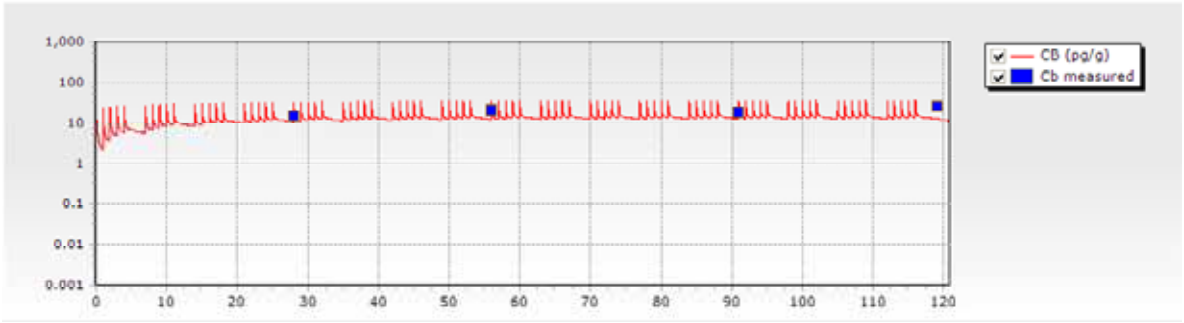


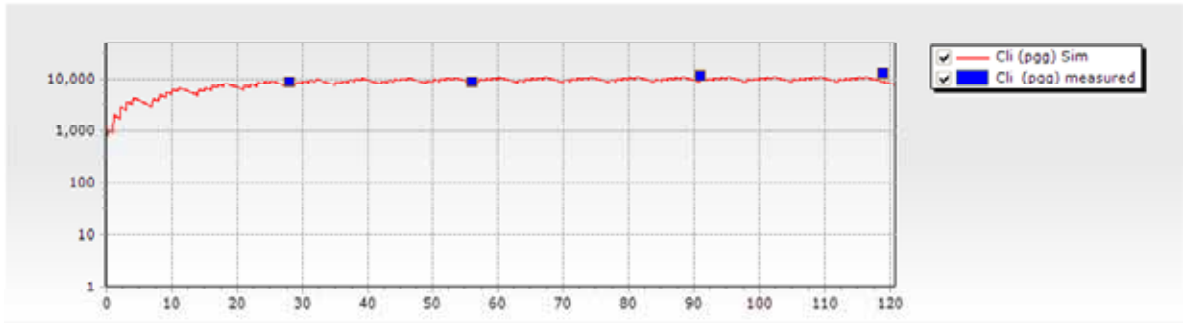
Figure 3-20. Sensitivity analysis was performed on the inducible elimination rate. The analysis was performed at 0.001 µg/kg (A) and at 10 µg/kg (B).

Source: Emond et al. (2006).

A



B



C

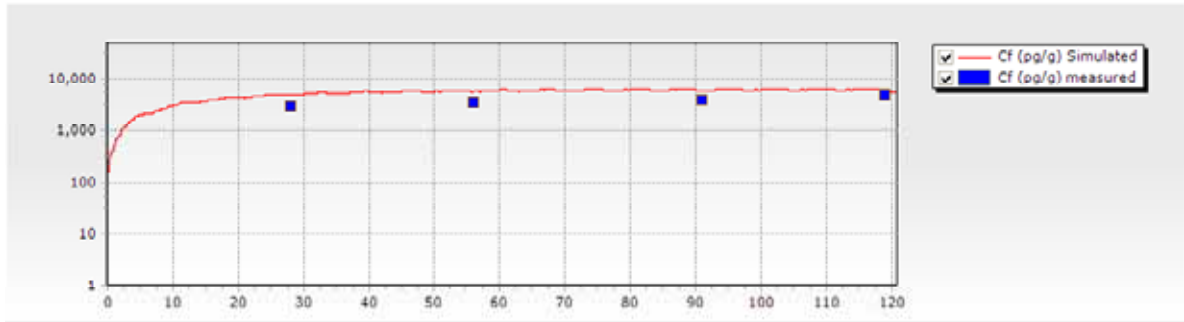


Figure 3-21. Experimental data (symbols) and model simulations (solid lines) of (A) blood, (B) liver and (C) adipose tissue concentrations of TCDD after oral exposure to 150 ng/kg-day, 5 days/week for 17 weeks in mice. Y-axis represents concentration in pg/g and X-axis represents time in days.

Source: Experimental data were obtained form Diliberto et al (2001).

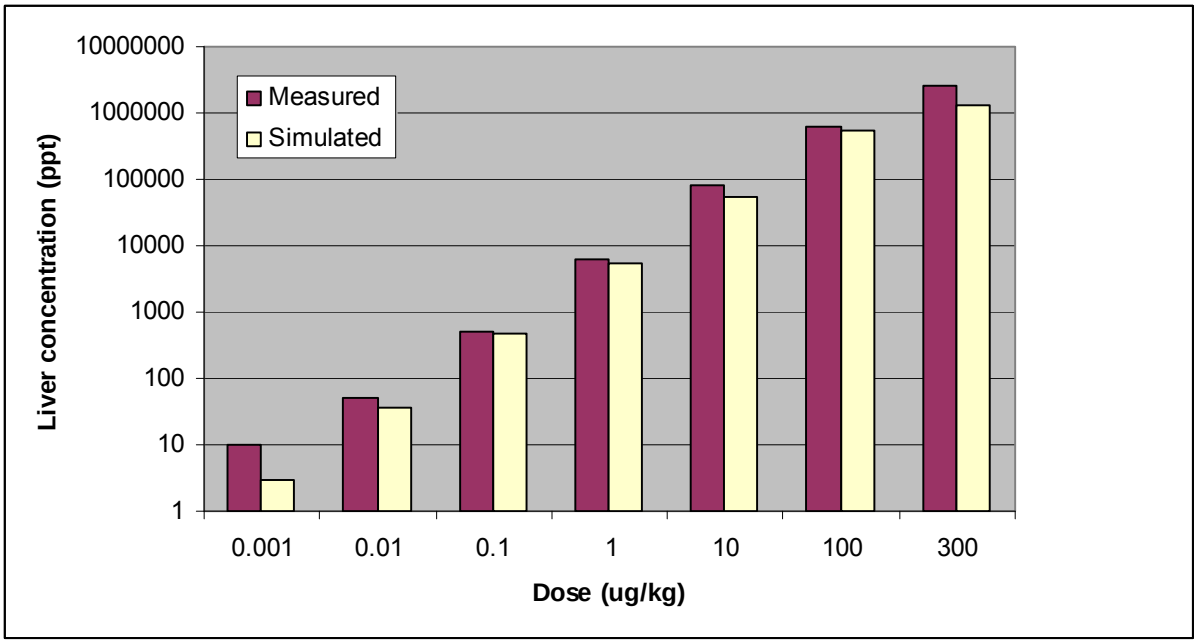


Figure 3-22. Comparison of PBPK model simulations with experimental data on liver concentrations in mice administered a single oral dose of 0.001−300 µg TCDD/kg. The simulations and experimental data were obtained 24 hour post-exposure.

Source: Data obtained from Boverhoff et al. (2005).

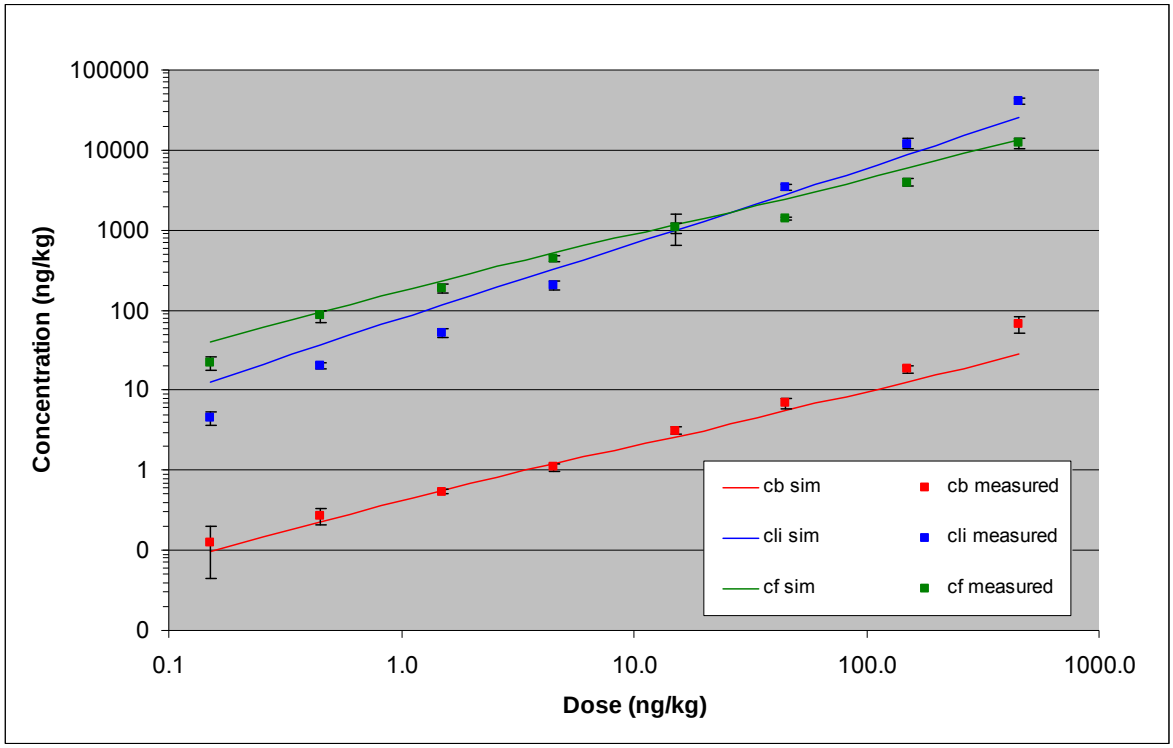
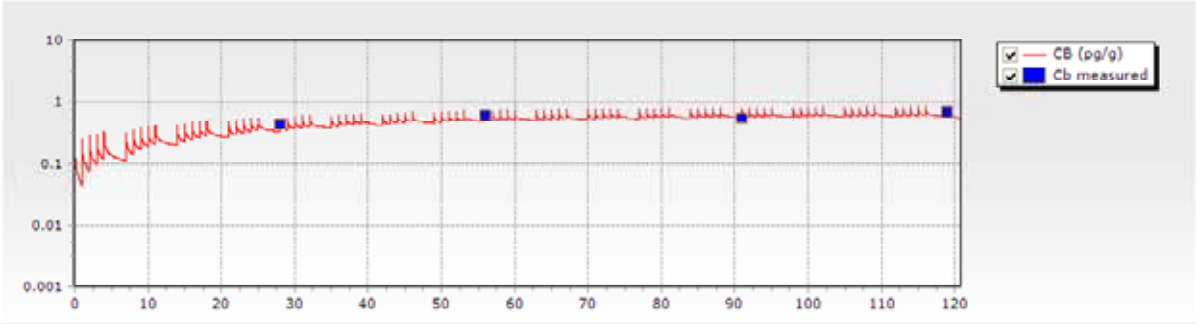


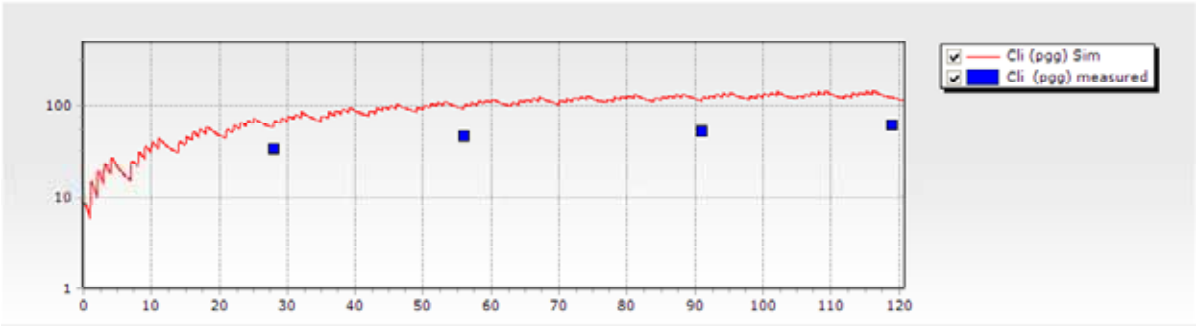
Figure 3-23. Comparison of model simulations (solid lines) with experimental data (symbols) on the effect of dose on blood (cb), liver (cli) and fat (cf) concentrations following repetitive exposure to 0.1−450 ng TCDD/kg, 5 days/week for 13 weeks in mice.

Source: Data obtained from Diliberto et al. (2001).

A



B



C

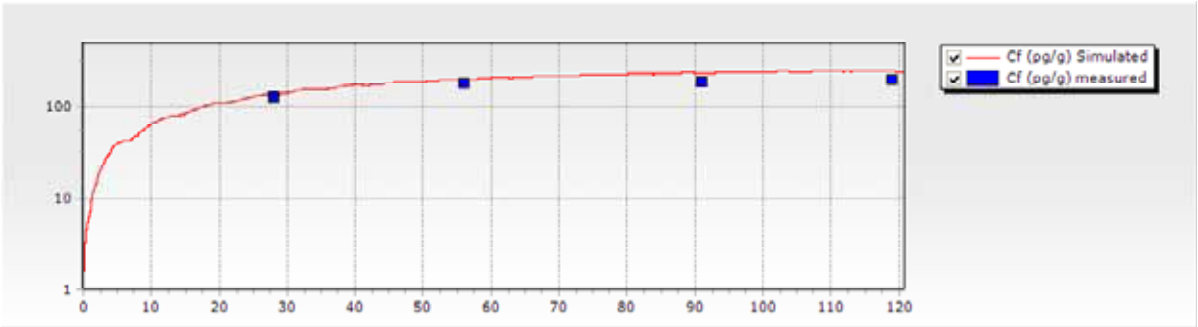


Figure 3-24. Comparison of experimental data (symbols) and model predictions (solid lines) of (A) blood, (B) liver and (C) adipose tissue concentrations of TCDD after oral exposure to 1.5 ng/kg-day, 5 days/week for 17 weeks in mice. Y-axis represents concentration in pg/g and X-axis represents time in days.

Source: Experimental data were obtained form Diliberto et al. (2001).

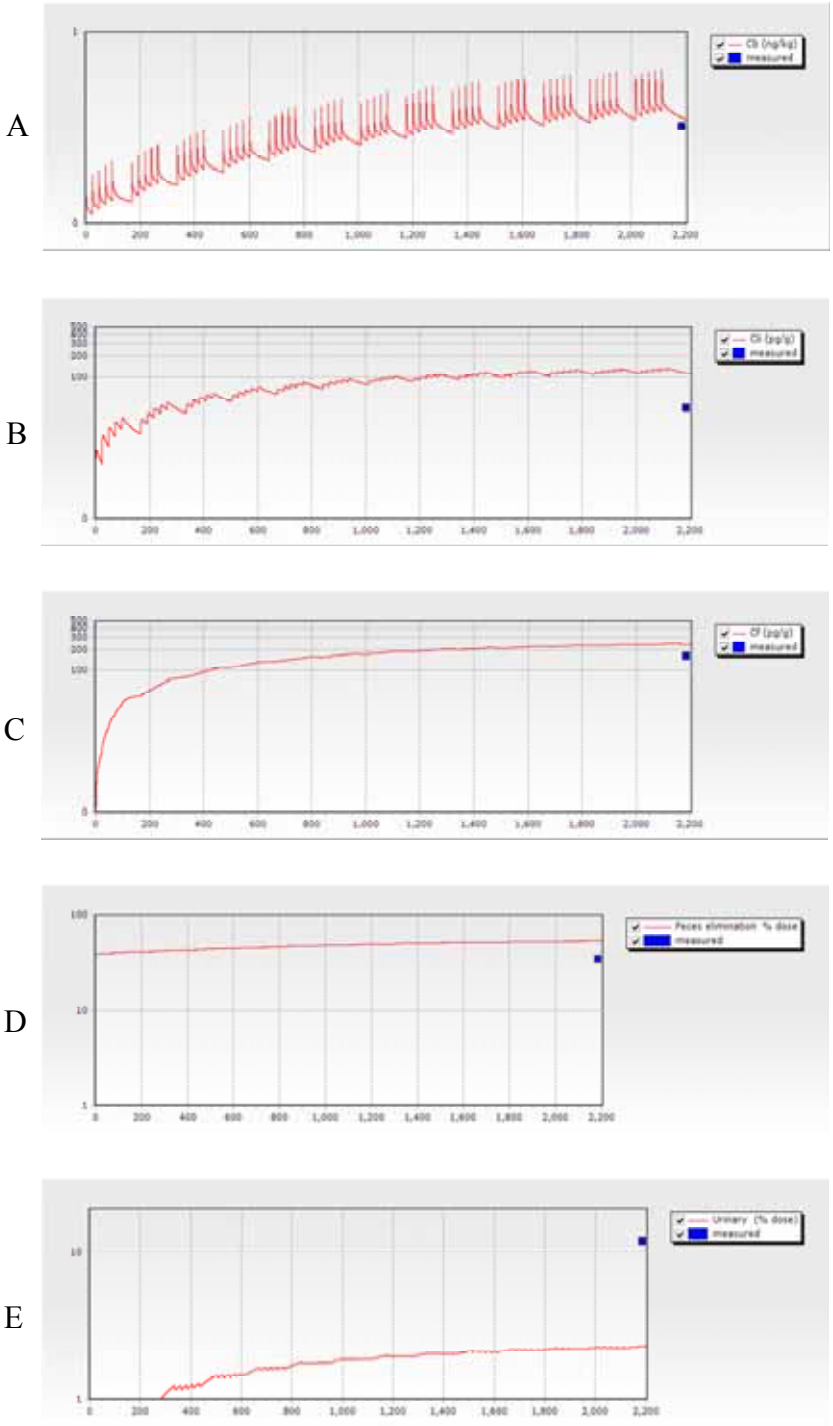


Figure 3-25. Comparison of experimental data (symbols) and model predictions (solid lines) of (A) blood concentration, (B) liver concentration, (C) adipose tissue concentration (D) feces excretion (% dose) and (E) urinary elimination (% dose) of TCDD after oral exposure to 1.5 ng/kg-day, 5 days/week for 13 weeks in mice. Y-axis represents concentration in pg/g and X-axis represents time in days.

Source: Experimental data were obtained form Diliberto et al. (2001).

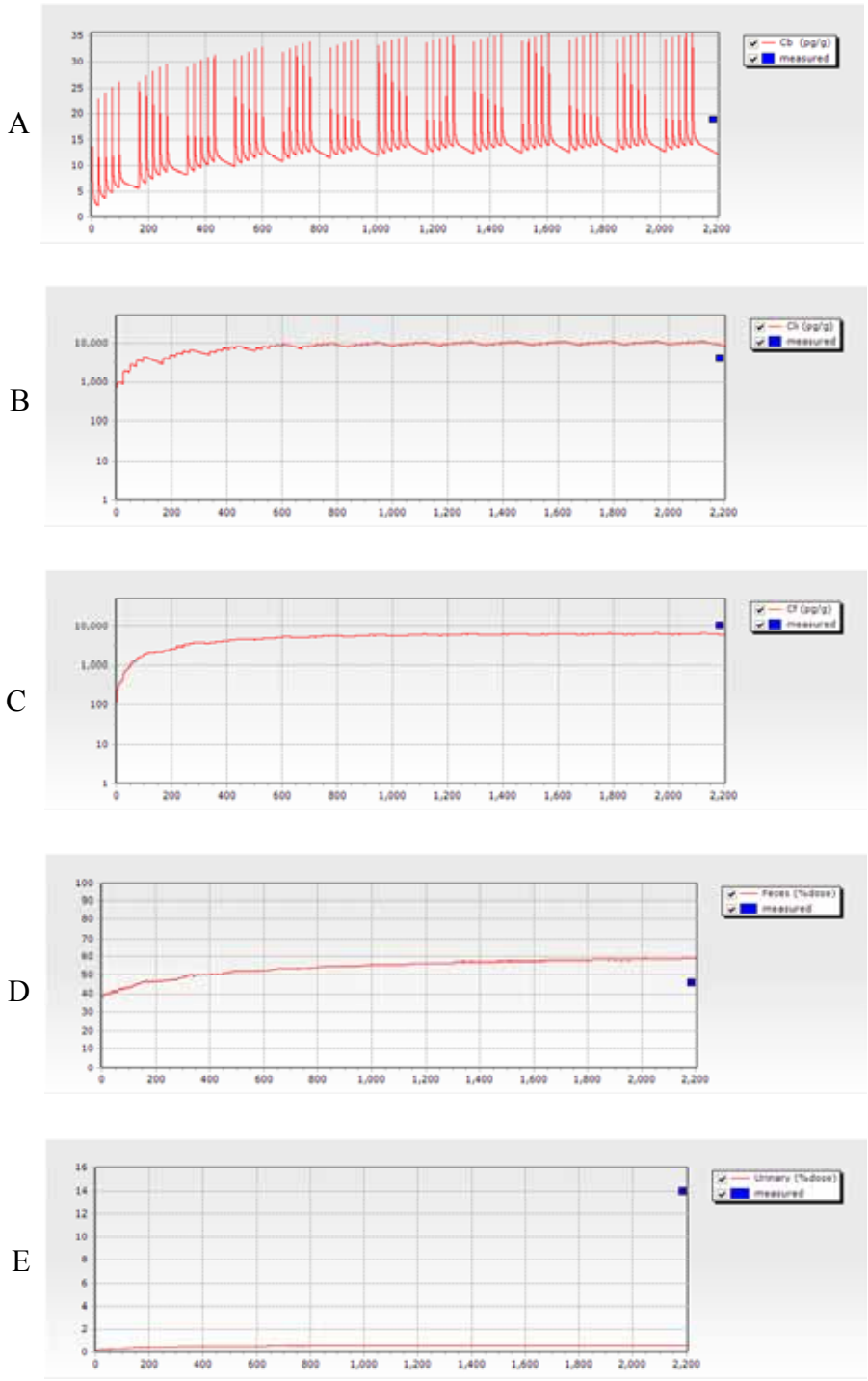


Figure 3-26. Comparison of experimental data (symbols) and model predictions (solid lines) of (A) blood concentration, (B) liver concentration, (C) adipose tissue concentration (D) feces excretion (% dose) and (E) urinary elimination (% dose) of TCDD after oral exposure to 150 ng/kg-day, 5 days/week for 13 weeks in mice. Y-axis represents concentration in pg/g and X-axis represents time in days.
Source: Experimental data were obtained form Diliberto et al. (2001).

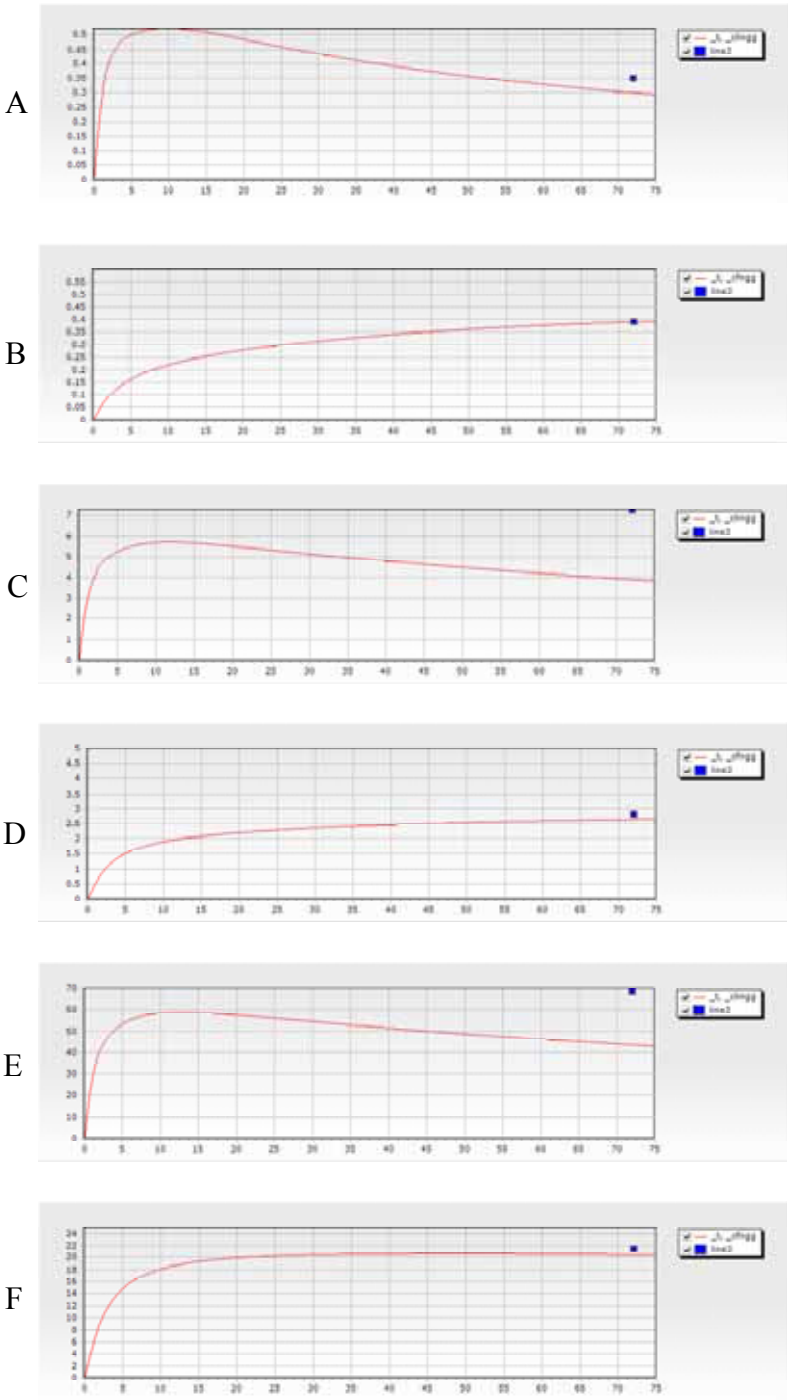
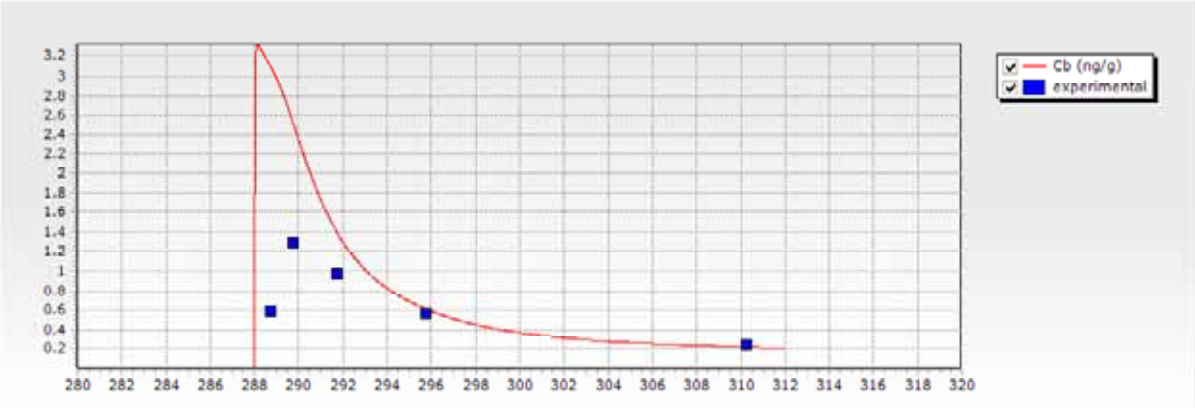
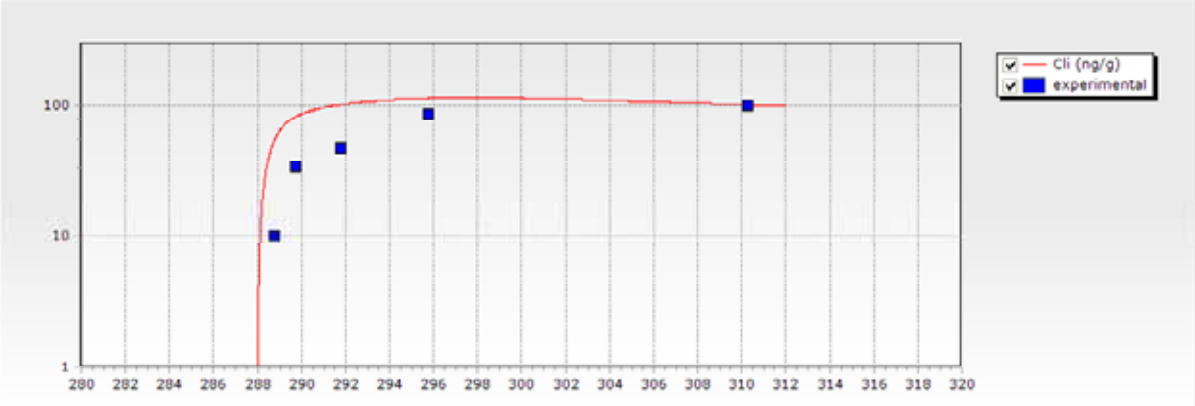


Figure 3-27. PBPK model simulations (solid lines) vs. experimental data (symbols) on the distribution of TCDD after a single acute oral exposure to A−B) 0.1, C−D) 1.0 and E−F) 10 µg of TCDD/kg of body weight in mice.
Liver and adipose concentration for each dose was measured after 72 hours. Y-axis represents the concentration in tissues (ng/g); insets A, C, and E represent liver tissue, whereas B, D, and F correspond to adipose tissue. X-axis represents the time in hours.
Source: experimental data were obtained from Santostefano et al. (1996).

A



B



C

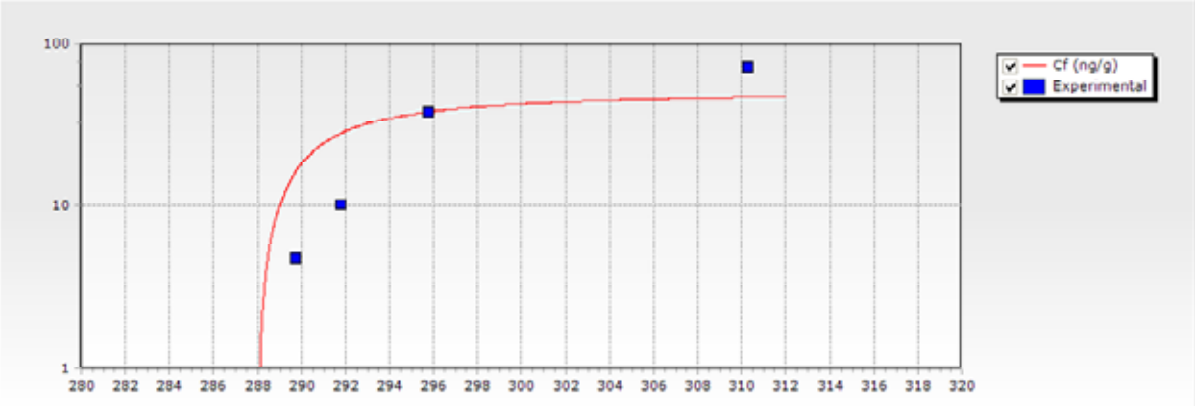


Figure 3-28. PBPK model simulation (solid lines) vs. experimental data (symbols) on the distribution of TCDD after a single dose of 24 µg/kgBW on GD 12 in mice. Concentrations expressed as ng TCDD/g tissue. (A) maternal blood, (B) maternal liver and (C) maternal adipose tissue. Y-axis represents the tissue concentration whereas X-axis represents the time in hours.

.

Source: Experimental data were obtained from (Abbott et al., 1996).

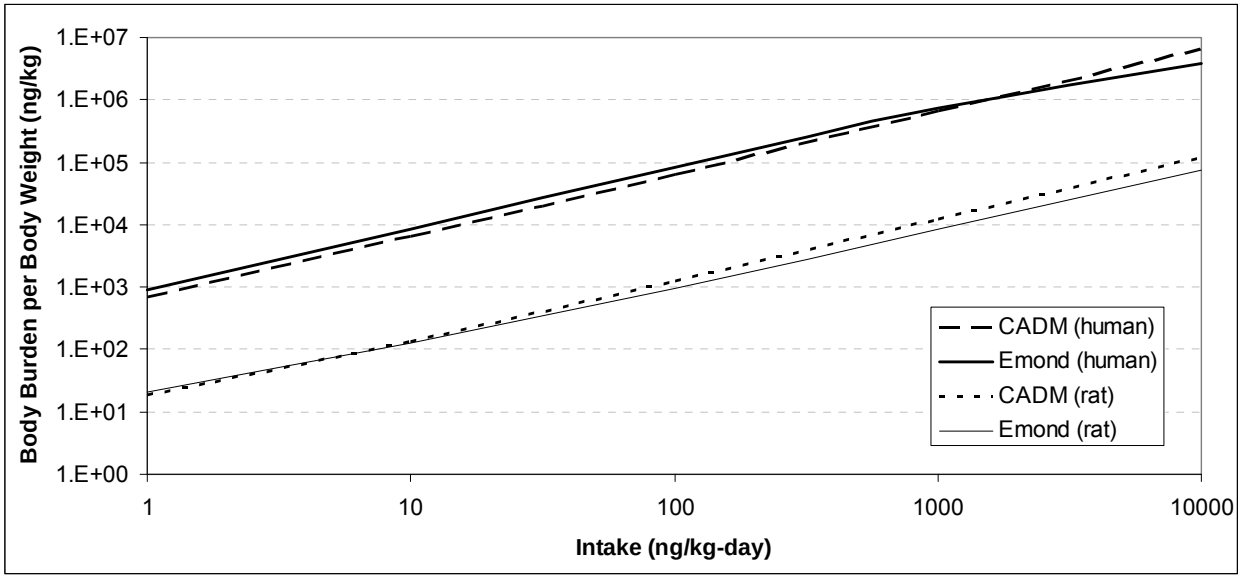


Figure 3-29. Comparison of the near-steady-state body burden simulated with CADM and Emond models for a daily dose ranging from 1 to 10,000 ng/kg-day in rats and humans. The rat model was run for 13 weeks and the human model was run from age 20 to 30. The time-averaged concentration was used for each.

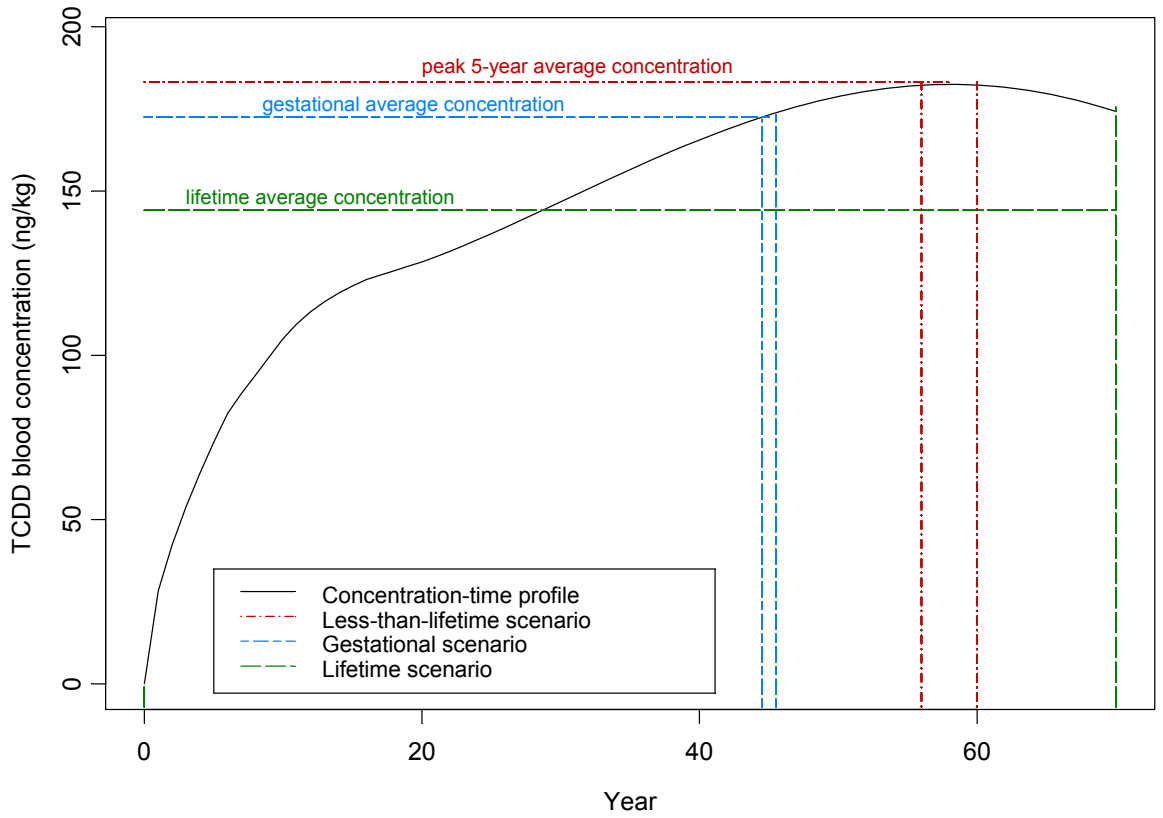


Figure 3-30. TCDD serum concentration-time profile for lifetime, less-than-lifetime and gestational exposure scenarios, with target concentrations shown for each; profiles generated with Emond human PBPK model.

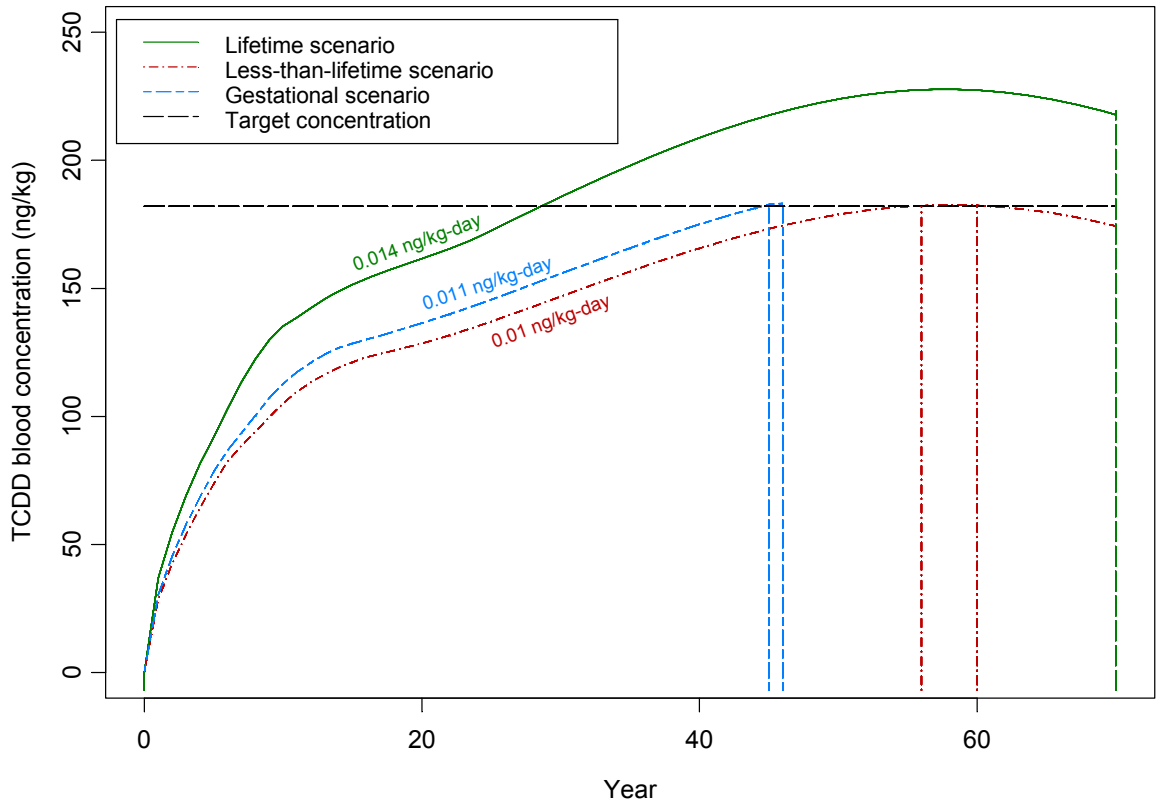


Figure 3-31. TCDD serum concentration-time profile for lifetime, less-than-lifetime and gestational exposure scenarios, showing continuous intake levels to fixed target concentration; profiles generated with Emond human PBPK model.

4. CHRONIC ORAL REFERENCE DOSE

This section presents U.S. Environmental Protection Agency (EPA)'s response to the National Academy of Sciences (NAS) recommendations that EPA more explicitly discuss the modeling of noncancer endpoints and develop a reference dose (RfD) to address noncancer effects associated with oral 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) exposures. Section 2 details the selection of the animal studies with the lowest TCDD doses associated with the development of adverse noncancer effects and the selection of relevant epidemiologic studies of adverse noncancer health effects. Section 3 discusses the kinetic modeling and estimation of human equivalent daily oral doses used in TCDD RfD development in this section. This section discusses the modeling of noncancer health effect data associated with TCDD exposure and the derivation of an RfD. Specifically, Section 4.1 summarizes the NAS comments on TCDD dose-response modeling and EPA's response, including justification of selected noncancer effects and statistical characterization of modeling results. Section 4.2 presents the TCDD dose-response modeling undertaken for identification of candidate points of departure (PODs) for derivation of an RfD. Finally, in Section 4.3, EPA develops an RfD for TCDD.

4.1. NAS COMMENTS AND EPA'S RESPONSE ON IDENTIFYING NONCANCER EFFECTS OBSERVED AT LOWEST DOSES

The NAS recommended that EPA identify the noncancer effects associated with low dose TCDD exposures and discuss its strategy for identifying and selecting PODs for noncancer endpoints, including biological significance of the effects.

> With respect to noncancer end points, the committee notes that EPA does not use a rigorous approach for evaluating evidence from studies. . . (NAS, 2006a, p. 47).
>
> The Reassessment should describe clearly the following aspects:
> 1. The effects seen at the lowest body burdens that are the primary focus for any risk assessment—the "critical effects."
> 2. The modeling strategy used for each noncancer effect, paying particular attention to the critical effects, and the selection of a point of comparison based on the biological significance of the effect; if the ED_{01} is retained, then the biological significance of the response should be defined and the precision of the estimate given... (NAS, 2006a, p. 187).

In this document, EPA has developed a strategy for identifying the noncancer data sets and PODs that represent the most sensitive and biologically relevant endpoints for derivation of an RfD for TCDD. EPA began this process by using the animal bioassays and human epidemiologic studies that met its study inclusion criteria as sources of these data sets. For all epidemiologic studies that were identified as suitable for further quantitative dose-response analyses in Section 2.4.3, EPA has chosen to identify PODs (i.e., estimates of a no-observed-adverse-effect level [NOAEL], lowest-observed-adverse-effect level [LOAEL], or benchmark dose lower confidence bound [BMDL] when possible). Figure 4-1 shows EPA's process to select and identify candidate PODs from these key epidemiologic studies. EPA first evaluated the dose-response information in the study to determine whether it provided an estimate of TCDD dose and an observed noncancer effect that was relevant for RfD derivation. If such data were available, then EPA identified a NOAEL or LOAEL as a candidate POD. For each of these, EPA applied a human kinetic model to estimate the continuous oral daily intake (ng/kg-day) associated with the POD that could be used in the derivation of an RfD (see Section 4.2). If all of this information was available, then the result was included as a candidate POD.

Figure 4-2 summarizes the strategy employed for identifying and selecting candidate PODs from the key animal bioassays identified in Section 2.4.3 for use in noncancer dose-response analysis of TCDD. For each noncancer endpoint, EPA first evaluated the toxicologic relevance of each endpoint, rejecting those judged not to be relevant for RfD derivation. Next, initial PODs (NOAELs, LOAELs, and BMDLs) based on the first-order body burden metric (see Section 3.3.4.2) and expressed as human-equivalent doses (HEDs) were determined for all relevant endpoints (summarized in Table 4-3). Because there were very few NOAELs and BMDL modeling was largely unsuccessful due to data limitations, the next stage of evaluation was carried out using LOAELs only. Endpoints not observed at the LOAEL (i.e., reported at higher doses) with BMDLs greater than the LOAEL were eliminated from further analysis, as they would not be considered as candidates for the final POD on either a BMDL or NOAEL/LOAEL basis (i.e. the POD would be higher than the PODs of other relevant endpoints). In addition, all endpoints with HED estimates based on LOAELs ($LOAEL_{HED}$s) beyond a 100-fold range of the lowest identified $LOAEL_{HED}$ were eliminated from further consideration, as they would not be potential POD candidates either (i.e., the POD would be higher than the PODs of other relevant endpoints). For the remaining endpoints, EPA then

determined final potential PODs (NOAELs, LOAELs and BMDLs) based on TCDD blood concentrations obtained from the Emond rodent physiologically based pharmacokinetic (PBPK) models. HEDs were then estimated for each of these PODs using the Emond human PBPK model. From these HEDs, a POD_{HED} was selected[14] for each study as the basis for the candidate RfD, to which appropriate uncertainty factors (UFs) were applied following EPA guidelines. The resulting candidate RfDs were then considered in the final selection process for the RfD. Other endpoints occurring at slightly higher doses representing additional effects associated with TCDD exposure (beyond the 100-fold LOAEL range) were evaluated, modeled, and included in the final candidate RfD array[15] to examine endpoints not evaluated by studies with lower PODs. In addition, BMD modeling based on administered dose was performed on all endpoints for comparison purposes. The final array of selected endpoints is shown in Table 4-4 (summary of BMD analysis) and Table 4-5 (candidate RfDs).

The NAS recommended that EPA better justify the selection of response levels for endpoints used to develop risk estimates. The NAS commented on EPA's decision to estimate an ED_{01} (effective dose eliciting x percent response) for noncancer bioassay/data set combinations as a comparative tool across studies, suggesting that EPA identify and evaluate the levels of change associated with adverse effects to define the benchmark response (BMR) level for continuous noncancer endpoints.

> The committee notes that the choice of the 1% response level as the POD substantially affects … the noncancer analyses…. The committee recommends that the Reassessment use levels of change that represent clinical adverse effects to define the BMR level for noncancer continuous end points as the basis for an appropriate POD in the assessment of noncancer effects (NAS, 2006a, p. 72).
>
> The committee concludes that EPA did not adequately justify the use of the 1% response level (the ED_{01}) as the POD for analyzing epidemiological or animal bioassay data for … noncancer effects (NAS, 2006a, p. 18).

In the 2003 Reassessment (U.S. EPA, 2003), EPA was not attempting to derive an RfD when it conducted TCDD dose-response modeling. The 2003 Reassessment developed ED_{01}

[14]In the standard order of consideration: BMDL, NOAEL, and LOAEL.
[15]However, studies with a lowest dose tested greater than 30 ng/kg-day were not included in the expanded evaluation.

estimates for noncancer effects in an attempt to compare disparate endpoints on a consistent response scale. Importantly, the 2003 Reassessment defined the ED_{01} as 1% of the maximal response for a given endpoint, not as a 1% change from control. Because RfD derivation is one goal of this document, the noncancer modeling effort undertaken here differs substantially from the modeling in the 2003 Reassessment.

The NAS committee was concerned with the statistical power to determine the shape of the dose-response curve at doses far below observed dose-response information. EPA agrees that the shape of the dose-response curve in the low-dose region cannot be determined confidently when based on higher-dose information. An observed response above background near (or below) the BMR level is needed for discrimination of the shape of the curve and for accurate estimation of an ED_x or BMDL. Although many of the ED_{01}s presented in the 2003 Reassessment were near the lowest dose tested, responses at the lowest doses were often high and much greater than a 1% response (i.e., 1% of the maximum response). The lack of an observed response near the BMR level is often a problem in interpretation of BMD modeling results.

In this document, EPA has used a 10% BMR for dichotomous data for all endpoints except for developmental study designs that incorporate litter effects, for which a 5% BMR is used (U.S. EPA, 2000a). For continuous endpoints in this document, EPA has used a BMR of 1 standard deviation from the control mean whenever a specific toxicologically-relevant BMR could not be defined. For the vast majority of continuous endpoints, EPA could not establish unambiguous levels of change representative of adversity, which EPA defines as "a biochemical change, functional impairment, or pathologic lesion that affects the performance of the whole organism, or reduces an organism's ability to respond to an additional environmental challenge" (U.S. EPA, 2009b). For body and organ weight change, EPA has previously established a BMR of 10% change, which also is used in this document.

The NAS commented on EPA's development of ED_{01} estimates for numerous study/data set combinations in the 2003 Reassessment, suggesting that EPA had not appropriately characterized the statistical confidence around such model predictions in the low-response region of the model.

> It is critical that the model used for determining a POD fits the data well, especially at the lower end of the observed responses. Whenever feasible, mechanistic and statistical information should be used to estimate the shape of the dose-response curve at lower doses. At a minimum, EPA should use rigorous statistical methods to assess model fit and to control and reduce the uncertainty of the POD caused by a poorly fitted model. The overall quality of the study design is also a critical element in deciding which data sets to use for quantitative modeling (NAS, 2006a, p. 18).
>
> EPA should … assess goodness-of-fit of dose-response models for data sets and provide both upper and lower bounds on central estimates for all statistical estimates. When quantitation is not possible, EPA should clearly state it and explain what would be required to achieve quantitation (NAS, 2006a, p. 10).

The NAS also commented that EPA report information describing the adequacy of dose-response model fits, particularly in the low response region. For those cases where biostatistical modeling was not possible, NAS recommended that EPA identify the reasons.

> The Reassessment should also explicitly address the importance of statistical assessment of model fit at the lower end and the difficulties in such assessments, particularly when using summary data from the literature instead of the raw data, although estimates of the impacts of different choices of models would provide valuable information about the role of this uncertainty in driving the risk estimates (NAS, 2006a, p. 73).

To address this concern, in this document EPA has reported the standard suite of goodness-of-fit measures from the benchmark dose modeling software (BMDS 2.1). These include chi-square *p*-values, Akaike's Information Criterion (AIC), scaled residuals at each dose level and plots of the fitted models. In some cases, when restricted parameters hit a bound, EPA used likelihood ratio tests to evaluate whether the improvement in fit afforded by estimating additional parameters could be justified. Goodness-of-fit measures are reported for all key data sets in Appendix E.

4.2. NONCANCER DOSE-RESPONSE ASSESSMENT OF TCDD

This section describes EPA's current effort to conduct an evaluation of TCDD dose-response for the noncancer endpoints from studies that met the study inclusion criteria. Discussions include benchmark dose modeling procedures, kinetic modeling, and POD candidates for derivation of the RfD. Section 4.2.1 describes how EPA has used physiologically-

based pharmacokinetic (PBPK) modeling to estimate effective internal exposures as an alternative to using administered doses or body burdens based on first-order kinetics. Section 4.2.2 details the dose-response analysis of the epidemiologic data, with supporting information on kinetic modeling in Appendix D. Section 4.2.3 details the dose-response analysis for the animal bioassay data; Appendix E provides the BMDS input tables (see Section E.1) and output for all modeling, including blood concentrations (see Section E.2) and administered dose (see Section E.3).

4.2.1.1. *Determination of Toxicologically Relevant Endpoints*

The NAS committee commented on the low dose model predictions and the need to discuss the biological significance of the noncancer health effects modeled in the 2003 Reassessment. In selecting POD candidates from the animal bioassays for derivation of the candidate RfDs, EPA had to consider the toxicological relevance of the identified endpoint(s) from any given study. Some endpoints/effects may be sensitive, but lack general toxicological significance due to not being clearly adverse (defined in EPA's Integrated Risk Information System glossary as "a biochemical change, functional impairment, or pathologic lesion that affects the performance of the whole organism, or reduces an organism's ability to respond to an additional environmental challenge" (U.S. EPA, 2009)), being an adaptive response or not being clearly linked to downstream functional or pathological alterations. It is standard EPA RfD derivation practice not to base a reference value on endpoints that are not adverse or not precursors to an adverse effect. Studies meeting the study selection criteria with endpoints that were not considered for derivation of a candidate RfD (because they were not considered to be toxicologically relevant noncancer effects) are: Kitchin and Woods (1979) , Hassoun et al. (1998, 2000, 2002, 2003), Burleson et al. (1996), Kuchiiwa et al. (2002), Mally and Chipman (2002), Vanden Heuvel et al. (1994), Devito et al. (1994), Lucier et al. (1986), Sugita-Konishi et al. (2003), and Sewall et al. (1993). Appendix G identifies the endpoints from these studies that were not considered to be toxicologically relevant (e.g., cytochrome P450 induction, oxidative stress measures, gap junction disruption, mRNA induction, brain serotonin levels) and provides the rationales for the toxicological relevance decisions on the endpoints. Note that for many of these studies, other endpoints were examined that are toxicologically relevant and were considered in the RfD derivation process.

4.2.2. Use of Toxicokinetic Modeling for TCDD Dose-Response Assessment

Given that TCDD accumulates in fat with continuous exposure and is eliminated slowly from the body, but at very different rates across species, EPA has determined that the standard UF approach or allometric scaling of body weight for interspecies extrapolation is not appropriate. Therefore, EPA has decided to use toxicokinetic modeling to estimate an effective internal dose for equivalence across species. The toxicokinetic models chosen by EPA are the rodent and human PBPK models described by Emond et al. (2004, 2006) and modified by EPA for this assessment as described in Section 3.3.4 (hereafter referred to as the "Emond [rodent or human] PBPK model"). Both the rodent and human models have a gestational component, which allow for more relevant exposure comparisons between general adult exposures and the numerous gestational exposure studies. Ideally, a relevant tissue concentration for each effect would be estimated. However, no models exist for estimation of all relevant tissue concentrations. Therefore, EPA has decided to use the concentration of TCDD in blood as a surrogate for tissue concentrations, assuming that tissue concentrations are proportional to blood concentrations.[16] Furthermore, because the RfD is necessarily expressed in terms of average daily exposure, the blood concentrations are expressed as averages over the relevant period of exposure for each endpoint. For the animal bioassay studies, the relevant period of exposure is the duration of dosing, starting at the age of the animals at the beginning of the study. For humans, the relevant period of exposure is generally lifetime, which is defined as 70 years by convention. However, EPA varied the averaging time for the equivalent human blood concentrations to correspond to the test-animal exposure duration in the following manner.

- For correspondence with animal chronic exposures,[17] the human-equivalent TCDD blood concentration is assumed to be the 70-year average.
- For correspondence with animal gestational exposures, the human-equivalent TCDD blood concentration is assumed to be the average over 45 years for a female, beginning at birth, plus 9 months of gestational exposure. The choice of 45 years to beginning of pregnancy is health protective of the population in that

[16]As virtually all TCDD is found in the adipose fraction of tissues, or bound to specific proteins, an ideal better approach would be to account for the fat fraction of each tissue and protein binding; EPA has not found sufficient data to implement this approach.

[17]Assumed to be ≥75% of nominal lifetime, or about 550 days in rodents.

the daily exposure achieving the target blood concentration is smaller than for shorter averaging times.[18]

- For correspondence with any other animal exposure duration, the human-equivalent TCDD blood concentration is assumed to be the average over the equivalent human exposure duration calculated backward from the peak exposure plateau at or near the end of the 70-year scenario. The average is determined from the terminal end of the human exposure period because the daily exposure achieving the target blood concentration is smaller than for the same exposure period beginning at birth and is health protective for effects occurring after shorter-term exposure. The determination of equivalent exposure durations across species is problematic and somewhat arbitrary, so EPA uses the average peak blood concentration as the human equivalent for all less-than-chronic animal exposures (other than gestational).[19] For the first-order kinetics model, the average peak exposure is close to the theoretical steady-state asymptote (see Section 3.3.4.2). However, for the Emond human PBPK model used by EPA in this assessment, the timing of the peak exposure is dose-dependent and tends to decline after 60 years in some cases. Therefore, the 5-year average TCDD blood concentration that includes the peak ("5-year peak") is used as the relevant dose-metric for the PBPK model applications.

4.2.3. Noncancer Dose-Response Assessment of Epidemiological Data

The following four epidemiologic studies describing noncancer endpoints were identified in Section 2.4.3 as studies to be evaluated for development of PODs for derivation of candidate RfDs: Baccarelli et al. (2008), Mocarelli et al. (2008), Alaluusua et al. (2004) and Eskenazi et al. (2002). Each of these studies described effects observed in the Seveso cohort (see study summaries in Section 2.4.1 and Table 2-5). Each study provided individual-level human exposure measures and an exposure window over which kinetic models could be used to quantify TCDD exposures for dose-response assessment. EPA used kinetic information to estimate group-mean daily TCDD intake rates for the exposure groups presented in these studies (see Appendix D for details). EPA focused on identifying NOAELs and LOAELs for these studies; EPA did not conduct Benchmark Dose modeling because the covariates identified by the study authors could not be incorporated by modeling the grouped response data. EPA's development of PODs for these studies is described in this section and shown in Table 4-1.

[18]See Section 3.3.4.2 for a discussion of this issue.

[19]By comparison to a half-lifetime equivalent (1 year in rodents, 35 years in humans), the ratio of body burden (1st-order kinetic model) to oral intake does not differ significantly from the average-peak scenario; all shorter-term scenarios differ even less (see Section 3.3.4.2). These relationships, with respect to the 5-year peak, hold for the PBPK model results, as well (see Section 3).

4.2.3.1. *Baccarelli et al. (2008)*

For Baccarelli et al. (2008), EPA was able to define a LOAEL as the group mean of 39 ppt TCDD in neonatal plasma for thyroid stimulating hormone (TSH) values above 5 μU/mL. Baccarelli et al. (2008) did not estimate the equivalent oral intake associated with TCDD serum concentrations and gave only neonatal serum TCDD concentrations for the groups above and below the 5 μU/mL standard. The study authors, however, developed a regression model relating the level of TSH in 3-day-old neonates to TCDD concentrations in maternal plasma at birth (given as lipid-adjusted serum concentrations, LASC). The authors extrapolated maternal plasma concentrations from previous measurements using simple first-order pharmacokinetic model. Because there is limited information regarding the relationship between maternal and neonatal serum TCDD levels, EPA determined that there was too much uncertainty in estimating maternal intake from neonatal TCDD serum concentrations, directly. Therefore, EPA determined the maternal intake at the LOAEL from the maternal serum-TCDD/TSH regression model by finding the maternal TCDD LASC at which neonatal TSH exceeded 5 μU/mL. EPA then used the Emond PBPK model under the human gestational scenario (see Section 4.2.1) to estimate the continuous daily TCDD intake that would result in a TCDD LASC corresponding to a neonatal TSH of 5 μU/mL at the end of gestation, with the resulting maternal intake established as the LOAEL (0.024 ng/kg-day), shown in Table 4-1 as a candidate POD for derivation of candidate RfDs. The results of the PBPK modeling are shown in Appendix D.

4.2.3.2. *Mocarelli et al. (2008)*

Mocarelli et al. (2008) reported decreased sperm concentrations (20%), decreased motile sperm counts (11%), and decreased serum estradiol (23%) in men who were 1–9 years old in 1976 at the time of the accident (initial TCDD exposure event). Men who were 10–17 years old in 1976 were not affected, with the possible exception of reduced serum estradiol. Serum (LASC) TCDD levels were measured within one year of the initial exposure. Serum TCDD levels and corresponding responses were reported by quartile, with a reference group of less-exposed individuals assigned a TCDD LASC value of 15 ppt (which was the mean of individuals outside the contaminated area). The lowest exposed group mean was 68 ppt (1st quartile). Because effects were detected only among boys under the age of 10, EPA assumes there is a maximum 10-year critical exposure window for elicitation of these effects. However, for the

exposure profile, with a high initial pulse followed by an extended period of elimination with only background exposure, the estimation of an average exposure resulting in the effect is problematic. Therefore, EPA implemented a procedure for the estimation of the continuous daily TCDD intake associated with the LOAEL in the Mocarelli et al. (2008) study using a 5-step process.

1. Using the Emond human PBPK model, the initial (peak) blood TCDD concentrations associated with the accident were back-calculated based on the time that had elapsed between the explosion and the serum collection. As serum measurements were taken within 1 year after the event, a lag time of 0.5 years was assumed.
2. The oral exposure associated with the peak blood TCDD concentration (peak exposure) was calculated using the Emond PBPK model.
3. Starting with the peak exposure and accounting for background TCDD intake, the average daily blood TCDD concentration experienced by a representative individual in the susceptible population (boys under 10 years old) was estimated using the Emond PBPK model. Assuming a random distribution of ages at the time of the event, the average age of the exposed male children would be 5 years. Consequently, a critical exposure window for the cohort was estimated to be, on average, 5 years (i.e., a boy aged 5 years would remain in this exposure window for 5 more years until he was 10 years of age).
4. Using the Emond PBPK model, the average daily TCDD intake rate needed to attain the 5-year average blood TCDD concentration in a boy 10 years old was calculated.
5. The LOAEL POD was calculated as the average of the peak exposure (0.032 ng/kg-day) and the 5-year average exposure (0.0080 ng/kg-day), resulting in LOAEL of 0.020 ng/kg-day, shown in Table 4-1 as a candidate POD for derivation of a candidate RfD. However, neither of the extremes was used because (1) the peak exposure does not account for the continuing internal exposure from TCDD given its slow elimination, and (2) the 5-year average does not reflect the influence of the much higher peak exposure, which may be a significant factor in TCDD toxicity (Kim et al., 2003).

The results of the modeling are shown in Appendix D.

4.2.3.3. *Alaluusua et al. (2004)*

For Alaluusua et al. (2004), the approach for estimation of daily TCDD intake is virtually identical to the approach used for the Mocarelli et al. (2008) data. Alaluusua et al. (2004) reported dental effects in male and female adults who were less than 5 years of age at the time of the initial exposure (1976). For the 75 boys and girls who were less than 5 years old at the time

of the accident, 25 (33%) were subsequently diagnosed with some form of dental enamel defect. For the 38 individuals who were older than 5, only 2 (5.3%) suffered dental enamel defects at a later date. A window of susceptibility of approximately 5 years is established. Serum measurements for this cohort were taken within a year of the accident. Serum TCDD levels and corresponding responses were reported by tertile, with a reference group of less-exposed individuals assigned a TCDD LASC value of 15 ppt (ng/kg); the tertile group means were 130, 383, and 1830 ppt. The incidence of dental effects for the reference group was 26% (10/39). The incidence of dental effects in the 1st, 2nd and 3rd tertile exposure groups was 10% (1/10), 45% (5/11) and 60% (9/15), respectively. EPA judged that the NOAEL and LOAEL were 130 and 383 ppt TCDD in serum. Following the same procedure used for the Mocarelli et al. (2008) study (see Section 4.2.3.2), EPA estimated the continuous daily human TCDD intake associated with each of the tertiles for both peak and average exposure across the critical exposure window, assuming that the average age of the susceptible cohort at the time of the accident was 2.5 years. Separate estimates for boys and girls were developed based on both the peak intake and average intake across the critical exposure window (see Appendix D for details of the PBPK modeling are shown in Appendix D). The estimated averaged daily intakes for the tertiles, averaged for boys and girls, are 0.20, 1.7 and 30 ng/kg-day for the peak exposure and 0.033, 0.15 and 1.5 ng/kg-day for the critical exposure window average. A study NOAEL at the second tertile of 0.12 ng/kg-day was identified as a candidate POD for derivation of a candidate RfD in Table 4-1.

4.2.3.4. *Eskenazi et al. (2002)*

The approach used to estimate daily TCDD intake in Eskenazi et al. (2002) combines the approaches EPA used for Baccarelli et al (2008), Mocarelli et al. (2008) and Alaluusua et al. (2004). Eskenazi et al. (2002) reported menstrual effects in female adults who were premenarcheal at the time of the initial exposure (1976). In Rigon et al. (2009), the median age at menarche was shown to be 12.4 in Italian females with intergenerational decreases in age at menarche. Thus, EPA established a window of susceptibility of approximately 13 years for this analysis. The average age of the premenarcheal girls at the time of the initial exposure in 1976 was 6.8 years, establishing an average critical-window exposure duration of 6.2 years for this cohort. Serum samples were collected within a year of the accident from this cohort. However, serum TCDD levels and corresponding responses were not reported by percentile and no internal

reference group was identified. As for Baccarelli et al (2008), Eskenazi et al. (2002) developed a regression model relating menstrual cycle length to plasma TCDD concentrations (LASC) measured in 1976. The model estimated that menstrual cycle length was increased 0.93 days for each 10-fold increase in TCDD LASC, with a 95% confidence interval of −0.01 to 1.86 days. EPA judged a 1-day increase in menstrual cycle length to be adverse; a normal menstrual cycle length is 28 days. EPA then determined the 1976 TCDD serum level corresponding to a 29-day menstrual cycle length in the exposed cohort from the regression model developed by Eskenazi et al. (2002). Using this serum level, the peak initial exposure and average exposure over the 6.2 year window were calculated using the Emond human PBPK model, in the same manner as for Mocarelli et al. (2008) and Alaluusua et al. (2004). The resulting peak TCDD intake is 3.2 ng/kg-day. The average exposure experienced by this cohort over the critical exposure window is estimated to be 0.12 ng/kg-day. The average of these two estimates is 1.64 ng/kg-day, which is designated as a LOAEL and shown in Table 4-1. Because the LOAEL is almost 2 orders of magnitude higher than the LOAELs for Baccarelli et al (2008) and Mocarelli et al. (2008), it was not considered further as a candidate POD for derivation of the RfD. The results of the PBPK modeling are shown in Appendix D.

4.2.4. Noncancer Dose-Response Assessment of Animal Bioassay Data

EPA followed the strategy illustrated in Figure 4-2 to evaluate the animal bioassay data for TCDD dose-response. For the administered average daily doses (ng/kg-day) in each animal bioassay, EPA identified NOAELs and/or LOAELs based on the original data presented by the study author. Section 2.4.2 identifies these values in the study summaries and in Table 2-7. These became candidate PODs for consideration in the derivation of an RfD for TCDD. The candidate RfD values associated with these candidate PODs are presented in Table 4-5. Additional PODs were identified using BMD modeling. All PODs were converted to HEDs using the Emond PBPK models. The remainder of this Section describes the steps in this process and concludes with the POD candidates from the animal bioassay data that were considered for derivation of the RfD.

4.2.4.1. *Use of Kinetic Modeling for Animal Bioassay Data*

Blood concentrations corresponding to the administered doses in each mouse or rat bioassay qualifying as a final RfD POD candidate were estimated using the appropriate Emond rodent PBPK model. In each case, the simulation was performed using the exposure and observation durations, body weights, and average daily doses from the original studies. For all multiple exposure protocols, the time-weighted average blood TCDD concentrations over the exposure period were used as the relevant dose metric. For single (gestational and nongestational) exposures, the initial peak blood TCDD concentrations were considered to be the most relevant exposure metric. Gestational exposures were modeled using the species-specific gestational component of the Emond rodent PBPK model. Bioassays employing exposure protocols spanning gestational and postpartum life stages were modeled by sequential application of the gestational and nongestational models.

The Emond PBPK models do not contain a lactation component, so exposure during lactation was not modeled explicitly. Only one bioassay (Shi et al., 2007) considered as a POD candidate for RfD derivation included exposure during lactation. In Shi et al. (2007) pregnant animals were exposed weekly to TCDD throughout gestation and lactation. Exposure was continued in the offspring following weaning for 10 months. For assessment of maternal effects, the Emond gestational model was used, terminating at parturition. For assessment of long-term exposure in the offspring, the Emond nongestational model was used, ignoring prior gestational and lactational exposure, with the assumption that the total exposure during these periods was small relative to exposure in the following 10 months. The assumption is conservative in that effects observed in the offspring would be attributed entirely to adult exposures, which is somewhat less than the actual total exposure.

The model code, input files and PBPK modeling results for each bioassay are reported in Appendix C. Note that the modeled output is given in terms of LASC. The corresponding TCDD whole blood concentrations, which are used directly as the toxicologically-equivalent dose metric, are derived by multiplying the LASC by the lipid fraction in rodent serum (0.0033) and the fraction of blood that is serum (0.55). For the rat gestational model, only, the lipid fraction in serum is 0.0023, rather than 0.0033.

These predicted TCDD blood concentrations were used for benchmark dose modeling of bioassay response data and determination of NOAELs and LOAELs. BMD modeling was

performed, as described in Section 3.5.2.2.1, by substituting the modeled blood concentrations for the administered doses and calculating the corresponding BMDL. For each of these LOAEL, NOAEL, or BMDL blood-concentration equivalents, corresponding HEDs were calculated using the Emond human PBPK model for the appropriate gestational or nongestational scenario as described previously (see Section 4.2.1).

4.2.4.2. *Benchmark Dose Modeling of the Animal Bioassay Data*

Benchmark dose modeling was performed using BMDS 2.1, Build 06/16/09 to estimate BMDs and BMDLs for each study/endpoint combination. The input data tables for these noncancer studies are shown in Appendix E, Section E1, including both administered doses (ng/kg-day) and blood concentrations (ng/kg) and either incidence data for the dichotomous endpoints or mean and standard deviations for the continuous endpoints. (See Section 4.3.1 [below] and Sections 3.3.4 and 3.3.5 for a description of the development of TCDD blood concentrations using kinetic modeling.) For the continuous endpoints, all available models were run separately using both the assumption of constant variance and the assumption of modeled variance. Saturated (0 degrees of freedom) model fits were rejected from consideration. Parameters in models with power or slope parameters were constrained to prevent supralinear fits, which EPA considers not to be biologically plausible and which often have undesirable statistical properties (i.e., the BMDL diverges towards zero). However, if the constrained parameters were estimated at their lower bounds, the unrestricted model was fit to the data, primarily for elucidation of the degree of supralinearity present in the data. Depending on the latter and the magnitude of the BMDL relative to the BMD, unrestricted model fits were occasionally deemed acceptable. Table 4-2 shows each model and any restrictions imposed. For the quantal/dichotomous endpoints, all primary BMDS dichotomous models were run. The alternative dichotomous models were fit to several data sets, but the results were very sensitive to the assumed independent background response and the fits were not accepted. The confidence level was set to 95% and all initial parameter values were set to their defaults in BMDS. For the continuous endpoints, one standard deviation was chosen as the default for the BMR when a

specific toxicologically-relevant BMR could not be defined. For the dichotomous endpoints, a BMR of 10% extra risk was used for all endpoints.[20]

The model output tables in Appendix E show all of the models that were run, both restricted and nonrestricted, goodness-of-fit statistics, BMD and BMDL estimates, and whether bounds were hit for constrained parameters. After all models were run, the one giving the best fit was selected using the selection criteria in the current BMDS guidance (U.S. EPA, 2000a) where possible. Acceptable model fits were those with chi-square goodness-of-fit *p*-values greater than 0.1. For the continuous endpoints, the *p*-value for the homogenous variance test (Test 2) was used to determine whether the constant variance ($p > 0.1$) model or modeled (nonconstant) variance ($p < 0.1$) model should be used. As BMDS offers only one variance model, model fits for nonconstant variance models were not necessarily rejected if the variance model did not fit well (Test 3 *p*-value < 0.05). Within the group of models with acceptable fits, the nominal best fitting model was selected by determining the model with the lowest AIC value if the BMDLs were within a factor of 3; otherwise, the model with the lowest BMDL was chosen (U.S. EPA, 2000). However, particularly for continuous models, the fit of the model to the control mean and standard deviation and in the lower response range was assessed. Models with higher AICs but much better fit to the lower response data were often chosen over the nominally best-fitting model. Judgment was also used to assess the plausibility of BMDLs far below the BMD. In most of these cases, deficiencies in the response data resulted in rejection of the modeling results.

For many data sets, no models satisfied the acceptance criteria and no clear BMD/BMDL selection could be made. In theses cases, model fits were examined on an individual basis to determine the reasons for the poor fits. On occasion, high doses were dropped and the models were refit. If the fit was still not acceptable, the NOAEL/LOAEL approach was applied to the study/data set combination.

4.2.4.3. *POD Candidates from Animal Bioassays Based on HED*

Table 4-3 summaries the PODs that EPA estimated for each key animal study included for TCDD dose-response modeling. After estimating the blood TCDD concentration associated

[20] There were no developmental studies that accounted for litter effects, for which a 5% BMR would be used.

with a particular toxicity measure (NOAEL, LOAEL, or BMDL) obtained from a rodent bioassay, EPA estimated a corresponding HED using the Emond human PBPK model (described in Section 3). Table 4-3 summarizes the NOAEL, LOAEL, or BMDL (ng/kg) based on the administered animal doses for each key bioassay/data set combination. Table 4-3 also summarizes the continuous daily HED corresponding to these administered doses. The doses in Table 4-3 are defined as follows:

- Administered Dose NOAEL: Average daily dose (ng/kg-day) from the study
- Administered Dose LOAEL: Average daily dose (ng/kd-day) from the study
- Administered Dose BMDL: Dose from BMD modeling of the administered doses converted to a blood concentration (ng/kg) using the Emond rodent PBPK models
- Emond Model NOAEL: Administered Dose NOAEL converted to an HED (ng/kg-day) using the Emond human PBPK model
- Emond Model LOAEL: Administered Dose LOAEL converted to an HED (ng/kg-day) using the Emond human PBPK model
- Emond Model BMDL: Dose from BMD modeling of the blood concentrations (ng/kg) converted to an HED (ng/kg-day) using the Emond human PBPK model

Tables showing the best model fit for each study/endpoint combination and the associated BMD/BMDL are shown in Appendix E.

An evaluation of key BMD analyses is presented in Table 4-4. The BMD modeling was largely unsuccessful, primarily because of a lack of response data near the BMR, poor modeled representation of control values, or nonmonotonic responses yielding poor fits. BMDLs were often implausibly low (orders of magnitude below the LOAEL) or implausibly high. The comments column in Table 4-4 lists reasons for poor or implausible results.

4.3. RfD DERIVATION

Table 4-5 lists all the studies and endpoints considered for derivation of the RfD. These studies were chosen from the entire list of candidate study/data set combinations (see Section 4.2) based on the toxicologic relevance of the endpoints and covering a range that includes three of the four human studies[21]. Figure 4-3 (exposure-response array) shows all of the

[21] The RfD derived from the study of Eskenazi et al. (2002) was outside the RfD range presented in Table 4-5.

endpoints listed in Table 4-5 graphically in terms of PODs in human-equivalent intake units (ng/kg-day). The human study endpoints are shown at the far left of the figure and the rodent endpoints are arranged by category to the right. Figure 4-4 demonstrates the same endpoints arrayed by RfD value, showing the POD, applicable UFs and candidate RfD.

Table 4-5 illustrates the study, species, strain and sex, study protocol, and toxicologic endpoints observed at the lowest TCDD doses. The table also identifies the human-equivalent BMDLs (when applicable), NOAELs and LOAELs, as well as the composite UF that applies to the specific endpoint, and finally, the corresponding candidate RfD.[22] The NOAELS, LOAELs, and BMDLs are presented as HEDs, based on the assumption that blood concentration is the toxicokinetically-equivalent TCDD dose metric across species and serves as a surrogate for tissue concentration.[23] For rats and mice, these estimates relied on the two Emond PBPK models—one for the relevant rodent species and one for the human—as described previously (see Sections 3.3.4.3). The two guinea pig studies that are included in Table 4-5 are given in HED units based on the first-order body burden model described in Section 3.3.4.2; there is currently no TCDD PBPK model for the guinea pig. The values listed for guinea pigs are not directly comparable to those for rats and mice but are probably biased low, as first-order body burden HED estimates for rats and mice are generally 2- to 5-fold lower than the corresponding PBPK model estimates. The LOAELs for the human studies also rely on the Emond PBPK model, as described in Sections 4.2.2 and 4.2.3.

As is evident from the table, very few NOAELs and even fewer BMDLs have been established for low-dose TCDD studies. BMD modeling was unsuccessful for all of the endpoints without a NOAEL, primarily because of the lack of dose-response data near the BMR (see discussion in Section 4.2). Therefore, the RfD assessment rests largely on evaluation of LOAELs to determine the POD.

The rows in Table 4-5 are arranged in order of increasing candidate RfD magnitude. Endpoints projected to occur at higher exposure levels are still considered for qualitative support of the effects shown in Table 4-5.

[22]Extra significant digits are retained for comparison prior to rounding to one significant digit for the final RfD.
[23]The procedures for estimating HEDs based on TCDD blood concentration are described in the preceding section.

4.3.1. Toxicological Endpoints

As can be seen in Table 4-5, a wide array of toxicological endpoints has been observed following TCDD exposure, ranging from subtle developmental effects to overt chronic liver toxicity. Developmental effects in rodents include dental defects, delayed puberty in males, and several neurobehavioral effects. Reproductive effects reported in rodents include altered hormone levels in females and decreased sperm production in males. Immunotoxicity endpoints such as decreased response to SRBC challenge in mice and decreased delayed-type hypersensitivity response in guinea pigs are also observed. Longer durations of TCDD exposure in rodents elicit results such as organ and body weight changes, renal toxicity, and liver and lung lesions. Adverse effects in human studies are also observed, which include male reproductive effects, increased TSH in neonates, and dental defects in children. Analogous results have been observed in animal bioassays for each of these human endpoints.

All but two of the study/endpoint combinations from animal bioassays listed in Table 4-5 are on TCDD-induced toxicity observed in mice and rats; the other two study/endpoint combinations are effects in guinea pigs. Although the effects of TCDD have been investigated in several other species (i.e., hamsters, monkeys, and mink), those studies were not included for final POD consideration because the effect levels were greater than those in Table 4-5, or because the effects could not be attributed solely to TCDD exposure (i.e., confounding by dioxin-like compounds [DLCs]).

Three human studies were also included for final POD consideration in the derivation of an RfD and are presented in Table 4-5 as candidate RfDs. All three human study/endpoint combinations are from studies on the Seveso cohort. The developmental effects observed in these studies were associated with TCDD exposures either *in utero* or in early childhood between 1 and 10 years of age. Baccarelli et al. (2008) reported increased levels of TSH in newborns exposed to TCDD *in utero*, indicating a possible dysregulation of thyroid hormone metabolism. Mocarelli et al. (2008) reported decreased sperm concentrations, decreased motile sperm counts, and decreased serum estradiol in men who were 1–9 years old in 1976 at the time of the Seveso accident (initial TCDD exposure event). Alaluusua et al. (2004) reported dental effects in adults who were less than 9.5 years of age at the time of the initial exposure (1976).

4.3.2. Exposure Protocols of Candidate PODs

The studies in Table 4-5 represent a wide variety of exposure protocols, involving different methods of administration and exposure patterns across virtually all exposure durations and life stages. Both dietary and gavage administration have been used in rodent studies, with gavage being the predominant method. Gavage dosing protocols vary quite widely and include single gestational exposures, multiple daily exposures (for up to 2 weeks, intermittent schedules that include 5 days/week, once weekly, or once every 2 weeks), and loading/maintenance dose protocols, in which a relatively high dose is initially given followed by much smaller weekly doses. The intermittent dosing schedules require dose-averaging over time periods as long as 2 weeks, which introduces uncertainty in the effective exposures. In other words, the high unit dose may be more of a factor in eliciting the effect than the average TCDD tissue levels over time. Although the loading/maintenance dose protocols are designed to maintain a constant internal exposure, these protocols are somewhat inconsistent with the constant daily TCDD dietary exposures associated with human ingestion patterns.

The epidemiologic studies conducted in the Seveso cohort represent exposures over different life stages including gestation, childhood, and young adulthood. The Seveso exposure profile is essentially a high initial pulse TCDD exposure followed by a 20−30 year period of elimination. Effects are realized, or measured, 10−20 years following the initial exposure; the critical exposure window for susceptibility varies with effect and is often unknown. Therefore, the effective exposure profiles for the Seveso cohort studies vary considerably. For the Mocarelli et al. (2008) and Alaluusua et al. (2004) studies where early childhood exposures proximate to the initial event are associated with the outcomes, there is some uncertainty as to the magnitude of the effective doses. Although the effects are associated with TCDD exposure in the first 10 years of life, it is not clear to what extent the initial peak exposure is primarily responsible for the effects. It is also not clear if averaging exposure over the critical window is appropriate given the large difference between initial TCDD body burden and body burden at the end of the critical exposure window. The LOAELs for both Mocarelli et al. (2008) and Alaluusua et al. (2004) are calculated as the average of the peak exposure and average exposure across the critical exposure window (see Section 4.2 for details). For the gestational exposure study (Baccarelli et al., 2008), the critical exposure window is strictly defined and relatively short (9 months) and occurs long after the initial exposure (15−20 years). In addition, the

maternal serum TCDD measurements were taken 10−15 years after the initial exposure and are proximate to the actual pregnancies, allowing for less uncertainty in the kinetic extrapolation to time of birth. The narrow critical exposure window at a much later time than the initial exposure (where the TCDD elimination curve is flattening) is assumed to lead to a relatively steady-state exposure over the critical time period with much less uncertainty in the magnitude of the effective dose. With the exception of Eskenazi et al (2002) (see section 4.2), the effective doses for other effects reported for the Seveso cohort (see Section 2.4.1.1.1.4) have not been quantified and are not represented in Table 4-5 because no critical exposure windows can be identified or individual exposure estimates were not reported.

4.3.3. Uncertainty Factors (UFs)

The UF column in Table 4-5 shows the composite (total) UF that would be applied to the POD for each endpoint. For the animal bioassays, a UF of 3 for the toxicodynamic component of the interspecies extrapolation factor (UF_A) was applied to all PODs. For both animal and human studies, when a NOAEL was used as the POD, a factor of 10 was applied for human interindividual variability (UF_H). For all of the animal bioassay endpoints lacking a NOAEL, a UF of 10 for the LOAEL-to-NOAEL UF (UF_L) was included. For the human LOAELs, a UF_L of 3 was applied because sensitive populations were identified. A subchronic-to-chronic UF (UF_S) of 1 and a database factor (UF_D) of 1 are applied to all endpoints. A rationale for each UF is provided for the derivation of the RfD below.

4.3.4. Human Studies

For selection of the POD, the human studies are given the highest consideration, as quality human data are always preferred. Although the lower end of the candidate RfD distribution is dominated by mouse studies (comprising 6 of the 7 lowest rodent-based RfDs), EPA considers these candidate RfD estimates to be much more uncertain than human candidate RfD estimates. The $LOAEL_{HED}$s identified in mouse bioassays are low primarily because of the large toxicokinetic interspecies extrapolation factor used for mice. The ratio of administered dose to HED (D_a:HED) ranges from 65 to 1227 depending on the duration of exposure. The D_a:HED for mice is, on average, about 4 times larger than that used for rats. In addition, each

one of the mouse studies has other qualitative limitations and uncertainties (discussed above and in Table 4-4) that make them poor candidates as the basis for the RfD.

Most of the other rodent studies between the first six mouse studies and the human studies of Mocarelli et al (2008) and Baccarelli et al. (2008) are of small size, using 10 or fewer animals per dose group and are considered too uncertain on which to base the final RfD. Two of the rat bioassays–Bell et al (2007) and NTP (2006)--however, are of particular note. Both of these studies were very well designed and conducted, using 30 or more animals per dose group (see Table 4-6). Bell et al (2007) evaluated several reproductive and developmental endpoints starting exposure well before mating and continuing through gestation. NTP (2006) is the most comprehensive evaluation of TCDD chronic toxicity in rodents to date, evaluating dozens of endpoints at several time points in all major tissues. The toxicokinetic extrapolation to humans, however, is still very uncertain. Despite the overall strength of the Bell and NTP studies, EPA still considers the human data to be a better basis for the RfD.

Baccarelli et al. (2008) reported increased levels of TSH in newborns exposed to TCDD *in utero*, indicating a possible dysregulation of thyroid hormone metabolism. The study authors related TCDD concentrations in neonatal blood to TSH levels, reporting group mean TCDD concentrations associated with TSH levels above or below 5 μ-Units TSH per mL of serum (5 μU/mL). The 5 μU/mL standard was established by the World Health Organization (WHO, 1994) as an indicator of potential iodine deficiency (and potential thyroid problems) in neonates. Baccarelli et al. (2008) also showed, in graphical form, how the TSH distribution in each of three categorical exposure groups (reference, zone A, and zone B—representing increasing TCDD exposure) shifted to higher TSH values with increasing exposure. The individuals comprising the above 5 μU/mL group were from all three categorical exposure groups, not just from the highest exposure group. Therefore, EPA was able to designate a LOAEL independently of the nominal categorical exposure groups; the LOAEL is designated as the group mean of 39 ppt TCDD in neonatal plasma as a LOAEL for TSH values above 5 μU/mL. The daily oral intake at the LOAEL is estimated to be 0.024 ng/kg-day (see Section 4.2.3.1). A NOAEL is not defined because it is not clear what maternal intake should be assigned to the group below 5 μU/mL.

Mocarelli et al. (2008) reported decreased sperm concentrations (20%), decreased motile sperm counts (11%), and decreased serum estradiol (23%) in men who were 1–9 years old in 1976 at the time of the Seveso accident (initial TCDD exposure event). The sperm

concentrations and motile sperm counts in men who were 10–17 years old in 1976 were not affected; marginal serum estradiol reductions were reported in this group. Serum (LASC) TCDD levels were measured within one year of the initial exposure. Serum TCDD levels and corresponding responses were reported by quartile, with a reference group of less-exposed individuals assigned a TCDD LASC value of 15 ppt (which was the mean of the TCDD LASC reported in individuals outside the contaminated area). The lowest exposed group mean was 68 ppt (1^{st}-quartile). Mean sperm concentrations and motile sperm counts were reduced about 20% from the reference group. Further decrease in these values was slight and reached a maximum of about 33%. Although a decrease in sperm production of 20% would not have clinical significance for an individual, EPA considers a 20% shift in the population mean to be of biological significance. Therefore, EPA has designated the lowest exposure group (68 ppt) as a LOAEL, which translates to a continuous daily oral intake of 0.020 ng/kg-day (see Section 4.2.3.2). The reference group is not designated as a NOAEL because there is no clear zero-exposure measurement for any of these endpoints, particularly considering the contribution of background exposure to DLCs, which futher complicates the interpretation of the reference group response as a true "control" response (see discussion in Section 4.4). However, males less than 10 years old can be designated as a sensitive population by comparison to older males who were not affected.

Alaluusua et al. (2004) reported dental effects in male and female adults were were less than 9.5 years of age, but not older, at the time of the initial exposure (1976) in Seveso. EPA used the same approach to estimate daily TCDD intake as was used for the Mocarelli et al. (2008) data; a window of susceptibility of about 5 years was established. Serum measurements for this cohort were taken within a year of the accident. Serum TCDD levels and corresponding responses were reported by tertile, with a reference group of less-exposed individuals assigned a TCDD LASC value of 15 ppt (ng/kg); the tertile group means were 130, 383, and 1,830 ppt. Both a NOAEL and LOAEL can be defined for this study. The NOAEL is 0.12 ng/kg-day, corresponding to the TCDD LASC of 130 ppt at the first tertile. The LOAEL is 0.93 ng/kg-day at the second tertile. The children in this cohort less than 5 years old can be designated as a sensitive population by comparison to older individuals who were not affected relative to the reference group.

4.3.5. Derivation of the RfD

The two human studies, Baccarelli et al. (2008) and Mocarelli et al. (2008), have similar LOAELs of 0.024 and 0.020 ng/kg-day, respectively. Together, these two studies constitute the best foundation for establishing a POD for the RfD, and are designated as co-principal studies. Therefore, increased TSH in neonates (Baccarelli et al., 2008) and male reproductive effects (decreased sperm count and motility, increased estradiol) are designated as cocritical effects. Although the exposure estimate used in determination of the LOAEL for Mocarelli et al. (2008) is more uncertain than the Baccarelli et al. (2008) exposure estimate, the slightly lower LOAEL of 0.020 ng/kg-day from Mocarelli et al. is designated as the POD. A composite UF of 30 is applied to account for lack of a NOAEL ($UF_L = 10$) and human interindividual variability ($UF_H = 3$); the resulting RfD in standard units is 7×10^{-10} mg/kg-day. Table 4-7 presents the details of the RfD derivation.

4.4. UNCERTAINTY IN THE RfD

Exposure assessment is a key limitation of the epidemiologic studies (of the Seveso cohort) used to derive the RfD. The Seveso cohort exposure profile consists of an initial high dose followed by a drop in body burden to background levels over a period of about 20 years, at which time the effects were observed. This exposure scenario is a mismatch with the constant daily intake scenario addressed by the RfD methodology. The determination of an effective average daily dose from the Seveso exposure scenario requires an understanding of the critical time-window of susceptibility and the influence of the peak exposure on the occurrence of the observed effects, particularly when the peak exposure is high relative to the average exposure over the critical exposure window. For one of the principal studies (Mocarelli et al., 2008), a maximum susceptibility exposure window can be identified based on the age of the population at risk. However, the influence of the peak exposure on the effects observed 20 years later is unknown and the biological significance of averaging the exposure over several years, with internal exposure measures spanning a 4.5-fold range, is unknown. EPA, in this assessment, has averaged intermittent exposures for rodent bioassays over weekly dosing intervals, but the peak and average body burdens varied by less than 50%. EPA has not developed guidance for larger-interval averaging. Furthermore, because there is an assumption of a threshold level of exposure below which the effects are not expected to occur, averaging over large intervals could include

below-threshold exposures. The process used by EPA to estimate the LOAEL exposure for the Mocarelli study is a compromise between the extremes; as such, there is some uncertainty in the estimate, perhaps in the range of 3- to 10-fold in either direction. This uncertainty also holds for the LOAEL determined for the dental effects reported in Alaluusua et al. (2004) and the increased menstrual cycle length reported in Eskenazi et al. (2002a; see Section 4.2.3.4); in both of those studies, the uncertainty is greater, as the difference between peak and average internal exposures is an order of magnitude or more. The LOAEL for increased TSH in neonates (Baccarelli et al., 2008), however, is less uncertain because the critical exposure window is much narrower (9 months) and the developmental exposures occurred 10 to 15 years after the initial exposure, when internal TCDD concentrations for the pregnant women were leveling off; that is, exposure over the critical window was more constant. However, there is uncertainty in the magnitude of the exposures because they were extrapolated from serum measurements taken several years earlier.

Another source of uncertainty using human epidemiologic data is the lack of completely unexposed populations. The available TCDD epidemiologic data were obtained by comparing populations that experienced elevated TCDD exposures to populations that experienced lower exposures, rather than to a population with no TCDD exposure. An additional complicating factor is coexposure to DLCs, which can behave in the same way as TCDD. Although the accidental exposure to the Seveso women's cohort was virtually 100% TCDD, background exposure was largely to DLCs. Eskenazi et al. (2004) reported that TCDD comprised only 20% of the total toxicity equivalence (TEQ) in the serum of the reference group that was not exposed as a result of the factory explosion, which implies that the effective background exposure could have been as much as 5-fold higher. The higher background exposure could be significant at the lower TCDD exposure levels, with the effect diminishing as TCDD exposure increased. For dose-response modeling, the effect of a higher background dose (i.e., total TEQ), if included, would be to shift the response curve to the right (responses associated with higher exposures) but, primarily, would reduce the spread of the exposures, which would tend to alter the shape of the dose response towards sublinear. Both the right shift and the more sublinear shape would result in higher ED_x estimates, such as BMDs and BMDLs, from fitting dose-response models. However, for determination of a LOAEL, which is the case for all the human studies in Table 4-5, the impact may be minimal, as the LOAEL depends only on establishing that an effect

of sufficient magnitude was observed at some TCDD exposure level. In this case, the effect of the increased effective background exposure would be to inflate the "control" (zero-TEQ) response, providing the threshold for the response had been exceeded. The potential impact of an inflated control response would be to mask a significant effect of the added TCDD exposure, when the latter effect is determined by comparison to the reference group response. To compensate for this, EPA has been somewhat conservative in interpreting the magnitude of responses defining LOAELs for the Seveso cohort studies. The actual magnitude of the impact of the DLC background exposure is impossible to assess without knowing the true (TEQ-free) background response.

A primary strength of the TCDD database is that analogous effects have been observed in animal bioassays for most of the human endpoints, increasing the overall confidence in the relevance to humans of the effects reported in rodents and the association of TCDD exposure with the effects reported in humans. Table 4-5 shows that low dose TCDD exposures are associated with a wide array of toxicological endpoints in rodents including developmental effects, reproductive effects, immunotoxicity and chronic toxicity. Effects reported in human studies are similar, including male reproductive effects, increased TSH in neonates and dental defects in children; other human health effects such as female reproductive effects and chloracne have been observed at higher exposures (see Section 2.4.1). Other effects reported in rodent studies such as liver toxicity and overt immunological endpoints have not been reported in human studies. However, with respect to immunological effects, Baccarelli et al (2002, 2004) evaluated immunoglobin and complement levels in the sera of TCDD-exposed individuals from the Seveso cohort and found slightly reduced immunoglobulin in the highest exposure groups but no effect on other immunoglobulins or on C3 or C4 complement levels. The latter finding indicates that at least one immunological measure in humans is not a sensitive endpoint, as it is for mice, with large reductions in serum complement at low exposure levels (White et al., 1986).

Although there is a substantial amount of qualitative concordance of effects between rodents and humans, quantitative concordance is not evident in Table 4-5. The differential sensitivity of mice and humans for the serum complement endpoint is one example. Other examples of differential sensitivity are developmental dental effects and thyroid hormonal dysregulation. Developmental dental defects are relatively sensitive effects in rodents, appearing at exposure levels in mice (Keller et al., 2007, 2008a,b) more than an order of magnitude lower

than effect levels in humans (Alaluusua et al., 2004). In contrast, thyroid hormone effects are seen in rats (Crofton et al., 2005) at 30-fold higher exposures than for humans (Baccarelli et al., 2008). Male reproductive effects (sperm production) occur in rats (Latchoumycandane and Mathur, 2002) and humans (Mocarelli et al., 2008) at about the same dose. To what extent these differential sensitivities depend on specifics of the comparison, such as species (mouse vs. rat), life-stage (e.g., fetal vs. adult), endpoint measure (e.g., thyroxine [T4] vs. TSH) or magnitude of the lowest dose tested, cannot be determined, so strong conclusions about quantitative concordance cannot be made.

Considering the issue of lowest tested dose, the general lack of NOAELs and acceptable BMDLs is a primary weakness of the rodent bioassay database. None of the 6 most sensitive rodent studies in Table 4-5, spanning a 30-fold range of LOAELs, had defined NOAELs or BMDLs. For only 4 of the next 10 rodent studies were NOAELs or BMDLs established. In addition, many of these LOAELs are characterized by relatively high responses with respect to the control population, so it is not certain that a 10-fold lower dose (based on the application of UF_L of 10) would be approximately equivalent to a NOAEL. A major reason for the failure of BMD modeling was that the responses were not "anchored" at the low end; first response levels were far from the BMR (see Table 4-4). Another major problem with the data was non-monotone and flat response profiles. The small dose-group sizes and large dose intervals probably contributed to many of these response characteristics that prevented successful BMD modeling. Larger study sizes with narrower dose intervals at lower doses are still needed to clarify rodent response to TCDD.

Lower TCDD doses have been tested in rodents but almost entirely for investigation of specialized biochemical endpoints[24] that would not be considered to be adverse health effects (see Appendix G). There is, however, a fundamental limit to the lowest dose of TCDD that can be tested meaningfully, as TCDD is present in feed stock and accumulates in unexposed animals prior to the start of a study. This issue is illustrated by the presence of TCDD in tissues of unexposed control animals, often at significant levels relative to the lowest tested dose in low dose studies (Vanden Heuvel et al., 1994; Ohsako et al., 2001; Bell et al., 2007a; see

[24]Enzyme induction, oxidative stress indicators, mRNA levels, etc.

Text Box 4-1). Some DLCs have been measured in animal feeds and are anticipated to accumulate in unexposed test animals further complicating the interpretation of low dose studies.

Text Box 4-1.

TCDD tissue levels in control animals are rarely reported either explicitly or implicitly. Vanden Heuvel et al. (1994), however, reported TCDD concentrations in livers of control animals (10-week-old female Sprague-Dawley rats) of 0.43 ppt (ng/kg) compared to 0.49 ppt in the livers of animals given a single oral TCDD dose of 0.1 ng/kg. Assuming proportionality of liver concentration to total body burden, the body burden of untreated animals was 87.8% of that of treated animals. The equivalent administered dose for untreated animals (d_0) can be calculated as equal to $0.878 \times (0.1 + d_0)$, assuming proportionality of body burden to administered dose and that all animals started with the same TCDD body burdens. The calculation yields a value of 0.72 ng/kg for d_0, which represents the accumulated TCDD from all sources in these animals prior to being put on test. This value would raise the nominal 0.1 ng/kg TCDD dose 8-fold to 0.82 ng/kg. The next higher dose of 1 ng/kg would be nearly doubled to 1.72 ng/kg. The impact on higher doses would be negligible. Bell et al. (2007a) reported slightly higher levels (0.66 ppt) in the livers of slightly older untreated pregnant female Sprague-Dawley rats (mated at 16–18 weeks of age and tested 17 days later).

Ohsako et al. (2001) reported TCDD concentrations in the fat of offspring of untreated pregnant Holtzman rats that were 46% of the TCDD fat concentrations in animals exposed in utero to 12.5 ng/kg (single exposure on GD 15). This level of TCDD would imply a very large background exposure, but quantitation based on simple kinetic assumptions probably would not reflect the more complicated indirect exposure scenario

Bell et al. (2007a) also reported concentrations of 0.1 and 0.6 ppt TCDD measured in two samples of feed stock. Assuming that the average of 0.35 ppt is representative of the entire supply of feed stock and a food consumption factor of 10% of body weight per day, the average daily oral exposure from feed to these animals would be 0.035 ng/kg. Discrimination of outcomes from longer-term repeated exposures might be problematic at exposure levels around 0.1 ng/kg-day. Background exposure was not much of an issue for Bell et al. (2007a), as the lowest TCDD exposure level was 2.4 ng/kg-day (28-day dietary exposure).

NTP (2006), however, found virtually no TCDD in the tissues of untreated animals or in the feed stock. In all of these studies, except the 28-day exposure in Bell et al. (2007a), control animals were gavaged with corn oil vehicle. TCDD concentrations in corn oil were not reported in any of the studies.

Table 4-6 compares the qualitative strengths and limitations/uncertainties associated with the top animal bioassays listed in Table 4-5.

Table 4-1. POD candidates for epidemiologic studies of TCDD

Study	POD (ng/kg-day)	Critical effects
Alaluusua et al., 2004	7.20E-02[a] (NOAEL)	Dental effects in adults exposed to TCDD in childhood
Baccarelli et al., 2008	2.40E-2[b] (LOAEL)	Elevated TSH in neonates
Eskenazi et al., 2002	1.66E+0[c] (LOAEL)	Increased length of menstrual cycle in women exposed to TCDD in childhood
Mocarelli et al., 2008	2.00E-2[d] (LOAEL)	Decreased sperm count and motility and increased estradiol in men exposed to TCDD in childhood

[a]Mean of peak exposure (0.15 ng/kg-day) and average exposure over 10-year critical window (0.0093 ng/kg-day).
[b]Maternal exposure corresponding to neonatal TSH concentration exceeding 5 μU/mL.
[c]Mean of peak exposure (3.2 ng/kg-day) and average exposure over 10-year critical window (0.12 ng/kg-day).
[d]Mean of peak exposure (0.035 ng/kg-day) and average exposure over 10-year critical window (0.0078 ng/kg-day).

Table 4-2. Models run for each study/endpoint combination in the animal bioassay benchmark dose modeling

Model	Restrictions imposed
Continuous models	
Exponential M2-M5, not grouped	Adverse direction specified according to the response data; power ≥1
Hill	Adverse direction is automatic; $n > 1$
Linear	Adverse direction is automatic; degree of polynomial = 1
Polynomial	Adverse direction is automatic; degree of polynomial = 2; restrict the sign of the power to non-negative or non-positive, depending on the direction of the responses
Power	Adverse direction is automatic; power ≥1
Dichotomous models	
Gamma	Power ≥1
Logistic	None
Log-Logistic	Slope ≥1
Log-Probit	None
Multistage	Beta ≥0, 2nd degree polynomial
Probit	None
Weibull	Power ≥1

Table 4-3. Summary of key animal study NOAELs, LOAELs, and BMDLs for different dose metrics (ng/kg-day)

Study	Endpoint	Administered dose[a]			1st-order body burden HED[b]			Blood concentration HED[c]		
		NOAEL	LOAEL	BMDL[d]	NOAEL	LOAEL	BMDL[d]	NOAEL	LOAEL	BMDL[d]
Amin et al., 2000	Saccharin preference ratio, female	-	2.50E+01	5.15E+01	-	2.51E+04	5.17E+04	-	1.71E-01	3.19E-01
Bell et al., 2007a	Balano-preputial separation in male pups	-	2.40E+00	1.25E-01	-	1.41E+02	7.36E+00	-	1.10E-01	7.40E-03
Cantoni et al., 1981	Urinary coproporhyrins	-	1.43E+00	2.82E-01	-	1.65E+02	3.25E+01	-	6.51E-02	1.60E-03
Chu et al., 2001	Tissue weight changes	2.50E+02	1.00E+03	-	8.28E+04	3.31E+05	-	-	-	-
Chu et al., 2007	Liver lesions	2.50E+00	2.50E+01	-	8.28E+02	8.28E+03	-	3.56E-02	5.76E-01	-
Crofton et al., 2005	Serum T4	3.00E+01	1.00E+02	3.01E+01	4.69E+04	1.56E+05	4.70E+04	1.73E-01	7.62E-01	1.41E-01
Croutch et al., 2005	Decreased body weight	5.43E+01	2.17E+02	-	1.33E+04	5.30E+04	-	-	-	-
DeCaprio et al., 1986	Decreased body weight	6.10E-01	4.90E+00	-	9.05E+01	7.27E+02	-	-	-	-
Fattore et al., 2000	Decreased hepatic retinol	-	2.00E+01	-	-	3.25E+03	-	-	8.01E-01	-
Fox et al., 1993	Increased liver weight	5.73E-01	3.27E+02	-	2.31E+02	1.32E+05	-	-	-	-
Franczak et al., 2006	Abnormal estrous cycle	-	7.14E+00	-	-	8.57E+02	-	-	0.326716	-
Hojo et al., 2002	DRL response per min	-	2.00E+01	1.10E+02	-	7.60E+04	4.18E+05	-	5.51E-02	2.59E-05
Ikeda et al., 2005a	Sex ratio	-	1.65E+01	-	-	2.60E+03	-	-	2.7500259	-
Ishihara et al., 2007	Sex ratio	1.00E-01	1.00E+02	-	3.15E+01	3.15E+04	-	-	-	-
Kattainen et al., 2001	3rd molar length	-	3.00E+01	2.14E+00	-	1.14E+05	8.15E+03	-	9.00E-02	1.71E-03
Keller et al., 2007a, 2008a, 2008b	Missing mandibular molars	-	1.00E+01	9.17E+00	-	3.88E+04	3.56E+04	-	9.81E-03	1.70E-02
Kociba et al., 1976	Liver and hematologic effects and body weight changes	7.14E+00	7.14E+01	-	1.13E+03	1.13E+04	-	1.60E-01	1.71E+00	-

Table 4-3. Summary of key study NOAELs, LOAELs, and BMDLs for different dose metrics (ng/kg-day) (continued)

Study	Endpoint	Administered dose[a]			1st-order body burden HED[b]			Blood concentration HED[c]		
		NOAEL	LOAEL	BMDL[d]	NOAEL	LOAEL	BMDL[d]	NOAEL	LOAEL	BMDL[d]
Kociba et al., 1978	Liver and lung lesions, increased urinary porphyrins	1.00E+00	1.00E+01	7.30E-01	9.31E+01	9.31E+02	6.80E+01	6.46E-02	6.46E-01	2.00E-02
Latchoumycandane and Mathur., 2002	Sperm production	-	1.00E+00	1.56E-02	-	2.59E+02	4.03E+00	-	1.70E-02	3.90E-05
Li et al., 1997	Increased serum FSH	3.00E+00	1.00E+01	-	1.14E+04	3.80E+04	-	2.97E-03	1.72E-02	-
Li et al., 2006	Hormone levels (serum estradiol)	-	2.00E+00	1.11E+02	-	4.06E+03	2.27E+05	-	1.57E-03	3.46E-01
Markowski et al., 2001	FR2 revolutions	-	2.00E+01	5.81E+01	-	6.40E+04	1.86E+05	-	5.17E-02	1.34E-01
Maronpot et al., 1993	Increased relative liver weight	1.07E+01	3.50E+01	-	1.28E+03	4.18E+03	-	-	-	-
Miettinen et al., 2006	Cariogenic lesions in pups	-	3.00E+01	4.99E+00	-	1.14E+05	1.90E+04	-	8.94E-02	9.32E-03
Murray et al., 1979	Fertility index in f2 generation	1.00E+00	1.00E+01	-	1.06E+02	1.06E+03	-	2.96E-02	3.88E-01	-
NTP, 1982	Liver lesions	-	1.39E+00	1.38E+01	-	2.98E+02	2.97E+03	-	2.19E-02	1.78E-02
NTP, 2006	Liver and lung lesions	-	2.14E+00	1.40E+00	-	1.96E+02	1.28E+02	-	1.39E-01	8.76E-02
Nohara et al., 2000	Decreased spleen cellularity	8.00E+02	-	-	3.04E+06	-	-	5.34E+00	-	-
Ohsako et al., 2001	Anogenital distance in pups	1.25E+01	5.00E+01	1.22E+01	4.75E+04	1.90E+05	4.62E+04	2.87E-02	1.80E-01	2.67E-02
Seo et al., 1995	Decreased thymus weight	2.50E+01	1.00E+02	-	2.51E+04	1.00E+05	-	1.69E-01	9.31E-01	-
Sewall et al., 1995	Serum T4	1.07E+01	3.50E+01	5.20E+00	1.28E+03	4.18E+03	6.20E+02	5.15E-01	1.76E+00	7.25E-02
Shi et al., 2007	Serum estradiol in female pups	1.43E-01	7.14E-01	2.24E-01	1.67E+01	8.32E+01	2.61E+01	4.71E-03	2.75E-02	4.95E-03

Table 4-3. Summary of key study NOAELs, LOAELs, and BMDLs for different dose metrics (ng/kg-day) (continued)

Study	Endpoint	Administered dose[a]			1st-order body burden HED[b]			Blood concentration HED[c]		
		NOAEL	LOAEL	BMDL[d]	NOAEL	LOAEL	BMDL[d]	NOAEL	LOAEL	BMDL[d]
Simanainen et al., 2002	Decreased serum T4	1.00E+02	3.00E+02	-	3.80E+05	1.14E+06	-	-	-	-
Simanainen et al., 2003	Decreased thymus weight and change in EROD activity	1.00E+02	3.00E+02	-	3.80E+05	1.14E+06	-	-	-	-
Simanainen et al., 2004	Decreased daily sperm production	1.00E+02	3.00E+02	-	3.80E+05	1.14E+06	-	-	-	-
Smialowicz et al., 2004	Decreased antibody response to SRBCs	3.00E+02	1.00E+03	-	1.16E+06	3.88E+06	-	-	-	-
Smialowicz et al., 2008	PFC per 10^6 cells	-	1.07E+00	7.19E+01	-	2.29E+02	1.54E+04	-	6.38E-03	2.00E-03
Toth et al., 1979	Skin lesions	-	1.00E+00	6.85E-02	-	2.70E+02	1.85E+01	-	1.00E-02	8.61E-01
VanBirgelen et al., 1995a,b	Decreased liver retinyl palmitate	-	1.40E+01	9.89E+02	-	2.27E+03	1.61E+05	-	5.25E-01	5.00E+00
Vos et al., 1973	Decreased delayed-type hypersensitivity response to tuberculin	1.14E+00	5.71E+00	-	2.02E+02	1.01E+03	-	-	-	-
White et al., 1986	Decreased serum complement	-	1.00E+01	2.89E+01	-	4.49E+03	1.30E+04	-	2.83E-02	4.65E-02
Yang et al., 2000	Increased endometrial implant survival	1.79E+01	-	-	4.73E+02	-	-	-	-	-

[a]Average administered daily dose over the experimental exposure period.
[b]HED based on 1st-order body burden model described in Section 3.2.4.4.
[c]HED based on Emond rodent and human PBPK models described in Section 3.3.6.
[d]BMR = 0.1 for quantal endpoints and 1 standard deviation control mean for continuous endpoints, except for body and organ weights, where BMR = 10% relative deviation from control mean.

- = value not established or not modeled.

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as LASC[a])

Study	NOAEL/ LOAEL	Endpoint	Control response	First response[b]	Max response[c]	Model fit detail	BMD/ BMDL	Comments
Smialowicz et al., 2008 (mouse)	None/ 2.41E+2	PFC per spleen (n = 15)	—	24% ↓ (0.5 SD)	89% ↓	Continuous power, unrestricted (p = 0.27)	6.54E+3 2.07E+3	BMDL > LOAEL; fit at control and low dose inconsistent with data; constrained parameters in other models hit lower bounds
		PFC per 10^6 cells (n = 8−15)	—	24% ↓ (0.5 SD)	9.3-fold ↓	Continuous power unrestricted, constant variance (p = 0.48)	1.05E+3 1.19E+2	Constant variance test failed; observed control variance underestimated by 35%; poor fits for all non-constant variance models
Li et al., 2006 (mouse)	None/ 8.75E+1	serum estradiol (n = 10)	—	2.0-fold ↑ (0.8 SD)	2.4-fold ↑	Continuous linear (p = 0.16)	8.85E+3 2.96E+3	BMDL > LOAEL; high control CV (1.25); near maximal response at low dose; nonmonotonic response; other model fits are step-function-like
		serum progesterone (n = 10)	—	33% ↓ (2.0 SD)	61% ↓	Continuous Hill (p = 0.44)	8.95E-1 6.31E-3	No response data near BMR; large CVs (>1) for treatment groups; poor fit for variance model; Hill coefficient at lower bound (step-function); implausible BMDL
Toth et al., 1979 (mouse)	None/ 3.15E+2	skin lesions (n = 38−44)	0/38	5/44	25/43	Dichotomous log-logistic (p = 0.67)	1.1E-2 4.7E-4	Constrained parameter lower bound hit; implausible BMDL
						Dichotomous log-logistic, unrestricted (p = 0.98)	2.60E+2 3.18E+0	Supralinear fit (slope = 0.48); implausible BMDL
		dermal amyloidosis (n = 38−44)	0/38	5/44	17/43	Dichotomous log-logistic (p = 0.02)	8.3E+3 4.8E+3	Poor fit; constrained parameter lower bound hit; BMDL > LOAEL
						Dichotomous log-logistic, unrestricted (p = 0.90)	2.67E+2 2.93E+0	Supralinear fit (slope = 0.33); implausible BMDL

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as LASC[a]) (continued)

Study	NOAEL/ LOAEL	Endpoint	Control response	First response	Max response	Model fit detail	BMD/ BMDL	Comments
Latchoumy-candane and Mathur, 2002 (rat)	None/ 4.37E+2	daily sperm production (n = 6)	—	29% ↓ (1.0 SD)	41% ↓	Continuous Hill, restricted (p = 0.98)	3.39E+2 1.51E-2	Near maximal response at LOAEL; constrained parameter bound hit; implausible BMDL; standard deviations given in paper interpreted as standard errors
						Continuous Hill, unrestricted (p = 0.96)	3.32E+2 8.77E-3	Slightly supralinear fit (***n*** = 0.91); implausible BMDL
NTP, 1982 (mouse)	None/ 4.20E+2	Toxic hepatitis; males (n = 50)	1/73	5/49	44/50	Dichotomous multistage (p = 0.01)	8.3E+2 3.7E+2	No acceptable model fits; lowest BMDL shown
White et al., 1986 (mouse)	- 6.03E+2	Total hemolytic complement activity (CH50) (n = 8)	—	41% ↓ (2.6 SD)	81% ↓	Continuous Hill, restricted (p = 0.002)	4.76E+3 8.56E+2	Poor fit; no response near BMR; constrained parameter bound hit; BMDL > LOAEL
						Continuous Hill, unrestricted (p = 0.07)	8.17E+1 7.55E+1	Supralinear fit (***n*** = 0.25); implausible BMDL
Keller et al., 2007, 2008a, b (mouse)	None/ 2.96E+2	Missing molars (n = 23−36)	0/29	2/23	30/30	Dichotomous 1° multi-stage (p = 0.26)	6.01E+2 4.20E+2	Poor fit at first response level; not most sensitive endpoint; other endpoints not amenable to BMD modeling
Shi et al., 2007 (rat)	1.88E+2 5.92E+2	Serum estradiol in female pups (n = 10)	—	38% ↓ (0.4 SD)	62% ↓	Continuous exponential (M4) (p = 0.69)	4.45E+2 1.95E+2	Adequate fit; selected
Cantoni et al., 1981 (rat)	None/ 1.02E+3	Urinary uro-porhyrins (n = 4)	—	2.4-fold ↑ (5.7 SD)	87-fold ↑	Continuous exponential (M2) (p = 0.0003)	2.07E+3 1.52E+3	No response near BMR; poor fits for all non-constant variance models; constant variance poor representation of control SD; BMDL > LOAEL

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as LASC[a]) (continued)

Study	NOAEL/ LOAEL	Endpoint	Control response	First response	Max response	Model fit detail	BMD/ BMDL	Comments
Cantoni et al., 1981 (rat, continued)	None/ 1.02E+3 (cont.)	Urinary copro-porhyrins (n = 4)	—	2.4-fold ↑ (3.1 SD)	4.0-fold ↑	Continuous exponential (M4) (p = 0.49)	2.94E+2 9.93E+1	No response near BMR
						Continuous power, unrestricted (p = 0.61)	1.52E+1 2.31E-6	Supralinear fit (n = 0.30); implausible BMDL; poor model choice for plateau effect
NTP, 2006 (rat)	None/ 1.41E+3	Hepatocyte hypertrophy (n = 53−54)	0/53	19/54	52/53	Dichotomous multistage (p = 0.03)	5.01E+2 4.33E+2	Poor fits for all models
		Alveolar metaplasia (n = 52−54)	2/53	19/54	46/52	Dichotomous log-logistic (p = 0.72)	3.58E+2 2.07E+2	No response near BMR
		Oval cell hyperplasia (n = 53−54)	0/53	4/54	53/53	Dichotomous probit (p = 0.23)	3.13E+3 2.64E+3	Relatively poor fit for control and low dose groups; negative response intercept (same for logistic); implausible model; BMDL > LOAEL
						Dichotomous Weibull (p = 0.08)	3.15E+3 2.25E+3	Marginal fit; BMDL > LOAEL
		Gingival hyperplasia (n = 53−54)	1/53	7/54	16/53	Dichotomous log-logistic (p = 0.06)	3.22E+3 2.25E+3	Poor fit; constrained parameter bound hit; BMDL > LOAEL
						Dichotomous log-logistic, unrestricted (p = 0.66)	3.88E+2 6.95E-3	Supralinear fit (slope = 0.33); implausible BMDL
		Eosinophilic focus, multiple (n = 53−54)	3/53	8/54	42/53	Dichotomous probit (p = 0.46)	3.08E+3 2.68E+3	Relatively poor fit to control response; BMDL > LOAEL
		Liver fatty change, diffuse (n = 53−54)	0/53	2/54	48/53	Dichotomous Weibull (p = 0.72)	2.16E+3 1.57E+3	BMDL > LOAEL; otherwise adequate fit

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as LASC[a]) (continued)

Study	NOAEL/ LOAEL	Endpoint	Control response	First response	Max response	Model fit detail	BMD/ BMDL	Comments
NTP, 2006 (rat, continued)	None/ 1.41E+3 (cont.)	Liver necrosis (n = 53−54)	1/53	4/54	17/53	Dichotomous log-probit, unrestricted (p = 0.80)	4.13E+3 1.93E+3	Adequate fit; slightly supralinear; BMDL > LOAEL
		Liver pigmentation (n = 53−54)	4/53	9/54	53/53	Dichotomous log-probit (p = 0.96)	1.36E+3 1.04E+3	Adequate fit
		toxic hepatopathy (n = 53−54)	0/53	2/54	53/53	Dichotomous multistage, restricted (p = 0.81)	2.11E+3 1.70E+3	BMDL > LOAEL; otherwise adequate fit
Bell et al., 2007a (rat)	None/ 2.00E+3	Balano-preputial separation in male pups (n = 30 [dams])	1/30	5/30	15/30	Dichotomous log-logistic, restricted (p = 0.79)	1.96E+3 1.22E+3	Adequate fit; constrained parameter bound hit; not litter based
						Dichotomous log-logistic, unrestricted (p = 0.51)	1.81E+3 2.64E+2	Adequate fit; slightly supralinear (slope = 0.95); selected
Kociba et al., 1978 (rat)	8.53E+2 3.94E+3	Uroporphyrin per creatinine, females (n = 5)	—	15% ↑ (0.48 SD)	89% ↑	Continuous linear (p = 0.79)	7.20E+3 5.12E+3	BMDL > LOAEL; otherwise adequate fit
		Urinary coproporphyrins, females (n = 5)	—	67% ↑ (5.1 SD)	78% ↑	Continuous exponential (M4, non-constant var) (p = 0.01)	8.6E+2 4.0E+2	Poor fit; no response near BMR; BMDL > LOAEL
		Liver lesions (n = 50)						No data presented
		Lung lesions (n = 50)						No data presented
Markowski et al., 2001 (rat)	None/ 1.23E+3	FR5 run opportunities (n = 4−7)	—	10% ↓ (0.21 SD)	51% ↓	Continuous Hill (p = 0.94)	1.37E+3 7.21E+2	Constrained parameter upper bound hit;

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as LASC[a]) (continued)

Study	NOAEL/ LOAEL	Endpoint	Control response	First response	Max response	Model fit detail	BMD/ BMDL	Comments
Markowski et al., 2001 (rat, continued)	None/ 1.23E+3 (cont.)					Continuous power, unrestricted (p = 0.13)	2.11E+3 8.15E-12	Saturated model; supralinear fit (power = 0.39); BMD/BMDL ratio >> 100
		FR2 revolutions (n = 4−7)	–	9% ↓ (0.15 SD)	43% ↓	Continuous Hill (p = 0.65)	1.46E+3 4.76E+2	Constrained parameter bound hit (upper bound)
						Continuous power, unrestricted (p = 0.16)	4.54E+3 8.15E-12	Supralinear fit (power = 0.32); implausible BMDL
		FR10 run opportunities (n = 4−7)	–	15% ↓ (0.24 SD)	57% ↓	Continuous exponential (M2) (p = 0.30)	6.77E+3 2.28E+3	BMDL > LOAEL
Hojo et al., 2002 (rat)	None/ 1.29E+3	DRL reinforce per min (n = 12)	–	55% ↑ (1.0 SD)	80% ↑	Continuous exponential (M4) (p = 0.054)	1.04E+3 4.94E+0	Poor fit; near maximal response at lowest dose, BMD/BMDL ratio >> 100
		DRL response per min (n = 12)	–	105% ↓ (2.4 SD)	105% ↓	Continuous exponential (M4) (p = 0.48)	3.02E+2 7.55E+0	No response data near BMR; maximal response at lowest dose, BMD/BMDL ratio >> 20
Murray et al., 1979 (rat)	6.19E+2/ 3.24E+3	fertility in f2 gen. (no litters) (n = 20)	4/32	0/20	9/20	Dichotomous multistage (p = 0.08)	1.50E+3 7.50E+2	Poor fit; non-monotonic response; no response data near BMR
Kattainen et al., 2001 (rat)	None/ 1.76E+3	3rd molar length in pups (n = 4−8)	–	15% ↓ (4.2 SD)	27% ↓	Continuous Hill, restricted (p < 0.01)	2.48E+2 1.33E+2	No response data near BMR; Constrained parameter lower bound hit
						Continuous Hill, unrestricted (p < 0.01)	2.01E+5 -	BMDL could not be calculated

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as LASC[a]) (continued)

Study	NOAEL/ LOAEL	Endpoint	Control response	First response	Max response	Model fit detail	BMD/ BMDL	Comments
Kattainen et al., 2001 (rat, continued)	None/ 1.76E+3 (cont.)	3rd molar eruption in pups (n = 4−8)	1/16	3/17	13/19	Dichotomous log-logistic, restricted (p = 0.98)	1.90E+3 1.05E+3	Constrained parameter lower bound hit
						Dichotomous log-logistic, unrestricted (p = 0.95)	1.53E+3 1.46E+2	Supralinear fit (slope = 0.91)
Miettinen et al., 2006 (rat)	None/ 1.76E+3	Cariogenic lesions in pups (n = 4−8)	25/42	23/29	29/32	Dichotomous log-logistic, restricted (p = 0.60)	1.13E+3 4.09E+2	Constrained parameter lower bound hit; near maximal response at LOAEL; high control response
						Dichotomous log-logistic, unrestricted (p = 0.73)	3.91E+1 -	Supralinear fit (slope = 0.47); BMDL could not be calculated
Ohsako et al., 2001 (rat)	8.45E+2/ 2.76E+3	Ano-genital distance in male pups (n = 5)	—	12% ↓ (1.0 SD)	17% ↓	Continuous Hill, restricted (p = 0.26)	3.63E+3 8.05E+2	Constrained parameter lower bound hit; near maximal response at LOAEL
						Continuous Hill, unrestricted (p = 0.11)	4.74E+3 4.52E+2	Supralinear fit (***n*** = 0.62)
Amin et al., 2000 (rat)	None/ 2.67E+3	Saccharin consumed, female, (0.25%) (n = 10)	—	22% ↓ (0.3 SD)	66% ↓	Continuous linear (p = 0.55)	7.22E+3 4.81E+3	BMDL > LOAEL; restricted power model, constrained parameter hit lower bound
						Continuous power, unrestricted (p = NA)	6.61E+3 2.70E+3	Saturated model; supralinear fit (power = 0.74)
		Saccharin consumed, female (0.50%) (n = 10)	—	49% ↓ (0.7 SD)	80% ↓	Continuous linear (p = 0.06)	8.02E+3 5.18E+3	Restricted power model, constrained parameter hit lower bound

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as LASC[a]) (continued)

Study	NOAEL/ LOAEL	Endpoint	Control response	First response	Max response	Model fit detail	BMD/ BMDL	Comments
Amin et al., 2000 (continued)	None/ 2.67E+3 (cont.)					Continuous power, unrestricted (p = NA)	5.19E+3 9.41E+2	Saturated model; supralinear fit (power = 0.40)
		Saccharin preference ratio, female (0.25%) (n = 10)	–	29% ↓ (1.8 SD)	33% ↓	Continuous linear (p = 0.002)	9.17E+3 4.39E+3	BMDL > LOAEL; no response near BMR; near maximal response at LOAEL
		Saccharin preference ratio, female (0.50%) (n = 10)	–	39% ↓ (1.1 SD)	54% ↓	Continuous linear (p = 0.14)	6.43E+3 4.03E+3	BMDL > LOAEL; near maximal response at LOAEL; restricted power model, constrained parameter hit lower bound
						Continuous power, unrestricted (p = NA)	2.05E+3 1.26E-5	Saturated model; supralinear fit (power = 0.28)
Crofton et al., 2005 (rat)	1.91E+3/ 5.10E+3	Serum T4, (n = 4−14)	–	29% ↓ (1.9 SD)	51% ↓	Continuous exponential (M4) (p = 0.94)	2.86E+3 1.67E+3	No response near BMR
Sewall et al., 1995 (rat)	3.92E+3 9.16E+3	Serum T4 (n = 9)	–	9.1% ↓ (0.6 SD)	40% ↓	Continuous Hill (p = 0.90)	5.68E+3 1.98E+3	Constrained parameter hit lower bound; otherwise acceptable fit
						Continuous Hill, unrestricted (p = 0.90)	5.35E+3 1.09E+3	Supralinear fit (power = 0.57); otherwise acceptable fit
Schantz et al., 1996 (rat)	None/ None	Maze errors per block, female (n = 10)	–	22% ↓ (0.2 SD)	34% ↓	Continuous linear (p = 0.16)	5.53E+3 3.63E+3	BMDL > LOAEL; near maximal response at LOAEL; restricted power model, constrained parameter hit lower bound
						Continuous power, unrestricted (p = NA)	2.02E+3 8.11E-6	Saturated model; supralinear fit (power = 0.37); implausible BMDL
Li et al., 1997 (rat)	1.46E+2/ 4.40E+2	FSH in female rats (n = 10)	–	3.6-fold ↑ (2.0 SD)	19-fold ↑	Continuous power, restricted (p < 0.01)	1.11E+5 7.50E+4	Power hit lower bound

Table 4-4. TCDD BMDL analysis (NOAEL, LOAEL, BMD, and BMDL values given as LASC[a]) (continued)

Study	NOAEL/ LOAEL	Endpoint	Control response	First response	Max response	Model fit detail	BMD/ BMDL	Comments
Van Birgelen et al., 1995ab (rat)	None/ 3.97E+3	Hepatitic retinol (n = 8)	–	44% ↓ (0.74 SD)	96% ↓	Continuous exponential (M4) ($p < 0.01$)	1.37E+4 1.85E+3	Poor fit
						Continuous power, unrestricted ($p = 0.01$)	1.03E+5 7.92E+4	Poor fit; supralinear fit (power = 0.14),
		Hepatitic retinyl palmitate (n = 8)	–	80% ↓ (1.4 SD)	99% ↓	Continuous exponential (M4) ($p < 0.01$)	7.79E+4 2.01E+4	Poor fit; no response near BMR
						Continuous power, unrestricted ($p = 0.24$)	2.90E+1 3.25E–2	Supralinear fit (power = 0.06); implausible BMDL

[a]Converted to whole blood concentrations as described in Section 3 prior to determining HEDs in Table 4-5.
[b]Magnitude of response at first dose where response differs from control value (in the adverse direction); continuous response magnitudes given as relative to control; quantal response given as number affected/total number.
[c]Magnitude of response maximally differing from control value (in the adverse direction).

SD = standard deviation.

Table 4-5. Candidate points of departure for the TCDD RfD using blood-concentration-based human equivalent doses

Study	Species, strain (sex, if not both)	Protocol	Endpoint	$NOAEL_{HED}$ (N) or $BMDL_{HED}$ (B) (ng/kg-day)	$LOAEL_{HED}$ (ng/kg-day)	UF[a]	RfD (mg/kg-day)
Li et al., 2006	Mouse, NIH (F)	Gavage GD 1–3; *n* = 10	Hormone levels in pregnant dams (decreased progesterone, increased estradiol)	-	1.6E-03	300	5.3E-12
Smialowicz et al., 2008	Mouse, B6C3F1 (F)	90-day gavage; *n* = 8–15	Decreased SRBC response	-	6.4E-03	300	2.1E-11
Keller et al., 2007, 2008a, b[b]	Mouse, CBA/J and C3H/HeJ	Gavage GD 13; *n* = 23–36 (pups)	Missing molars, mandibular shape changes in pups	-	9.8E-03	300	3.3E-11
Toth et al., 1979	Mouse, Swiss/ H/Riop (M)	1-year gavage; *n* = 38–44	Dermal amyloidosis, skin lesions	-	1.0E-02	300	3.3E-11
Latchoumy-candane and Mathur, 2002	Rat, Wistar (M)	45-day oral pipetting; *n* = 6	Decreased sperm production	-	1.7E-02	300	5.7E-11
NTP, 1982	Mouse, B6C3F1 (M)	2-year gavage; *n* = 50	Liver lesions	-	2.2E-02	300	7.3E-11
White et al., 1986	Mouse, B6C3F1 (F)	14-day gavage; *n* = 6–8	Decreased serum complement	-	2.8E-02	300	9.4E-11
Li et al., 1997	Rat, S-D (F, 22 day-old)	Single gavage; n = 10	Increased serum FSH	3.0E-03 (N)	1.7E-02	30[c]	1.0E-10
DeCaprio et al., 1986	Guinea pig, Hartley	90-day dietary; *n* = 10	Decreased body weight, organ weight changes (liver, kidney, thymus, brain)	4.1E-03[d] (N)	3.3E-02[d]	30[c]	1.4E-10
Shi et al., 2007	Rat, S-D (F)	11-month gavage; *n* = 10	Decreased serum estradiol	4.7E-03 (N) 5.0E-03 (B)	2.8E-02	30[c]	1.6E-10
Markowski et al., 2001	Rat, Holtzman	Gavage GD 18; *n* = 4–7	Neurobehavioral effects in pups (running, lever press, wheel spinning)	-	5.2E-02	300	1.7E-10
Hojo et al., 2002	Rat, S-D	Gavage GD 8; *n* = 12	Food-reinforced operant behavior in pups	-	5.5E-02	300	1.8E-10
Vos et al., 1973	Guinea pig, Hartley (F)	8-week gavage; *n* = 10	Decreased delayed-type hypersensitivity response to tuberculin	6.4E-03[d] (N)	3.2E-02[d]	30[c]	2.1E-10

Table 4-5. Candidate points of departure for the TCDD RfD using blood-concentration-based human equivalent doses (continued)

Study	Species, strain (sex, if not both)	Protocol	Endpoint	$NOAEL_{HED}$ (N) or $BMDL_{HED}$ (B) (ng/kg-day)	$LOAEL_{HED}$ (ng/kg-day)	UF[a]	RfD (mg/kg-day)
Cantoni et al., 1981	Rat, CD-COBS (F)	45-week gavage; *n* = 4	Increased urinary porhyrins	-	6.5E-02	300	2.2E-10
Bell et al., 2007	Rat, CRL:WI (Han) (M)	17 week dietary; *n* = 30	Delay in onset of puberty	7.4E-03 (B)	1.1E-01	30[c]	2.5E-10
Miettinen et al., 2006	Rat, Line C	Gavage GD 15; *n* = 3–10	Cariogenic lesions in pups	-	8.9E-02	300	3.0E-10
Kattainen et al., 2001	Rat, Line C	Gavage GD 15; *n* = 4–8	Inhibited molar development in pups	-	9.0E-02	300	3.0E-10
NTP, 2006	Rat, S-D (F)	2-year gavage; *n* = 53	Liver and lung lesions	-	1.4E-01	300	4.7E-10
Amin et al., 2000	Rat, S-D	Gavage GD 10–16; *n* = 10	Reduced saccharin consumption and preference	-	1.7E-01	300	5.7E-10
Schantz et al., 1996	Rat, S-D (F)	Gavage GD 10–16; *n* = 10	Altered maze performance	-	1.7E-01	300	5.7E-10
Mocarelli et al., 2008	**Human (M)**	**Childhood exposure; *n* = 157**	**Decreased sperm conc, sperm motility and increased estradiol, as adults**	**-**	**2.0E-02[e]**	**30[f]**	**6.7E-10**
Baccarelli et al., 2008	**Human infants**	**Gestational exposure; *n* = 51**	**Increased TSH in newborn infants**	**-**	**2.4E-02[g]**	**30[f]**	**8.2E-10**
Ohsako et al., 2001	Rat, Holtzman	Gavage GD 15; *n* = 5	Decreased ano-genital distance in male pups	2.9E-02 (N)	1.8E-01	30[c]	9.6E-10
Murray et al., 1979	Rat, S-D	3-generation dietary	Reduced fertility and neonatal survival (f 0 and f 1)	3.0E-02 (N)	3.9E-01	30[c]	1.0E-09
Chu et al., 2007	Rat, S-D (F)	28-day gavage, *n* = 5	Liver lesions	3.6E-02 (N)	5.8E-01	30[c]	1.2E-09
Van Birgelen et al., 1995	Rat, S-D (F)	13-week dietary; *n* = 8	Decreased liver retinyl palmitate	-	5.3E-01	300	1.8E-09

Table 4-5. Candidate points of departure for the TCDD RfD using blood-concentration-based human equivalent doses (continued)

Study	Species, strain (sex, if not both)	Protocol	Endpoint	$NOAEL_{HED}$ (N) or $BMDL_{HED}$ (B) (ng/kg-day)	$LOAEL_{HED}$ (ng/kg-day)	UF[a]	RfD (mg/kg-day)
Kociba et al., 1978	Rat, S-D (F)	2-year dietary; *n* = 50	Liver and lung lesions, increased urinary porhyrins	6.5E-02 (N)	6.5E-01	30[c]	2.2E-09
Sewall et al., 1995	Rat, S-D (F)	30-week gavage; *n* = 9	Decreased serum T4	5.2E-01 (N) 7.3E-02 (B)	1.8E+00	30[c]	2.4E-09
Fattore et al., 2000	Rat, S-D	13-week dietary; *n* = 6	Decreased hepatic retinol	-	8.01E-1	300	2.7E-09
Seo et al., 1995	Rat, S-D	Gavage GD 10–16; *n* = 10	Decreased serum T4 and thymus weight	1.7E-01 (N)	9.3E-01	30[c]	5.6E-09
Crofton et al., 2005	Rat, Long-Evans (F)	4-day gavage; *n* = 4−14	Decreased serum T4	1.7E-01 (N)	7.6E-01	30[c]	5.7E-09
Alaluusua et al., 2004	Human	Childhood exposure; *n* = 48	Dental defects	1.2E-01[h] (N)	9.3E-01[i]	[j]	3.9E-08

3

[a]Except where indicated, $UF_A = 3$ (for dynamics), $UF_H = 10$, $UF_L = 10$.
[b]Results from 3 separate studies with identical designs combined.
[c]$UF_L = 1$ (NOAEL or BMDL).
[d]HED determined from 1st-order body burden model; no PBPK model available for guinea pigs.
[e] Mean of peak exposure (0.0319 ng/kg-day) and average exposure over 10-year critical window (0.00802 ng/kg-day).
[f]$UF_H = 3$, $UF_L = 10$.
[g]Maternal exposure corresponding to neonatal TSH concentration exceeding 5 µU/mL.
[h]Mean of peak exposure (0.200 ng/kg-day) and average exposure over 10-year critical window (0.0335 ng/kg-day).
[i] Mean of peak exposure (1.71 ng/kg-day) and average exposure over 10-year critical window (0.153 ng/kg-day).
[j]$UF_H = 3$.

S-D = Sprague-Dawley.

Table 4-6. Qualitative analysis of the strengths and limitations/uncertainties associated with animal bioassays possessing candidate points-of-departure for the TCDD RfD

Study	Strengths	Limitations	Remarks
Li et al., 2006	• Female reproductive effects (i.e., early embryo loss and changes in serum progesterone and estradiol) were tested at multiple exposure times—early gestation, preimplantation, and peri- to postimplantation.	• Relatively small sample size of mouse dams/dose employed (n = 10)	Study may have human relevance based on perceived TCDD-induced female reproductive effects; uterine sequestration of TCDD observed.
Smialowicz et al., 2008	• Sheep red blood cell (SRBC) plaque forming cell assay is highly sensitive and reproducible across laboratories when examining TCDD	• Small sample size of mice/dose employed (n = 8) • Only female mice were tested • Thymus and spleen weights were only other immune response-related endpoints tested	Study adds to a substantial database on the immunotoxicity of TCDD in laboratory animals
Toth et al., 1979	• Large sample size of mice/dose employed • One-year exposure duration	• Reporting of findings is terse and lacks sufficient detail (e.g., materials and methods, thorough description of pathological findings, etc.) • Only male mice were tested for amyloidosis and skin lesions	Study has human relevance based on similarity of ulcerous skin lesions and amyloidosis in mice to chloracne observed in humans
Latchoumy-candane and Mathur, 2002	• Compared to epididymal sperm counts, the testicular spermatid head count provides better quantitation of acute changes in sperm production and can indicate pathology	• Small sample size of rats/dose employed (n = 6) • Oral pipette administration of TCDD may be a less efficient dosing method than gavage	Study has human relevance based on observed TCDD-induced male reproductive effects

Table 4-6. Qualitative analysis of the strengths and limitations/uncertainties associated with animal bioassays possessing candidate points-of-departure for the TCDD RfD (continued)

Study	Strengths	Limitations	Remarks
Keller et al. 2007, 2008a,b	• Six different inbred mouse strains were utilized • Large sample size of mouse offspring/dose/strain evaluated • Low TCDD dose levels used compared to typical mouse studies allowed for identification of subtle sensitivity differences in presence of absence of third molars, variant molar morphology, and mandible structure in offspring	• Unknown sample size of mouse dams/dose/strain employed • All inbred strains possessed sensitive *b* allele at the *Ahr* locus (i.e., a potentially resistant sub-population was not evaluated for comparison purposes) • Morphological dental and mandibular changes induced by TCDD at earlier or later gestational dosing dates are unknown because of single gavage administration on GD 13 • Difficulties breeding A/J mice led to abandonment of that strain in the analysis (Keller et al., 2008a, b)	Studies have human relevance based on observed TCDD-induced developmental dental effects.
NTP, 1982	• Large sample size of mice and rats/dose employed • Comprehensive 2-year bioassay that assessed body weights, clinical signs, and pathological changes in multiple tissues and organs	• Elevated background levels of hepatocellular tumors in untreated male mice • Gavage exposure was only 2 days/week	One of two comprehensive chronic toxicity evaluations of TCDD in rodents
White et al., 1986	• Total hemolytic complement (CH50) is representative functional assay of the complete complement sequence	• Small sample size of rats/dose employed (n = 6−8) • Individual complement factors may be significantly depleted without affecting CH50 activity (only C3 is measured) • TCDD used for dosing was of unknown purity	First report on TCDD-induced effects on serum complement
Shi et al., 2007	• Study design evaluated TCDD effects on aging female reproductive system (i.e., exposure began *in utero* and spanned across reproductive lifespan) • Several female reproductive endpoints were evaluated, including cyclicity, endocrinology, serum hormone levels, and follicular reserves	• Relatively small sample size of rats/dose employed (n = 10)	Study may have human relevance based on perceived TCDD-induced female reproductive effects

Table 4-6. Qualitative analysis of the strengths and limitations/uncertainties associated with animal bioassays possessing candidate points-of-departure for the TCDD RfD (continued)

Study	Strengths	Limitations	Remarks
Markowski et al., 2001	• Low TCDD dose levels used allowed for subtle behavioral deficits to be identified in rat offspring. • Several training sessions on wheel apparatuses were extensive. • Neurobehavioral effects are exposure-related and cannot be attributed to motor or sensory deficits.	• Unknown sample size of rat dams/dose employed. • Small sample size of rat offspring/dose evaluated (n = 4-7). • TCDD used for dosing was of unknown purity and origin. • Neurobehavioral effects induced by TCDD at earlier or later gestational dosing dates are unknown because of single gavage administration on GD 18. • Although BMD analysis was conducted, the model parameters were not constrained according to EPA guidance, so the results cannot be used.	Study informs a dearth of information on neurobehavioral toxicity of TCDD; observed neurobehavioral changes signify meaningful developmental outcome.
DeCaprio et al., 1986	• Subchronic oral dosing duration up to 90 days. • Male and female guinea pigs tested	• Relatively small sample size of guinea pigs/dose employed (n = 10) • No histopathological analyses performed • TCDD used for dosing was of unknown purity	Standard subchronic oral study
Vos et al., 1973	• Three different animal species tested (guinea pigs, mice, rats) • Effects of TCDD tested on both cell-mediated and humoral immunity	• Small sample size of animals/dose employed in each experiment (n = 5−10) • Only female guinea pigs and rats were tested, and only male mice were tested • Only one experimental assay was utilized to assess cell-mediated and humoral immunity in each animal species; humoral immunity was only investigated in guinea pigs • TCDD used for dosing was of unknown purity	Early study on the immunotoxicity of TCDD in laboratory animals
Cantoni et al., 1981	• Experiments were designed to test qualitative and quantitative composition and the course of urinary excretion in TCDD-induced porphyria	• Small sample size of rats/dose employed (n = 4) • Concurrent histological changes with tissue porphyrin levels were not examined • TCDD used for dosing was of unknown purity	Early study on porphyrogenic effects of TCDD

Table 4-6. Qualitative analysis of the strengths and limitations/uncertainties associated with animal bioassays possessing candidate points-of-departure for the TCDD RfD (continued)

Study	Strengths	Limitations	Remarks
Bell et al., 2007	• Large sample size of both rat dams and offspring/dose employed • Several developmental effects tested	• Batch-to-batch variation of up to 30% in TCDD concentration in the diet • Chronic dosing of dams does not accurately define gestational period when fetus is especially sensitive to TCDD-induced toxicity	Study adds to a substantial database on the developmental toxicity of TCDD in laboratory animals
Hojo et al., 2002	• Low TCDD dose levels used allowed for subtle behavioral deficits to be identified in rat offspring. • Preliminary training sessions in operant chamber apparatuses were extensive. • Neurobehavioral effects are exposure-related and cannot be attributed to presence of learning or discrimination deficits.	• Relatively small sample size of rat dams/dose employed (n = 12). • Small sample size of rat offspring/dose evaluated (n = 5-6). • Neurobehavioral effects induced by TCDD at earlier or later gestational dosing dates are unknown because of single gavage administration on GD 8. • Although BMD analysis was conducted, the model parameters were not constrained according to EPA guidance, so the results cannot be used.	Study informs a dearth of information on neurobehavioral toxicity of TCDD; observed neurobehavioral changes signify meaningful developmental outcome.
NTP, 2006	• Chronic exposure duration with several interim sacrifices • Large number of dose groups with close spacing • Large number of animals per dose group • Comprehensive suite of endpoints evaluated • Comprehensive biochemical, clinical and histopathological tests and measures • Detailed reporting of results, with individual animal data presented as well as group summaries • Evaluation of background exposure to TCDD and DLCs in feed	• Single species, strain and sex • Lowest dose tested too high for establishing NOAEL	Study is the most comprehensive chronic TCDD toxicity evaluation in rats to date

Table 4-7. Basis and derivation of the TCDD reference dose

Principal study detail		
Study	**POD (ng/kg-day)**	**Critical effects**
Mocarelli et al., 2008	0.020 (LOAEL)	Decreased sperm count (20%) and motility (11%) and decreased estradiol (23%) in men exposed to TCDD during childhood
Baccarelli et al., 2008	0.024 (LOAEL)	Elevated TSH (> 5 μU/mL) in neonates

RfD derivation	
POD	0.020 ng/kg-day (2.0E-8 mg/kg-day)
UF	30 (UF_L = 10, UF_H = 3)
RfD	7×10^{-10} (7E-10) mg/kg-day (2.0E-8 ÷ 30)

Uncertainty factors		
LOAEL-to-NOAEL (UF_L)	10	No NOAEL established; cannot quantify lower exposure group in Baccarelli et al. (2008); magnitude of effects at LOAEL sufficient to require a 10-fold factor.
Human interindividual variability (UF_H)	3	A factor of 3 ($10^{0.5}$) is used because the effects were elicited in sensitive populations. A further reduction to 1 was not made because the sample sizes were relatively small, which, combined with uncertainty in exposure estimation, may not fully capture the range of interindividual variability.
Interspecies extrapolation (UF_A)	1	Human study.
Subchronic-to-chronic (UF_S)	1	Chronic effect levels are not well defined for humans; however, animal bioassays indicate that developmental effects are the most sensitive, occurring at doses lower than other effects noted in chronic studies. Considering that exposure in the principal studies encompasses the critical window of susceptibility associated with development, an UF to account for exposure duration is not warranted.
Database sufficiency (UF_D)	1	The database for TCDD contains an extensive range of human and animal studies that examine a comprehensive set of endpoints. There is no evidence to suggest that additional data would result in a lower reference dose.

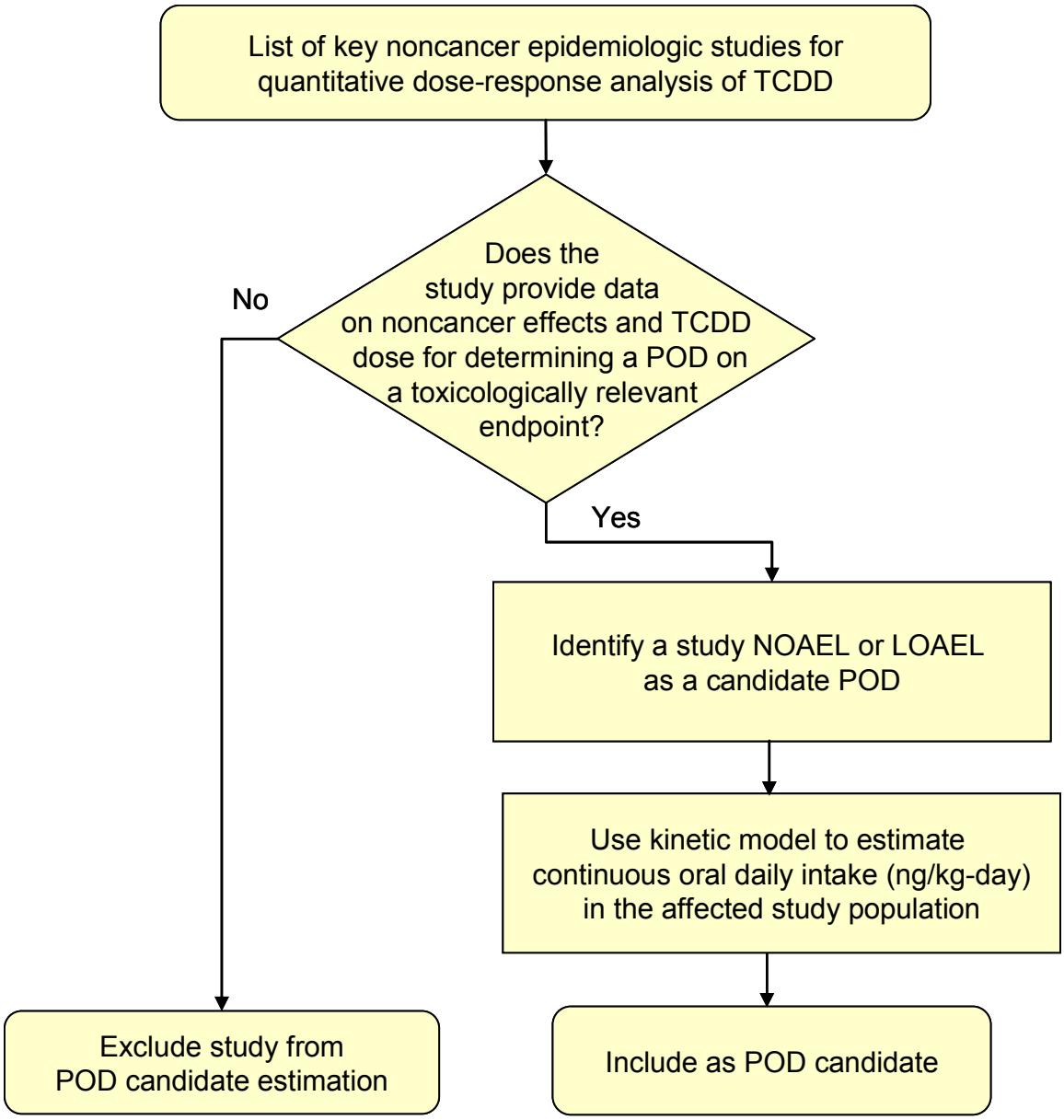


Figure 4-1. EPA's process to select and identify candidate PODs from key epidemiologic studies for use in the noncancer risk assessment of TCDD. For each noncancer study that qualified for TCDD dose-response assessment using the study inclusion criteria, EPA first evaluated the dose-response information developed by the study authors for whether the study provided noncancer effects and TCDD dose data for a toxicologically relevant endpoint. If such data were available, then EPA identified a NOAEL or LOAEL as a candidate POD. Then, EPA used a human kinetic model to estimate the continuous oral daily intake (ng/kg-day) for the candidate POD that could be used in the derivation of an RfD based on the study data. If all of this information was available, then the result was included as a candidate POD.

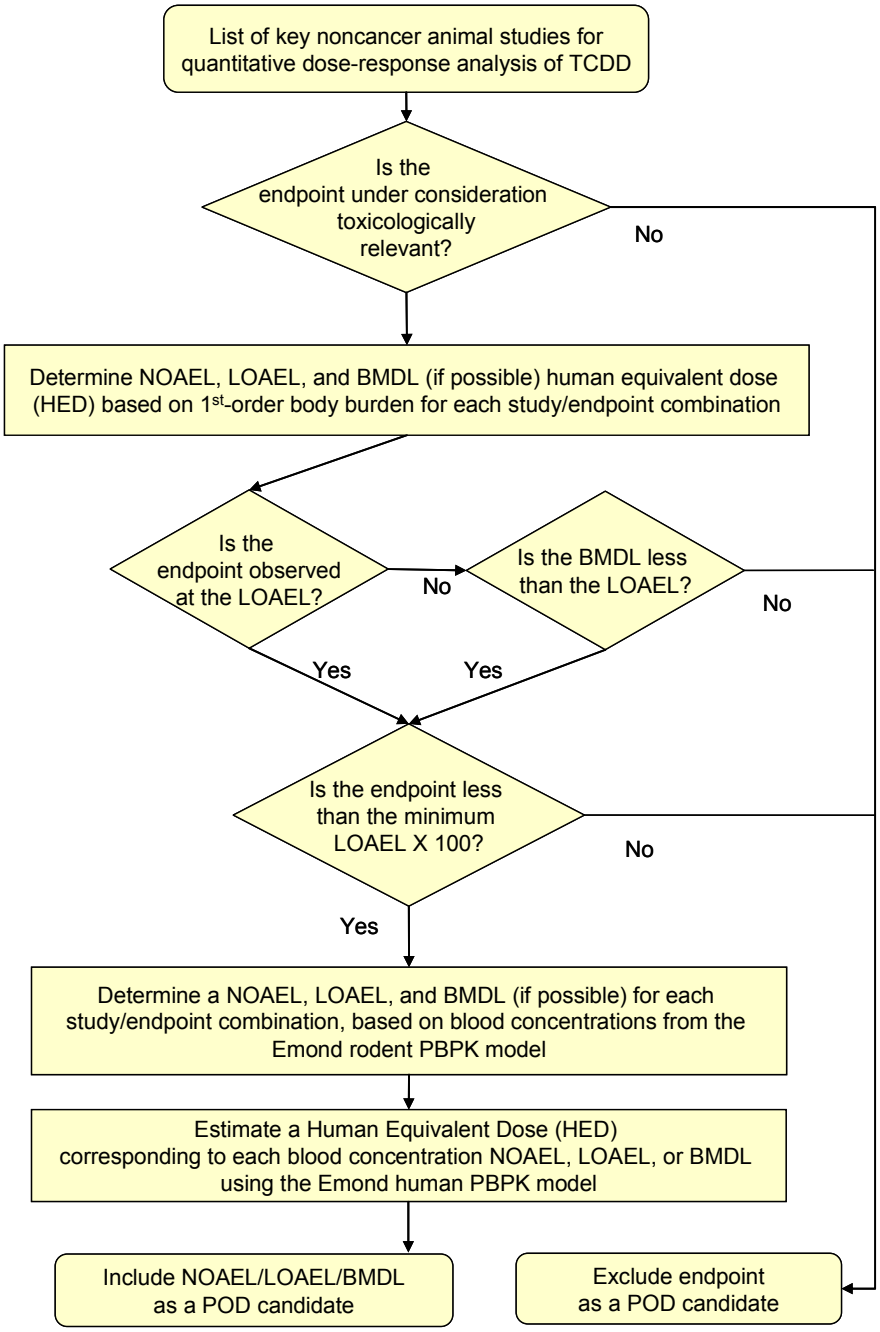


Figure 4-2. EPA's process to select and identify candidate PODs from key animal bioassays for use in noncancer dose-response analysis of TCDD. For each noncancer endpoint found in the studies that qualified for TCDD dose-response assessment using the study inclusion criteria, EPA attempted to estimate six candidate PODs, i.e., estimates of a NOAEL, LOAEL, and BMDL using administered average daily doses (ADD) and using blood concentrations from the animal kinetic model. Benchmark dose modeling was not always possible due to poor model fits. Then for all of the candidate PODs that were estimated, HEDs were estimated using the human kinetic model. Next, the toxicological relevance of each candidate POD was evaluated relative to its usefulness in human health risk assessment. Finally, the lowest group of the toxicologically relevant candidate PODs are selected for final use in derivation of an RfD.

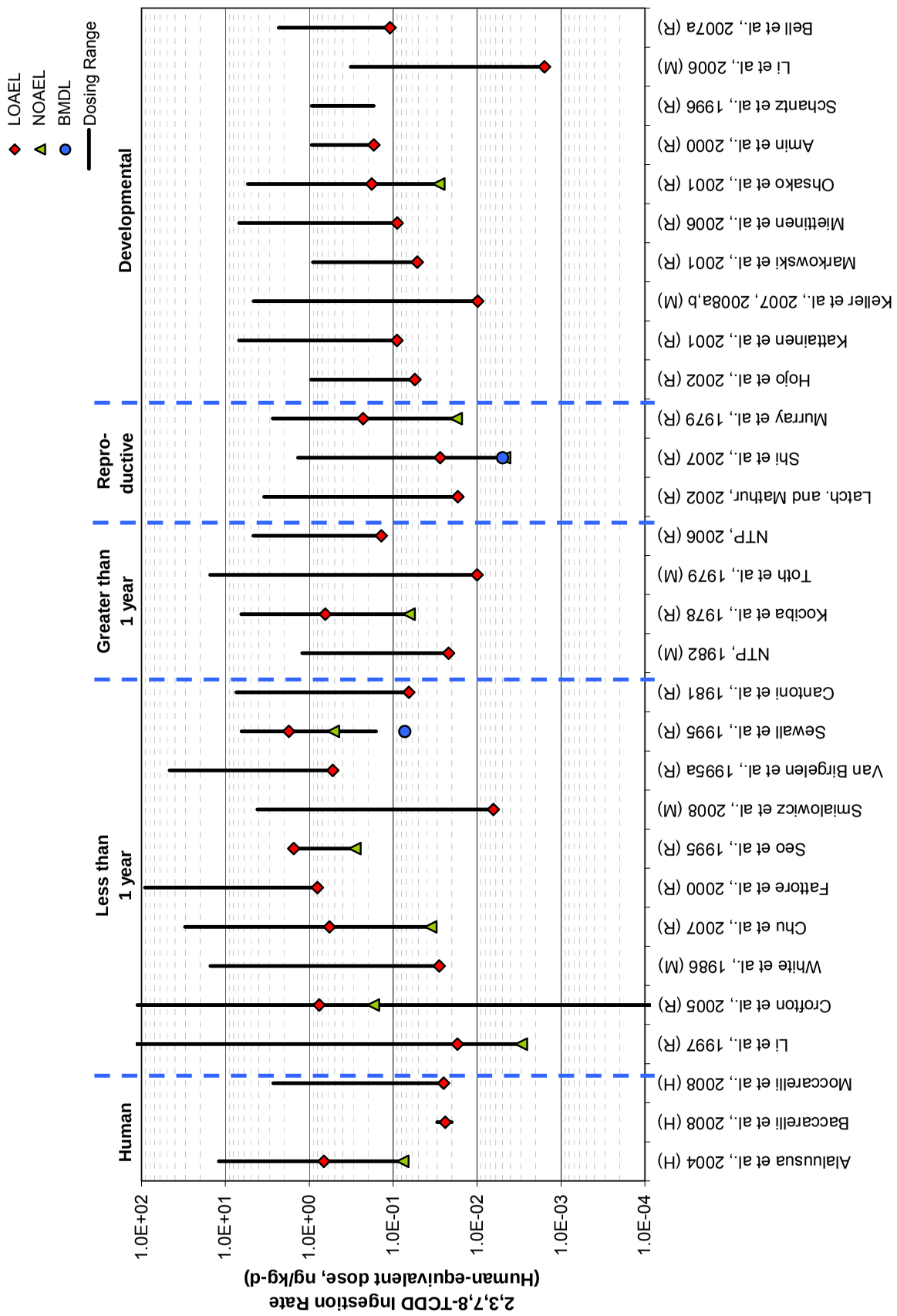


Figure 4-3. Exposure-response array for ingestion exposures to TCDD.

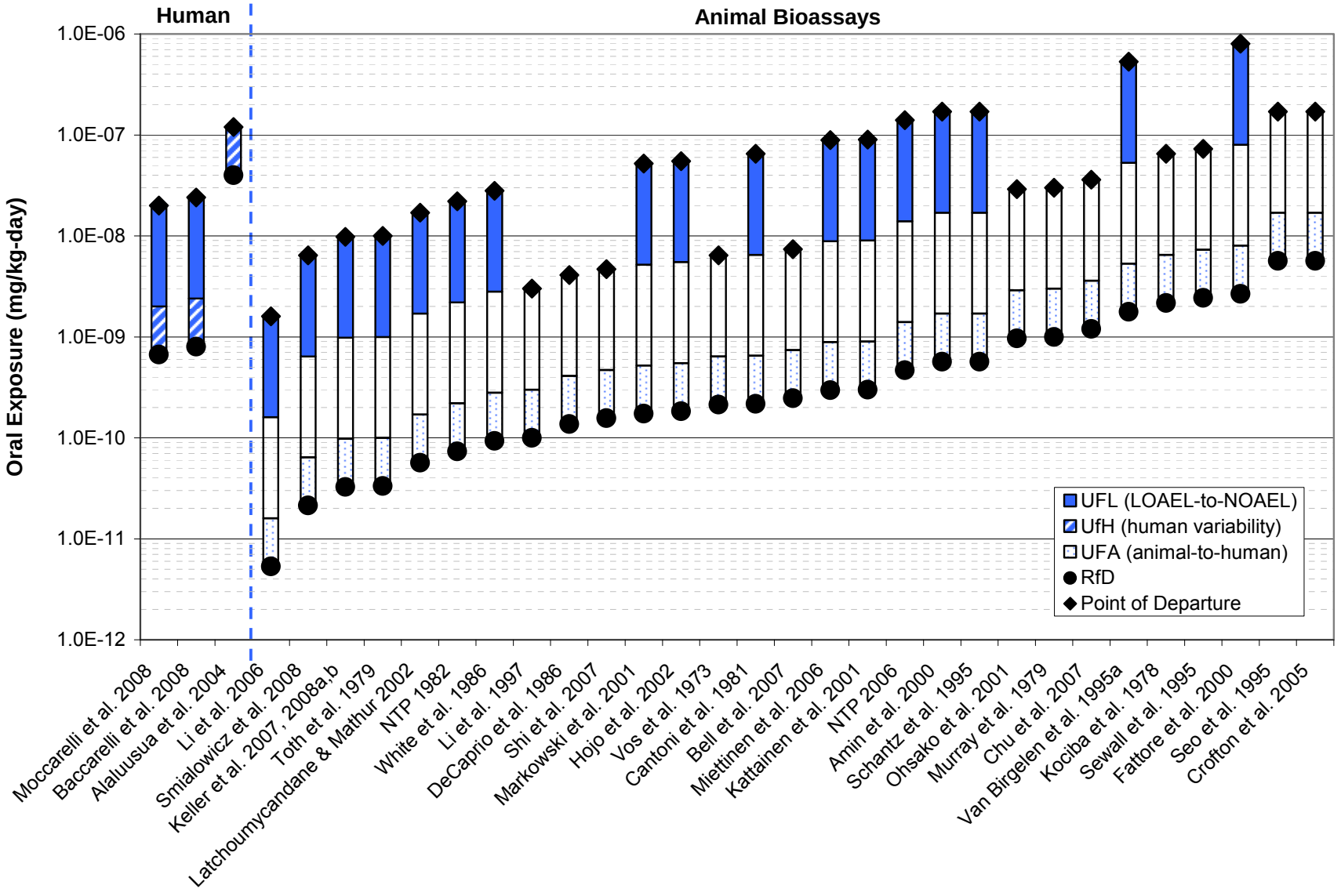


Figure 4-4. Candidate RfD array.

5. CANCER ASSESSMENT

5.1. QUALITATIVE WEIGHT-OF-EVIDENCE CARCINOGEN CLASSIFICATION FOR 2,3,7,8-TETRACHLORODIBENZO-*p*-DIOXIN (TCDD)

5.1.1. Summary of National Academy of Sciences (NAS) Comments on the Qualitative Weight-of-Evidence Carcinogen Classification for 2,3,7,8-Tetrachlorodibenzo-*p*-Dioxin (TCDD)

In its charge, the National Academy of Scinces (NAS) was requested to comment specifically on U.S. Environmental Protection Agency (EPA)'s conclusion that 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) is best characterized as "carcinogenic to humans." While indicating that distinction between the categories of "carcinogenic to humans" and "likely to be carcinogenic to humans" is "…based more on semantics than on science…" (NAS, 2006a, p. 141) and recommending that EPA "…spend its energies and resources on more carefully delineating the assumptions used in quantitative risk estimates for TCDD…" (NAS, 2006a, p. 141) rather than on the qualitative cancer descriptor for TCDD, the NAS provided the following comments:

> …the classification of dioxin as "carcinogenic to humans" versus "likely to be carcinogenic to humans" depends greatly on the definition and interpretation of the specific criteria used for classification, with the explicit recognition that the true weight of evidence lies on a continuum with no bright line that easily distinguishes between these two categories. The committee agreed that, although the weight of epidemiological evidence that dioxin is a human carcinogen is not strong, the human data available from occupational cohorts are consistent with a modest positive association between relatively high body burdens of dioxin and increased mortality from all cancers. Positive animal studies and mechanistic data provide additional support for classification of dioxin as a human carcinogen. However, the committee was split on whether the weight of evidence met all the necessary criteria described in the cancer guidelines for classification of dioxin as "carcinogenic to humans." EPA should summarize its rationale for concluding that dioxin satisfies the criteria set out in the most recent cancer guidelines for designation as either "carcinogenic to humans" or "likely to be carcinogenic to humans (NAS, 2006a, p. 140).
>
> If EPA continues to designate dioxin as "carcinogenic to humans," it should explain whether this conclusion reflects a finding that there is a strong association between dioxin exposure and human cancer or between dioxin exposure and a key precursor event of dioxin's mode of action (presumably AhR binding). If EPA's

finding reflects the latter association, EPA should explain why that end point (e.g., AhR binding) represents a "key precursor event (NAS, 2006a, p. 141).

5.1.2. EPA's Response to the NAS Comments on the Qualitative Weight-of-Evidence Carcinogen Classification for TCDD

A cancer descriptor is used to express the conclusion of the weight of evidence regarding the carcinogenic hazard potential of a compound. EPA agrees with the NAS committee that cancer descriptors represent points along a continuum of evidence. Relatedly, EPA acknowledges that there are gradations and borderline situations that cannot be communicated through a descriptor and are best clarified by a full weight of evidence narrative.

The 2003 Reassessment contains a detailed discussion of TCDD carcinogenicity in both humans (Part II, Chapter 7a; 8) and animals (Part II, Chapter 6; 8) as well as an overall summary of TCDD carcinogenicity (Part III, Chapter 2.2.1). Since the release of the 2003 Reassessment, the database pertaining to TCDD carcinogenicity has been strengthened and expanded by numerous publications (see U.S. EPA, 2008b), including a new chronic bioassay in female rats (NTP, 2006) and several new follow-up epidemiological investigations (see Section 2.4.1 and references therein). Many of these studies have been published subsequent to the NAS review. These new data are summarized and evaluated in Section 2.4 of this document.

As noted by the NAS, the 2003 Reassessment was released prior to EPA's publication of the U.S. EPA *Guidelines for Carcinogen Risk Assessment* ("2005 Cancer Guidelines"; U.S. EPA, 2005). Using EPA's guidance at the time of its release (U.S. EPA, 1996a), the 2003 Reassessment determined that the available evidence was sufficient to classify TCDD as a "human carcinogen." The 1996 guidance suggested "human carcinogen" to be an appropriate descriptor of carcinogenic potential when there is an absence of conclusive epidemiologic evidence to clearly establish a cause-and-effect relationship between human exposure and cancer, but there are compelling carcinogenicity data in animals and mechanistic information in animals and humans demonstrating similar modes of carcinogenic action.

The 2005 Cancer Guidelines (U.S. EPA, 2005) are intended to promote greater use of the increasing scientific understanding of the mechanisms that underlie the carcinogenic process. The 2005 Cancer Guidelines expand upon earlier guidance applied in the 2003 Reassessment and encourage the use of chemical- and site-specific data versus default options, the consideration of mode of action information and understanding of biological changes, fuller characterization of

carcinogenic potential, and consideration of differences in susceptibility. The 2005 Cancer Guidelines also emphasize the importance of weighing all of available evidence in reaching conclusions about the human carcinogenic potential of an agent. As noted above, additional information on TCDD carcinogenicity has been published since the release of the 2003 Reassessment. This information has expanded the TCDD database and provided additional support for conclusions made in the 2003 Reassessment regarding the carcinogenic potential of TCDD.

Under the 2005 Cancer Guidelines (U.S. EPA, 2005), TCDD is characterized as *carcinogenic to humans*, based on the available data as of 2009. The 2005 Cancer Guidelines indicate that this descriptor is appropriate when there is convincing epidemiologic evidence of a causal association between human exposure and cancer or when all of the following conditions are met (a) there is strong evidence of an association between human exposure and either cancer or the key precursor events of the agent's mode of action, but not enough for a causal association, and (b) there is extensive evidence of carcinogenicity in animals, and (c) the mode(s) of carcinogenic action and associated key precursor events have been identified in animals, and (d) there is strong evidence that the key precursor events that precede the cancer response in animals are anticipated to occur in humans and progress to tumors, based on available biological information.

As noted above, the NAS commented that EPA should "…explain whether this conclusion reflects a finding that there is a strong association between dioxin exposure and human cancer or between dioxin exposure and a key precursor event of dioxin's mode of action (presumably AhR binding)" (NAS, 2006a). When evaluating the carcinogenic potential of a compound, EPA employs a weight of evidence approach in which all available information is evaluated and considered in reaching a conclusion. The following sections provide a summary of EPA's weight of evidence evaluation for TCDD.

5.1.2.1. *Summary Evaluation of Epidemiologic Evidence of TCDD and Cancer*

The available occupational epidemiologic studies provide convincing evidence of an association between TCDD exposure and all cancer mortality. Among the strongest of these are the studies of over 5,000 U.S. chemical manufacturing workers (the National Institute for Occupational Safety and Health [NIOSH] cohort) (Fingerhut et al., 1991; Steenland et al., 1999,

2001; Aylward et al., 1996; Cheng et al., 2006; Collins et al., 2009); a study of nearly 2,500 German workers involved in the production of phenoxy herbicides and chlorophenols (the Hamburg cohort) (Becher et al., 1996, 1998; Manz et al., 1991; Nagel et al., 1994; Flesch-Janys et al., 1995, 1998); a study of more than 2,000 Dutch workers in two plants involved in the synthesis and formulation of phenoxy herbicides and chlorophenols (the Dutch cohort) (Bueno de Mesquita et al., 1993; Hooiveld et al., 1998); a smaller study of roughly 250 workers involved in a chemical accident cleanup (the BASF cohort) (Thiess et al., 1982; Zober et al., 1990; Ott and Zober, 1996); and an international study of more than 18,000 workers exposed to phenoxy herbicides and chlorophenols (Saracci et al., 1991; Kogevinas et al., 1997) including newer studies of smaller subsets of these workers (t'Mannetje et al., 2005; McBride et al., 2009a, b). The findings from these studies have been thoroughly described either in the 2003 Reassessment or in Section 2.4.1 of this document.

As noted in Section 2.4, there are considerable challenges inherent in addressing potential sources of confounding from smoking and co-exposure to other carcinogens, (which could produce inflated or spurious associations), the healthy worker effect, (which could result in attenuated effects through comparison with a referent background with an inappropriately high background risk), and quantifying exposure to the populations included in many of these retrospective studies. The more recent studies of these cohorts have made significant advances in reducing the potential for bias from the healthy worker effect through use of internal cohort analyses and/or controlling for potential confounders through statistical adjustment, restriction, and use of internal comparisons. Although some exposure assessment uncertainties remain, some of these studies have also collected individual-level TCDD exposure estimates that allow quantification of effective dose necessary for dose-response modeling. Overall, the occupational data provide consistent support for an association between exposure to TCDD and increased cancer mortality.

Additional epidemiologic evidence supporting an association between TCDD exposure and cancer comes from studies investigating the morbidity and mortality of residents exposed to TCDD following an accidental release from a chemical plant near Seveso, Italy (the Seveso cohort) (Bertazzi et al., 1989, 1993, 1997, 2001; Pesatori et al., 1998, 2003; Consonni et al., 2008; Warner et al., 2002). Pesatori et al. (2003) and Consonni et al. (2008) were not available at the time the 2003 Reassessment was released. Among individuals with relatively high

exposure at Seveso (Zones A and B combined), all-cancer mortality in the 20-year post-accident period and all-cancer incidence in the 15-year post-accident period failed to exhibit significant departures from the expected (Bertazzi et al., 2001; Pesatori et al., 2003; Consonni et al., 2008). However, an increased risk of all-cancer mortality was noted among men 15−20 years after first exposure; not only is the association similar in magnitude to other studies (relative risk [RR] = 1.3; 95% confidence interval [CI] = 1.0−1.7) but also emphasizes the importance of consideration of latency (Bertazzi et al., 2001). Furthermore, associations between TCDD and some specific cancer sites were detected in this cohort, including increased incidence (based on 15 years of follow-up) and mortality (based on 20 years follow-up) from lymphatic and hematopoietic neoplasms in both males and females from Zones A and B (Consonni et al., 2008). This excess was primarily due to non-Hodgkin's lymphoma. Additionally, there was an increase in lung and rectal cancer mortality in men (Bertazzi et al., 2001) and limited evidence of increased liver cancer incidence in women based on the 15-year follow-up study (Bertazzi et al., 1993). In a separate analysis of 981 women in Zone A, breast cancer incidence (n = 15) was associated (a 2-fold increase for a 10-fold increase in serum TCDD) with TCDD measurements first collected in 1976 and 1977 (Warner et al., 2002). The authors also reported a 2−3-fold increase in all cancer incidence (n = 21) for the two upper quartiles of TCDD exposure.

Overall, the newer studies of the Seveso cohort have reported significant increases in cancer incidence and elevations in cancer mortality that were not evident in earlier studies of this cohort. While these studies demonstrate an association between TCDD exposure and different types of cancer, one of the main limitations is the small number of cancer cases to assess site-specific associations with TCDD exposure. Ongoing studies in that cohort should help further elucidate potential risk for specific cancer types (and other endpoints) associated with TCDD exposures among this population.

5.1.2.1.1. *Evidence for causality.*

The evidence for causality for cancer from the human studies is briefly summarized in the paragraphs that follow and is based on recommendations from the 2005 Cancer Guidelines. It should be noted that there are methodological limitations of the epidemiologic studies that may temper some of the conclusions regarding causality. These limitations include limited statistical power, exposure assessment uncertainty, and lack of control of confounders (e.g., dioxin-like

compounds and smoking) in some studies. There is additional uncertainty regarding lack of organ specificity in TCDD associated cancers, as the most consistent results occurred for all-cancer mortality. Despite these uncertainties, many of the more recent studies have greatly improved exposure assessments compared to earlier studies of the same cohorts and have addressed the potential for confounding and other types of biases.

Temporality—exposure must precede the effect for causal inference. Given the long induction period for many types of cancers, exposure must precede the effect with a sufficient latency (i.e., typically 15−20 years for environmental carcinogens). In all the occupational studies reviewed (with the exception of McBride et al., 2009a), TCDD exposure has preceded the effect with sufficient latency to be considered causally associated. In the studies of the Seveso cohort, the follow-up exposure period has now reached 20 years, a latency sufficient to address some carcinogenic endpoints. Since most of the studies are based on occupational exposures or accidental releases into the environment, temporality is more readily established due to the obvious determination of the specific exposure windows prior to disease onset.

Strength of Association—refers to the magnitude of measures of association such as the ratio of incidence or mortality (e.g., standardized mortality ratio [SMRs], standardized incidence ratios, RRs, or odds ratios) irrespective of statistical significance. Effect estimates that are large in magnitude are less likely to be due to chance, bias, or confounding. Reports of modest risk, however, do not preclude a causal association and may reflect an agent of lower potency, lower levels of exposure or attenuation due to nondifferential exposure misclassification. The four occupational cohorts with the highest exposures (NIOSH, Hamburg, Dutch, and BASF) consistently showed statistically significant, although moderate, elevations in cancer mortality. When the data were combined, the SMR for all four subcohorts was 1.4 [95% CI = 1.2−1.6] (IARC, 1997). Based on findings from the International Agency for Research on Cancer (IARC) Working Group, increases in all cancer (combined) mortality of the magnitude reported for TCDD have rarely been found in occupational cohort studies (IARC, 1997). Although these estimates are higher than the all-cancer mortality results among Seveso men (RR = 1.1; 95% CI = 1.0−1.3), they are comparable to the risk estimated in this population (RR = 1.3; 95% CI = 1.0-1.7) 15−20 years after first exposure. These consistent results comparable in magnitude from the occupational cohorts and Seveso population are not likely due to chance.

The occupational cohort studies also show an increased risk for lung cancer in the previously mentioned four subcohorts. The relative risk for lung cancer in the combined highly exposed subcohorts was estimated to be 1.4 (95% CI = 1.1−1.7) (IARC, 1997). This is consistent with the lung cancer mortality findings for the highest exposed group of men in Seveso (RR = 1.3; 95% CI = 1.0−1.7). Additionally, there was an increase in rectal cancer mortality in the Seveso cohort (RR = 2.4; 95% CI = 1.2−4.6) (Bertazzi et al., 2001) with a corresponding increase in incidence. Consistent relative risks of more than two were also detected for rectal cancer in the Hamburg and New Zealand cohorts, but increased risks were not found in the other cohorts. Although there was limited evidence of increased incidence or morality from hepatobiliary cancers across the cohorts, liver cancer incidence was elevated in the 15-year postaccident period among women in the Seveso cohort (RR = 2.4; 95% CI = 1.1−5.1, Warner et al., 2002). An association in this population was also detected for between breast cancer incidence (RR = 2.1; 95% CI = 1.0–4.6) and serum TCDD levels (per a 10-fold increase in serum TCDD). Although findings were based on small numbers, three- and four-fold increased risks of soft tissue sarcoma were detected among the NIOSH (Collins et al., 2009) and New Zealand cohorts (McBride et al., 2009a). No other cases of this very rare cancer were detected in the exposed populations from the other cohorts.

Elevated risk of lymphohemopoietic cancer mortality was noted among the Seveso cohort (RR = 1.7; 95% CI = 1.2, 2.5) (Consonni et al., 2008). Increased SMRs for lymphohemopoietic cancer comparable in magnitude (range: 1.6−2.2) were also detected among the Hamburg and New Zealand occupational cohorts, but limited evidence (range: 1.0 to 1.2) of increased mortality was found in the BASF, NIOSH and Ranch Hands employees (Ott and Zober, 1996; Steenland et al., 1999; Akhtar et al., 2004). Most of the lymphohemopoietic cancer mortality risk was reportedly due to non-Hodgkin's lymphoma in most of the cohorts. Relative risks for non-Hodgkin's lymphoma among TCDD exposed populations from the NIOSH, Hamburg, New Zealand, Dutch, and Seveso cohorts ranged from 1.2 to 3.8. Although statistical power was limited in most of these studies, relative risks exceeded 3.0 for non-Hodgkin's lymphoma in three of these cohorts (Flesch-Janys et al., 1998; Consonni et al., 2008; Hooiveld et al., 1998).

Consistency—the observation of the same site-specific effect across several independent study populations strengthens an inference of causality. Despite differences across occupational cohorts, most studies have consistently reported increases in all-cancer mortality with TCDD

exposure. Several of these studies have also reported increases in lung cancer related to TCDD exposure. As noted above, there is also suggestive evidence of an increased risk in all-cancer and lung cancer mortality among the Seveso cohort consistent in magnitude to the occupational cohorts. Elevated risk of lymphohemopoietic cancer mortality consistent in magnitude (range: 1.6−2.2) was also detected among the Seveso, Hamburg and New Zealand cohorts. An increased risk for non-Hodgkin's lymphoma was found in two of the occupational cohorts as well as in the Seveso cohort, although the relative risks largely did not achieve statistical significance. Among those studies detecting an association, consistent two-fold relative risks were found for rectal cancer (McBride et al., 2009a; Flesch-Janys et al., 1998; Bertazzi et al., 2001) and relative risks in excess of three were detected for soft tissue sarcoma (McBride et al., 2009a; Collins et al., 2009).

Biological Gradient—refers to the presence of a dose-response and/or duration-response between a health outcome and exposure of interest. Several of the occupational cohort studies (Ott and Zober, 1996; Flesch-Janys et al., 1998; Manz et al., 1991; Steenland et al., 1999; Michalek and Pavuk, 2008) found evidence of a dose-response relationship for all cancers and various TCDD exposure measures. The SMR analyses based on internal comparisons within the occupational cohorts show a biological gradient by comparing highly TCDD exposed workers to low or unexposed workers. A biological gradient was also demonstrated in the Seveso cohort by comparing highly exposed individuals (Zones A and B) to individuals in lower exposure zones (Zones C and R). Warner et al. (2002) also reported evidence of a dose-response trend for breast cancer and increasing TCDD exposures.

Biological Plausibility—refers to the observed effect having some biological link to the exposure. Most evidence suggests that toxic effects of TCDD are mediated by interaction with the aryl hydrocarbon receptor (AhR). AhR is a highly conserved protein among mammals, including humans (Fujii-Kuriyama et al., 1995; Harper et al., 2002; Nebert et al., 1991). Several hypothesized modes of action have been presented for TCDD-induced tumors in rodents, all involving AhR activation. The available evidence does not preclude the relevance of these hypothesized modes of action to humans.

Specificity—as originally intended, this refers to increased inference of causation if a single site effect, as opposed to multiple effects, is observed and associated with exposure. Based on current biological understanding, this is now considered one of the weaker guidelines

for causality. As stated in the 2005 Cancer Guidelines, given the current understanding that many agents cause cancer at multiple sites, and cancers have multiple causes, the absence of specificity does not detract from evidence for a causal effect. Given that the most consistent findings associating TCDD and cancer are for all-cause cancer mortality, epidemiological evidence suggests that TCDD lacks specificity for particular tumor sites. A key event in TCDD's mechanism of action is binding to and activating AhR; however, downstream events leading to tumor formation are uncertain and may likely be tissue specific. Given that the AhR is highly conserved within species and is expressed in various human tissues, the lack of tumor site specificity does not preclude a determination of causality.

In summary, EPA finds the available epidemiological information provides strong evidence of an association between TCDD exposure and human cancer that cannot be reasonably attributed to chance or confounding and other types of bias, and with a demonstration of temporality, strength of association, consistency, biological plausibility, and a biological gradient. Additional evidence from animal studies and from mechanistic studies (described below) provides additional support for the classification of TCDD as carcinogenic to humans.

5.1.2.2. *Summary of Evidence for TCDD Carcinogenicity in Experimental Animals*

An extensive database on the carcinogenicity of TCDD in experimental animals is described in detail in Part II, Chapter 6 of the 2003 Reassessment. There is substantial evidence that TCDD is carcinogenic in experimental animals based on long-term bioassays conducted in both sexes of rats and mice (Kociba et al., 1978; NTP, 1982, 2006) and in male hamsters (Rao et al., 1988). Additionally, National Toxicology Program (NTP, 2006) has completed a new chronic bioassay in female Sprague Dawley rats. These studies are summarized in Section 2.4.2 of this document. All studies have produced positive results, with TCDD increasing the incidence of tumors at sites distant from the site of treatment and at doses well below the maximum tolerated dose. In both sexes of rodents, when administered by different routes and at low doses, TCDD caused tumors at multiple sites; tumors were observed in liver, lung, lymphatic system, soft tissue, nasal turbinates, hard palate, thyroid, adrenal, pancreas, and tongue. The most consistent and best characterized carcinogenic responses to TCDD are in the rodent liver, lung, and thyroid (discussed below in Section 5.1.2.3).

5.1.2.3. *TCDD Mode of Action*

The 2005 Cancer Guidelines defines the term "mode of action" as "a sequence of key events and processes, starting with interaction of an agent with a cell, proceeding through operational and anatomical changes, and resulting in cancer formation." A "key event" is an empirically observable precursor step that is itself a necessary element of the mode of action or is a biologically based marker for such an element. Mode of action is contrasted with "mechanism of action," which implies a more detailed understanding and description of events, often at the molecular level. In the case of TCDD, the terms 'mechanism of action' and 'mode of action' are often used interchangeably in the scientific literature in reference to TCDD's interaction with the AhR. A thorough discussion of TCDD's interaction with the AhR can be found in the 2003 Reassessment (Part II, Chapter 2; Part III, Chapter 3), and is summarized below (see Section 5.1.2.3.1).

Most evidence suggests that the majority of toxic effects of TCDD are mediated by interaction with the AhR. EPA considers interaction with the AhR to be a necessary, but not sufficient, event in TCDD carcinogenesis. The sequence of key events following binding of TCDD to the AhR and that ultimately leads to the development of cancer is unknown. Therefore, in the strictest sense, TCDD's interaction with the AhR does not constitute a "mode of action" as defined by the 2005 Cancer Guidelines because information about the progression of necessary events is lacking. However, AhR binding and activation by TCDD is considered to be a key event in TCDD carcinogenesis.

5.1.2.3.1. *The aryl hydrocarbon receptor (AhR).*

While substantial evidence suggests that most toxic effects of TCDD are mediated by interaction with the AhR, less is known about the complex responses that result in tumor formation. Nonetheless, a picture is emerging wherein TCDD is considered a "receptor-mediated carcinogen" in laboratory animals (see Figure 5-1), acting in a manner similar to peroxisome proliferators, phorbol esters, or estrogen (Woods et al., 2007).

The mechanism of action of TCDD has been extensively studied. TCDD activates the AhR, a member of the basic helix-loop-helix, Per-Arnt-Sim (bHLH-PAS) family of transcription factors. AhR is present in most cell types and in the unactivated state is cytosolic and exists in a complex with chaperone proteins, such as heat shock protein 90 (Hsp90). Binding of TCDD to

AhR leads to nuclear translocation and heterodimerization with its partner protein Arnt, another bHLH-PAS family member. The AhR:Arnt heterodimer binds to specific cognate DNA sequence elements known as dioxin/xenobiotic response elements (DRE/XRE) present in the regulatory region of specific genes. Binding of the AhR:Arnt heterodimer to these elements, and subsequent recruitment of transcription coactivator complexes, leads to increased transcription of specific genes, known as "target genes." There is a battery of genes affected in this manner and targets include certain xenobiotic-metabolizing enzymes, such as cytochrome P450 (CYP)1A1, CYP1A2, CYP2B1, and UDP-glucuronosyltransferase (UGT)1A6 (reviewed in Schwartz and Appel [2005]). In addition, genes affected by the TCDD/AhR-complex code for both inhibitory and stimulatory growth factors and their gene products affect cellular growth, differentiation and homeostasis and have been shown to contribute to carcinogenicity as well as other forms of toxicity (reviewed in Popp et al. [2006]).

Detailed molecular biology research has been performed to identity the extent of the genes regulated by AhR (Woods et al., 2007); however a complex and still ill-defined profile remains. Additionally, it is important to note that the extent of the response of individual genes may not correlate with site-specific tumorigenicity. For example, while TCDD is ineffective as a tumor promoter in ovariectomized rats and does not stimulate liver cell proliferation in these animals, it is still capable of inducing CYP1A2 in roughly the same magnitude as in the intact female rats (Lucier et al., 1991). Similarly, CYP1A1 induction by TCDD is very similar in male and female rats even though males are almost completely resistant to TCDD carcinogenicity (Wyde et al., 2002).

Some of AhR's effects on gene expression may be the result of interaction with other transcription factors (such as the retinoblastoma protein [Ge and Elferink, 1998], NF-κB [Tian et al., 1999] or with the tyrosine kinase c-Src (Blankenship and Matsumura, 1997]) rather than via direct interaction with DNA. By far the most extensive studies involving cross talk between AhR and another transcription factor are those involving the estrogen receptor alpha (ERα). The anti-estrogenic properties of TCDD have been well-documented, beginning with the observations that TCDD repressed estradiol function in rat uterus and liver. The AhR-ERα cross-talk can be manifested at several levels including direct interaction, association of the receptors with the other's response element and altered metabolism of estradiol by AhR ligand (Takemoto et al., 2004). The interactions between AhR/Arnt- and estrogen receptor-dependent

signaling pathways, which mediate anti-estrogenic effects of dioxins and dioxin-like polychlorinated biphenyls (PCBs; Bock, 1994), is probably causal for the well-documented gender-specificity of the carcinogenic effects of these agents (e.g., hepatocarcinogenicity of TCCD in female as opposed to male rats) (Lucier et al., 1991). In addition, cross-talk between AhR/Arnt and other nuclear receptors, their coactivators, and corepressors, has been described. In fact, cross-talk has been reported for AhR and numerous signaling pathways involved in a broad range of physiological processes. The molecular mechanisms by which the AhR interferes with these signaling networks are multifaceted and occur at multiple levels of regulation (many beyond transcriptional control) (Haarmann-Stemmann et al., 2009).

Pertinent to human risk assessment, there are wide inter- and intraspecies differences in the toxicological responses to TCDD (Ema et al., 1994; Poland and Glover, 1990; Poland et al., 1994) some of which can be explained by polymorphisms in AhR. For instance there is a 10-fold difference in susceptibility to TCDD-induced toxicity between the TCDD-sensitive C57BL/6 and the resistant DBA/2 strains of mice (Poland and Glover, 1980) that can be explained by polymorphic variations in the ligand-binding domain and in the C-terminal region of the AhR molecule of each strain (Dolwick et al., 1993). Depending on the system examined, the estimated affinity of binding of TCDD (and related compounds) to the human AhR is about 10-fold lower than that observed to the AhR from "responsive" rodent species and is comparable to that observed to the AhR from "nonresponsive" mouse strains (Ramadoss and Perdew, 2004). This reduced affinity is due, in part, to a single amino acid substitution within the ligand binding domain of the human and "nonresponsive" mouse AhRs (Ramadoss and Perdew, 2004). However, there is considerable tissue and species variability in response to TCDD that cannot be ascribed solely to polymorphisms of the AhR gene (Pohjanvirta and Tuomisto, 1994; Geyer et al., 1997). Although the affinity of binding of TCDD and related compounds to the human AhR is reduced compared with rodent AhRs, the qualitative and quantitative rank-order potency of these chemicals is similar.

5.1.2.3.1.1. ***Other AhR considerations.***

In addition to the potent agonist TCDD, there are many other exogenous ligands for the AhR, including certain polycyclic aromatic hydrocarbons, polychlorinated dibenzofurans, and PCBs (Bock, 1994). Several natural and endogenous compounds are also regulators of AhR

(Chiaro et al., 2008). The classes of endogenous compounds that have been shown to induce CYP1 and/or activate AhR include: (a) tryptophan metabolites, other indole-containing molecules, and phenylethylamines (Gielen and Nebert, 1971); (b) tetrapyrroles such as bilirubin and biliverdin; (c) sterols such as 7-ketocholesterol and the horse steroid equilenin; (d) fatty acid metabolites, including at least six different prostaglandins (Seidel et al., 2001) and lipoxin A4; and (e) the ubiquitous second messenger cAMP (reviewed in McMillan and Bradfield [2007] and Barouki et al. [2007]). The existence of multiple ligands with varying affinity and responses suggests that "selective receptor modulators" (or SRMs) of the AhR exist. SRMs are ligands for a receptor that, upon binding, elicit a conformational change in the receptor that results in differential recruitment of coregulatory molecules to the target gene promoter region, thereby imparting a different biological activity relative to the prototypical ligand. This phenomenon has been most studied for nuclear receptors such as the ERα with the classic example being tamoxifen, which has estrogen-like activity in the uterus but anti-estrogen-like effects in the breast. Thus, the relative abilities of compounds to stimulate gene expression or other effects vary in promoter- and cell type-specific manners. It is now apparent that SRMs exist for the AhR as well (SAhRMs, Fretland et al. [2004]). For example, 6-methyl-1,3,8-trichlorodibenzofuran (6-MCDF), a SAhRM whose structure is similar to that of the prototypical AhR agonist TCDD, can induce CYP1A1 gene expression in liver but does not lead to the toxic responses associated with TCDD (Fritz et al., 2009).

It has been demonstrated that AhR-deficient (AhR-/-) mice show no signs of toxicity at doses of TCDD approximating the lethal dose eliciting 50% response (LD_{50}) dose (200 μg/kg) in mice expressing the AhR (Fernandez-Salguero et al., 1996). However, a single high exposure of 2,000 μg/kg to AhR-deficient mice produced several minor lesions including scattered necrosis and vasculitis in the liver and lungs. These data suggest that a pathway leading to toxicity exists, albeit at very high doses, that is independent of the AhR. However, these data also indicate that, at least in mice, the major in vivo effects of TCDD are mediated through the AhR. The finding of carcinogenicity in hamsters (Rao et al., 1988) is of special interest since hamsters have been found to be relatively resistant to the lethal effects of TCDD (Henck et al., 1981; Olson et al., 1980). To date, there have been no chronic bioassay studies of TCDD carcinogenicity in AhR-deficient transgenic animals.

There are additional insights into the complexity of TCDD's mechanism of action involving AhR. Some biochemical responses to TCDD treatment in isolated cells have been reported in cells lacking Arnt, in cells expressing a mutated Arnt protein and in cells with highly reduced levels of AhR (Puga et al., 1992; Kolluri et al., 1999), implying either a non nuclear role of the AhR in mediating these events or an AhR-independent process.

Additionally, recent studies have linked AhR activation in the absence of exogenous ligand to a multitude of biological effects, ranging from control of mammary tumorigenesis to regulation of autoimmunity (reviewed in Hahn et al., 2009). Finally, constitutively activated AhR in rodents has been shown to induce stomach tumors (Andersson et al., 2002). This indicates that AhR activation alone (i.e., in the absence of ligand) is sufficient to induce tumors.

5.1.2.3.2. *TCDD as a tumor promoter.*

The role of TCDD as a tumor promoter is discussed in the 2003 Reassessment (Part II, Chapter 6). The following is a brief summary of the information regarding TCDD as a tumor promoter.

TCDD is typically designated as a nongenotoxic carcinogen because it does not damage DNA directly through the formation of DNA adducts, is negative in most short-term assays for genotoxicity, and is a potent tumor promoter and a weak initiator or noninitiator in multistage models for chemical carcinogenesis (Pitot et al., 1980; Graham et al., 1988; Lucier et al., 1991; Clark et al., 1991; Flodstrom and Ahlborg, 1991; Poland et al., 1982). However, mechanisms have been proposed that support the possibility that TCDD might be indirectly genotoxic, either through the induction of oxidative stress or by altering the DNA-damaging potential of exogenous and endogenous compounds, such as estrogens. In addition, there have been numerous reports demonstrating TCDD-induced modifications of growth factor signaling pathways and cytokines in experimental animals and cell culture systems. Some of the altered signaling pathways include those for epidermal growth factor, transforming growth factor alpha, glucocorticoids, estrogen, tumor necrosis factor-alpha, interleukin 1-beta, plasminogen inactivating factor-2, and gastrin. Many of these pathways are involved in cell homeostasis, proliferation, and differentiation and provide plausible mechanisms responsible for the carcinogenic actions of TCDD.

5.1.2.3.3. *Hypothesized modes of action of TCDD in rodents.*

TCDD has been shown to consistently induce multiple tumors in both sexes in several rodent species. These tumors are observed in various tissues, including (but not limited to): liver, lung, thyroid, lymphatic system, soft tissue, nasal turbinates, hard palate, adrenal, pancreas, and tongue. While the mode of action of TCDD in producing cancer has not been elucidated for any tumor type, the best characterized carcinogenic actions of TCDD are in rodent liver, lung, and thyroid. The hypothesized mode of action for each of these three tumor types is briefly discussed below and is described in Figure 5-2. The hypothesized sequence of events following TCDD interaction with the AhR is markedly different for each of these three tumor types. No detailed hypothesized mode of action information exists for any of the other reported tumor types. Further, no single definitive mode of action of TCDD-mediated carcinogenicity has been identified.

5.1.2.3.3.1. ***Liver tumors.***

The mode of action of TCDD in producing liver cancer in rodents has not been elucidated. One hypothesized mode of carcinogenic action of TCDD in the liver is mediated through hepatotoxicity. Generically speaking, TCDD activation of the AhR leads to a variety of changes in gene expression, which then lead to hepatotoxicity, followed by compensatory regenerative cellular proliferation and subsequent tumor development. The details of the mechanism of TCDD-induced hepatotoxicity have not been fully determined but both CYP induction and oxidative stress have been postulated to be involved (Maronpot et al., 1993; Viluksela et al., 2000). The enhanced cell proliferation arising from either altered gene expression or hepatotoxicity, or both, could be the principal factor leading to promotion of hepatocellular tumors (Whysner and Williams, 1996). The sensitivity of female rat liver to TCDD, which apparently does not extend to the mouse, depends on ovarian hormones (Lucier et al., 1991; Wyde et al., 2001). This sensitivity has been ascribed to induction of estradiol metabolizing enzymes (Graham et al., 1988) and is hypothesized to lead either to generation of reactive metabolites of endogenous estrogen or to active oxygen species of estrogens. Oxidative DNA damage has been implicated in liver tumor promotion (Umemura et al., 1999).

A dose-response relationship exists for TCDD-mediated hepatotoxicity, and this parallels the dose-response relationship for tumor formation (or formation of foci of cellular alteration as a

surrogate of tumor formation). However, the dose-response relationship for other TCCD-induced responses such as enhanced gene expression is different from the dose-response for tumor formation in terms of both efficacy and potency (see Popp et al. [2006] for review). It is important to note that differences in potency between events (i.e., gene expression versus cell proliferation) does not necessary imply alternative mechanisms of action.

5.1.2.3.3.2. *Lung tumors.*

The mode of action of TCDD in producing lung cancer in rodents (predominantly keratinizing squamous cell carcinoma, [Larsen, 2006]) has not been elucidated. One hypothesized mechanism of the carcinogenic action of TCDD in the lung involves disruption of retinoid homeostasis in the liver. Retinoic acids and their corresponding nuclear receptors, the retinoic acid receptors (RARs) and the retinoid X receptors (RXRs), work together to regulate cell growth, differentiation, and apoptosis. It is hypothesized that TCDD, through activation of the AhR, can affect parts of the complex retinoid system and/or other signaling systems regulated by, and/or cross-talking with, the retinoid system (reviewed in Nilsson and Hakansson, 2002). These effects are then hypothesized to lead to lung tumor development; however, the mechanisms underlying this hypothesis are not well-defined. Pulmonary squamous proliferative lesions have been reported following oral exposure to TCDD in rats (Tritscher et al., 2000). In general, squamous metaplasia with some inflammation is associated with significant forms of injury via inhalation of toxic compounds but is also seen with vitamin A deficiency (Tritscher et al., 2000) and gives some credence to this hypothesis.

Another hypothesized mechanism for the carcinogenic action of TCDD in the lung is through induction of metabolic enzymes. Through activation of AhR and subsequent induction of metabolizing enzymes (such as CYP1A1), TCDD may enhance bioactivation of other carcinogens in lung (Tritscher et al., 2000). There have been few studies to support this hypothesis; however, in a long-term continuous-application study of carcinogenesis using airborne particulate extract (APE), squamous cell carcinoma occurred in 8 of 17 AhR+/+ mice (47%) while no tumors were found in AhR-/- mice (Matsumoto et al., 2007). In addition CYP1A1 was induced in AhR+/+ mice but not in AhR-/- mice in this study. These results suggest that AhR plays a significant role in APE-induced carcinogenesis in AhR+/+ mice and

CYP1A1 activation of carcinogenic polycyclic aromatic hydrocarbons (the primary carcinogenic component of APE) is also of importance.

5.1.2.3.3.3. ***Thyroid tumors.***

The mode of action of TCDD in producing thyroid cancer in rodents has not been elucidated. It is hypothesized that TCDD increases the incidence of thyroid tumors through an extrathyroidal mechanism. The prevailing hypothesis for the induction of thyroid tumors by TCDD involves the disruption of thyroid hormone homeostasis via induction of Phase II enzymes UGTs in the liver (reviewed in Brouwer et al., 1998) by an AhR-dependent transcriptional mechanism (Bock et al., 1998; Nebert et al., 1990). This induction of hepatic UGT results in increased conjugation and elimination of thyroxine (T4), leading to reduced serum T4 concentrations. T4 synthesis is controlled by the thyroid stimulating hormone (TSH) which is under negative and positive regulation from the hypothalamus, pituitary, and thyroid via thyrotrophin-releasing hormone, TSH, T4, and triiodothyronine. Consequently, the reduced serum T4 concentrations lead to a decrease in the negative feedback inhibition on the pituitary gland. This would then lead to a rise in secreted TSH and stimulation of the thyroid. The persistent induction of UGT by TCDD and the subsequent prolonged stimulation of the thyroid could result in thyroid follicular cell hyperplasia and hypertrophy of the thyroid, thereby increasing the risk of progression to neoplasia. Increases in blood TSH levels are consistent with prolonged stimulation of the thyroid and may represent an early stage in the induction of thyroid tumors identified in animal bioassays. Statistically significant increases in neonatal blood TSH levels have been recently been reported in children born to TCDD-exposed mothers in the Seveso cohort (Baccarelli et al., 2008; discussed in Section 2.4.1.1.1.4.4). Support for this hypothesis comes from several studies showing that TCDD decreases serum total thyroxine and free thyroxine concentrations in rats following both single dose and repeated dose exposures (Bastomsky, 1977; Brouwer et al., 1998; Pohjanvirta et al., 1989; Potter et al., 1983, 1986; Sewall et al., 1995; Van Birgelen et al., 1995). Further support comes from studies of transgenic animals in which TCDD exposure resulted in a marked reduction of total thyroxin and free T4 levels in the serum of AhR+/- mice but not AhR-/- mice (Nishimura et al., 2005). Additionally, gene expression of UGT1A6, CYP1A1, and CYP1A2 in the liver was markedly induced by TCDD in AhR+/- but not AhR-/- mice (Nishimura et al., 2005).

5.1.2.3.4. *Summary of TCDD mode of action in rodents.*

Overall, there are inadequate data to support the conclusion that any of the particular mode of action hypotheses described above is operant in TCDD-induced carcinogenesis. However, the wealth of scientific evidence available indicates that most, if not all, of the biological and toxic effects of TCDD are mediated by the AhR. Although the receptor may be necessary for the occurrence of these events, it is not sufficient because other proteins and conditions are known to affect the activity of the receptor and its ability to alter gene expression or to induce other effects. Certain studies could be interpreted to indicate AhR-independent mechanisms, although these studies have not clearly ruled out involvement of the AhR. The only consistent, but limited, evidence for TCDD-induced effects that do not involve the AhR comes from studies using AhR-deficient transgenic animals. Here however, only minor effects occurred following treatment with extremely high doses of TCDD. Thus, a toxic response to TCDD has AhR interaction as a key event, but there are various species-, cell-, development-, gender-, and disease-dependent differences in the cellular milieu that can affect the nature and extent of the response observed.

The findings that many AhR-modulated effects are regulated with distinct specificity supports the understanding that the molecular and cellular pathways leading to any particular toxic event are extremely complex. Precise dissection of these events represents a considerable challenge, especially in that a toxic response may depend on timely modulation of several genes rather than of just one particular gene, and possibly modulation of these genes in several rather than just one cell type or tissue.

While a defined mechanism at the molecular level or a defined mode of action for TCDD-induced carcinogenicity is lacking, EPA concludes the following

- interaction with the AhR is a necessary early event in TCDD carcinogenicity in experimental animals.
- through interaction with the AhR, TCDD modifies one or more of a number of cellular processes, such as induction of enzymes, changes in growth factor and/or hormone regulation, and/or alterations in cellular proliferation and differentiation.

- AhR activation is anticipated to occur in humans and to progress to tumors. AhR is present in human cells and tissues, studies using human cells are consistent with the hypothesis that the AhR mediates TCDD toxicity and no data exist to suggest that the biological effects of AhR activation by TCDD are precluded in humans.
- non-AhR mediated carcinogenic effects of TCDD are possible.

5.1.3. Summary of the Qualitative Weight of Evidence Classification for TCDD

Under the 2005 Cancer Guidelines (U.S. EPA, 2005), TCDD is characterized as carcinogenic to humans, based on the available data as of 2009. This conclusion is based on

- Multiple occupational epidemiologic studies showing strong evidence of an association between TCDD exposure and increased mortality from all cancers.
- Epidemiological studies showing an association between TCDD exposure and certain cancers in individuals accidentally exposed to TCDD in Seveso, Italy.
- Extensive evidence of carcinogenicity at multiple tumor sites in both sexes of multiple species of experimental animals.
- General scientific consensus that the mode of TCDD's carcinogenic action in animals involves AhR-dependent key precursor events and proceeds through modification of one or more of a number of cellular processes, such as induction of enzymes, changes in growth factor and/or hormone regulation, and/or alterations in cellular proliferation and differentiation.
- The human AhR and rodent AhR are similar in structure and function and human and rodent tissue and organ cultures respond to TCDD in a similar manner and at similar concentrations.
- General scientific consensus that AhR activation is anticipated to occur in humans and to progress to tumors.

5.2. QUANTITATIVE CANCER ASSESSMENT

5.2.1. Summary of NAS Comments on Cancer Dose-Response Modeling

5.2.1.1. *Choice of Response Level and Characterization of the Statistical Confidence Around Low Dose Model Predictions*

The NAS commented on the low dose model predictions in the 2003 Reassessment, including EPA's development of ED_{01} (effective dose eliciting x percent response) estimates for numerous study/endpoint combinations. The committee also suggested that EPA had not appropriately characterized the statistical confidence around such model predictions in the low-response region of the model.

> The committee concludes that EPA did not adequately justify the use of the 1% response level (the ED_{01}) as the POD for analyzing epidemiological or animal bioassay data for both cancer and noncancer effects. The committee recommends that EPA more explicitly address the importance of the selection of the POD and its impact on risk estimates by calculating risk estimates using alternative assumptions (e.g., the ED_{05}) (NAS, 2006a, p. 18)
>
> It is critical that the model used for determining a POD fits the data well, especially at the lower end of the observed responses. Whenever feasible, mechanistic and statistical information should be used to estimate the shape of the dose-response curve at lower doses. At a minimum, EPA should use rigorous statistical methods to assess model fit, and to control and reduce the uncertainty of the POD caused by a poorly fitted model. The overall quality of the study design is also a critical element in deciding which data sets to use for quantitative modeling (NAS, 2006a, p. 18).
>
> EPA should … assess goodness-of-fit of dose-response models for data sets and provide both upper and lower bounds on central estimates for all statistical estimates. When quantitation is not possible, EPA should clearly state it and explain what would be required to achieve quantitation (NAS, 2006a, p. 10).

The NAS also suggested that EPA report information describing the adequacy of dose-response model fits, particularly in the low-response region. For those cases where biostatistical modeling was not possible, the NAS recommended that EPA identify the reasons.

> The Reassessment should also explicitly address the importance of statistical assessment of model fit at the lower end and the difficulties in such assessments, particularly when using summary data from the literature instead of the raw data, although estimates of the impacts of different choices of models would provide valuable information about the role of this uncertainty in driving the risk estimates (NAS, 2006a, p. 73).

5.2.1.2. *Model Forms for Predicting Cancer Risks Below the Point of Departure (POD)*

The NAS focused much of its review on EPA's derivation of a cancer slope factor. Specifically, the NAS commented extensively on the selection of the appropriate point of departure (POD) and the extrapolation of dose response modeling below the POD.

The NAS questioned EPA's choice of a linear, nonthreshold model for extrapolating risk associated with exposure levels below the POD, concluding that the current scientific evidence was sufficient to justify the use of nonlinear methods when extrapolating below the POD for

TCDD carcinogenicity. The committee further recommended that EPA include a nonlinear model for low dose cancer risk estimates as a comparison to the results from the linear model.

> The committee concludes that EPA's decision to rely solely on a default linear model lacked adequate scientific support. The report recommends that EPA provide risk estimates using both nonlinear and linear methods to extrapolate below PODs (NAS, 2006a, p. 5).
>
> After reviewing EPA's 2003 Reassessment and additional scientific data published since completion of the Reassessment, the committee unanimously agreed that the current weight of scientific evidence on the carcinogenicity of dioxin is adequate to justify the use of nonlinear methods consistent with a receptor-mediated response to extrapolate below the POD. The committee points out that data from NTP released after EPA generated the 2003 Reassessment provide the most extensive information collected to date about TCDD carcinogenicity in test animals, and the committee found the NTP results to be compelling. The committee concludes that EPA should reevaluate how it models the dose-response relationships for TCDD… (NAS, 2006a, p. 16).
>
> Because EPA's assumption of linearity at doses below the 1% excess risk level for carcinogenic effects of TCDD, other dioxins, and DLCs is central to the ultimate determination of regulatory values, it is important to critically address the available scientific evidence on the most plausible shape of the dose-response relationship at doses below the POD (LED_{01}). On the basis of a review of the literature, including the detailed review prepared by EPA and presented in Part II of EPA's Dioxin Risk Assessment and new literature available since the last EPA review, the committee concludes that, although it is not possible to scientifically prove the absence of linearity at low doses, the scientific evidence, based largely on mode of action, is adequate to favor the use of a nonlinear model that would include a threshold response over the use of the default linear assumption (NAS, 2006a, p. 122).
>
> On the whole, the committee concluded that the empirical evidence supports a nonlinear dose-response below the ED_{01}, while acknowledging that the possibility of a linear response cannot be completely ruled out. The Reassessment emphasizes the lack of such nonlinear models, hence its adoption of the approach of linear extrapolation below the POD level. Although this approach remains consistent with the cancer guidelines (EPA 2005, see also Appendix B), EPA should acknowledge the qualitative evidence of nonlinear dose response in a more balanced way, continue to fill in the quantitative data gaps, and look for opportunities to incorporate mechanistic information as it becomes available. The committee recommends adopting both linear and nonlinear methods of risk characterization to account for the uncertainty of dose-response relationship shape below ED_{01} (NAS, 2006a, p. 72).

5.2.2. Overview of EPA Response to NAS Comments on Cancer Dose-Response Modeling

EPA agrees with the NAS that the approaches to cancer dose-response modeling for TCDD should be clearly communicated and justified. Furthermore, due to the abundance of new information on TCDD carcinogenicity published since the 2003 Reassessment, EPA has reevaluated the cancer dose-response modeling for TCDD presented in the 2003 Reassessment. As detailed below in Section 5.2.3, EPA has conducted an updated cancer dose-response assessment for TCDD that incorporates key NAS recommendations discussed in this document, reflects the current state-of-the science in cancer dose-response modeling and integrates new TCDD carcinogenic information. Detailed responses to the NAS comments summarized above are found in Section 5.2.3.3.

The 2003 Reassessment presents an extensive dose-response assessment of TCDD and provides a comprehensive summary of dose-response relationships. The analyses and discussions synthesized a considerable breadth of data and model types, highlighting the strengths and weaknesses of the then-available scientific information. Modeling included both administered dose and steady state body burden dose metrics, taking into account variation in half-lives of TCDD across species. These body burden calculations used a simple one-compartment kinetic model based on the assumption of a first-order decrease in the levels of administered dose as a function of time. An excess risk of 1% was chosen to model the cancer data, but comparative results were also shown for 5% and 10% excess risk (see Table 8-2 of the 2003 Reassessment). Dose response was also explored thoroughly for a number of in vitro and biochemical endpoints in addition to the in vivo data analyses, and ranges of these values were presented (see Figures 8-1, 8-2 and 8-3 of the 2003 Reassessment). Thus, the 2003 Reassessment provides an initial evaluation of the carcinogenic database for TCDD and serves as the foundation for the analyses presented below.

5.2.3. Updated Cancer Dose-Response Modeling for Derivation of Oral Slope Factor

The following sections describe the dose-response analysis of the cancer data from epidemiologic cohort studies (see Section 2.4.1) and rodent bioassays (see Section 2.4.2), concluding with the derivation of oral slope factors for TCDD based on epidemiologic data (see Section 5.2.3.1) and rodent bioassay data (see Section 5.2.3.2).

5.2.3.1. *Dose-Response Modeling Based on Epidemiologic Cohort Data*

The 2003 Reassessment included dose-response analyses and the development of oral slope factors from the following three occupational cohorts: the NIOSH cohort, the Hamburg cohort, and the BASF cohort. In this document, EPA determined that specific studies from each of these cohorts (Steenland et al., 2001; Ott and Zober, 1996; Becher et al., 1998) met the epidemiologic study inclusion criteria (see Section 2.3.1 and Section 2.4.1). In Section 5.2.3.1.1, the oral slope factors derived from these studies in the 2003 Reassessment are reviewed. Since the publication of the 2003 Reassessment, additional cancer epidemiology studies based on these cohorts have been published in the peer-reviewed literature. Of these, Cheng et al. (2006) met the epidemiologic study inclusion criteria (see Section 2.3.1 and Section 2.4.1). In Section 5.2.3.1.2, EPA evaluates the suitability of deriving an oral slope factor from this study and derives oral slope factor estimates.

Another study that met the current epidemiologic study inclusion criteria (Warner et al., 2002) was also briefly discussed in the 2003 Reassessment, but an oral slope factor (OSF) was not derived from that study. In Section 5.2.3.1.2.2, EPA discusses its unsuccessful attempt to use the categorical results published by Warner et al. (2002) to develop an oral cancer risk estimate.

5.2.3.1.1. *Evaluation of Epidemiologic Studies in the 2003 Reassessment for OSF Derivation.*

In the 2003 Reassessment, EPA reported dose-response modeling results for three epidemiologic studies of human occupational cohorts: the NIOSH cohort with data published by Steenland et al. (2001); the Hamburg cohort with data published by Becher et al. (1998); and the BASF cohort with data published by Ott and Zober (1996). Each of these studies is summarized in Section 2.4.1 of this document and in the 2003 Reassessment (Part II, Chapter 8; Part III, Chapter 5). Furthermore, EPA has evaluated the suitability of these studies for use in TCDD dose-response modeling and concluded that each of these studies meet the inclusion criteria for epidemiology studies presented in Section 2.3.1.

Each of these studies reports all cancer mortality as an outcome. Steenland et al. (2001) and Becher et al. (1998) analyzed cohorts of primarily male workers who experienced occupational exposures to TCDD over long periods of time, while Ott and Zober (1996) studied a cohort of primarily male workers who were exposed to high TCDD concentrations at a single point in time due to an industrial accident.

The authors of all three of these studies measured and then back-extrapolated TCDD levels in a subset of workers to estimate exposures during employment and then the authors used this information to estimate exposures in the remainder of the cohort. These measured TCDD samples generally were collected decades after the last known occupational exposure. In each study, the authors relied on TCDD measures in the cohort to back-calculate serum lipid or body fat levels of TCDD using a simple one-compartment kinetic model based on the assumption of a first-order decrease in the levels of exposure dose as a function of time. The assumed half-life of TCDD used in the models varied from study to study. None of the studies sampled TCDD levels from the entire cohort; for example, Ott and Zober collected samples from 138/243 workers (57% of the cohort), which was the highest percentage of workers sampled among the three studies. Steenland et al. (2001) and Becher et al. (1998) used the measured and back-extrapolated TCDD concentrations to estimate the exposures that were associated with various occupations within the cohort, and subsequently used this information to develop exposure matrices (i.e., the TCDD load per unit time for an occupation) that then could be used to estimate the cumulative dioxin dose for each cohort member. Ott and Zober (1996) used regression procedures with data on time spent at various occupational tasks to estimate TCDD levels for all members of the cohort. Following the estimation of worker exposures in each of these three studies, the studies' authors divided these cohorts into exposure subgroups based on the estimated TCDD levels.

In the 2003 Reassessment, EPA identified a POD based on a 1% response in cancer mortality (ED_{01}) for the Steenland et al. (2001), and the Ott and Zober (1996) studies. EPA extrapolated from this POD to lower doses using a straight line drawn from the POD to the origin—zero incremental dose, zero incremental response—to give a probability of extra risk. Because there was insufficient evidence to support an assumption of nonlinearity, EPA chose to develop these models using a linear model.

5.2.3.1.1.1. ***Steenland et al. (2001).***

Steenland et al. (2001) developed dose-response models based on TCDD exposures and all cancer mortalities from eight plants in the NIOSH cohort. Serum lipid levels of TCDD in 1988 were measured in 193 workers at one of these plants. Steenland and coauthors relied on a first-order kinetic model (assuming a constant 8.7 year half-life) to back-extrapolate to serum

TCDD levels at the time of the last occupational exposure. The study authors assigned exposure estimates to each of the 3,538 workers in the cohort based on a job-exposure matrix. This matrix was based on (1) an estimated level of contact with TCDD, (2) the degree of TCDD contamination of the products the workers produced, and (3) the fraction of a workday during which the worker likely contacted the TCDD-contaminated products. They then estimated each worker's serum TCDD levels as an area under the concentration curve (AUC) for lipid-adjusted serum levels over time. The mortality analysis was conducted on 256 cancer decedents.

Several different dose-response models were fit to these data to provide estimates of fatal cancer risk. The best-fitting model was a Cox regression exposure-response model using the log(AUC) of TCDD lipid concentration (ppt-year) lagged by 15 years as the exposure metric. Steenland and colleagues also developed a piecewise linear regression model with no lag, in which two separate linear slopes were estimated. This analysis assumed a background exposure of 0.5 pg/kg-day. The lipid concentrations were converted to body burdens by dividing by 4. The central tendency estimate and lower bound ED_{01}s from the piecewise linear model and their associated cancer slope factors for the most sensitive endpoint (male cancer mortality) are presented in Table 5-1.

5.2.3.1.1.2. ***Becher et al. (1998).***

Based on the Hamburg cohort, Becher et al. (1998) reported a dose-response analysis for all fatal cancers combined. The mortality analysis was conducted in 1992 on 124 cancer decedents. The exposure variable in the study was the integrated blood levels for TCDD concentration over time (AUC, ng/kg-years), as estimated by Flesch-Janys et al. (1998); these were converted to body burdens by dividing by 4. Estimates of the half-life of TCDD, based on the sample of 48 individuals with repeated measures, were incorporated into a model that back-extrapolated TCDD exposures to the end of the employment after accounting for the workers' ages and body fat percentages. These estimated exposure measures were then applied to the entire cohort, which consisted of all 1,189 regular male employees who were employed for at least 3 months between 1952 and 1984 at the Boehringer chemical plant in Hamburg, Germany.

Becher et al. (1998) used a Cox regression approach for the dose-response modeling and developed three models: a multiplicative model, an additive model, and a power model.[25] The response variable in each model was the SMR for total cancer mortality. The models were calculated with lag times of up to 20 years. The multiplicative model provided the best fit; however, the study authors judged the fits for all three models to be acceptable. The model results were used to calculate unit risk estimates derived as the risk of cancer death through age 70 given a daily dose of 1 pg/kg body weight of TCDD minus the risk given no exposure to TCDD. These calculations were based on background German cancer mortality rates. The model results were used to calculate cancer risk estimates. The lower bound estimates on the dose were not available for models published by Becher et al. due to the absence of statistical parameter measures. The central tendency estimate ED_{01}s from the three statistical models and their associated cancer slope factors are presented in Table 5-1.

5.2.3.1.1.3. *Ott and Zober (1996).*

In the 2003 Reassessment, EPA also developed a dose-response analysis based on a study reported by Ott and Zober (1996) for cancer incidence and mortality experienced by 243 men, who were exposed to TCDD in 1953 during an accident at the BASF plant in Germany. The cohort was followed through 1992. TCDD blood lipid levels were available for 138 of these men 30 years after the accident. These levels were back-extrapolated and used to estimate the AUC for TCDD. Body burdens (ng/kg) were estimated by dividing AUC by 4, and steady-state

[25]The "multiplicative model" set relative risk (RR) equal to exp(βd), where the dose d is the AUC. The "additive model" set RR = 1+βd, and the "power model" set RR = exp(β log (kd+1)). The values β and k are estimated parameters.

body burdens were estimated assuming a constant half-life of approximately 7.1 years.[26] Ott and Zober (1996) used Cox proportional hazard approaches to estimate cancer mortality risk per unit TCDD dose.[27] The central tendency estimate and lower bound ED_{01}s from the modeling and their associated cancer slope factors are presented in Table 5-1.

5.2.3.1.2. *Evaluation of Epidemiologic Studies Published Since the 2003 Reassessment for OSF Derivation.*

Two additional epidemiological studies that met the study inclusion criteria (see Section 2.3) for use in dose response modeling as set forth in this document are evaluated in this section for the estimation of cancer risk estimates. Each of these studies is summarized in Section 2.4.1 of this document.

5.2.3.1.2.1. ***Cheng et al. (2006).***

As discussed in Section 2.4.1.1.1, Cheng et al. (2006) analyzed the relationship between TCDD dose and all cancer mortality for the same subset of NIOSH workers as analyzed previously by Steenland et al. (2001). In contrast to Steenland et al., Cheng et al. (2006) used the "concentration- and age-dependent elimination model" (concentration- and age-dependent elimination [CADM], discussed in Section 3.3; see also Aylward et al. [2005a]), rather than a constant 8.7-year half-life, and calculated serum-derived TCDD estimates for use in dose-response analysis. An important feature of CADM is that it incorporates concentration- and age-dependent elimination of TCDD from the body, meaning that the effective half-life of TCDD

[26]Based on the initial body burden (B_0) EPA estimated the body burden at time t using the following formula: $B(t) = B_0 e^{-k_e t}$, where k_e is an elimination constant equal to ln(2)/(half-life in years). This implies that the AUC at time T after initial exposure is $AUC = \frac{B_0}{k_e}(1 - e^{-k_e T})$. T in this case was 39 years (time from the accident in 1953 to the follow-up in 1992). Dividing by a lifetime of 71 years (mean age in 1954, 33 years, plus 38 years from 1954 to the followup in 1992) yields the lifetime mean body burden as: $B_{mean} = \frac{B_o}{71 k_e}(1 - e^{-k_e T})$. In the 2003 Reassessment, EPA converted the steady-state body burden to units of equivalent initial dose by dividing by the constant $\frac{1}{71 k_e}(1 - e^{-k_e T})$. With the given values for half-life and T, that constant is 0.1411 and 1/(the constant) is 7.09.

[27]The model from Ott and Zober has risk proportional to $e^{\beta \times dose}$ with $\beta = \ln(1.22)$. The corresponding slope for the mean (steady-state) body burden is 7.0851 * log(1.22) * 0.001 (the 0.001 converts nanograms to micrograms).

elimination varies based on exposure history, body burden, and age of the exposed individuals. As discussed in Section 3.3, the use of the CADM model to simulate TCDD kinetics in the NIOSH cohort results in time-integrated body burden estimates four to five times greater than those obtained with the simple first-order model, with smaller differences between the two methods at lower exposures.

Following the estimation of dose using the CADM-derived AUC values, Cheng and colleagues derived dose-response estimates for the NIOSH cohort using Cox regression and piecewise linear modeling. Their results are summarized in Table 5-2. For comparison, the Cox regression coefficient from the analysis conducted by Steenland et al. (2001), which relied on a first-order elimination rate model assuming a constant 8.7-year half-life, is also shown on the first line of the table. As in Steenland et al. (2001), Cheng et al. (2006) found a much stronger relationship between cancer mortality and exposure metrics lagged 15 years compared to the relationships for unlagged exposures. Cheng et al. (2006) also noted that the dose-response relationship plateaued above the 95th percentile of exposure. For exposures lagged 15 years, the regression coefficient of the linear slope derived by Cheng et al. (2006) was 3.3×10^{-6} per ppt-year lipid-adjusted serum TCDD (in Table III in their analysis—the standard error of this regression coefficient was 1.4×10^{-6}). The upper 5% of the exposure range (individuals ≥252,950 ppt-year lipid adjusted serum TCDD) was excluded in estimating this slope. Because this exclusion reduces the upper portion of the response where the slope is shallow, this likely better represents the slope in the region of the curve where the fatal cancer risk is increasing with dose, which is the equivalent of dropping the highest dose in an animal bioassay.

To develop an OSF for TCDD, EPA used data and equations from the Cheng et al. (2006) in its calculations as follows

- Upper 95th percentile slope (β_{95}). To estimate β_{95} of this regression coefficient, EPA summed the regression coefficient (β) and the product of 1.96 and the standard error of the regression coefficient, yielding an estimate of 6.0×10^{-6} per ppt-year lipid adjusted serum TCDD, as follows

$$\beta_{95} = \beta + 1.96 * SE$$

- Background cancer mortality risk estimate (R_0). EPA used an R_0 of 0.112 as reported by Cheng et al. (2006)[28]

- Risk in the exposed group associated with a 1% extra risk of fatal cancer (R_{exp}). EPA estimated R_{exp} to be 0.12088, using the following relationship for extra risk:

$$0.01 = \frac{R_{\text{exp}} - R_0}{1 - R_0}$$

- Incremental cancer mortality risk in the exposed population based on a 1% extra risk (R_D). R_D, was then calculated to be 8.9×10^{-3} using the following equation:

$$R_D = R_{exp} - R_0$$

- Cumulative TCDD concentration in the fat compartment for a 1% extra risk (AUC_{01}). EPA then estimated AUC_{01} using Equation 3 from Cheng et al. (2006):

$$AUC_{01} = \ln(R_D + R_0/R_0)/\beta_{95}$$

  Cheng et al. (2006) assume that the TCDD concentration in fat is the same as ppt-yr lipid adjusted serum concentration. The AUC_{01} was calculated to be 1.26×10^4 ppt-yr.

- Oral slope factor associated with 1% extra risk [$OSF(AUC_{01})$]. AUC_{01} of 1.26×10^4 ppt-yr is used by EPA as a POD for linear extrapolation to zero. $OSF(AUC_{01})$ is calculated to be 7.92×10^{-7} by dividing 0.01 by the AUC_{01}. $OSF(AUC_{01})$ is linear with the TCDD concentration in fat. Ingested TCDD doses, however, are not linear with the predicted TCDD fat concentrations in the Emond pharmacokinetic model. Thus, the $OSF(AUC_{01})$ that is linear with TCDD in the fat is not linear with ingested TCDD dose.

Thus, to estimate the fatal cancer risk associated with an oral intake of TCDD, estimates of both the average TCDD concentration in the fat resulting from the oral intake and the risk using the $OSF(AUC_{01})$ are needed. Next in this report, EPA presents estimates of OSFs for specific TCDD intake rates based on target risk levels (RLs) of 1×10^{-3}, 1×10^{-4}, 1×10^{-5}, 1×10^{-6}, and 1×10^{-7}, using the following calculations:

- Area under the TCDD fat concentration curve associated with a target risk level (AUC_{RL}) (ppt-yr). For each target risk level, EPA calculated an AUC_{RL} by dividing the target risk level by the $OSF(AUC_{01})$ [i.e., $RL/7.92 \times 10^{-7}$ $(\text{ppt-yr})^{-1}$].

- Lifetime averaged concentration of TCDD in the fat compartment associated with the target risk level (FAT_{RL}) (ng/kg). The AUC_{RL} estimates were then further divided by

[28]In Table IV, Cheng et al. (2006) report two estimates of background fatal cancer risk, R_0, for males aged 75 years: 0.112 and 0.124. A R_0 estimate of 12.4% was used by Steenland et al. (2001), and 11.2%, as estimated for all males in the 1999–2001 Surveillance Epidemiology and End Result data set. EPA chose to use 11.2% as this reflects more current cancer mortality rates in the U.S. population.

70 years to identify (FAT_{RL}) (ng/kg). This step essentially reverses the integration undertaken to calculate AUC_{01}.

- Continuous daily TCDD intake D_{RL} (ng/kg-day) associated with a target risk level over a lifetime. Using the Emond human physiologically based pharmacokinetic (PBPK) model, EPA estimated the D_{RL} necessary to achieve the FAT_{RL}. Table 5-3 shows the ingested TCDD doses (D_{RL}) corresponding to these target risk levels.
- Oral slope factor at the target risk level (OSF_{RL}) (per mg/kg-day). Table 5-3 also shows that at the target risk levels, the associated OSF_{RL}s range from 3.7×10^5 to 1.3×10^6 per mg/kg-day. These are calculated as $OSF_{RL} = RL/D_{RL} \times 10^6$.

Table 5-4 shows analogous results based on the best estimate of regression coefficient (3.3×10^{-6}) for all fatal cancers as reported by Cheng et al. (2006) by comparing lipid-adjusted serum concentrations, fat concentrations, risk specific dose estimates and equivalent oral slope factors for risk levels of 1×10^{-3}, 1×10^{-4}, 1×10^{-5}, 1×10^{-6}, and 1×10^{-7}.

5.2.3.1.2.2. *Warner et al. (2002).*

Warner et al. (2002) is a study of 981 females exposed to elevated TCDD levels following the Seveso accident of 1976. The TCDD exposure pattern involving a single period of elevated TCDD exposures followed by an extended period of lower ambient level TCDD exposures and elimination is similar to that of the BASF cohort (Ott and Zober, 1996). TCDD levels, measured or estimated in blood lipids shortly after the time of the accident, were available for all women. These women were divided into four exposure groups of <20, 20−44, 44.1−100, and >100 ppt. In this cohort, 21 total cancers have been observed; 15 of these were breast cancer cases and 3 were thyroid cancer cases. Cox proportional hazards modeling showed that the hazard ratio for breast cancer associated with a 10-fold increase in serum TCDD levels ($\log_{10}$ (TCDD)) was significantly increased to 2.1 (95% CI = 1.0−4.6). Rate ratios (95% CI) for cancer incidence in these 4 groups were 1.0, 1.0 (0.2−5.5), 2.2 (0.5−10.8) and 2.5 (0.5−11.8). Using a Cox proportional hazards model and assuming continuous exposure, the rate ratio was 1.7 (0.9−3.4) for each 10-fold increase in serum TCDD; that is, a $\log_{10}$ transformation of the exposure estimates in their analysis was presented. They reported a test for trend of $p = 0.09$.

EPA attempted to estimate an ED_{01} from the modeled results of Warner et al. (2002) from the statistically significant hazard ratio for breast cancer. However, EPA had to estimate the slope of the tangent to the log-linear relationship. Because the slope of a log-linear relationship

is not constant but varies with dose, and because the lowest exposure measure was well-above the 1% response region of interest, EPA could not confidently estimate the tangent to the log-linear relationship. The transformation of the $\log_{10}$ dose units to linear units of TCDD yielded an implausibly low ED_{01} and correspondingly high cancer risk that was inconsistent with a visual inspection of the untransformed plot. EPA was not confident in these values for health risk assessment because of uncertainties in the transformation in the low response region of the original model. Thus, EPA did not derive an ED_{01} or oral slope factor for this study.

5.2.3.2. *Dose-Response Modeling Based on Animal Bioassay Data*

Figure 5-3 provides a summary of the process EPA has utilized to select and identify candidate TCDD OSFs from key animal bioassays that were identified in Section 2.4.3 of this document. For each in vivo animal cancer study that qualified for TCDD dose-response assessment using the study inclusion criteria specified in Section 2.3.2, EPA first selected the species/sex/tumor data set combinations that had been characterized as having statistically significant increases in tumor incidence by either a pair-wise test between the treated group and the controls or by a trend test showing increases in tumors with increases in dose. Next, EPA used the Emond animal kinetic model discussed in Section 3 to estimate blood concentrations corresponding to each study's average daily administered doses for use in dose response modeling. Benchmark dose lower confidence bounds ($BMDL_{01}$s) were then estimated for the blood concentrations by (1) using the multistage cancer model for each species/sex/tumor combination within each study and (2) using a Bayesian Markov Chain Monte Carlo framework that assumes independence of tumors, modeling all tumors together for each species/sex combination within each study. The final selected models were subjected to goodness-of-fit tests and visual inspections of fit to the raw data. Thus, for each sex/species combination within each study, this process generated a $BMDL_{01}$ for each single tumor type and another $BMDL_{01}$ for the combined tumors. Finally, using the Emond human kinetic model discussed in Section 3, human equivalent doses ($BMDL_{HED}$s) were then estimated for each of the $BMDL_{01}$s and, using a linear extrapolation, OSFs were calculated by $OSF = 0.01/BMDL_{HED}$. The lowest OSF for a species/sex combination for either a single tumor type or all combined tumors was selected as a candidate OSF for TCDD cancer assessment. These steps in Figure 5-3 are further described in detail in the following sections.

5.2.3.2.1. *Selection of key data sets.*

Based on the study selection criteria outlined in Section 2.3.2 (see Figure 2-3), EPA selected four animal bioassays for use in the cancer dose-response assessment for TCDD (see Table 2-6 and Section 2.4.2 for detailed study descriptions). Three of these studies (Kociba et al., 1978; NTP, 1982; Toth et al., 1979) were evaluated in the 2003 Reassessment, while one study (NTP, 2006) was published after the 2003 Reassessment was released. The NTP (2006) study was specifically called out by the NAS (2006a) report as cancer bioassay data that EPA should evaluate prior to completing its TCDD dose-response assessment. As discussed below, EPA has chosen to conduct dose-response modeling for a number of tumor types from each of the sex/species combinations in these studies in order to maximize the amount of information available to support OSF derivation. Because tumors occurred in multiple sites in the exposed animals, each tumor type was considered separately (individual tumor models) and were also combined into composite tumor incidence dose estimates (multiple tumor models).

The tumor incidence tables for these four bioassays are shown in Tables 5-5 through 5-8. The data in these tables are summarized from each study's reference publication and are the species/sex/tumor incidence data used for TCDD dose-response modeling in this report. In the case of the Kociba et al. (1978) female rat combined hepatocellular adenomas and carcinomas only, EPA used data from a reanalysis of the pathology slides that was published by Goodman and Sauer (1992). The data sets in Tables 5-5 through 5-12 were selected because they had been characterized by the study authors as having statistically significant increases in tumor incidence by either a pair-wise test between the treated group and the controls or by a trend test showing increases in tumors with increases in dose. One exception is that Goodman and Sauer (1992) did not statistically analyze the revised tumor incidence data from their reanalysis of the Kociba et al. (1978) female rat combined hepatocellular adenomas and carcinomas, so a Fischer's Exact Test to evaluate the statistical significance of those data was used in this document. In the case of the NTP (2006) study only, information was available regarding the length of time that the animals stayed on test (105 weeks); animals who died within the first year were censored from analysis in this document because animals who died within the first year were not considered to have been alive long enough to develop tumors. Therefore, those animals were not included in the denominators in Table 5-11. These adjusted incidence data were used in the analysis of

tumor dose-response for NTP (2006) in this document. The tumor incidence data in Tables 5-5 through 5-12 include

- nasal, tongue and adrenal tumors in males, and liver, nasal and lung tumors in females from the Kociba et al. (1978) 2-year study of Sprague-Dawley rats,
- liver, thyroid and adrenal tumors in males, and subcutaneous tissue, liver, adrenal and thyroid tumors in females and from the NTP (1982) 2-year study of Osborne-Mendel rats,
- lung and liver tumors in males, and subcutaneous tissue, hematopoietic system, liver and thyroid tumors in females from the NTP (1982) 2-year study of $B6C3F_1$ mice
- liver, oral mucosa, pancreas and lung tumors in females from the NTP (2006) 2-year study of Sprague-Dawley rats, and
- liver tumors in males from the Toth et al. (1979) 1-year study of Swiss/H/Riop mice.

For each cancer endpoint, the reported (administered) doses from each study were converted, where necessary, to average daily doses in ng/kg-day (e.g., doses administered 5 days/week were adjusted by multiplying by 5 and dividing by 7 to get average daily doses). These doses were then subjected to kinetic modeling to generate blood concentrations for use in TCDD dose-response modeling.

5.2.3.2.2. *Dose adjustment and extrapolation methods for selected data sets.*

5.2.3.2.2.1. ***Dose metric estimation for dose-response modeling.***

Tables 5-5 through 5-12 show the blood concentrations that were used in TCDD dose-response modeling of the animal bioassay data. Based on kinetic analysis (see Section 3), a choice for whole blood concentration of TCDD was made for the purpose of dose extrapolation between animals and humans. In order to estimate blood concentrations for each study selected, the Emond PBPK model was run using ACSLX® software, version 2.5.0.6 (see Section 3). Depending on the selected study, either rat or mouse versions of the model were used. In each case, the simulation was performed using the exposure and observation durations, the body weights, and the adjusted doses from the original studies. Details of PBPK model input parameters are given for each study's m-file in Appendix C.2. In the case of Toth et al. (1979) study, which dosed the animals for a year and then followed up for the lifetime of the animal, only the one-year simulation was performed. The m-files were used to run the appropriate PBPK

model to estimate time-averaged, maximum, and terminal (end of exposure) lipid adjusted blood serum concentration (see Appendix C.3). Other model simulated dose metrics such as concentrations for liver, fat, Ah-receptor bound in liver, body burden, and the time at which the maximum concentration was reached for each dose metric are also reported for illustrative purposes in Appendix C.3. The complete results for each study modeled are shown in Appendix C.3.

Applying the Emond PBPK model to each study/endpoint combinations in Tables 5-5 through 5-12, whole blood concentrations were derived by converting the model estimated lipid-adjusted to whole blood concentration using model constants for the lipid fraction in blood serum for rodents (0.0033) and the blood serum fraction (0.55). Once the Emond modeling runs were performed, the model outputs of whole blood concentrations were used as dose metrics to estimate BMDLs (in ng/kg). To obtain BMDLs, benchmark dose modeling was performed as described below by substituting the model simulated dose metrics (whole blood concentrations) for the original study doses and calculating the corresponding BMDL (results appear below and in Appendix F).

5.2.3.2.2.2. ***Calculation of human equivalent doses (HEDs).***

Human equivalent doses (ng/kg-day), corresponding to each BMDL (ng/kg) were calculated using the Emond human PBPK model (see Section 3) and are denoted as $BMDL_{HED}$s. The Emond human PBPK model was run for 70 years assuming a constant daily dose starting from birth. The model concentrations were averaged over both the entire 70 year lifetime (lifetime average) and over the five years surrounding the peak concentration (five-year average) (see Section 3.3.1, describing first order body burden estimation). The human equivalent doses were estimated by adjusting the daily dose model input until the time-averaged whole blood concentration matched the associated alternative dose BMDL (derived earlier from animal PBPK model). For animal studies which lasted longer than 540 days, the lifetime average was used; for studies lasting less than 540 days, the five year average was used. The process was iterative and continued until the modeled human concentration was within 1% of the BMDL. In general, however, the concentrations matched to within 0.1%.

5.2.3.2.3. *Dose-response modeling approaches for rodent bioassays.*

5.2.3.2.3.1. ***Modeling of individual tumors.***

EPA's BMDS Software, version 2.1 was used to estimate the $BMDL_{01}$s for each of the species/sex/tumor combinations, using the blood concentrations and incidence data shown in Tables 5-5 through 5-12. Each data set was modeled using the multistage cancer model, and a $BMDL_{01}$ in blood concentration was estimated. The multistage model has been used by EPA in the majority of its quantitative cancer assessments because it is statistically robust and able to provide good fits to a wide range of dose-response patterns. It is also consistent with the multistage nature of the carcinogenic process. The mathematical form of the multistage model is

$$P(d) = 1 - \exp[-(q_0 + q_1 d + q_2 d^2 + ... + q_k d^k)] \qquad \text{(Eq. 5-3)}$$

where

$P(d)$ = lifetime excess risk (probability) of cancer at dose d
q_i = parameters estimated in fitting the model, $i = 1, \ldots, k$.

To estimate the BMD_{01}s and $BMDL_{01}$s, BMDS was run with all parameters set to their defaults; up to three degrees of freedom were specified for the dichotomous, multistage cancer model; and a 95% confidence level. A 1% benchmark response (BMR) was used for each tumor type to ensure the PODs were estimated in the linear portion of the dose-response curve. The $BMDL_{01}$ (ng/kg) was then converted to a $BMDL_{HED}$ (ng/kg-day) using the Emond human model, and an OSF in units of $(\text{mg/kg-day})^{-1}$ was calculated by, $OSF = 0.01/BMDL_{HED} \times 10^6$. It may be noted that the OSF is linear with the blood concentration, and is not linear with ingested TCDD doses; thus, using the Emond et al. pharmacokinetic model, risk-specific doses of TCDD intake (ng/kg-day) corresponding to the target risk levels would need to be provided for use in human health risk assessment. In the following sections, results are presented for the models that provided the best overall fit to the data as judged by comparison of likelihood ratios for models that had an acceptable fit (chi-squared goodness of fit statistic $p > 0.05$).

5.2.3.2.3.2. ***Multiple tumor (Bayesian) models.***

Statistically significant increased tumor incidences were observed at multiple sites in male and/or female rats (Kociba et al., 1978; NTP, 1982, 2006) and male and female mice (NTP,

1982) following oral exposures to TCDD. With this multiplicity of tumors, the concern is that a potency or risk estimate based solely on one tumor site (e.g., the most sensitive site) may underestimate the overall cancer risk associated with exposure to this chemical. Relevant approaches in the 2005 Cancer Guidelines (U.S. EPA, 2005) for characterizing total risk include the following: (1) analyze the incidence of tumor-bearing animals, or (2) combine the potencies associated with significantly elevated tumors at each site. The NRC (1994) concluded that an approach based on counts of animals with one or more tumors (tumor-bearing animals) would tend to underestimate overall risk when tumor types occur independently, and thus an approach based on combining the risk estimates from each separate tumor type should be used. On independence of tumors, NRC (1994) stated "…a general assumption of statistical independence of tumor-type occurrences within animals is not likely to introduce substantial error in assessing carcinogenic potency."

Because potencies are typically upper bound estimates, summing such upper bound estimates across tumor sites is likely to overstate the overall risk. Therefore, following the recommendations of the NRC (1994) and the 2005 Cancer Guidelines (U.S. EPA, 2005), a statistically valid upper bound on combined risk was derived, assuming independence, in order to gain some understanding of the overall risk resulting from tumors occurring at multiple sites. In the case of TCDD, tumors are thought to be independent across the sites found in these three studies because: (1) they are in different organs and tissues, specifically liver, lung, thyroid, subcutaneous tissue, oral cavity, tongue, pancreas, adrenal cortex and the hematopoietic system; (2) different kinds of tumors were found, even within the same organ (e.g., both cholangiocarcinomas and hepatocellular adenomas were found in female rat livers in NTP, 2006); and (3) the tumors found in these studies were not progressive (i.e., they did not metastasize to other sites in the body). It is important to note that this estimate of overall potency describes the risk of developing tumors at any combination of the sites and is not the risk of developing tumors at all sites simultaneously.

For modeling individual tumor data, the multistage model is specified as shown in the previous section (see Eq. 5-3). Under the assumption of independence, the model for the combined (or composite) tumor risk is still multistage, with a functional form that has the sum of stage-specific multistage coefficients as the corresponding multistage coefficient.

$$P_c(d) = 1 - exp[-(\Sigma q_{0i} + d\Sigma q_{1i} + d^2\Sigma q_{2i} + ... + d^m\Sigma q_{mi})], \text{ for } i = 1,\ldots, k, \qquad \text{(Eq. 5-4)}$$

where k = total number of sites.

The resulting equation for fixed extra risk (BMR) is polynomial in dose (when logarithms of both sides are taken) and can be solved in a straightforward manner for the combined BMD. However, the current version of BMDS cannot estimate confidence bounds for this combined BMD.

Therefore, a Bayesian approach to finding confidence bounds on the combined BMD was implemented using WinBUGS (Spiegelhalter et al., 2003). WinBUGS software is freely available and implements Markov Chain Monte Carlo (MCMC) computations. Use of WinBUGS has been demonstrated for derivation of a distribution of BMDs for a single multistage model (Kopylev et al., 2007) and is easily generalized (Kopylev at al., 2009) to derive the distribution of BMDs for the combined tumor load, following the NRC (1994) methodology described above. The advantage of a Bayesian approach is that it produces a distribution of BMDs that allows better characterization of statistical uncertainty. For the current analysis, a diffuse (high variance or low tolerance) Gaussian prior restricted to be nonnegative was used. The posterior distribution was based on three simulation chains with 50,000 burn-in (i.e., the initial 50,000 iterations were dropped) and a thinning rate of 20, resulting in 150,000 interactions total. The median and 5th percentile of the posterior distribution provided the BMD_{01} (central estimate) and $BMDL_{01}$ (lower bound) for combined tumor load, respectively.

The methodology above was applied to the statistically significant dose-response data from Kociba et al. (1978), NTP (1982), and NTP (2006) (see Section 2.3.2 for data set selection criteria). As with the risk estimates generated for individual tumor sites, the combined analysis used the internal dose metric, whole blood concentration (see Section 3). For the combined tumors for each sex/species combination, a $BMDL_{01}$ in blood concentrations was estimated. The $BMDL_{01}$ (ng/kg) was then converted to a $BMDL_{HED}$ (ng/kg-day) using the Emond human model, and an OSF in units of $(mg/kg\text{-}day)^{-1}$ was calculated by, $OSF = 0.01/BMDL_{HED} \times 10^6$. It may be noted that the OSF factor is linear with the blood concentration, and is not linear with ingested TCDD doses; thus, using the Emond et al. pharmacokinetic model, risk-specific doses of TCDD intake (ng/kg-day) corresponding to the target risk levels would need to be provided for use in human health risk assessment.

5.2.3.2.4. *Results of dose-response modeling for rodent bioassays.*

Table 5-13 presents the benchmark dose modeling results for both the individual tumors and the combined tumors based on blood concentrations. The *p*-values in the table are for a chi-square goodness of fit statistic with significance of $p > 0.05$. Goodness of fit was acceptable at $p > 0.05$ for all models. The difference in log likelihood (dLL) statistic documents the difference in log likelihoods between stages of the models in cases where the stage is above 1; it shows the difference between the stage in the table and the lower stage. For example, for the NTP (2006) liver cholangiocarcinomas, twice the difference of 2.92 would be >3.84, the test statistic from the assumed chi-square distribution,[29] with $p = 0.95$, justifying the choice of 3 stages over 2 stages. The best fitting multistage models include a 1-stage (linear) model for all of the individual tumor data sets from Kociba et al. (1978), NTP (1982), and Toth et al. (1979), as well as for the pancreatic and oral mucosa tumors in NTP (2006), a 2-stage model for the lung tumors in NTP (2006), and a 3-stage model for the liver cholangiocarcinoma and liver adenoma data sets from NTP (2006). For the Toth et al. (1979) liver tumors, the model fit to all of the data was poor, and the highest dose group was dropped in order to achieve an acceptable fit ($p = 0.29$). The BMD_{01}s and $BMDL_{01}$s (ng/kg) were estimated from these multistage models for the individual tumors. BMD_{01}s and $BMDL_{01}$s (ng/kg) were also provided in Table 5-13 for the combined tumors for each sex/species combination within a study. These were estimated from the distributions of BMD_{01}s produced by the Bayesian MCMC simulation (see Section 5.2.3.1.2.3.2). The BMD_{01}s and $BMDL_{01}$s (ng/kg) for the combined tumors in Table 5-13 are the mean and lower 95% percentile values from these distributions, respectively.

5.2.3.2.4.1. ***Individual tumor models.***

Table 5-14 shows the $BMDL_{HED}$s (ng/kg-day) that were estimated from the $BMDL_{01}$s in Table 5-13 using the Emond human model (see Section 5.2.3.1.2.2.2) and the OSFs calculated by, $OSF = 0.01/BMDL_{HED} \times 10^6$ to convert the OSF to $(mg/kg\text{-}day)^{-1}$ units. BMDS results, details of the model fits and dose-response graphics for all endpoints are shown in Appendix F.

[29]The chi-square distribution with 1 degree of freedom is the correct distribution only under standard conditions (e.g., no boundary parameters in null hypothesis). Thus, the correct distribution for the situation where the parameter of interest is on the boundary, as happens with testing for the order of the multistage model, and, possibly nuisance parameters (estimated parameters of the model), is very difficult to derive (Self and Liang, 1987). Therefore the *p*-value of chi-square with one degree of freedom is used as the best available choice.

Although only the blood concentration results are presented in this section, for comparison purposes, Appendix F also provides modeling results for the studies' administered average daily doses. Table 5-14 lists the OSFs in decreasing value. It can be seen that liver tumors in male mice yield the highest slope factors; OSF values are 5.9×10^6 and 5.2×10^6 per mg/kg-day in NTP (1982) and Toth et al. (1979), respectively. The OSFs for the new NTP (2006) study in female rats are two orders of magnitude lower, ranging from 1.8×10^4 to 1.8×10^5 per mg/kg-day, representing the lowest OSFs for TCDD from the individual tumor models.

5.2.3.2.4.2. ***Multiple tumor (Bayesian) models.***

Table 5-15 shows the $BMDL_{HED}$s (mg/kg-day) that were estimated from the $BMDL_{01}$s in Table 5-13 using the Emond human model (see Section 5.2.3.1.2.2.2) and the OSFs calculated by, $OSF = 0.01/BMDL_{HED} \times 10^6$ to convert the OSF to $(mg/kg\text{-}day)^{-1}$ units. Table 5-15 lists the OSFs in decreasing value. It can be seen that the combined liver and lung tumors in male mice yield the highest OSF value of 9.4×10^6 per mg/kg-day from NTP (1982), and the combined adrenal, tongue and nasal tumors in male rats yield the lowest OSF value of 3.2×10^5 from Kociba et al. (1978). The OSF for the combined liver, oral mucosa, lung, and pancreatic tumors in female rats from the newer NTP (2006) study is 4.4×10^5.

5.2.3.2.5. ***Summary evaluation of slope factor estimates from rodent bioassays.***

To estimate a range of candidate TCDD OSFs from the animal data, dose-response modeling of the four chronic rodent bioassays identified in Section 2.4.3 was conducted. Dose-response modeling was performed using whole blood concentrations, and $BMDL_{HED}$ values (ng/kg-day) were derived for the 25 species/sex/endpoint data sets individually (see Table 5-14) and for seven species/sex combined tumor data sets (see Table 5-15).

For each sex/species combination within a study, the combined tumor OSFs presented in Table 5-15 represent the highest OSFs that have been derived from the animal cancer bioassay multistage models. The most sensitive species and sex is male mice, for which the estimated $BMDL_{HED}$ for combined tumors is 1.1×10^{-3} ng/kg-day. This result, which is derived under the assumption that multiple tumor types occur independently in the exposed animals, is, as expected, lower than the $BMDL_{HED}$ for the most sensitive individual tumor.

Based on these results, EPA believes that a credible value for the $BMDL_{HED}$ derived from the animal studies lies in the range shown in Table 5-15 between 3.1×10^{-2} and 1.1×10^{-3} ng/kg-day. These values, which correspond to oral slope factor values of 3.2×10^{5} and 9.4×10^{6} per mg/kg-day, respectively, encompass the range at which elevated cancer risks can be detected for the most sensitive species, sex, and endpoints in the animal bioassay data.

As noted above in Sections 5.2.3.1.2.2 and 5.2.3.1.2.3, the OSFs shown in this section are linear with blood concentration. The OSFs shown in Tables 5-14 and 5-15 correctly depict the cancer risks generated from the multistage models for their specific associated $BMDL_{HED}$s. However, ingested TCDD doses are not linear with the predicted TCDD blood concentrations generated by the Emond pharmacokinetic model. Thus, the OSFs associated with the ingested doses are not linear with ingested dose. If the OSFs derived in this section were to be used in human health risk assessment, target risk levels (e.g., 1×10^{-3}, 1×10^{-4}, 1×10^{-5}, 1×10^{-6}, etc.) would need to be identified. Then, the Emond et al. pharmacokinetic model could be used to generate risk-specific doses of TCDD intake (ng/kg-day) corresponding to the target risk levels.

5.2.3.2.6. *Qualitative uncertainties in slope factor estimates from rodent bioassays.*

This section presents a qualitative discussion of the uncertainties associated with calculating the OSF for TCDD from chronic animal bioassay data. Discussions on the feasibility of conducting a quantitative uncertainty analysis for TCDD using dose-response methods are provided in Section 6.4.2 of this document.

5.2.3.2.6.1. ***Quality of studies relied upon for determining POD.***

EPA considers the overall quality and breadth of the studies used for the cancer dose-response analysis to be excellent. All of the studies were published in the peer-reviewed literature, and two of them were conducted by NTP (1982, 2006). Kociba et al. (1978) and Toth et al. (1979) are older studies, but appear to have been conducted according to good laboratory practice standards. The control and dose group sample sizes were relatively large, ~50 animals or more per group for all of the studies except for Toth et al. (1979), where the dose group sizes were ~40 animals per group. All four studies exposed the test animals via the oral route to TCDD alone, as was stipulated in EPA's study inclusion criteria. Collectively, these four studies reported development of numerous cancer endpoints (tumors) in both sexes in two strains of rats

(Sprague-Dawley and Osborne-Mendel) and two strains of mice (i.e., $B6C3F_1$, Swiss/H/Riop). The overall high quality of these studies and the strong, positive association between TCDD exposure and cancer suggests that study quality is not a major contributing factor to uncertainty in the risk estimates.

5.2.3.2.6.2. ***Interpretation of results from studies relied upon for determining POD.***

As discussed in Section 3.4.3.2.1, questions arose about the interpretation of liver tumor responses in female rats in the Kociba et al. (1978) study. Three re-evaluations of the slides have been reported (Kociba et al., 1978; Squire, 1980; Goodman and Sauer, 1992). The decision to use the Goodman and Sauer (1992) evaluation was based on the fact that the authors used the most current tumor classification procedures. The incidence of hepatocellular adenomas and carcinomas (individually and combined), however, did vary (sometimes widely) for each dose group across the three evaluations. Although the state-of-the-science is reflected in the incidence data used for analysis, a small degree of additional uncertainty (i.e., above that associated with data collection and reporting errors) could be associated with the female liver tumor risk estimates from the Kociba et al. (1978) study due to this variability. No controversy has arisen with regard to the interpretation of the NTP (1982, 2006), or Toth et al. (1979) tumor identification and classification.

5.2.3.2.6.3. ***Consistency of results across chronic rodent bioassays.***

The existence of four high-quality chronic bioassays for TCDD increases confidence and reduces uncertainty in the cancer OSFs. Considered together, these studies tested two species and both sexes of mice and rats, and a wide range of well-characterized tumor types. All four studies were consistent in observing increases (at some dose level) in rates of liver tumors (in both species and sexes). While tumors at other sites were observed (and those sites varied across study, species, and sex), the liver tumors were consistently the most sensitive indicators of carcinogenic response (with respect to $BMDL_{HED}$ estimates). Lung tumors were also consistently observed across three of the studies, in male mice in the NTP (1982) study and in female rats in Kociba et al. (1978) and NTP (2006). As discussed above, the two most sensitive single-tumor endpoints as judged by $BMDL_{01}$ values were associated with elevated liver tumor risks, followed by lung, lymphoma or leukemia, thyroid and adrenal cancers. The consistency of

tumor types and sensitivities across endpoints and studies lends confidence to the multistage modeling results.

5.2.3.2.6.4. ***Human relevance of rodent tumor data.***

There is some concordance in the tumor responses observed in the rodent test species and humans, however, the most sensitive tumor site in the animals, the liver, has not been associated with cancer from TCDD exposures in humans. On the other hand, lung cancer and leukemia have been are found both in the animal studies and in epidemiologic studies of exposed workers. The consistency across sex, species and strains in the animal studies suggests that the occurrence of several of these tumors, in particular, liver and lung tumors is not an idiosyncratic response of a particular combination of species, strain, or sex. As discussed in Section 5.2.1, the likely AhR related carcinogenic mechanism is credible for humans as well as for rodent species.

5.2.3.2.6.5. ***Relevance of rodent exposure scenario.***

The four chronic rodent bioassays exposed the test animals for ~2 years, the majority of their lifespans. The exception is the Toth et al. (1979) study, where the animals were exposed only for one year, but were kept on the study for a second year before they were evaluated for cancer. These lifetime bioassays in animals have long been used by EPA to assess potential lifetime exposures and effects in humans. However, in the case of TCDD, the half life of TCDD in the body for rats, mice, and humans is very different (see Section 3). Thus, there is a significant amount of uncertainty in the use of rat and mouse data to develop OSFs for human cancer risk assessment of TCDD.

5.2.3.2.6.6. ***Impact of background TCDD exposures.***

It is known that TCDD has been found in the feed used in animal bioassays, and that this is a confounding factor, particularly in older studies. The effect of TCDD in the diets of test species has the potential to be quite significant given the low levels of TCDD at which adverse effects have been observed. Insofar as that is an issue, the risks associated with TCDD exposures in the animal bioassays, and therefore the OSFs, would be biased high, which could be the case for the NTP (1982), Kociba et al. (1978) and Toth et al. (1979) studies. The impact of this issue is that the newer study, NTP (2006), accounted for TCDD exposures in the animal

feed. Thus, there is likely to be less uncertainty in the TCDD dose-response information presented in NTP (2006) than in the other three studies conducted before 1990.

5.2.3.2.6.7. ***Choice of endpoint for POD derivation.***

As noted above, the liver tumor PODs represent the most sensitive single-tumor endpoint across the four cancer bioassays. Thus, the liver cancer endpoints must be seriously considered for derivation of a TCDD OSF. As discussed in the previous section, EPA has also developed Bayesian dose-response estimates for combined tumors, which yield $BMDL_{01}$ values slightly lower than those for any individual tumor type. Although it is the most conservative choice to select the lowest combined tumor POD for OSF derivation, there are uncertainties associated with the multiple tumor analysis. The assumption of independence of tumors across sites is reasonable, particularly since the tumors from TCDD do not metastasize. However, the independence assumption lacks hard evidence and needs further laboratory confirmation.

5.2.3.2.6.8. ***Choice of animal-to-human extrapolation method.***

The analyses presented here have used the Emond human kinetic model for extrapolating dose from animals to humans (as discussed in Section 3.4.2). The rationale for this choice is that the blood concentration metric most accurately reflects the concentration of TCDD in the various tissues. As discussed in Section 3.4.3.2.4, use of the blood concentration dose metric results in critical dose estimates (HEDs) that are considerably lower (10- to more than 100-fold) than those derived based on administered dose. This does not reflect bias in the blood-based measure; rather it is a reflection of the highly nonlinear biokinetics of TCDD in the body. EPA has also explored the impacts of using other dose metrics, including liver bound TCDD concentration calculated based on the Emond model. As discussed in Section 3.4.3.2.6.2, this also results in HED estimates much lower than those obtained based on administered dose.

5.2.3.2.6.9. ***Choice of model for POD and model uncertainty for POD derivation.***

The bioassay-based cancer dose-response assessment in this section has used the multistage model which is the standard model choice for such assessments and has been the basis for most of EPA's cancer risk assessments. In that sense, there is no associated uncertainty for model choice.

There is some model choice uncertainty associated with instances of lack of fit. When the multistage model does not adequately describe the observed pattern of responses (typically determined by examining the *p*-value for lack of fit), a decision must be made about possible adjustments, including the dropping of higher dose groups thought to be less relevant to the estimation of low-dose slopes. In this analysis, poorer fits (*p*-values less than 0.10) were observed in four cases, all from NTP (1982) (see Table 5-13). The lowest $BMDL_{01}$ associated a low *p*-value ($p = 0.09$) was for the lung tumors in the NTP (1982) male mouse, the third lowest POD behind the liver PODs in the individual tumor data sets. The other instances were for adrenal cortex and thyroid follicular cell adenomas in male rats and for subcutaneous tissue in female mice in the NTP (1982) study. In those instances, the *p*-values were 0.06, 0.06, and 0.09, respectively; the fits were considered adequate for describing the low-dose response patterns and estimating slope factors. These poorly fit data sets contribute to uncertainty in the combined tumor PODs. The lowest $BMDL_{01}$ in the combined tumors is for the male mice combined liver and lung tumors, thus estimates from this sex/species combination from NTP (1982) is highly uncertain and impacts its choice as a POD.

5.2.3.2.6.10. ***Statistical uncertainty in model fits.***

Every model fit to a data set is associated with some inherent statistical uncertainty. For this reason, bounds were calculated and used for OSF derivation (e.g., lower bounds on benchmark doses, in this case the $BMDL_{01}$s). Those bounds account for uncertainties associated with finite samples of test animals, both in terms of the number of dose groups and of the number of animals per dose group. Valid and accepted statistical procedures have been applied to ascertain the impact of those limitations on the estimates of interest. That being the case, the statistical uncertainties associated with finite samples have been adequately addressed.

5.2.3.2.6.11. ***Choice of risk level for POD derivation.***

The BMR level that has been used for the POD in deriving the cancer OSF is one percent extra risk. This choice is supported by the fact that risks need to be low enough so that the linear extrapolation from them reflects low-dose behavior, reducing the impact of any higher-dose changes (inflexions) in the dose-response. The multistage model has the characteristic that it can be approximated by a linear model at low enough doses. Thus, a risk level in the range of one

percent is generally suitable for characterizing that approximate linear function, while not falling too far below the observed range of the data.

5.2.3.3. *EPA's Response to the NAS Comments on Choice of Response Level and Characterization of the Statistical Confidence Around Low Dose Model Predictions*

The NAS was concerned with the statistical power to determine the shape of the dose response curve at low doses, well below observed dose-response information. EPA shares this concern in that the shape of the dose-response curve in the low-dose region cannot be determined with confidence when based on higher dose information.

When tumor data are used for dose-response modeling, a POD is obtained from the modeled tumor incidences. When assessing carcinogenicity using a linear extrapolation approach from a POD, a balance must be struck between staying within the range of the observations and obtaining a representative estimate of the low-dose slope. Traditional cancer bioassays, with approximately 50 animals per group, can typically support modeling down to an increased incidence of 1–10%; epidemiologic studies, with larger sample sizes, below 1%. For the TCDD animal cancer bioassay data, most of the low-dose tumor incidence responses are under 10% (relative to controls), with some as low as 2%. After evaluating the magnitude of the uncertainty in $BMDL_{01}$s against the impact of using $BMDL_{10}$s, EPA has chosen to use a 1% BMR in all cases, determining that the uncertainty bounds on the $BMDL_{01}$ values are reasonable (see Appendix F for details). For comparison purposes, $BMDL_{01}$s, $BMDL_{05}$s, and $BMDL_{10}$s are presented for all modeled tumor incidences in Appendix F.

In the analysis of the animal cancer bioassays presented in this document, the multistage cancer model was applied with a linear dose extrapolation to zero. EPA used a 1% excess risk estimate, i.e., a $BMDL_{01}$, as the POD for development of candidate TCDD cancer oral slope factors using a Bayesian multitumor approach (see Section 5.2.3.2. The advantage of a Bayesian approach is that it produces a distribution of BMDs that allows better characterization of statistical uncertainty.

Central tendency and lower bound slope factor estimates are part of the standard BMDS multistage cancer model and are included in each output file for the animal bioassay single tumor analyses in Appendix F. Central tendency BMDs are also reported for the results of the animal

bioassay multitumor analysis (see Table 5-13). Central tendency slope factor estimates are given for all the qualifying epidemiological studies as well (see Tables 5-1 and 5-4).

5.2.3.4. *EPA's Response to the NAS Comments on Model Forms for Predicting Cancer Risks Below the POD*

The NAS offered extensive comments on the cancer dose-response modeling in the 2003 Reassessment. Although epidemiologic and rodent bioassay data are useful for the evaluation of the dose-response curve within the range of the observed response data, they have traditionally not been useful sources of information for identifying a threshold or for estimating the shape of the dose-response curve below the POD. Rather, mechanistic toxicological data have been the evidentiary sources of choice for those types of analyses. As noted above, any quantitative estimation of carcinogenic risk associated with TCDD exposure requires low-dose extrapolation of experimental data. Unfortunately, the shape of the dose-response curve in the low dose region is unknown.

Several of the analyses of epidemiological cohort data evaluated the fit of different dose-response models to the data. Log-dose models accentuate the importance of low-dose low-magnitude responses and can yield implausible results. The most relevant models used in these studies are the untransformed-dose Cox regression models, which are the most similar to the multistage (1-degree) model used for animal bioassay analysis. Better results have been obtained in the cohort analyses when the flattening of the hazard-ratio curve is taken into account. The latter has been modeled explicitly by Steenland et al. (2001), who use a piecewise linear model and implicitly by Cheng et al. (2006), who drop out a percentage of the high-dose response data and fit a linear model to the remainder. Importantly, the analyses of the epidemiologic cohorts presented in Section 5.2.3.1 are limited to evaluation and reanalyses of published data as reported by the study authors. EPA does not have access to the raw data from these epidemiologic studies and, therefore, could not conduct *de novo* analyses.

5.2.3.4.1. *Choice of extrapolation approach*

5.2.3.4.1.1. ***TCDD and receptor theory.***

TCDD is considered to be a receptor-mediated carcinogen in animals. Nearly all TCDD experimental data are consistent with the hypothesis that the binding of TCDD to the AhR is the

first step in a series of biochemical, cellular, and tissue changes that ultimately lead to toxic responses observed in both experimental animals and humans (Part II, Chapter 2 of the 2003 Reassessment). Ligand-receptor binding, like any bimolecular interaction, obeys the law of mass action as originally formulated by A.J. Clark (Limbird, 1996). The law of mass action predicts the fractional receptor occupancy at equilibrium as a function of ligand concentration. Fractional occupancy (Y) is defined as the fraction of all receptors that are bound to ligand:

$$Y = \frac{[TCDD\ -\ AhR\ ]}{[\ AhR\ ]_{TOT}} = \frac{[TCDD\ -\ AhR\ ]}{[\ AhR\ ] + [TCDD\ -\ AhR\ ]} = \frac{[TCDD\ ]}{[TCDD\ ] + K_d} \qquad \text{(Eq. 5-5)}$$

where [TCDD] is the concentration of the ligand, [AhR] is the concentration of the receptor and [TCDD-AhR] is the amount of liganded receptor. The equilibrium dissociation constant K_d describes the affinity of the interaction and is the concentration of TCDD that results in 50% receptor occupancy. This simple equation defines a rectangular hyperbola, which is the characteristic shape of the vast majority of biological dose-response relationships.

In certain cases, no response occurs even when there is some receptor occupancy. This suggests that there may be a threshold phenomenon that reflects the biological "inertia" of the response (Ariens et al., 1960). In other cases, a maximal response occurs well before all receptors are occupied, a phenomenon that reflects receptor "reserve" (Stephenson, 1956). Therefore, the law of mass action cannot by itself fully explain the effect or response observed after TCDD interacts with AhR. The ligand-receptor complex is associated with a signal transduction or effector system. In the case of the AhR, this effector system can be considered to be the transcriptional machinery itself. The key feature of this formulation is that a response is proportional, or a function of, the number of receptors occupied.

Furthermore, for a ligand such as TCDD that elicits multiple receptor-mediated effects, one cannot assume that the binding-response relationship for a simple effect (such as enzyme induction) will necessarily be identical to that for a different and more complex effect (such as cancer). The cellular cascades of events leading to different complex responses (e.g., altered immune function, developmental effects, or cancer) are different, and other rate-limiting events likely influence the final biological outcome resulting in different dose-response curves. Thus, even though TCDD binding to AhR is assumed to be the initial event leading to a spectrum of

biological responses, TCDD-AhR binding data may not always correlate with the dose-response relationship observed for particular effects.

A receptor-based mechanism would predict that, except in cases where the concentration of TCDD is already high (i.e., [TCDD]~K_d), incremental exposure to TCDD will lead to some increase in the fractional occupancy of AhR. However, as discussed above, it cannot be assumed that an increase in receptor occupancy will necessarily elicit a proportional increase in all biological response(s), because numerous molecular events contributing to the biological endpoint are integrated into the overall response. That is, the final biological response could be considered as an integration of a series of interdependent dose-response curves with each curve dependent on the molecular dosimetry for each particular step. Dose-response relationships that will be specific for each endpoint must be considered when using mathematical models to estimate the risk associated with exposure to TCDD. It remains a challenge to develop models that incorporate all the complexities associated with each biological response as the modes of action for various toxicological endpoints appear to vary greatly. For TCDD, extensive experimental data from studies using animal and human tissues indicate that cell- or tissue-specific factors determine the quantitative relationship between receptor occupancy and the ultimate biological response. This would suggest that the parameters for each mathematical model might only apply to a single biological response within a given tissue and species, making extrapolation to other systems challenging.

5.2.3.4.1.2. ***Low-dose extrapolation: threshold or no threshold?***

As indicated in the 2005 Cancer Guidelines,[30] toxicity reference values for human noncancer endpoints have historically been estimated based on a no-observed-adverse-effect level (NOAEL) or lowest-observed-adverse-effect level (LOAEL) from animal bioassay studies. This terminology suggests a biological population threshold beneath which no harm is anticipated. Reference values such as the oral reference dose (RfD) or inhalation reference concentration (RfC) are derived by applying uncertainty factors (UFs) to a POD. Depending on the nature of available data and modeling choice, a POD can be selected from values other than

[30]As stated in the 2005 Cancer Guidelines (U.S. EPA, 2005): "For effects other than cancer, reference values have been described as being based on the assumption of biological thresholds. The Agency's more current guidelines for these effects (U.S. EPA, 1996a, 1998b), however, do not use this assumption, citing the difficulty of empirically distinguishing a true threshold from a dose-response curve that is nonlinear at low doses."

an NOAEL or LOAEL, such as an ED_x, or a benchmark dose (BMD) or its BMDL. An RfD is described as "likely to be without appreciable risk" but the probabilistic language has not as yet been operationalized. There is no quantitative definition of "appreciable" and no mechanism to compute risk as a function of dose, so as to ascertain that the risk is indeed not appreciable. The risk at the RfD is not calculated, and it cannot be calculated within the current UF framework. Instead, a hazard quotient is computed as the ratio of a given exposure to the RfD, or a margin of exposure is estimated as the ratio of the POD to the human exposure level.

Cancer endpoints are predominantly thought to have no population biological threshold. Although the terminology "threshold/nonthreshold" is still common in cancer dose-response discussions, the 2005 Cancer Guidelines propose a different terminology, whereby "nonlinear models" are those whose dose-response *slope* is zero at or above zero. In the natural language, and indeed in data analysis, it is difficult to distinguish the following situations:

- The response approaches zero as dose goes to zero, versus
- The response *slope* goes to zero as dose goes to zero (nonlinear model).

This use of "nonlinear" is acknowledged to be idiosyncratic.[31] The NAS review (NAS, 2006a) does not consistently apply the terminology from the 2005 Cancer Guidelines, nor does it consistently distinguish the above two circumstances: "…the observed data are more consistent with a sublinear response that approaches zero at low doses rather than a linear dose response" (NAS, 2006a). The point of a nonlinear model in the sense of the 2005 Cancer Guidelines is that the response *slope* approaches zero. Both linear and nonlinear *responses* approach zero at low dose (in the absence of background). Since the terms "linear," "sublinear," and "nonlinear" invite confusion in this context, the following terminology is used in this document:

[31]From the 2005 Cancer Guidelines (U.S. EPA, 2005): "The term '*nonlinear'* is used here in a narrower sense than its usual meaning in the field of mathematical modeling. In these cancer guidelines, the term '*nonlinear'* refers to threshold models (which show no response over a range of low doses that include zero) and some nonthreshold models (e.g., a quadratic model, which shows some response at all doses above zero). In these cancer guidelines, a nonlinear model is one whose slope is zero at (and perhaps above) a dose of zero. …. Use of nonlinear approaches does not imply a biological threshold dose below which the response is zero."

Threshold Model: There is some threshold T > 0 such that the probability of response for any dose less than or equal to T is zero, and the probability is nonzero for any dose greater than T.

Linear/ Linear above Threshold Model: For the linear model, the probability of response is proportional to the dose. For the linear over threshold model, the probability of response is zero for a dose below the threshold, and it is proportional to the excess dose over the threshold otherwise.

Nonlinear Model: Any model that is not linear.

Supralinear/ Supralinear above Threshold Model: For the supralinear model, the slope of the probability of response decreases as dose increases; in other words, the second derivative of the response curve is negative. For the supralinear above threshold model, the second derivative is negative above the threshold, and the response probability is zero below the threshold.

Sublinear/Sublinear above Threshold Model: For the sublinear model, the slope of the probability of response increases as dose increases; in other words, the second derivative of the response curve is positive. For the sublinear above threshold model, the second derivative is positive above the threshold, and the response probability is zero below the threshold.

Zero Slope at Zero Model: The slope of the response curve is zero at or above dose zero.

All of these models may be understood in an individual or population sense. According to the 2005 Cancer Guidelines, the trigger for applying the basic RfD methodology for cancer endpoints is sufficient evidence for the "zero slope at zero" model for the population. By definition, any sublinear, supralinear, or linear model *above the threshold* is a zero slope at zero ("ZS@Z") model.

The relation between individual and population models is not simple. Suppose for purposes of illustration that each individual has a threshold, and that the threshold values are uniformly distributed in the population, for some neighborhood of zero. (Note that "in the neighborhood of zero" has the mathematical meaning that all values are sufficiently close to zero; i.e., for some $\delta > 0$, the thresholds are uniformly distributed over the interval $[0,\delta]$.)

Then (see Text Box 5-1)

A. If each person's response function is linear for a dose above the threshold, then the population dose-response curve is quadratic, and the slope of the population curve is zero for dose zero (ZS@Z).

B. If each person's response function is log linear, for a dose above the threshold, then the population dose-response curve is linear without a threshold, and the slope of the population curve is positive for dose zero.

Of course these are not the only possibilities; in general, the population dose-response curve depends on (1) the distribution of individual thresholds in the neighborhood of zero, (2) the dose-response curve for each individual, and (3) the dose metric.

On the nature or the distribution of individual thresholds, often referred to as the population tolerance distribution, there is ongoing debate as to how receptor kinetics influence the shape of that distribution. Even within an individual, there is a lack of consensus as to whether receptor kinetics confer linear or sublinear attributes to downstream events, or whether receptor kinetics, themselves, are linear, sublinear, or supralinear. Whatever the nature of the form of receptor kinetics, it may have little or no influence on the ultimate population response. The kinetics of receptors is in the domain of the individual, rather than the population. As described previously, receptor kinetics are governed by the law of mass action, which leads to a low-dose proportional (linear) response model, generally modeled by some form of Hill function, the linear form being Michaelis-Menten kinetics. Linearity can be violated by a number of inhibitory or stimulatory processes, but is a fundamental conserved characteristic in living systems. There is no *a priori* reason to believe that the shape of the dose-response curve in an individual has any relationship to the shape of the population response, particularly for quantal endpoints. Lutz and Gaylor (2008) present a strong argument for considering the population response in terms of the more

Text Box 5-1. Individual vs. Population Thresholds

Suppose each individual has a threshold T, below dose T the probability of response is zero, and suppose that the values for T are uniformly distributed in some neighborhood of zero.

A. Suppose that for dose $\delta > T$, the probability of response is proportional to $(\delta - T)$. For dose δ small enough to be in this neighborhood, the population response probability is proportional to:

$$\int_{T\in(0,\delta)} (\delta - T)\, dT =$$
$$\delta\int_{T\in(0,\delta)} dT - \int_{T\in(0,\delta)} T\, dT = \delta^2 - \delta^2/2 = \delta^2/2.$$

The slope of the response function is found by taking the derivative with respect to δ:
$(d/d\delta)\ \delta^2/2 = \delta$, which goes to zero as $\delta \to 0$.

B. Suppose now that the response probability at dose δ is proportional to $\ln(\delta) - \ln(T)$, all else being the same. The population response probability at dose δ is:

$$\int_{T\in(0,\delta)} (\ln(\delta) - \ln(T))\, dT.$$

To evaluate this, note that
$\int_{T\in(0,\delta)} \ln(T)\, dT = \delta \ln(\delta) - \delta$.

Hence the integral is:

$\delta \ln(\delta) - [\ \delta \ln(\delta) - \delta] = \delta$.

The derivative of this with respect to δ is 1, which goes to 1 as $\delta \to 0$.

traditional tolerance distribution, which is likely the result of more variable factors than the shape of receptor kinetics. Perhaps more to the point, receptor activation is only the first of many events in the path to the apical event (a tumor in this example). Because there are undoubtedly numerous additional downstream events that must occur before the apical effect is observed, there are many opportunities for interindividual variability to become manifest in the tolerance distribution. Even at the first step, a more likely contributor to interindividual variability than the shape of the response is the dose resulting in the response, as measured by the ED_{50} (the aforementioned equilibrium dissociation constant K_d, K_m in the Michaelis-Menten formulation), which shifts the response curve. Factors that influence shifts in response curves are generally modeled as normal or log-normal distributions and will most likely confer a log-normal shape on the population tolerance distribution, particularly if there are a number of dependent sequential steps or distinct subpopulations (Lutz, 1999; Hattis et al., 1999; Hattis and Burmaster, 1994).

Pertaining to the shape of the dose-response curve, because the criterion for applying the RfD model is a zero slope at dose zero, the role of dose metric must be fully appreciated. A slope is simply the ratio of a change on the vertical axis (the probability of response) relative to a change on the horizontal axis (the dose). Changing the dose metric from the dose to the *logarithm* of the dose dramatically changes this ratio. As dose goes to zero, the rate of change of log(dose) becomes infinite. Therefore, ANY dose-response relation with a finite slope at zero will *appear* to have a zero slope when graphed against log(dose). Text Box 5-2 illustrates this point for the mass action dose response, which has been proposed for receptor-mediated modes of action.

To see how the discussion over threshold/nonthreshold might play out for TCDD, consider the equilibrium dissociation constant K_d for TCDD, which measures the binding affinity of TCDD to the AhR. Lower values indicate higher binding affinity and (other things being equal) greater risk. For Han/Wistar rats, the value $K_d = 3.9$ is reported; human values are reported as $K_d = 9.6 \pm 7.8 (0.3 – 38.8$ *with 15 of 67 donors without detectable binding)* (Connor and Aylward, 2006).

If AhR binding is necessary for carcinogenesis, then the majority of a human population may be much less susceptible than Han/Wistar rats, whereas a population threshold, if it exists, might be well below the Han/Wistar rat threshold (see Section 4.4.2.9). The NAS contends that

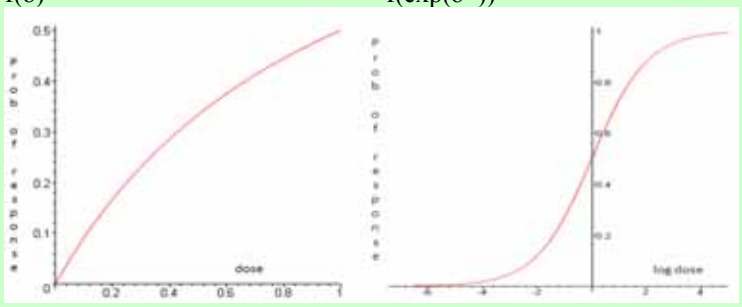
Text Box 5-2. Logarithmic Transform of Mass Action Hyperbolic Model

According to the operational model (Black and Leff, 1983; Motulsky and Christopoulos, 2003), the fractional occupancy of receptors by ligands follows a mass action law; the concentration of ligand-occupied receptors [LR] is equal to the product of the concentration of receptors [R] and the concentration of ligands [L], divided by the ligand concentration plus the binding affinity. This is a hyperbolic function written as [LR] = [R][L]/([L] + K_d).
The efficacy of a ligand can vary by tissue and endpoint;
this is usually expressed as a hyperbolic function of
[LR]; (TOP-BOTTOM) [LR]/([LR] + K_E) where K_E is the concentration of [LR] producing the effect halfway between TOP and BOTTOM. Combining these expressions gives:

$$\text{Response}([L]) = \frac{[L](\text{TOP-BOTTOM})\,(\tau/(\tau+1))}{[L] + K_d/(\tau+1)},$$

where τ is the tissue-/endpoint-specific 'transducer constant' [R]/KE. Writing δ for dose [L], this is proportional to the hyperbolic equation f(δ) = δ/(δ + K); $K = K_d/(\tau + 1)$.
Switching the dose metric to δ* = log(δ),
f(δ) becomes f(exp(δ*)) = exp(δ*)/[exp(δ*) + K].
f(δ) and f(δ*) are graphed below, for K = 1.

f(δ) f(exp(δ*))



Although these represent the same dose-response relation, the graph of f(exp(δ*)) has zero slope as dose goes to zero (δ* goes to $-\infty$). In fact, $df/d\delta^* = (df/d\delta)(d\delta/d\delta^*)$, and $d\delta/d\delta^* = e^{\delta^*} \to 0$ as $\delta^* \to -\infty$. The change in response relative to the change in log dose becomes infinitesimal, since log dose changes infinitely fast as dose goes to zero. f(δ*) is not a zero slope at zero model.

an AhR-mediated mode of action indicates a threshold dose-response relation (NAS, 2006a). Presumably, the value of the threshold, if it exists, depends on the AhR binding affinity.

Arguing for a population threshold in this case requires two types of information:

1. The distribution of the individual thresholds induced by, *inter alia* (among other things), the individual K_d values; and
2. The dose-response function for values above the threshold induced by K_d.

Without this information, the two possibilities {A,B} enumerated on Pages 5-50 through 5-51 cannot be distinguished, and the default linear relationship applies response probability is modeled as a linear function of dose, for dose near zero. However, from the 2005 Cancer Guidelines: "When adequate data on mode of action provide sufficient evidence to support a nonlinear mode of action *for the general population* (emphasis added) and/or any subpopulations of concern, a different approach—a reference dose/reference concentration that assumes that nonlinearity—is used." In current terminology, the reference dose methodology applies if there is sufficient evidence supporting a "zero slope at zero" model; otherwise, the linear nonthreshold model applies by default.

In principle, the choice between the above models could fall within the purview of dose-response modeling. However, standard statistical methods encounter well-known difficulties in detecting thresholds. Without going into detail, suffice to say that the maximum likelihood estimate of response probability when no responses are observed in a finite sample is always zero. That said, some researchers have attempted to identify thresholds (Mackie et al., 2002; Alyward et al., 2003) or nonlinearity (Hoel and Portier, 1994) by means of parameter estimation of appropriate models. A review of 344 rodent bioassays on 315 chemicals led to the following conclusion by Hoel and Portier (1994):

> We have also found that the oft-held belief that genotoxic compounds typically follow a linear dose-response pattern and that nongenotoxic compounds follow a nonlinear or threshold dose response pattern is not supported by the data. In fact we find the opposite with genotoxic compounds differing from linearity more often than nongenotoxic compounds.

The choice between a linear and "zero slope at zero" model in current practice does not fall under dose-response model fitting, it is made on the basis of a structured narrative as set forth in the 2005 Cancer Guidelines (U.S. EPA, 2005):

> In the absence of sufficiently, scientifically justifiable mode of action information, EPA generally takes public health-protective, default positions regarding the interpretation of toxicologic and epidemiologic data: animal tumor findings are judged to be relevant to humans, and cancer risks are assumed to conform with low dose linearity. … The linear approach is used when: (1) there is an absence of sufficient information on modes of action or (2) the mode of action information indicates that the dose-response curve at low dose is or is expected to be linear. Where alternative approaches have significant biological support, and no scientific consensus favors a single approach, an assessment may present results using alternative approaches. A nonlinear approach can be used to develop a reference dose or a reference concentration.

5.2.3.4.1.3. *Extrapolation method.*

The 2005 Cancer Guidelines (U.S. EPA, 2005) emphasize that the method used to characterize and quantify cancer risk from a chemical is determined by what is known about the MOA of the carcinogen and the shape of the cancer dose-response curve.

The NAS was critical of EPA's decision to apply linear low-dose extrapolation for TCDD cancer assessment in the 2003 Reassessment and encouraged the use of a nonlinear approach. The 2005 Cancer Guidelines state that a nonlinear approach should be used when "there are sufficient data to ascertain the mode of action and conclude that it is not linear at low doses and the agent does not demonstrate mutagenic or other activity consistent with linearity at low doses."

Receptor modeling theory (as outlined in the 2003 Reassessment, Part II, Chapter 8) indicates that exogenous compounds which operate through receptor binding mechanisms, such as TCDD, will follow a linear dose-response binding in the 1−10% receptor occupancy region. This theory has been supported by empirical findings and suggests that the proximal biochemical effects (such as enzyme induction) and transcriptional reactions for TCDD may also follow linear dose-response kinetics. More distal toxic effects could take any one of multiple forms (i.e., linear, sublinear, supralinear or threshold) depending on (1) the toxic mechanism; (2) location on the dose-response curve; and (3) interactions with other processes such as intracellular protein binding and cofactor induction/repression.

In the case of TCDD, many adverse effects experienced at low exposure levels have too much data variability to distinguish on a statistical basis (goodness-of-fit) between dose-response curve options, and whether the dose-response is linear, sublinear or supralinear. For tumor

responses, with the exception of squamous cell carcinoma of the oral mucosa and adenomas or carcinomas of the pancreas, which were fit with a linear multistage model, the tumor endpoints in the NTP (2006) study using female Sprague-Dawley (S-D) rats are all best fit with a sublinear model (i.e., the multistage model fits to tumor incidence data were second or third degree; see Table 5-13 and Appendix F). For all other tumor incidence data from all other cancer bioassays that met the study inclusion criteria (NTP, 1982, Toth et al., 1979, and Kociba et al., 1978), the multistage model fit was linear (first degree), when based on either administered dose or modeled blood concentrations (see Appendix F).

Another issue of potential importance when evaluating the shape of the dose-response curve for low dose effects is the concept of "interacting background." The concept of interacting background refers to a pathological process in the exposed population that shares a causal intermediate with the toxicant being evaluated. On this issue, a recent NAS committee (NAS, 2009) contended that

> …the current EPA practice of determining "nonlinear" MOAs does not account for mechanistic factors that can create linearity at low dose. The dose-response relationship can be linear at a low dose when an exposure contributes to an existing disease process (Crump et al.1976, Lutz 1990). Effects of exposures that add to background processes and background endogenous and exogenous exposures can lack a threshold if a baseline level of dysfunction occurs without the toxicant and the toxicant adds to or augments the background process. Thus, even small doses may have a relevant biologic effect. That may be difficult to measure because of background noise in the system but may be addressed through dose-response modeling procedures. Human variability with respect to the individual thresholds for a nongenotoxic cancer mechanism can result in linear dose-response relationships in the population (Lutz 2001) (NAS, 2009, p. 130).

AhR activation could be considered a causal intermediate in several disease processes. Recent studies have linked AhR activation in the absence of exogenous ligand to a multitude of biological effects, ranging from control of mammary tumorigenesis to regulation of autoimmunity (reviewed in Hahn et al., 2009). While the level of background activation of AhR by endogenous compounds (and/or exogenous compounds other than TCDD) in the human population is unknown, given the ubiquitous nature of several of the known endogenous and exogenous AhR ligands, it is reasonable to assume that a certain baseline level of AhR activation

exists in the population. The degree to which TCDD exposure augments this baseline level of AhR activation is unknown.

The 2005 Cancer Guidelines (U.S. EPA, 2005) recommend that the method used to characterize and quantify cancer risk from a chemical be determined by what is known about the mode of action of the compound and the shape of the cancer dose-response curve. The linear approach is used if the mode of action is not understood (U.S. EPA, 2005). In the case of TCDD, (1) the mode of action of TCDD-induced carcinogenesis beyond potential AhR activation is unknown; (2) information is lacking to determine confidently the shape of the dose-response curves for various adverse endpoints (including cancer) in humans or experimental animals; (3) there is undoubtedly a certain level of interacting background (i.e., AhR activation by endogenous ligands) in the human population; (4) many of the rodent cancer dose-response relationships (NTP, 1982; Toth et al., 1979; Kociba et al., 1978) are linear (first degree multistage model fit) when based on either administered dose or modeled blood concentrations; and (5) higher human interindividual variability compared to experimental rodents will tend to shift the shape of the dose-response towards linear (relative to rodents). Therefore, a linear low-dose extrapolation approach was used to estimate human carcinogenic risk associated with TCDD exposure.

5.2.3.4.1.4. ***Consideration of nonlinear methods.***

While the 2005 Cancer Guidelines deem linear extrapolation to be most appropriate for TCDD, EPA has carefully considered the NAS recommendation to provide risk estimates using both linear and nonlinear methods.

The 2005 Cancer Guidelines state

> For cases where the tumors arise through a nonlinear mode of action, an oral reference dose or an inhalation reference concentration, or both, should be developed in accordance with EPA's established practice for developing such values ... This approach expands the past focus of such reference values (previously reserved for effects other than cancer) to include carcinogenic effects determined to have a nonlinear mode of action.

In this section, EPA presents two illustrative examples of RfD development for carcinogenic effects of TCDD. Each of these examples focuses on data derived from animal bioassays as described in Section 2.4.2.

5.2.3.4.1.4.1. *Illustrative RfDs based on tumorigenesis in experimental animals.*

TCDD has been shown to be a multisite carcinogen in both sexes of several species of experimental animals. It also has been shown to be carcinogenic to humans. Most of the available quantitative human epidemiologic data related to TCDD carcinogenesis are for all cancer mortality. Mortality is a frank effect and is generally considered to be inappropriate for RfD development, therefore, the illustrative example below utilizes available evidence from experimental animals. Table 5-16 presents candidate PODs and RfDs for TCDD carcinogenicity based on combined tumor responses from the animal bioassays described in Section 2.4.2. The PODs from the NTP (1982, 2006) and Kociba et al. (1978) animal studies were derived from Bayesian multitumor dose-response modeling (as described in Section 5.2.3.2, Table 5-15) using a BMR of 1%. Because only TCDD-induced liver tumors were reported by Toth et al. (1979), the BMR of 1% (POD) from that study was generated using a first degree linear multistage model (see Table 5-14). Following BMD modeling, $BMDL_{HED}$s were then estimated using alternative dose metrics from the Emond model as described in Section 3. The illustrative RfDs were derived by dividing the $BMDL_{HED}$s by appropriate uncertainty factors. In each instance, a total UF of 30 was applied, comprising factors of 3 for the toxicodynamic component of the interspecies extrapolation factor (UF_A) and a factor of 10 for human interindividual variability (UF_H).

As shown in Table 5-16, the illustrative RfDs for TCDD-induced tumors range from 3.6E-11 for liver and lung tumors in male mice (NTP, 1982) to 1.0E-9 for adrenal cortex, tongue and nasal/palate tumors in male rats (Kociba et al., 1978). This illustrative RfD range for TCDD tumorigenesis falls within the range of candidate RfDs for noncancer TCDD effects presented in Table 4-5.

5.2.3.4.1.4.2. *Illustrative RfDs based on hypothesized key events in TCDD's MOAs for liver and lung tumors.*

As described in Section 5.1, most evidence suggests that the majority of toxic effects of TCDD are mediated by interaction with the AhR. EPA considers interaction with the AhR to be a necessary, but not sufficient, event in TCDD carcinogenesis. The sequence of key events following binding of TCDD to the AhR and that ultimately leads to the development of cancer is unknown. While the mode of action of TCDD in producing cancer has not been elucidated for any tumor type, the best characterized carcinogenic actions of TCDD are in rodent liver, lung, and thyroid. The hypothesized sequence of events following TCDD interaction with the AhR is markedly different for each of these three tumor types. Additionally, no detailed hypothesized mode of action information exists for any of the other reported tumor types.

The endpoints selected for this illustration were evaluated to provide insight into the quantitative relationships between tumor development and precursor events in TCDD-induced carcinogenesis. The endpoints described below may or may not be biologically adverse in themselves; the intent herein was to consider TCDD-induced biochemical and cellular changes that could lead to subsequent tumor development.

In the following exercise, illustrative RfDs were derived for key events in TCDD's hypothesized modes of action in the liver and lung. No appropriate dose-response data were identified for key events in TCDD's hypothesized MOA for thyroid tumors in a sex/species/strain that has been shown to develop thyroid tumors (i.e., female B6C3F1 mice and male and female Osborne-Mendel rats [NTP, 1982]).

As this is an illustrative exercise only, only studies that were originally identified in Section 2 for potential noncancer dose-response modeling were evaluated here (see Section 2.4.2 for study details). There may be additional studies available in the literature that would further inform the dose-response assessment of these endpoints.

Additionally, for animal model consistency, only results from studies conducted in female S-D rats are presented here. The majority of the available information on TCDD carcinogenicity (and TCDD carcinogenic precursor events) comes from studies conducted in female S-D rats and the most recent TCDD carcinogenicity study was conducted in female S-D rats (NTP, 2006). While both Kociba et al. (1978) and NTP (2006) have conducted TCDD carcinogenicity studies in female S-D rats, different substrains were used; this difference in

substrain may have resulted in the different carcinogenic responses reported from the two studies. While the carcinogenicity of TCDD in female S-D rats has been well characterized, this animal model does not exhibit the full suite of tumor responses reported for TCDD (for instance, female S-D rats have not been shown to develop thyroid tumors). Additionally, the most sensitive single tumor response in female S-D rats from NTP (2006) is squamous cell carcinoma of the oral mucosa (see Section 5.2.3.2), a tumor type for which no mode of action information exists. Therefore, the illustrative RfDs described below may not be protective against all tumor types.

For each endpoint, PODs for illustrative cancer RfD development were identified as described for the noncancer RfD derivation in Section 4. Briefly, for the endpoints identified below, the $NOAEL_{HED}$s and/or $LOAEL_{HED}$s were determined based on EPA analysis of the original data presented by the study author (see Section 2.4.2 for details) and by application of the Emond PBPK models as described in Section 3.3.4. $BMDL_{HED}$s were determined as described in Section 4.2 for all data sets amenable to BMD modeling. Modeling outputs for the endpoints are presented in Appendices E and G as noted in Table 5-17. The illustrative RfDs were derived by dividing the POD by appropriate uncertainty factors as indicated in Table 5-17.

5.2.3.4.1.4.2.1. *Liver tumors.*

Figure 5-4 presents one hypothesized mode of action for TCDD-induced liver tumors in rats. TCDD activation of the AhR leads to a variety of changes in gene expression, including increased CYP1A1 mRNA and subsequent increases in CYP1A1 activity. These alterations in gene expression are hypothesized to lead to hepatotoxicity, followed by compensatory regenerative cellular proliferation and subsequent tumor development. The details of the mechanism of TCDD-induced hepatotoxicity have not been fully determined but both CYP induction and oxidative stress have been postulated to be involved (Maronpot et al., 1993; Viluksela et al., 2000). Additionally, oxidative DNA damage has been implicated in liver tumor promotion (Umemura et al., 1999). The enhanced cell proliferation arising from either altered gene expression or hepatotoxicity, or both, could be the principal factor leading to promotion of hepatocellular tumors (Whysner and Williams, 1996).

A dose-response relationship exists for TCDD-mediated hepatotoxicity, and this parallels the dose-response relationship for tumor formation (or formation of foci of cellular alteration as a

surrogate of tumor formation). However, the dose-response relationship for other TCCD-induced responses such as enhanced gene expression is different from the dose-response for tumor formation in terms of both efficacy and potency (see Popp et al. [2006] for review).

A representative endpoint for each of the hypothesized key events following AhR activation for TCDD-induced liver tumors was identified and is shown in Figure 5-4. Illustrative RfDs based on each representative endpoint are shown in Table 5-17.

5.2.3.4.1.4.2.2. *Lung tumors.*

Far less is known about TCDD's mode of action in the lung. Figure 5-5 presents two hypothesized modes of action for TCDD-induced lung tumors in rats. The first hypothesized mode of action of TCDD in the lung involves disruption of retinoid homeostasis in the liver. Retinoic acids and their corresponding nuclear receptors, the RARs and the RXRs, work together to regulate cell growth, differentiation, and apoptosis. It is hypothesized that TCDD, through activation of the AhR, can affect parts of the complex retinoid system and/or other signaling systems regulated by, and/or cross-talking with, the retinoid system (reviewed in Nilsson and Hakansson, 2002). These effects are then hypothesized to lead to lung tumor development, however the mechanisms underlying this hypothesis are not well-defined. The second hypothesized mechanism for the carcinogenic action of TCDD in the lung is through induction of metabolic enzymes. Through activation of AhR and subsequent induction of metabolizing enzymes (such as CYP1A1), TCDD may enhance bioactivation of other carcinogens in lung (Tritscher et al., 2000). However, there are few studies to support this hypothesis.

Representative endpoints could only be identified for two of the hypothesized key events following AhR activation for TCDD-induced lung tumors. These endpoints are presented in Figure 5-5. Illustrative RfDs based on each of these two representative endpoints are shown in Table 5-17. There is insufficient information to form any conclusions on the quantitative progression to tumorigenicity or on the relative protection afforded by preventing the key events shown.

5.2.3.4.1.4.2.3. *Limitations of illustrative RfDs based on hypothesized key events in TCDD's MOAs for liver and lung tumors.*

A trend for increasing RfD values that follows the progression of endpoints towards the production of tumors is evident. However, there are a number of factors that prevent making strong conclusions based on this exercise. These limitations include the following

- This example addresses only two tumor types in one species, strain and sex (female S-D rats), with little information available on the hypothesized mode of action for lung tumors. No mode of action information is available for the most sensitive tumor type in this animal model (squamous cell carcinoma of the oral mucosa). Therefore, it is possible that the illustrative RfDs presented in this example would not be protective against all tumor types in female S-D rats. Importantly, other animal models have been shown to be more sensitive to TCDD-induced carcinogenesis based on combined tumor analysis (see Section 5.2.3.2); an RfD based on tumorigenesis in this animal model may not be protective against tumorigenesis in other, more sensitive, animal models (or, by extension, in humans).
- Several of the BMDLs are based on poorly-fitting models, such that the RfD is based on a LOAEL, which is not a particularly good measure for comparison across endpoints (e.g., LOAELs are subject to influences of dose spacing in bioassays). Furthermore, the hepatotoxicity BMDL based on a dichotomous 10% BMR, is not directly comparable to all the other BMDLs based on a continuous 1 standard-deviation BMR (Crump, 2002).
- The endpoints selected as representative of each hypothesized key event may not be the most appropriate choices. These particular endpoints were chosen because they were the most sensitive indicator (i.e., lowest POD) from the available data or were the only available choice based on a lack of data for other effects related to the hypothesized key event.
- The optimum timing of these events may not be reflected in the endpoints selected. Almost certainly, changes in gene expression are early events, such that a single exposure should be relevant, as in the mRNA changes reported after a single TCDD exposure (Vanden Heuvel et al., 1994), although it is not known whether the magnitude of these changes would be altered after longer-term exposure, or whether longer-term exposure would be more relevant to downstream events. Similarly, single exposures for induction of CYP enzymes would seem to relevant as a measure of the immediate effect, but it may be longer-term repeated CYP activity that is important for longer-term downstream events; Table 5-17 shows a nominal order-of-magnitude difference in effect levels for similar effect magnitudes (ca. 20-fold) from single exposures (Kitchin and Woods, 1979) and long-term exposures (53-weeks; NTP, 2006). The relevant exposure durations for oxidative stress and later effects are longer term, so a measurement of oxidative stress at 90-days in a rodent may be appropriate; Wyde et al. (2001) suggest that induction of 8-oxo-dG DNA adducts are a result of longer-term oxidative stress because of the lack of effect of single exposures. Hepatotoxicity and hepatocellular

proliferation events would appear at successively later times, but the effective exposure levels would depend heavily on the endpoints chosen to represent those events and the time at which they were measured. The toxic hepatopathy endpoint reported in NTP (2006), is a general measure of mild to moderate liver toxicity, but is measured only at the end of the study when tumors have already appeared. Hepatocyte hypertrophy, measured at 31weeks may be more duration-relevant, but may not indicate actual hepatocellular toxicity.

- The lowest of the tested doses may well be much higher, given that all animal diets are contaminated to a certain extent by TCDD, resulting in initial TCDD body burdens in all animals. Vanden Heuvel et al. (1994) reported TCDD liver concentrations in control animals almost as high as for the low-dose group, which could equate to a significant increase in the actual exposure experienced by the low-dose group. A similar effect on the low-dose group (0.45 ng/kg) in Kitchin and Woods (1979) is possible, although they did not report control animal tissue concentrations. Higher exposure levels or longer-term exposures would not be affected to the same degree, as administered TCDD levels would likely be large compared to initial body burden or low-level feed stock exposure.

Given the limitations described above, establishing an unambiguous progression of effects is extremely problematic given the lack of sufficient data. Identifying a RfD that could be considered to be protective against tumorigenesis in humans based on these data and models is subject not only to the determination of effective low doses for the RfDs in Table 5-17 but also to the determination of effective exposures that could be considered to be protective of all other tumor types in female S-D rats as well as all other animal models. The latter would entail identifying precursors that are sufficient in themselves for progression to tumorigenesis for all tumor types. Given the disparate sequence of hypothesized key events following TCDD-induced AhR activation for the tumor types for which some information is available, AhR binding/activation is the only key event that is likely to be shared across tumor types. No appropriate quantitative data on AhR binding/activation by TCDD in relevant animal models were located; therefore, an illustrative RfD based on TCDD AhR activation could not be developed.

Simon et al. (2009) present a similar analysis for the liver tumors observed in the NTP (2006) study, showing a progression of effects from early biochemical events to irreversible liver toxicity, culminating in tumorigenesis. While illustrative of the putative tumor-promoting MOA for TCDD, the limitations of using such an approach within the context of an assessment of the overall carcinogenic risk of TCDD as detailed above still apply. Simon and colleagues also

present RfDs for liver tumors and several precursor endpoints. All the RfDs presented in Simon et al. (2009) are essentially equivalent and are 1 to 3 orders of magnitude higher than the RfDs for equivalent endpoints presented in Table 5-17. These discrepancies are partly due to the fact that the Emond PBPK models (Emond et al., 2004, 2005, 2006; see also Section 3.3.4) used in this document predicts lower TCDD intakes for similar tissue concentrations than the CADM kinetic model (Carrier et al., 1995a, b; Aylward et al., 2005b) used by Simon and colleagues. However, a larger contributor to these discrepancies is the use of a chemical-specific adjustment factor (CSAF) of 0.1 for the toxicodynamic component of the interspecies uncertainty factor by Simon et al. (2009), while EPA used an uncertainty factor of 3. EPA does not find that the *in vitro* evidence presented by Simon et al. in support of a CSAF of 0.1 for interspecies toxicodynamics meets the burden of proof necessary for a reduction in this uncertainty factor.

5.2.3.4.1.4.3. *Effect of dose metric on linearity of response.*

EPA agrees that there is evidence for sublinearity in the TCDD carcinogenic response in rats. The lung and liver tumor endpoints in the NTP (2006) study using female S-D rats are best fit with a sublinear model. Except for squamous cell carcinoma of the oral mucosa and combined carcinomas and adenomas of the pancreas, the multistage model fits to tumor incidence data were second or third degree (see Table 5-13 and Appendix F). Figure 5-6 shows the multistage model fit for cholangiocarcinomas as a representative example. For all other tumor incidence data in all other cancer bioassays, the multistage model fit was linear (first degree), when based on either administered dose or modeled blood concentrations (see Appendix F). In further evaluating the combined liver tumor incidence for female S-D rats in Kociba et al. (1978) (as re-evaluated by Goodman and Sauer, 1992; "Kociba/G&S data"), using an unconstrained dichotomous Hill model, a progression from supralinearity to linearity to sublinearity is obtained when fitting the model to the data based on progressively more relevant dose metrics. Figure 5-7 shows the Hill model fits to the Kociba/G&S data based on administered dose, modeled blood concentrations and modeled liver AhR-bound concentrations.[32] Details of the model fits are presented in Table 5-18. The Hill coefficient parameter values are also given in the caption to Figure 5-7. A Hill coefficient of 1 indicates

[32]The modeled TCDD compartmental concentrations were obtained using the Emond rodent PBPK model described in Section 3.

linearity, while values less than 1 indicate supralinearity and values greater than 1 indicate sublinearity. The TCDD blood concentration metric is used in this document as it was determined to be the most relevant toxicokinetically-equivalent metric across species. However, as discussed in Section 3, the AhR-bound TCDD concentration may be the most relevant metric for liver effects, but this metric is also the most uncertain. Uncertainty aside, the tendency for the model fits to move towards sublinearity with increasing relevance of dose metric suggests that the "true" tumorigenic response may be sublinear in female S-D rat liver. The progression towards sublinearity is likely to be due to the failure of the less-relevant metrics to reflect the dose-dependent elimination of TCDD in the liver. The range of the administered doses in Kociba et al. (1978) is much larger than that of the effective internal TCDD concentrations. This "stretches out" the response relative to dose, resulting in a supralinear model fit. The blood-concentration metric (as determined using the Emond PBPK model) partially accounts for dose-dependent elimination, but does not account for the partition of effective (AhR-bound) and ineffective (CYP1A2-bound) TCDD in the liver[33], thus yielding only a partial correction to the shape of the dose-response curve (see Section 3). Thus, hypothesizing that the "true" response for liver tumors is sublinear in female S-D rats and assuming that the same mechanisms are operant in humans, a zero-slope at zero dose relationship could be proposed for liver tumors, allowing for an RfD approach for liver tumorigensis (as described above). This approach, however, would not be considered to be protective against all tumors given the lack of appropriate mode of action information for many of the known tumor types, particularly squamous cell carcinoma of the oral mucosa in this animal model.

On balance, however, the evidence is only suggestive and the arguments are speculative. Even if a sublinear response could be shown for rodent tumors, the greater interindividual variability in the human population would tend to shift the response towards linear (Lutz, 1999; Hattis, 1996). Hattis (1996), however, points out that, assuming a lognormal distribution of susceptibilities, increasing the variance may result in a linear or supralinear shape for a large portion of the response curve but never brings the relationship to a true low-dose linear form. How close to zero the linearity extends and how much it encompasses the risks of concern depends on the specifics of the particular human variability scenario.

[33] Modeling the response against the whole liver concentrations results in a supralinear fit similar to administered dose (Hill coefficient = 0.71).

5.3. DERIVATION OF THE TCDD ORAL SLOPE FACTOR AND CANCER RISK ESTIMATES

EPA was able to derive candidate OSFs for all cancer mortality from human epidemiologic studies as well as for individual and combined tumor incidence from rodent cancer bioassays. Each of these studies were selected for TCDD dose-response modeling using the study inclusion criteria outlined in Section 2. The derivation of these OSFs can be found for the epidemiologic data in Section 5.2.3.1 and for the rodent bioassay data in Section 5.2.3.2.

The OSFs based on epidemiologic studies from three cohorts ranged from 3.75×10^5 to 2.5×10^6 per mg/kg-day (see Tables 5-1 and 5-3). For the animal data, OSFs based on individual tumors were developed for 25 study/sex/endpoint combinations, and the results ranged from 1.8×10^4 to 5.9×10^6 per mg/kg-day (see Table 5-14). The OSFs based on combined tumors were developed for 7 study/sex combinations, and the results ranged from 3.2×10^5 to 9.4×10^6 per mg/kg-day (see Table 5-15).

As recommended by expert panelists at EPA's 2009 Dioxin Workshop (U.S. EPA, 2009c) and in the 2005 Cancer Guidelines (U.S. EPA, 2005), EPA has chosen to give higher consideration to the human epidemiologic data rather than the animal bioassay data in developing an OSF for TCDD. Candidate OSFs derived from the human data are consistent with the animal bioassay OSFs; specifically, the human OSFs fall within the same range as the animal bioassay OSFs. Because all the human and animal studies were considered to be of high quality and yielded similar ranges of OSFs, EPA has chosen to rely on the epidemiologic data for OSF derivation.

The strengths and limitations of the five epidemiological studies meeting the inclusion criteria for cancer dose-response modeling are summarized in Table 5-19. Among the human studies, the occupational TCDD exposures in the NIOSH and Hamburg cohorts are assumed to be reasonably constant over the duration of occupational exposure. In contrast, the TCDD exposure patterns in the Seveso and BASF cohorts are associated with industrial accidents; as a consequence, the exposure patterns are acute, high dose followed by low-level background exposure. Such exposure patterns similar to those experienced by the BASF and Seveso cohorts have been shown to yield higher estimates of risk when compared to constant exposure scenarios with similar total exposure magnitudes (Kim et al., 2003; Murdoch and Krewski, 1988; Murdoch et al., 1992). Thus, EPA has judged that the NIOSH and Hamburg cohort response data are more

relevant than the BASF and Seveso data for assessing cancer risks from continuous ambient TCDD exposure in the general population.

The NIOSH (Steenland et al., 2001; Cheng et al., 2006) and Hamburg (Becher et al., 1998) cohort studies report cumulative TCDD levels in the serum for cohort members. The most significant difference among the Cheng et al. (2006) analysis and those of Steenland et al. (2001) and Becher et al. (1998) is the method used to back-extrapolate exposure concentrations based on serum TCDD measurements. Steenland et al. (2001) and Becher et al. (1998) back-extrapolated exposures and body burdens using a first-order model with a constant half-life. In contrast, Cheng et al. (2006) back-extrapolated body burdens using a kinetic modeling approach that incorporated concentration- and age-dependent elimination kinetics.

Although all three of these are high-quality studies, the kinetic modeling used by Cheng et al. (2006) is judged to better reflect TCDD pharmacokinetics, as currently understood, than the first-order models used by Steenland et al. (2001) and Becher et al. (1998). EPA believes that the representation of physiological processes provided by Cheng et al. (2006) is more realistic than the assumption of simple first-order kinetics and this outweighs the attendant modeling uncertainties. Furthermore, the use of kinetic modeling is consistent with recommendations both by the NAS and the Dioxin Workshop panel.

However, as discussed in Section 3.3.2, the kinetic model that they employed does have certain limitations, including the fact that it has been calibrated based on a relatively small number of human subjects. In addition, their kinetic model does not allow body mass index (BMI; and hence fat content) to vary with age, which may bias the model results. Nonetheless, EPA prefers the increased technical sophistication of the dose estimates used in the cancer mortality risk estimates derived from Cheng et al. (2006) to those derived from Steenland et al. (2001).

EPA, therefore, has decided to use the results of the Cheng et al. (2006) study for derivation of the TCDD OSF based on total cancer mortality as calculated by EPA using data and models from the Cheng et al. (2006) study as described in Section 5.2.3.1.2. Table 5-3 shows the oral slope factors at specific target risk levels (OSF_{RL}s) which range from 3.7×10^{5} to 1.3×10^{6} per (mg/kg-day). EPA recommends the use of an OSF of 1.3×10^{6} per (mg/kg-day) when the target risk range is 10^{-5} to 10^{-7}.

5.3.1. Uncertainty in Estimation of Oral Slope Factors From Human Studies

A substantial degree of uncertainty is associated with the estimation of slope factor values and cancer risk specific doses for TCDD based on the epidemiological studies. In some instances, the influence of a given factor is theoretically amenable to analysis, but such investigation is limited by the availability of sufficiently detailed data to support such an analysis. In other cases, only very broad ranges can be placed on the uncertainty associated with a given feature of the analysis, or uncertainties must be discussed qualitatively.

5.3.1.1. *Uncertainty in Exposure Estimation*

The major technical challenge within each of the epidemiological studies was developing relevant and precise estimates of exposure. While Warner et al. (2002) collected blood samples relatively close to the time of the Sevesso accident and could reasonably estimate peak exposures based on these collected samples, in the case of the Becher et al. (1998), Ott and Zober (1996), Steenland et al. (2001), and Cheng et al. (2006) studies, the major exposure issues included the following

- Selecting (an) appropriate dose metric(s) for dose-response modeling,
- Estimating serum TCDD levels for the entire cohort based on measurements from a smaller number of the subjects in the cohort collected long after the occupational exposures had occurred, and then assigning exposures to the remaining members of the cohort based on qualitative job classifications.
- Estimating time-weighted average tissue doses (e.g., lipid-average serum concentration over time) based on single samples taken at one point in time. (Except for the Becher et al. [1998] analysis where one of the study strengths was their estimate of TCDD half life, which utilized repeated measurements from a subset of their cohort).

In the Becher et al. (1998), Steenland et al. (2001), and Cheng et al. (2006) studies, dose-response modeling was performed using ppt-years lipid-adjusted serum concentration as the primary dose metric for TCDD. The choice of serum concentration was natural in the sense that the serum TCDD was the only direct measurement of exposure or dose that was available. In addition, as discussed in Section 3.3.4, serum concentration is a reasonable index of total tissue concentration (target organ dose), and lipid-adjusted serum concentration provides a reasonable index of TCDD in the fatty components of tissues. Ott and Zober (1996) used ng/kg body

weight at the time of the accident as the primary dose metric, and EPA later converted these to units of ppt-years lipid-adjusted serum concentration.

The decision to use cumulative serum concentrations (ppt-years) as the primary dose metric for carcinogenicity is based on the understanding that time weighted concentrations (over a chronic exposure period) are the most appropriate dose measures for cancer risk assessment. This may not be strictly true if cancer induction by TCDD is considered to be a "threshold process." However, as discussed in Section 5.2, there are reasonable grounds to believe that the assumption of low-dose linearity is reasonable for TCDD, especially when calculating population risks where the effects of interindividual variability must be taken into account.

In addition to the issue of low-dose thresholds, the rationale for using cumulative dose metrics also can fail at high doses if the adverse response in question involves a step that is saturable (e.g., where there is a maximum level of response that cannot be exceeded owing to a rate-limited process). There is some evidence for such a phenomenon in the NIOSH cohort where cancer risks in the highest exposure group (>50,000 ppt-years) appear to saturate, and the response decreases at this level. Steenland et al. (2001) suggest that the apparent saturation of dose-response in this cohort may be due, at least partially, to exposure misclassification, rather than to an actual reduction in response per unit exposure.

The uncertainty associated with differences in the exposure patterns is important to consider across the five epidemiologic studies. Steenland et al. (2001), Cheng et al. (2006), and Becher et al. (1998) studied cohorts exposed to elevated TCDD levels over a long period of time, while Ott and Zober (1996) and Warner et al. (2002) studied cohorts exposed to TCDD levels significantly above background at one point in time but the exposures and likely the TCDD body burdens declined significantly following these periods of elevated exposure. Both these chronic and acute exposures can be analyzed in terms of cumulative exposure to TCDD. Use of such a metric requires an assumption that the "actual" cancer potency associated with a cumulative dose where much of the dose is received at a single point in time and then gradually eliminated would be similar to the cancer potency of the same cumulative dose received over a longer period of time and also gradually eliminated. While EPA believes that such an assumption is not unreasonable, the experiment of Kim et al. (2003), which showed statistically significant increase in liver effects due to a peak TCDD dose when compared to chronically-dosed Sprague-Dawley rats administered the same levels of TCDD when measured as a cumulative dose,

suggests that additional analyses of cumulative and peak TCDD dose measures may need to be conducted.

There are uncertainties associated with the approaches used to estimate TCDD exposures in the members of the occupational epidemiologic studies for which no measurement data were available. To impute TCDD levels for workers without measured samples, all four occupational epidemiologic studies matched workers for whom measured TCDD samples had never been reported to workers with measured TCDD levels based on job histories. The NIOSH cohort is used to illustrate some of the uncertainties. In the NIOSH cohort, the subset of workers (roughly five percent of the total cohort) with blood serum data comprised surviving members of the cohort (in 1988), and therefore, their age distribution would have differed from the rest of the cohort. For each worker in this subset, the following data were available: (1) job classification information, (2) employment history, and (3) serum TCDD measures. All of the workers in this subset were employed at a single plant where the work histories were less detailed than at other plants, and many of the workers at this plant had the same job title and were employed during the same calendar period. There is an assumption that workers with same job title and work history were exposed to the same TCDD levels within a plant and across plants; this obviously does not account for exposure heterogeneity.

Both Steenland et al. (2001) and Cheng et al. (2006) addressed the potential for exposure measurement error in TCDD estimates and possible exposure misclassification. For the highest exposure workers, Steenland et al. (2001) and Cheng et al. (2006) found weak, "noisy," and/or negative exposure-response relationships. Steenland et al. (2001) suggests that possible explanations for this observation include the saturation of effects at the upper end of the dose-response curve, instability of the TCDD exposure estimates based on the limited number of highly exposed individuals, and the increased probability of exposure misclassification for workers whose job histories indicate the highest exposures. As Steenland et al. (2001) reported, some of the highest exposures might have been inaccurately estimated because they occurred in workers exposed to short-term, high-dose exposures during spill clean-up. Cheng et al. (2006) used sensitivity analyses to examine this measurement error issue and evaluated the potential for exposure misclassification by using ln-transformed TCDD ppt-years. The authors also removed all observations with exposures within the lower and upper 1, 2.5, or 5th percentiles of the TCDD ppt-year distribution and also removed observations within just the upper 1, 2.5, or 5th

percentile of TCDD ppt-years. These sensitivity analyses provided similar results. An additional concern is that exposure errors might distort the exposure distribution in the population, which generally spreads the response out over a wider dose range. This serves to increase the variance of the regression model, altering both the POD and the corresponding OSF.

Becher et al. (1998) only considered workers from a single plant but their analysis included workers employed in five different job locations within the plant. The influence of worker location on slope factor estimates does not appear to be further explored and may represent a source of uncertainty.

To estimate long-term body burden metrics from the serum TCDD measurements, Steenland et al. (2001) employed simple first order kinetic elimination rate model with a half-life of 8.7 years. Limitations of this approach include (1) the average elimination half-life among the study subjects may not be 8.7 years given differences between the study population and the Ranch Hand population from which the value was estimated, (2) use of a single-value estimate fails to take into account the inherent variability in elimination half life among the individual workers, and (3) it fails to take into account variations in elimination kinetics throughout the lifetime of the exposed worker due to change in body fat, age, etc. The impact of these potential sources of bias on the estimates of time-integrated body burden cannot be quantitatively assessed. However, Steenland et al. (2001) noted that modest changes in elimination half-life (to 7.1 years) had only a very small impact on risk estimates.

Cheng et al. (2006) estimated past body burdens using the CADM approach (described in Section 3) (Aylward et al., 2005a, b) rather than a half-life estimate. As noted above, the incorporation of concentration- and age-dependent elimination into this approach has significant advantages over the use of a constant elimination half-life. However, as discussed in Section 3.3, the CADM has only been subject to limited testing against human validation data sets, so the degree to which its advantages are realized in practice cannot be easily assessed. There are no available human data in the low dose region, the region of interest to this assessment, to compare with the CADM (or Emond) model predictions.

Becher et al. (1998) developed half life estimates based on multiple TCDD blood measures in 48 individuals from this cohort. These half life estimates were then used to back calculate TCDD concentrations at the end of each worker's employment, accounting for age and percentage of body fat. This cohort-specific information may provide a better exposure estimate

than Steenland et al. (2001) or Ott and Zober (1996) who used similar kinetic approaches. However, the comparison of the accuracy of the exposure estimates across the cohorts is not easily assessed. There are several assumptions and important uncertainties involved in modeling TCDD exposures in these cohorts. The study authors have invoked different kinetic assumptions when extrapolating measured levels of TCDD in sera backward in time to estimate higher chronic or peak dosage (i.e., there is uncertainty in these back-calculations that includes assumptions regarding elimination kinetics). There is also uncertainty in applying such estimates to other members of the cohort based on similar characteristics (e.g., job category).

5.3.1.2. *Uncertainty in Shape of the Dose-Response Curve*

Another source of uncertainty is the nature of the dose-response curve in the low dose region of interest for risk assessment for environmental exposures (e.g., 5−10 ng/kg-day). Most of the dose-response curves appear reasonably linear in this region. These epidemiologic data, however, are based on occupational studies in which exposures were often several orders of magnitude higher than environmental exposures. Data from these studies are quite sparse in the low dose region, and only one study examined uncertainty due to the low dose region. Steenland and Deddens (2003) attempted to analyze this region specifically by fitting threshold curves to the NIOSH data in which there was no extra risk from exposure until some specific level. However, this model did not fit as well as models without a threshold. In general, the usual assumption of linearity in the low dose region seems reasonable when using epidemiologic data given the lack of data in this region that precludes the rejection of linearity.

There is uncertainty in the extrapolation of the OSF to the low dose region (e.g., <1 ng/kg-day). EPA developed the cancer assessment in this document assuming the slope in the low-dose region of the dose-response curve is linear; the decision was made due to the lack of sufficient evidence to support an assumption of nonlinearity as outlined in the 2005 Cancer Guidelines (U.S. EPA, 2005). Similarly, there is uncertainty as to whether a threshold exists for TCDD-induced toxicity leading to tumorigenesis and the dose associated with such a threshold, if it exists, is unknown. EPA chose to model this dose-response without a threshold because there is insufficient evidence to support an assumption of a threshold.

It also is noteworthy that the shapes of the exposure-response in several of these studies, based on the published statistical models, is indicative of a response that tends to tail off or

"plateau" at high cumulative exposures to TCDD. This phenomenon has been seen in many studies of occupational carcinogens, and may reflect a number of things including exhaustion of people susceptible to cancer, saturation of biological pathways which are part of the pathway to cancer, and increased error measurement of dose at high levels biasing dose-response towards the null (Stayner et al., 2003).

5.3.1.3. *Uncertainty in Defining the Reference Population*

Another source of uncertainty using human epidemiologic data is due to the lack of completely unexposed populations. Epidemiologic data in TCDD studies is based on comparing health outcomes from populations which experienced elevated TCDD exposures to outcomes from populations which experienced lower exposures, rather than to a population with zero TCDD exposure (They also compared outcomes in members of the cohort that were more highly exposed to TCDD to outcomes in members that were less exposed to TCDD). This lack of a completely unexposed population in epidemiology studies is inevitable given that truly unexposed human populations do not exist. Therefore, the extra risk calculated in the form of ED_{01}s, is inevitably the extra risk above a low background exposure due to dioxins in the general environment. Typically, the general population in western countries where the epidemiologic studies have been done have had serum levels on the order of 5 ppt. Hence, the extra risks may be considered as those incurred by added exposure above these background doses. For example, an ED_{01} of 18 ng/kg, which yields an excess risk of cancer mortality of 1%, would mean 18 ng/kg above a typical background level of 5 ng/kg (5 ppt).

5.3.1.4. *Uncertainty in Cancer Risk Estimates*

It is important to remember the overall uncertainty associated with cancer risk estimates for TCDD can be approximately bounded by evaluating the spread of estimates from the several epidemiological studies that have been conducted. In the 2003 Reassessment, EPA noted that the range of ED_{01}/lower bound of the 95% confidence interval on the dose that yields a 1% effect (LED_{01}) values identified in the then-available dose-response studies (Steenland et al., 2001; Becher et al., 1998; Ott and Zober, 1996) implied an approximate range of all-cancer mortality slope factors between 1.2 to 2.5 per ng/kg-day. The equivalent oral cancer slope factors resulting from the analysis of the newer Cheng et al. (2006) study are below and at the low end of this

range (0.37−1.2 per ng/kg-day for fatal cancer risks between 1×10^{-3} and 1×10^{-7}). This suggests that the highest quality epidemiological studies all yield risk estimates that agree to within about a factor of 7. This relatively close agreement between the studies of the different cohorts lends additional confidence to the slope factor derivation. EPA notes that each of these estimates is derived from an occupational cohort, and significant differences likely exist (e.g., exposure levels and routes) between these cohorts and those exposed to TCDD environmentally (i.e., primarily through the food chain).

5.3.2. Other Sources of Uncertainty in Risk Estimates From the Epidemiological Studies

There are a number of other aspects of the Steenland et al. (2001), Cheng et al. (2006), Becher et al. (1998), and Ott and Zober (1996) studies that may contribute uncertainty to the cancer risk estimates. First, all studies that meet the criteria (with the exception of Warner et al., 2002) measure cancer mortality rather than cancer incidence. This presumably biases the slope factor downward relative to that which would be calculated for cancer incidence, which would give a truer picture of the total health impacts of TCDD exposures on the general population. In the NIOSH cohort, roughly one-third of the fatal cancers were from lung cancer. Because of the high case mortality rate associated with lung cancer during the period of cohort evaluation, the slope factor estimated for cancer mortality might not be much lower than that calculated for cancer incidence. Estimation of cancer incidence in the general population associated with TCDD exposure would require assumptions related to the relative survival and age-specific cancer risks in the exposed population compared to the NIOSH cohort or the Hamburg cohort; insufficient data are available to support such an analysis.

The general issues associated with potential confounding effects were noted and discussed in the 2003 Reassessment. In addition to smoking and lifestyle factors that might affect cancer risks, intra-individual variation in TCDD kinetics and susceptibility could also affect the relationship between exposure and cancer risk. One specific example of such a confounding factor would be variation in TCDD elimination half-lives. For example, if a large proportion of the 256 observed cancer deaths occurred in NIOSH workers with longer half lives, this could bias the slope factor downward because higher doses would be associated with cancer cases in the deceased workers.

Another source of uncertainty in the exposure/dose assessment in the Steenland et al. (2001), Cheng et al. (2006), Becher et al. (1998), and Ott and Zober (1996) studies is the limited of information related to exposure to dioxin-like compounds (DLCs). In addition, the Hamburg cohort was also exposed to hexachlorocyclohexane (HCH) and lindane. The oral slope estimates derived by these studies attribute all the observed excess cancer mortality solely to TCDD exposures. This assumes that occupational exposure was entirely to TCDD, with no other occupational exposure to dioxins and furans. However, because TCDD typically occurs as a component of a mixture with other DLCs, this assumption could leads to a positive bias in the slope factors estimates derived from these epidemiologic studies if confounding is present. The magnitude of the potential bias can be estimated in a general way through the estimation of risks for plausible mixtures of DLCs and TCDD exposures in the cohort with the same composition as the Steenland et al. (2001) and Cheng et al. (2006) studies, but the detailed data required to perform such an analysis on the NIOSH cohort are not available.

Cheng et al. (2006) assessed the impact of possible confounding by conducting excluding individual plants in the modeling. If the estimated cancer risks as a function of exposure did not change too much when specific facilities were left out, then confounding was deemed unlikely. Cheng et al. (2006) likewise found little variation in risks based on these analyses. In addition to the slope factor estimated for TCDD, Becher et al. (1998) also evaluated the slope based on TEQs. They found a dose-response effect for TCDD but not for TEQ (excluding TCDD) which suggests that confounding by DLCs did not occur.

There is adequate evidence to believe age, gender, and body fat content all can have a significant impact on elimination kinetics and consequent cancer risks associated with TCDD exposure (U.S. EPA, 2003). The Hamburg cohort accounted for impacts in their kinetic analysis. There may be gender differences that affect susceptibility to TCDD exposure, and the NIOSH cohort and the Becher et al. (1998) analysis were comprised almost exclusively of men, so these differences were not systematically addressed. There are few quantitative data on which to examine the potential variation of these modifying factors on exposure estimates; however several of the studies have reported similar results based on variable half-life estimates.

Finally, none of these cancer cohorts contained any children, and the unique sensitivities of infants, toddlers, and children were not addressed. Aside from differences in exposure patterns and body fat content, the unique developmental status of children may result in a

substantially different profile of health risks (and magnitudes of those risks) than can be addressed by simply compensating on the basis of differences in body weight, food intake, etc. Further, because EPA could not develop an estimate from the Warner et al. (2002) cohort, none of the studies for cancer dose-response analyzed contained a significant percentage of women. Thus, the generalizability of the slope factor estimates to women and children is uncertain.

5.3.3. Approaches to Combining Estimates From Different Epidemiologic Studies

Meta-analyses and pooled analyses are two common approaches for combining epidemiologic study data. Meta-analyses are a useful way to combine epidemiologic data from different studies and derive a common estimate of effect, particularly when there are a large number of comparable studies that are fairly homogenous as to make them impossible to combine. A meta-analysis often involves a weighted average of effect measures, dose-response coefficients, or ED_{01}s.

Unlike a meta-analysis, a pooled analysis combines the original exposure and health outcome data across multiple studies, enabling a fit of new models to the data which were not used in the original publications. Whereas a pooled analysis of the four different cohorts considered here would be useful to explore the functional form and fit of models (either statistical or multistage) across all four cohorts, this would entail a lengthy undertaking and is not being contemplated here, due in part to concerns about the confidence in the results of such an undertaking.

5.3.3.1. *The Crump et al. (2003) Meta-analysis*

Crump et al. (2003) published a meta-analysis that incorporated data from the three studies EPA used in the quantitative dose-response modeling presented in the 2003 Reassessment (U.S. EPA, 2003). These three study populations were the NIOSH (Steenland et al., 2001), the Hamburg (Becher et al., 1998), and the BASF (Ott and Zober, 1996) cohorts. The data for the NIOSH study included six additional years of follow-up and improved TCDD exposure estimates that had not been applied to EPA's dose-response modeling in the 2003 Reassessment. This study examined the relationship between TCDD exposure and all-cancer mortality. SMR statistics that had been used in all three studies were applied.

The Crump et al. (2003) analysis was based on published data, and therefore, selection of the dose metric was limited to how aggregated data had been presented in the publications. For the NIOSH component of the analysis, the exposure data were based on worker-specific data and specific processes performed at each plant (Steenland et al., 2001). The previous approach assigned workers that had broad categories of exposure duration with the same cumulative serum level, and did not take into account the particular plant or the job assignment within the plant. The Crump et al. (2003) approach did take into account when exposure occurred in relation to the follow-up interval. The TCDD exposure metric used was a cumulative serum lipid concentration (CSLC). For the Hamburg cohort, Crump et al. (2003) used an average value from the exposure ranges provided in Flesch-Janys et al. (1998). For the BASF cohort, arithmetic averages for the dose categories were converted to TCDD CSLC intakes by dividing them by 0.25 (average body fat of 25%) and a decay rate that corresponded to a half-life of 7 years.

The outcome variable for the dose-response modeling was all cancer mortality, and CSLC was the independent variable. Crump et al. (2003) performed a series of trend tests to determine the lowest dose for which a statistically significant trend in SMR could be shown and all other lower doses. These tests also examined the highest dose in which there was no statistically significant trend using data from this dose and all other lower doses. Estimates of ED_{10}, ED_{05}, and ED_{01} for TEQ with respect to the lifetime probability of dying from cancer were calculated. This calculation assumed a first-order elimination process with a half-life of 7.6 years, a 50% systemic uptake of ingested dioxin, that dioxin concentration in serum lipid is a suitable measure for dioxin concentration in all lipid, and that all dioxin is sequestered in lipid (which comprises 25% of body weight). Age-specific mortality rates in the presence of dioxin exposure were then generated. Life-table methodology was used to calculate lifetime risks of cancer mortality.

Based on the modeling results, the hypothesis of a baseline SMR of 1.0 was rejected, and the linear model produced an SMR estimate of 1.17 (95% CI = 1.04−1.30) from these studies. The dose-response curves for the three studies were not homogeneous. Namely, the points from the BASF cohort fell below the predicted curve. Because the heterogeneity was not judged to be extreme by different statistical tests, however, the investigators used a common model in a combined analysis of the data from the three studies. The linear model provided an adequate fit of the data, and the slope associated with CSLC-ppt was 6.3×10^{-6} (95% CI = 8.8×10^{-7} to

1.3×10^{-5}). Based on goodness of fit analysis, the preferred estimate of ED_{01} was 7.7 pg/kg/day, which was six times higher than the estimate derived Steenland et al. (2001).

5.3.3.2. *EPA's Decision Not to Conduct a Meta-analysis*

From a statistical perspective, meta-analyses may not be very reliable when applied to a small number of studies. Crump et al. (2003) used only three studies. Had EPA undertaken a meta-analysis for the studies that met its criteria, most of the weight would come from the two large studies on the NIOSH and Hamburg cohorts. However, such an analysis relies on an assumption of a normally distributed between-study effect. This normality assumption cannot be assessed with only three observations, yet the meta-analysis estimate is highly sensitive to this distributional assumption (Higgens et al., 2009). Because of this limitation and the imprecision of the between-study variance estimate, statisticians often recommend forgoing meta-analysis in favor of discussing the individual studies when few studies are available (Cox, 2006; Higgens et al., 2009). Based on these considerations, EPA decided not to undertake a meta-analysis in this document.

As noted previously, Crump et al. (2003) has conducted a meta-analysis of the three cohorts considered here, i.e., the NIOSH, Hamburg, and BASF cohorts. However, Crump et al. modeled SMR data in which the cohorts were compared to the general population, rather than on internal exposure-response analyses as relied upon in this document. Their analysis included a total of 15 different SMRs from the three studies. A prior analysis of the dose-responses by Becher et al. (1998) was used (i.e., the categorical SMR analysis by Flesch-Janys et al. [1998]). Additionally, a prior analysis of the NIOSH cohort (Steenland et al., 1999) in which SMRs were calculated was used. Crump et al. (2003) found that a linear dose-response gave a good fit to the data, and used that for deriving an ED_{01}. However, they found that a supra-linear dose-response provided a better fit to the data, but rejected the supra-linear model (a power model) because of an infinite slope at zero dose. In the original publications by Becher et al. (1998) and Steenland et al. (2001), both observed a supra-linear dose-response trend. Crump et al. (2003) concluded that the ED_{01} was 45 pg/kg-day, similar to the ED_{01} of 8 pg/kg-day calculated by Steenland et al. (2001) using the same dietary units (pg/kg-day).

Table 5-1. Cancer slope factors calculated from Becher et al. (1998), Steenland et al. (2001) and Ott and Zober (1996) from 2003 Reassessment Table 5-4

Study	ED_{01} (LED_{01}) (ng/kg)	Cancer slope factor per ng/kg-day above background[a] (UCL)
Hamburg cohort Power model Becher et al. (1998)	6 (N.A.)	5.1 (N.A.)
Hamburg cohort Additive model Becher et al. (1998)	18.2 (N.A.)	1.6 (N.A.)
Hamburg cohort Multiplicative model Becher et al. (1998)	32.2 (N.A.)	0.89 (N.A.)
NIOSH cohort Piecewise linear model Steenland et al. (2001)	18.6 (11.5)	1.5 (2.5)
BASF cohort, from Ott and Zober (1996), multiplicative	50.9 (25.0)	0.57 (1.2)

[a]Assumes 25% of body weight is lipid; in humans 80% of dioxin dose is absorbed from the normal diet; the TCDD half-life is 7.1 years in humans. Background all cancer mortality rate calculated through lifetable analysis to 75 years. Summary results are for male all cancer risk, because the male lifetime (to 75 years) all cancer risk is greater than for females, leading to correspondingly higher cancer slope factors. As detailed in Part III, Chapter 8, $RelRisk(ED_{01}) = 0.99 + 0.01/Risk_{(0\ dose)}$. Based on the manner in which the dose-response data were calculated using Cox regression rate ratio analyses, risks are given as cancer slope factors for 1 pg/kg-day above background, assumed 5 ppt TCDD in lipid.

UCL = upper confidence limit.

Source: U.S. EPA (2003; Part III, Chapter 5, Table 5-4).

Table 5-2. Cox regression coefficient estimate and incremental risk for NIOSH cohort data as presented in Cheng et al. (2006) compared with Steenland et al. (2001)

Model	Cox regression coefficient estimate $(\text{ppt-year})^{-1}$	Incremental risk[a]	
		$R_0 = 0.124$	$R_0 = 0.112$
Steenland et al. (2001) result			
Piecewise linear[b]	1.5×10^{-5}	7.0×10^{-4}	6.3×10^{-4}
Unlagged exposure, Cheng et al. (2006)			
Piecewise linear[c]	1.4×10^{-6}	6.5×10^{-5}	5.9×10^{-5}
Linear, lower 95% of observations	1.6×10^{-6}	7.4×10^{-5}	6.7×10^{-5}
Linear, full data	-8.9×10^{-9}[d]	<0	<0
Lagged exposure (15 years), Cheng et al. (2006)			
Linear, lower 95% of observations	3.3×10^{-6}	1.2×10^{-4}	1.1×10^{-4}
Linear, full data	1.7×10^{-8}[d]	6.3×10^{-7}	5.7×10^{-7}

[a]Assumes 5 ppt serum lipid TCDD concentration for 75 years (unlagged) or 60 years (taking into account a 15-year lag).

[b]Cox regression coefficient from the Steenland et al. (2001) analysis as reported in the 2003 Reassessment, Part III, pp. 5–34, males, no lag. The value of this coefficient for the piecewise linear modeling was not explicitly reported by Steenland et al. (2001).

[c]Piecewise linear with cutpoint set at maximum likelihood for model fit, 452,000 ppt-years.

[d]Not statistically significantly different from 0.

Source: Cheng et al. (2006; Table IV).

Table 5-3. Comparison of lipid-adjusted serum concentrations, fat concentrations, risk specific dose estimates and equivalent oral slope factors based on upper 95th percentile estimate of regression coefficient[a] of all fatal cancers reported by Cheng et al. (2006) for risk levels of 1×10^{-3}, 1×10^{-4}, 1×10^{-5}, 1×10^{-6}, and 1×10^{-7}

Risk level (RL)	AUC_{RL}, (ppt-yr)	FAT_{RL} (ng/kg)	Risk specific dose (D_{RL}) (ng/kg-day)	Equivalent oral slope factors (OSF_{RL}) per (mg/kg-day)
1×10^{-3}	1.26×10^{-3}	1.803×10^{1}	2.73×10^{-3}	3.7×10^{5}
1×10^{-4}	1.26×10^{2}	1.803×10	1.23×10^{-4}	8.1×10^{5}
1×10^{-5}	1.26×10^{1}	1.803×10^{-1}	8.57×10^{-6}	1.2×10^{6}
1×10^{-6}	1.26×10	1.803×10^{-2}	7.77×10^{-7}	1.3×10^{6}
1×10^{-7}	1.26×10^{-1}	1.803×10^{-3}	7.62×10^{-8}	1.3×10^{6}

[a] Based on regression coefficient of Cheng et al. (2006; Table III), excluding observations in the upper 5% range (≥252,950 ppt-year lipid adjusted serum TCDD) of the exposures; where reported $\beta = 3.3 \times 10^{-6}$ ppt-years and standard error = 1.4×10^{-6}. Upper 95%tile estimate of regression coefficient (β_{95}) calculated to be $6.04 \times 10^{-6} = (3.3 \times 10^{-6}) + 1.96 \times (1.4 \times 10^{-6})$.

Table 5-4. Comparison of lipid-adjusted serum concentrations, fat concentrations, risk specific dose estimates and equivalent oral slope factors based on best estimate of regression coefficient[a] of all fatal cancers reported by Cheng et al. (2006) for risk levels of 1×10^{-3}, 1×10^{-4}, 1×10^{-5}, 1×10^{-6}, and 1×10^{-7}

Risk level (RL)	AUC_{RL}, (ppt-yr)	FAT_{RL} (ng/kg)	Risk specific dose (D_{RL}) (ng/kg-day)	Equivalent oral slope factors (OSF_{RL}) per (mg/kg-day)
1×10^{-3}	2.31×10^{3}	3.303×10^{1}	6.63×10^{-3}	1.5×10^{5}
1×10^{-4}	2.31×10^{2}	3.303×10	2.63×10^{-4}	3.8×10^{5}
1×10^{-5}	2.31×10^{1}	3.303×10^{-1}	1.67×10^{-5}	6.0×10^{5}
1×10^{-6}	2.31×10	3.303×10^{-2}	1.44×10^{-6}	6.9×10^{5}
1×10^{-7}	2.31×10^{-1}	3.303×10^{-3}	1.40×10^{-7}	7.1×10^{5}

[a]Based on regression coefficient of Cheng et al. (2006; Table III), excluding observations in the upper 5% range (≥252,950 ppt-year lipid adjusted serum TCDD) of the exposures; where reported $\beta = 3.3 \times 10^{-6}$ ppt-years.

Table 5-5. Kociba et al. (1978) male rat tumor incidence data[a] and blood concentrations for dose-response modeling

Morphology: topography	Vehicle control (ng/kg)	Low dose (ng/kg)	Medium dose (ng/kg)	High dose (ng/kg)
	0	**860**	**3,945**	**21,334**
Stratified squamous cell carcinoma of hard palate or nasal turbinates	0/85	0/50	0/50	4/50[b]
Stratified squamous cell carcinoma of tongue	0/85	1/50	1/50	3/50[b]
Adenoma of adrenal cortex	0/85	0/50	2/50	5/50[b]

[a]Source: Kociba et al. (1978; Table 4).
[b]Statistically significant by Fischer Exact Test ($p < 0.05$).

Table 5-6. Kociba et al. (1978) female rat tumor incidence data[a] and blood concentrations for dose-response modeling

Morphology: topography	Vehicle control (ng/kg)	Low dose (ng/kg)	Medium dose (ng/kg)	High dose (ng/kg)
	0	**853**	**3,942**	**21,246**
Hepatocellular adenoma(s) or carcinoma(s)	2/86	1/50	9/50[a]	18/45[b]
Stratified squamous cell carcinoma of hard palate or nasal turbinates	0/86	0/50	1/50	4/49[b]
Keratinizing squamous cell carcinoma of lung	0/86	0/50	0/50	7/49[b]

[a]Source: Kociba et al. (1978; Table 5). Incidence for Hepatocellular adenomas or carcinomas is from Goodman and Sauer (1992; Table 1); EPA calculated statistical significance as the study authors did not provide this.
[b]Statistically significant by Fischer Exact Test ($p < 0.05$).

Table 5-7. NTP (1982) female rat tumor incidence data[a] and blood concentrations for dose-response modeling

Morphology: topography	Vehicle control (ng/kg)	Low dose (ng/kg)	Medium dose (ng/kg)	High dose (ng/kg)
	0	**1,072**	**3,111**	**16,207**
Subcutaneous tissue: fibrosarcoma	0/75	2/50	3/50	4/49[b]
Liver: neoplastic nodule or hepatocellular carcinoma	5/75[c]	1/49	3/50	14/49[b]
Adrenal: cortical adenoma, or carcinoma or adenoma, NOS	11/73[c]	9/49	5/49	14/46[b]
Thyroid: follicular-cell adenoma	3/73[c]	2/45	1/49	6/47

[a]Source: NTP (1982; Table 10).
[b]Statistically significant by Fischer Exact Test ($p < 0.05$).
[c]Statistically significant trend by Chochran-Armitage test ($p < 0.05$).

Table 5-8. NTP (1982) male rat tumor incidence data[a] and blood concentrations for dose-response modeling

Morphology: topography	Vehicle control (ng/kg)	Low dose (ng/kg)	Medium dose (ng/kg)	High dose (ng/kg)
	0	**1,072**	**3,116**	**16,272**
Liver: neoplastic nodule or hepatocellular carcinoma	0/74[b]	0/50	0/50	3/50
Thyroid: follicular-cell adenoma or carcinoma	1/69[b]	5/48[c]	8/50[c]	11/50[c]
Adrenal cortex: adenoma	6/72	9/50	12/49[b]	9/49

[a]Source: NTP (1982; Table 9).
[b]Statistically significant trend by Chochran-Armitage test ($p < 0.05$).
[c]Statistically significant by Fischer Exact Test ($p < 0.05$).

Table 5-9. NTP (1982) female mouse tumor incidence data[a] and blood concentrations for dose-response modeling

Morphology: topography	Vehicle control (ng/kg)	Low dose (ng/kg)	Medium dose (ng/kg)	High dose (ng/kg)
	0	**1,064**	**3,184**	**17,406**
Subcutaneous tissue: fibrosarcoma	1/74[b]	1/50	1/48	5/47[c]
Hematopoietic system: lymphoma or leukemia	18/74[b]	12/50	13/48	20/47[c]
Liver: hepatooellular adenoma or carcinoma	3/73[b]	6/50	6/48	11/47[c]
Thyroid: follicular-cell adenoma	0/69[b]	3/50	1/47	5/46[c]

[a]Source: NTP (1982; Table 15).
[b]Statistically significant trend by Chochran-Armitage test ($p < 0.05$).
[c]Statistically significant by Fischer Exact Test ($p < 0.05$).

Table 5-10. NTP (1982) male mouse tumor incidence data[a] and blood concentrations for dose-response modeling

Morphology: topography	Vehicle control (ng/kg)	Low dose (ng/kg)	Medium dose (ng/kg)	High dose (ng/kg)
	0	**420**	**1,240**	**6,118**
Lung: alveolar/bronchiolar adenoma or carcinoma	10/71[b]	2/48	4/48	13/50
Liver: hepatocellular adenoma or carcinoma	15/73[b]	12/49	13/49	27/50[c]

[a]Source: NTP (1982; Table 14).
[b]Statistically significant trend by Chochran-Armitage test ($p < 0.05$).
[c]Statistically significant by Fischer Exact Test ($p < 0.05$).

Table 5-11. NTP (2006) female rat tumor incidence data[a] and blood concentrations for dose-response modeling[b]

System: morphology: topography	Vehicle control (ng/kg)	Low dose (ng/kg)	Low-med dose (ng/kg)	Median dose (ng/kg)	Med-high dose (ng/kg)	High dose (ng/kg)
	0	**1,408**	**3,137**	**5,393**	**9,129**	**16,361**
Liver: cholangiocarcinoma	0/49[c]	0/48	0/46	1/50	4/49	25/53[c]
Liver: hepatocellular adenoma	0/49[c]	0/48	0/46	0/50	1/49	13/53[c]
Oral mucosa: squamous cell carcinoma	1/49[c]	2/48	1/46	0/50	4/49	10/53[c]
Pancreas: adenoma or carcinoma	0/48[c]	0/48	0/46	0/50	0/48	3/51
Lung: cystic keratinizing epithelioma	0/49[c]	0/48	0/46	0/49	0/49	9/52[c]

[a]Source: NTP (2006; Table A3a).
[b]Incidence adjusted for animals <365 days on study.
[c]Statistically significant by Poly-3 Test ($p < 0.05$).

Table 5-12. Toth et al. (1979) male mouse tumor incidence data[a] and blood concentrations for dose-response modeling

Morphology: topography	Vehicle control (ng/kg)	Low dose (ng/kg)	Medium dose (ng/kg)	High dose (ng/kg)
	0	**315**	**7,814**	**50,105**
Liver tumors	7/38	13/44	21/44[b]	13/43

[a]Source: Toth et al. (1979; Table 1).
[b]Statistically significant by Chi^2 Test ($p < 0.01$).

Table 5-13. Comparison of multi-stage modeling results across cancer bioassays using blood concentrations

Study	Species	Sex	Morphology: topography	Multi-stage modeling*: stage, GoF *p*-value, LL difference	BMD_{01} (ng/kg)	$BMDL_{01}$ (ng/kg)
Kociba et al., 1978	Rat	Male	Stratified squamous cell carcinoma of hard palate or nasal turbinates	1, $p = 0.81$	3,175	1,540
			Stratified squamous cell carcinoma of tongue	1, $p = 0.47$	3356	1433
			Adenoma of adrenal cortex	1, $p = 0.78$	1,793	1,020
			Combined tumors Bayesian analysis		849	528
		Female	Hepatocellular adenoma(s) or carcinoma(s)	1, $p = 0.24$,	387	277
			Stratified squamous cell carcinoma of hard palate or nasal turbinates	1, $p = 0.97$	2,484	1,289
			Keratinizing squamous cell carcinoma of lung	1, $p = 0.63$	1,730	984
			Combined tumors Bayesian analysis		280	206
NTP, 1982	Rat	Female	Subcutaneous tissue: fibrosarcoma	1, $p = 0.18$	1,700	751
			Liver: neoplastic nodule or hepatocellular carcinoma	1, $p = 0.22$	638	402
			Adrenal: cortical adenoma, or carcinoma or adenoma, NOS	1, $p = 0.34$	878	444
			Thyroid: follicular-cell adenoma	1, $p = 0.57$	1,840	846
			Combined tumors Bayesian analysis		251	172
		Male	Liver: neoplastic nodule or hepatocellular carcinoma	1, $p = 0.85$	3,345	1,472
			Thyroid: follicular-cell adenoma or carcinoma	1, $p = 0.06$	657	380
			Adrenal cortex: adenoma	1, $p = 0.06$	2,161	665
			Combined tumors Bayesian analysis		410	243
	Mouse	Female	Subcutaneous tissue: fibrosarcoma	1, $p = 0.93$	1,849	916
			Hematopoietic system: lymphoma or leukemia	1, $p = 0.98$	622	331
			Liver: hepatocellular adenoma or carcinoma	1, $p = 0.34$	807	449
			Thyroid: follicular-cell adenoma	1, $p = 0.09$, no improvement with higher orders	1,653	779
			Combined tumors Bayesian analysis		244	162

Table 5-13. Comparison of multi-stage modeling results across cancer bioassays using blood concentrations (continued)

Study	Species	Sex	Morphology: topography	Multi-stage modeling*: stage, GoF *p*-value, LL difference	BMD_{01} (ng/kg)	$BMDL_{01}$ (ng/kg)
NTP, 1982 cont.	Mouse cont.	Male	Lung: alveolar/bronchiolar adenoma or carcinoma	1, $p = 0.09$	354	191
			Liver: hepatocellular adenoma or carcinoma	1, $p = 0.93$	115	75
			Combined tumors Bayesian analysis		88	58
NTP, 2006	Rat	Female	Liver: cholangiocarcinoma	3, $p = 0.99$, dLL = 2.93	4,173	2,277
			Liver: hepatocellular adenoma	3, $p = 0.93$, dLL = 2.10	5,631	3,596
			Oral mucosa: squamous cell carcinoma	1, $p = 0.27$	1,214	765
			Pancreas: adenoma or carcinoma	1, $p= 0.64$	5,794	2,551
			Lung: cystic keratinizing epithelioma	2, $p = 0.51$, dLL = 3.55	4,575	2,890
			Combined tumors Bayesian analysis		651	431
Toth et al., 1979	Mouse	Male	Liver: tumors	1, $p = 0.29$	203	115

*Analysis uses a chi-square goodness of fit statistic for differences in the log likelihoods ($p > 0.05$).

Table 5-14. Individual tumor points of departure and slope factors using blood concentrations

Study	Tumor Site (Sex/Species)	$BMDL_{HED}$ (ng/kg-day)	OSF (per mg/kg-day)
NTP, 1982	Liver: adenoma or carcinoma (male mice)	1.7E−03	5.9E+6
Toth et al., 1979	Liver tumors (male mice)	1.9E−03	5.2E+6
NTP, 1982	Lung: adenoma or carcinoma (male mice)	6.6E−03	1.5E+6
Kociba et al., 1978	Liver: adenoma or carcinoma (female rats)	1.2E−02	8.6E+5
NTP, 1982	Hematopoietic: lymphoma or leukemia (female mice)	1.5E−02	6.6E+5
NTP, 1982	Thyroid: follicular cell adenoma (male rats)	1.9E−02	5.3E+5
NTP, 1982	Liver: neoplastic nodule or hepatocellular carcinoma (female rats)	2.1E−02	4.9E+5
NTP, 1982	Adrenal: cortical adenoma or carinoma or adenoma, NOS (female rats)	2.4E−02	4.2E+5
NTP, 1982	Liver: adenoma or carcinoma (female mice)	2.4E−02	4.1E+5
NTP, 1982	Adrenal cortex: adenoma (male rats)	4.4E−02	2.3E+5
NTP, 1982	Subcutaneous fibrosarcoma (female rats)	5.3E−02	1.9E+5
NTP, 2006	Oral mucosa: squamous cell carcinoma (female rats)	5.5E−02	1.8E+5
NTP, 1982	Thyroid: adenoma (female mice)	5.6E−02	1.8E+5
NTP, 1982	Thyroid: follicular cell adenoma (female rats)	6.4E−02	1.6E+5
NTP, 1982	Subcutaneous fibrosarcoma (female mice)	7.2E−02	1.4E+5
Kociba et al., 1978	Lung: carcinoma (female rats)	8.0E−02	1.2E+5
Kociba et al., 1978	Adenoma of adrenal cortex (male rats)	8.5E−02	1.2E+5
Kociba et al., 1978	Nasal/Palate: carcinoma (female rats)	1.2E−01	8.2E+4
Kociba et al., 1978	Tongue: carcinoma (male rats)	1.4E−01	7.0E+4
NTP, 1982	Liver: neoplastic nodule or hepatocellular carcinoma (male rats)	1.5E−01	6.7E+4
Kociba et al., 1978	Nasal/Palate: carcinoma (male rats)	1.6E−01	6.3E+4
NTP, 2006	Liver: cholangiocarcinoma (female rats)	2.9E−01	3.5E+4
NTP, 2006	Pancreas: adenoma or carcinoma (female rats)	3.4E−01	2.9E+4
NTP, 2006	Lung: cystic keratinzing epithelioma (female rats)	4.1E−01	2.4E+4
NTP, 2006	Liver: hepatocellular adenoma (female rats)	5.6E−01	1.8E+4

Table 5-15. Multiple tumor points of departure and slope factors using blood concentrations

Study	Sex/species: tumor sites	$BMDL_{HED}$ (ng/kg-day)	OSF (per mg/kg-day)
NTP, 1982	Male mice: liver adenoma and carcinoma, lung	1.1E−03	9.4E+6
NTP, 1982	Female mice: liver adenoma and carcinoma, thyroid adenoma, subcutaneous fibrosarcoma, all lymphomas	5.3E−03	1.9E+6
NTP, 1982	Female rats: liver neoplasitc nodules, liver adenoma and carcinoma, thyroid follicular cell adenoma ,adrenal cortex adenoma or carcinoma	5.7E−03	1.8E+6
Kociba et al., 1978	Female rats: liver adenoma carcinoma, oral cavity, lung	7.3E−03	1.4E+6
NTP, 1982	Male rats: thyroid follicular cell adenoma, adrenal cortex adenoma	9.6E−03	1.0E+6
NTP, 2006	Female rats: liver cholangiocarcinoma, hepatocellular adenoma, oral mucosa squamous cell carcinoma, lung cystic keratinizing epithelioma, pancreas adenoma, carcinoma	2.3E−02	4.4E+5
Kociba et al., 1978	Male rats: adrenal cortex adenoma, tongue carcinoma, nasal/palate carcinoma	3.1E−02	3.2E+5

Table 5-16. Illustrative RfDs based on tumorigenesis in experimental animals

Study	Species, strain (sex)	Protocol	Endpoint	$BMDL_{HED}$[a] (ng/kg-day)	RfD[b] (mg/kg-day)
NTP, 1982	Mouse, B6C3F1, male	2-year gavage; *n* = 50	Liver adenoma and carcinoma, lung	1.1E-3	3.6E-11
NTP, 1982	Mouse, B6C3F1, female	2-year gavage; *n* = 50	Liver adenoma and carcinoma, thyroid adenoma, subcutaneous fibrosarcoma, all lymphomas	5.3E-3	1.7E-10
NTP, 1982	Rat, Osborne-Mendel, female	2-year gavage; *n* = 50	Liver neoplasitc nodules, thyroid follicular cell adenoma, liver adenoma and carcinoma, adrenal cortex adenoma or carcinoma	5.7E-3	1.9E-10
Toth et al., 1979	Mouse, Swiss/ H/Riop, male	1-year gavage (1-year average); *n* = 38−44	Liver tumors	6.1E-3	2.0E-10
Kociba et al., 1978	Rat, S-D, female	2-year dietary; *n* = 50	Liver adenoma carcinoma, oral cavity, lung	7.3E-3	2.4E-10
NTP, 1982	Rat, Osborne-Mendel, male	2-year gavage; *n* = 50	Thyroid follicular cell adenoma, adrenal cortex adenoma	9.6E-3	3.2E-10
NTP, 2006	Rat, S-D, female	2-year gavage; *n* = 53	Liver cholangiocarcinoma, hepatocellular adenoma, oral mucosa squamous cell carcinoma, lung cystic keratinizing epithelioma, pancreas adenoma, carcinoma	3.1E-2	1.0E-9
Kociba et al., 1978	Rat, S-D, male	2-year dietary; n = 50	Adrenal cortex adenoma, tongue carcinoma, nasal/palate carcinoma	3.1E-2	1.0E-9

[a]BMR = 0.01.
[b]UF = 30; UF_A = 3, UF_H = 10.

Table 5-17. Illustrative RfDs based on hypothesized key events in TCDD's MOAs for liver and lung tumors

Key event	Endpoint and exposure duration	$NO(A)EL_{HED}$ (ng/kg-day)	$LO(A)EL_{HED}$ (ng/kg-day)	$BMDL_{HED}$[a] (ng/kg-day)	RfD[b] (mg/kg-day)	Study
Liver tumors						
Changes in gene expression	CYP1A1 mRNA, 1 day	**1.8E-05**	3.4E-04	1.1E-01[c] (Appendix G)	6E-13[d]	Vanden Heuvel et al., 1994
Changes in gene expression	Benzo(a)pyrene hydroxylase (BPH) activity (CYP1A1), 1 day	9.2E-04	6.0E-03	**4.6E-04**[c] (Appendix G)	2E-11[d]	Kitchin and Woods, 1979
	EROD (CYP1A1), 53 weeks	none	1.4E-01	**9.5E-03**[c] (Appendix G)	3E-10[d]	NTP, 2006
Oxidative stress	DNA single-strand breaks, 90 days	none	3.3E-02	**2.2E-02**[c] (Appendix G)	7E-10[d]	Hassoun et al., 2000
	TBARS, 90 days	-	-	**4.4E-02** (Appendix G)	2E-09[d]	Hassoun et al., 2000
	Cytochrome C reductase, 90 days	-	-	**8.8E-02** (Appendix G)	3E-09[d]	Hassoun et al., 2000
Hepatotoxicity	Toxic hepatopathy, 2 years	none	**1.4E-01**	1.8E-01[c] (Appendix E)	5E-09[e]	NTP, 2006
	Hepatocyte hypertrophy, 31 weeks	9.3E-02	3.3E-01	**8.8E-03** (Appendix E)	3E-10[d]	NTP, 2006
Hepatocellular proliferation	Labeling index, 31 weeks	none	1.4E-01	**6.6E-02**[c] (Appendix G)	2E-09[d]	NTP, 2006
Lung tumors						
Metabolic enzyme induction	EROD (CYP1A1), 53 weeks	none	1.4E-01	**2.9E-04** (Appendix G)	1E-11[d]	NTP, 2006
Retinoid homeostatsis	Hepatic retinol and retinyl palmitate, 90 days	none	1.1E+00	**1.7E-01**[c] (Appendix E)	6E-09[d]	Van Birgelen et al., 1995

[a]BMR for continuous endpoints – 1 standard deviation; for quantal endpoints – 10%.
[b]Bolded NOAEL, LOAEL, or BMDL is selected POD; poorly-fitting BMDLs above the LOAEL not used.
[c]Poor BMD model fit or no good model fit.
[d]UF = 30; UF_A – 3, UF_H – 10.
[e]UF = 300; UFA – 3, UFH – 10; UFL – 10.

Table 5-18. Dichotomous Hill model fits to combined adenoma and carcinoma data from Kociba et al. (1987) as re-evaluated by Goodman and Sauer (1992)[a]

TCDD dose metric	Background response	ED_{50}[b]	Hill coefficient	Chi-square GOF p-value
Administered dose	0.0193	173	0.682	0.185
Blood concentration	0.0193	57.1	0.951	0.155
Liver AhR-bound concentration	0.0193	79.2	1.365	0.270
Whole liver concentration	0.0193	2.51×10^4	0.714	0.172

[a]Calculations performed in S-PLUS® 6.2 for Windows®.
[b]ng/kg tissue concentration except for administered dose, which is ng/kg BW per day (ng/kg-day).

Table 5-19. Comparison of principal epidemiological studies

Strengths	Weaknesses	Study
Cumulative TCDD levels in the serum were estimated on an individual-level basis for the 3,538 workers. Evaluated effect of lag periods (0 and 15 years). Measured and back-extrapolated TCDD concentrations to refine and quantify job exposure matrices, which were then used to estimate dioxin cumulative dose for each member of their entire cohort. Internal cohort comparisons (Cox regression model). Background exposure estimated.	• Exposure to other chlorinated hydrocarbons (dioxin like compounds). • Extrapolation of dose from a small subset (roughly 5%, *n* = 170) of the cohort. • Serum fat or body fat levels of TCDD were back-calculated using a simple first-order model. Half-life of TCDD is variable but simulated as a constant. Changes in the lipid fraction of body weight or presence/absence of genetic differences in humans that alter the distribution and metabolism of TCDD were not considered. • Serum lipid levels of TCDD in 1988 were measured only at one of the eight plants in the study. No follow-up measures. The estimates of dose are based on blood samples taken decades after exposure.	NIOSH cohort Steenland et al. (2001)
Cumulative TCDD levels in the serum were estimated on an individual-level basis for the 3,538 workers. TCCD dose estimates were simulated with a kinetic model that included considerations of exposure intensity and age-dependent body weight and fat levels. Evaluated effect of lag periods (0 and 15 years). Background exposure estimated. Stratified risk estimates for smoking and nonsmoking. Race and age adjustments. Internal cohort noted an inverse-dose response for high-exposure groups and thus excluded the data resulting in better associations.	• Extrapolation of dose from a small subset (roughly 5%, *n* = 170) of the cohort. • The authors reported the CADM model provided an improved fit over the one-compartmental model, but no evidence was reported regarding any formal test of statistical significance. • Serum lipid levels of TCDD in 1988 were measured only at one of the eight plants in the study. No follow-up measures. The estimates of dose are based on blood samples taken decades after exposure. • Exposure to other chlorinated hydrocarbons (dioxin like compounds). • No consideration for recent exposures to TCDD, changes in the lipid fraction of body weight or presence/absence of genetic differences in humans that alter the distribution and metabolism of TCDD could cause misclassification.	NIOSH cohort Cheng et al. (2006)

1

Table 5-19. Comparison of principal epidemiological studies (continued)

Strengths	Weaknesses	Study
• Repeated TCDD measures in serum in 48 individuals. Used to estimate half-life for study cohort. Took into account the age and body fat percentage of the workers. Measured and back-extrapolated TCDD concentrations to quantify exposures for the remaining cohort members using 5 different working areas of the plant. • Evaluated effect of lag periods up to 20 years. • Multiple statistical models used to evaluate fatal cancer slope estimates. • Background exposure estimated.	• Exposure to other chlorinated hydrocarbons (dioxin like compounds), HCH, and lindane. • Extrapolation of dose from a small subset (roughly 4%, n = 1,189) of the cohort. • Serum fat or body fat levels of TCDD were back-calculated using a simple first-order model. Presence/absence of genetic differences in humans that alter the distribution and metabolism of TCDD were not considered. • Serum lipid levels of TCDD for only 275 workers.	Becher et al. (1998); Hamburg Cohort
• Both internal and external analyses. • Adjustment for age, BMI, and smoking. • Both cancer incidence and cancer mortality data available, although results somewhat discordant, with steeper dose-response seen for cancer mortality.	• Acute dose due to accident may not be comparable to chronic dose accumulated over a long time, as in most environmental exposures. • Relatively small number of cancer deaths compared to NIOSH and Hamburg cohorts (n = 31). • Serum TCDD levels measured 30 years after accident, requiring extrapolation back in time to estimate cumulative dose over time. • Serum TCDD levels measured only on a sample of the cohort (138 out of 243), requiring assumptions about similarities in exposure scenario for other workers to estimate their exposure	Ott and Zober (1996)

Table 5-19. Comparison of principal epidemiological studies (continued)

Strengths	Weaknesses	Study
• TCDD levels measured in all 891 members of this female cohort. • Most TCDD measurements based on observed levels in stored serum at the time of the accident in 1976, no extrapolation needed to estimate past levels. • Internal analyses. • Evaluates female cancer incidence, other studies evaluate male cancer mortality. • Presumed adjustment for age and potential breast cancer confounders (15 of 21 cancers were breast cancer).	• Acute dose due to accident may not be comparable to chronic dose accumulated over a long time, which is typical of most environmental exposures. • Did not evaluate different lag periods. • Not clear if any adjustment for confounders. • Small number of cancers (n = 21). • Doses known in 1976, require assumptions about excretion over time to estimate cumulative dose (9 year half life assumed), presumed metric of primary interest. No more recent TCDD concentration data used. • Reported $\log_{10}$ transformation of the exposure estimates in their regression analysis.	Warner et al. (2002)

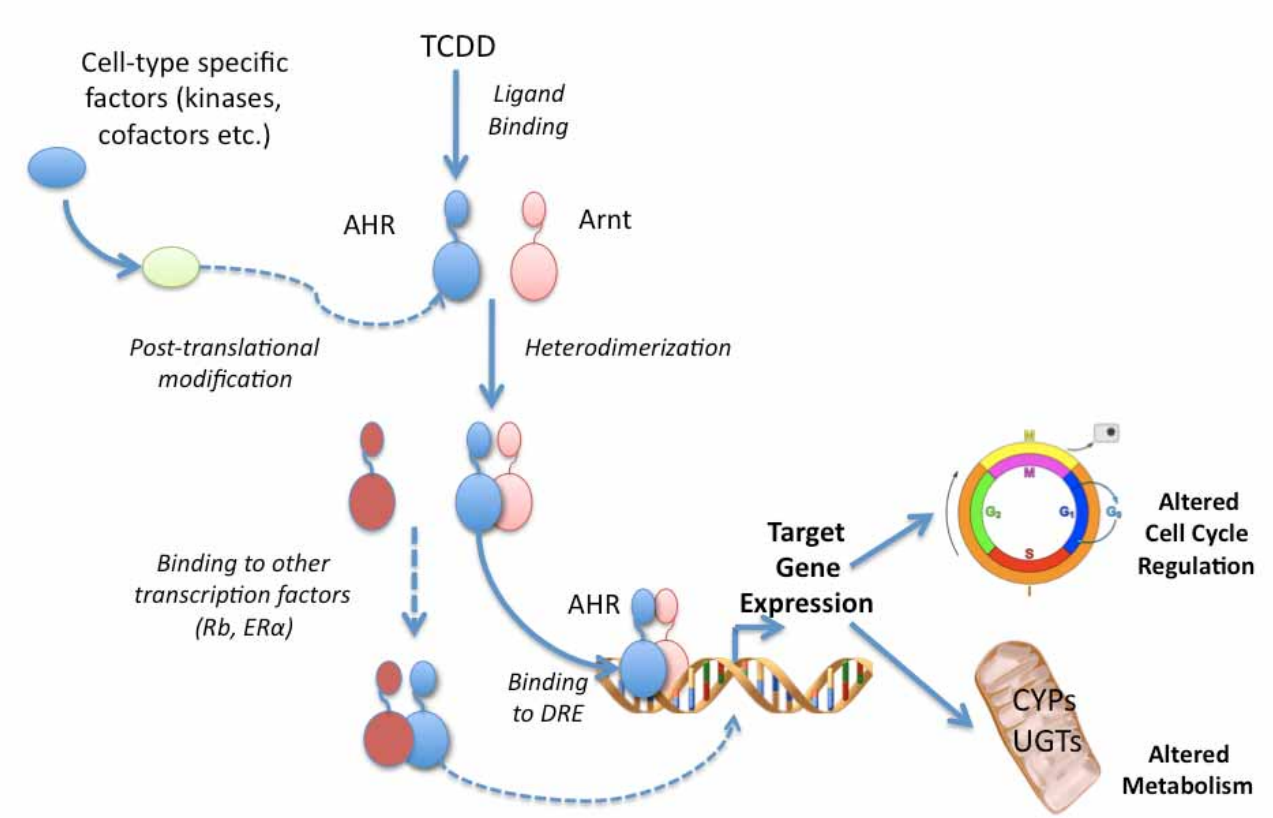


Figure 5-1. Mechanism of altered gene expression by AhR. The regulation of gene expression by TCDD in mammalian cells requires binding of the xenobiotic to the aryl hydrocarbon receptor (AhR). The AhR is part of a multi-protein complex that includes heat shock proteins and various kinases and other post-translational modifying factors. Upon ligand binding, the AhR heterodimerizes with the aryl hydrocarbon receptor nuclear translocator (Arnt) and binds to dioxin response elements (DREs) found in target genes. Alternatives to DRE-dependent gene expression exist whereby the AhR complex associates with other transcription factors and results in a cross-talk between these systems. The culmination of regulation of AhR targets genes (both increases ad decreases in transcription) results in an alteration in cellular phenotypes, including changes in intracellular metabolism and changes in cell cycle regulation.

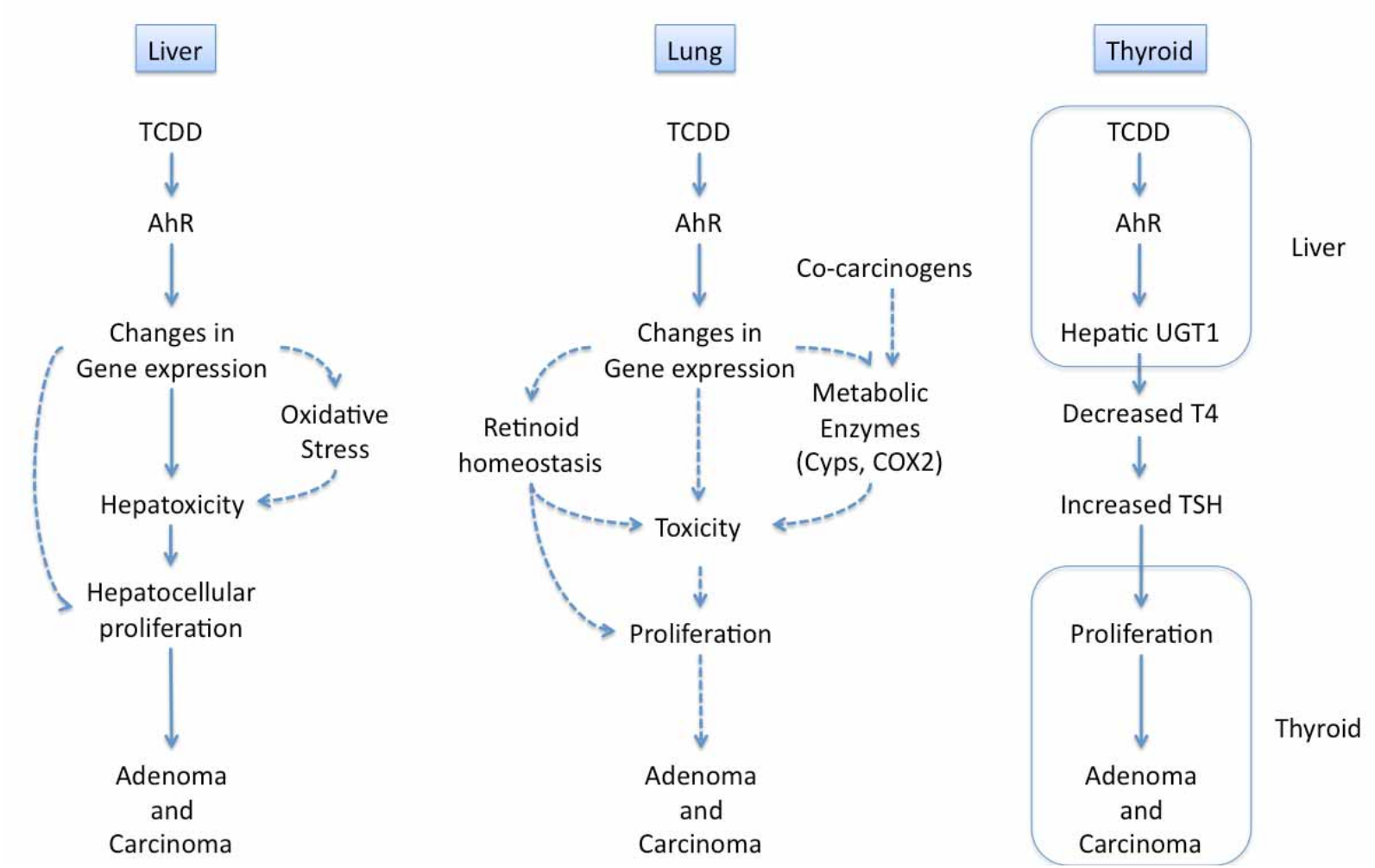


Figure 5-2. TCDD's hypothesized modes of action in site-specific carcinogenesis. See text for details. In each instance, the solid arrows depict pathways that are well-established and are associated with low uncertainty. The dashed arrows represent connections that are less established and are associated with higher uncertainty.

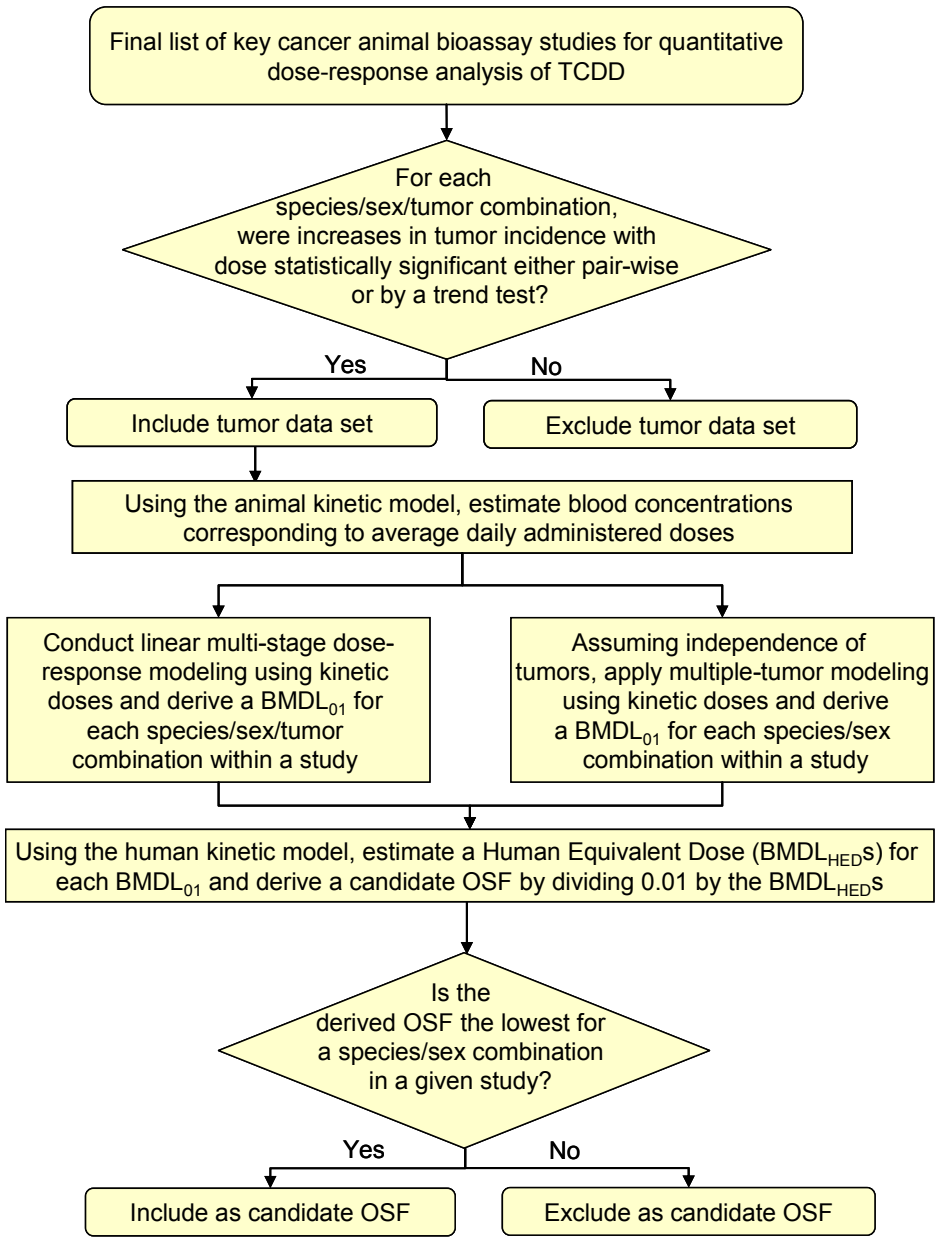


Figure 5-3. EPA's process to select and identify candidate OSFs from key animal bioassays for use in the cancer risk assessment of TCDD.

For each cancer study that qualified for TCDD dose-response assessment using the study inclusion criteria, EPA first selected the species/sex/tumor combinations with statistically significant increases in tumor incidence by either a pair-wise test between the treated group and the controls or by a trend test showing increases in tumors with increases in dose. Next, EPA used an animal kinetic model to estimate blood concentrations corresponding to the study average daily administered doses for use in dose response modeling. $BMDL_{01}$'s were then estimated for the blood concentrations by, (1) using the linearized multistage model for each species/sex/tumor combination within each study, and (2) using the linearized multistage model within a Bayesian Markov Chain Monte Carlo framework that assumes independence of tumors and modeling all tumors together for each species/sex combination within each study. Using the human kinetic model, human equivalent doses ($BMDL_{HED}$s) were then estimated for each of the $BMDL_{01}$s and oral slope factors were calculated by $OSF = 0.01/BMDL_{HED}$. The lowest OSF for a species/sex combination for either a single tumor type or all tumors combined was selected as a candidate OSF for TCDD risk assessment.

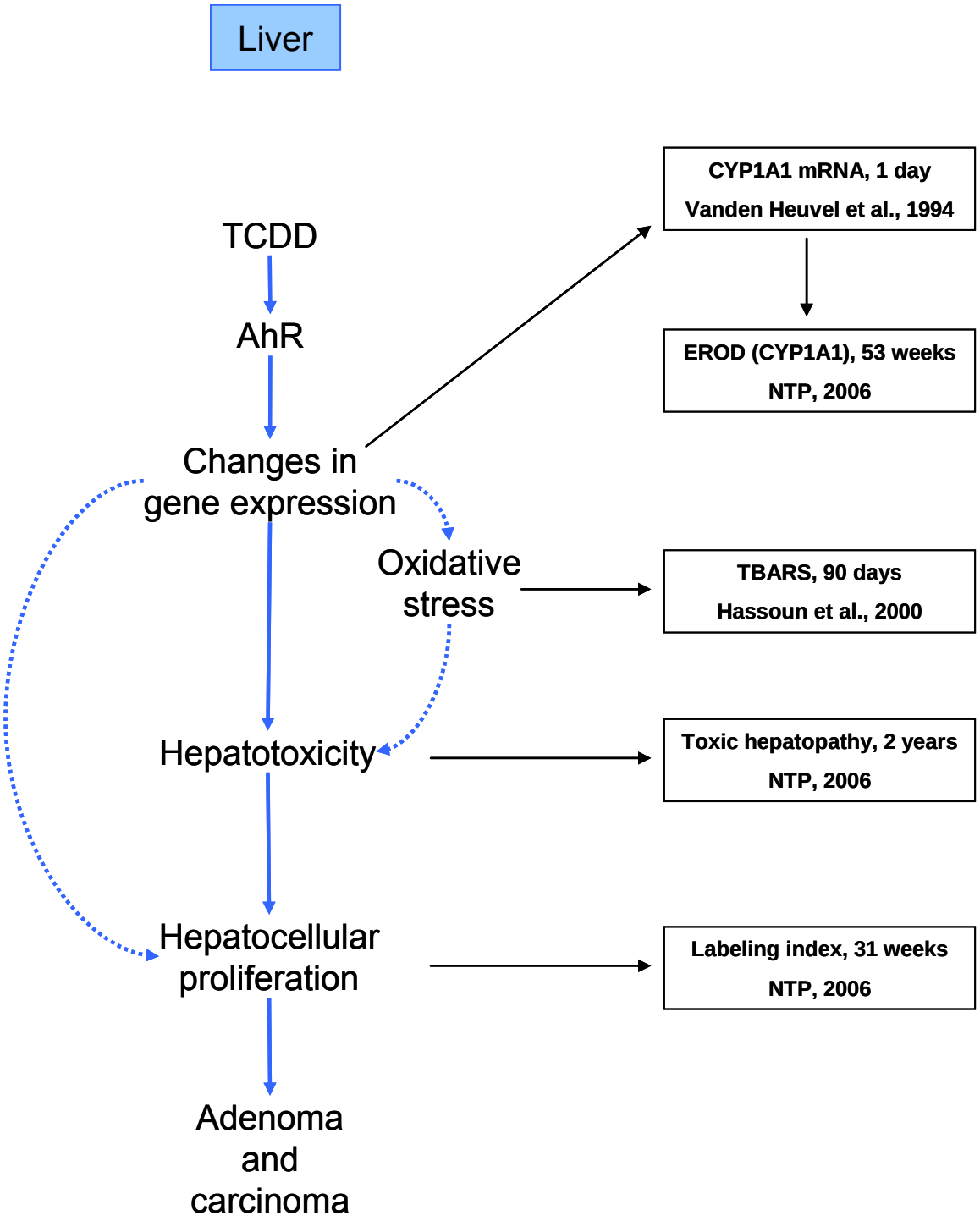


Figure 5-4. Representative endpoints for each of the hypothesized key events following AhR activation for TCDD-induced liver tumors.

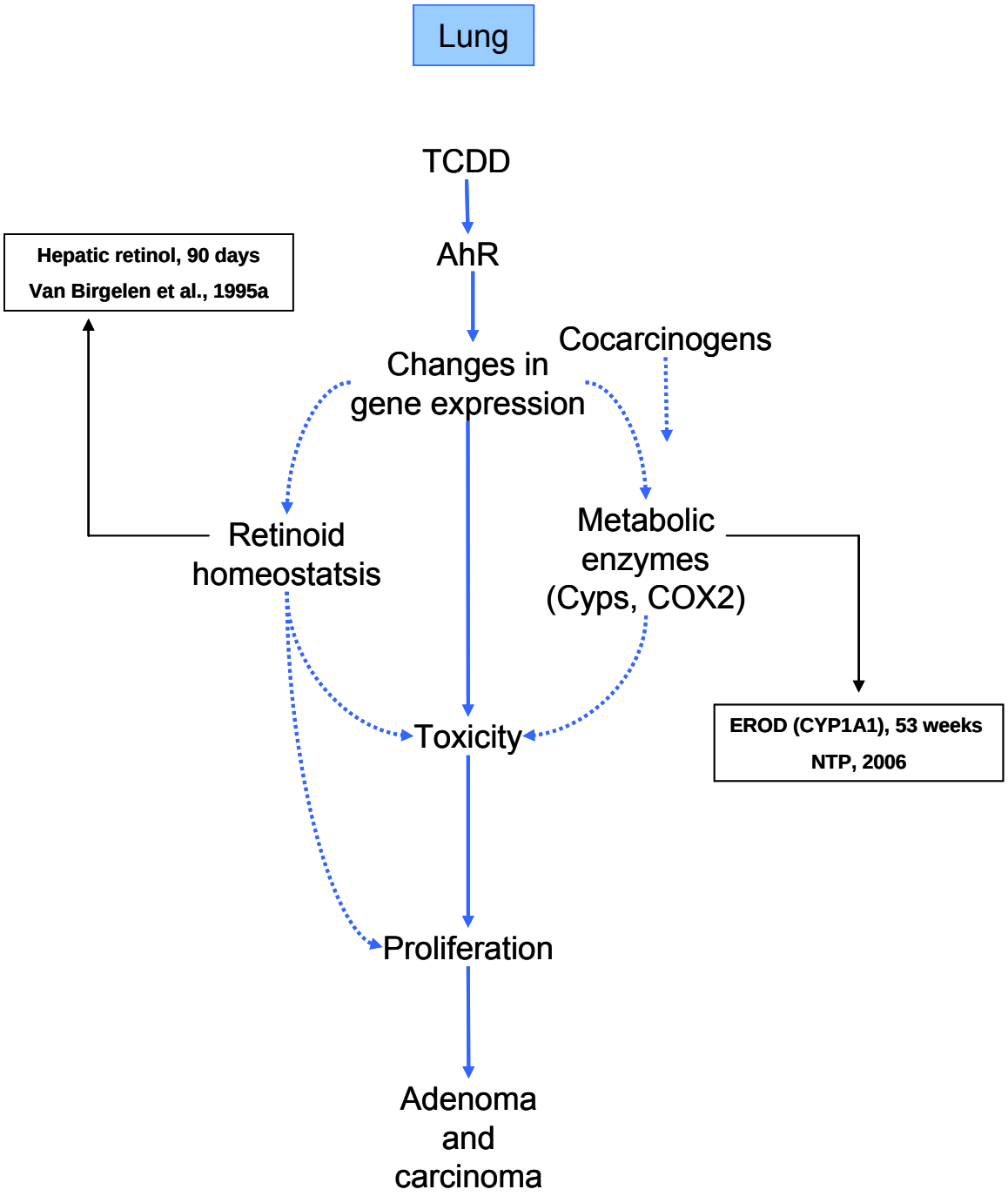


Figure 5-5. Representative endpoints for two hypothesized key events following AhR activation for TCDD-induced lung tumors.

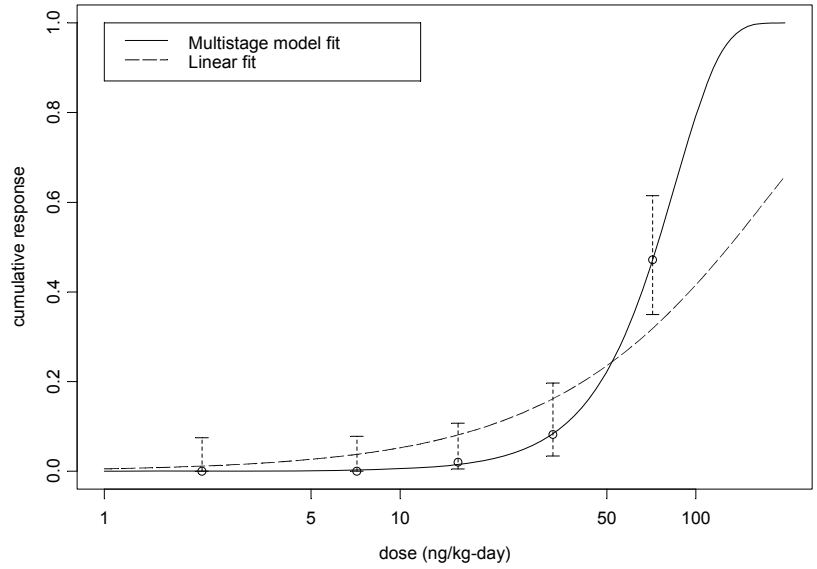


Figure 5-6. Multistage model fit to cholangiocarcinoma response data (NTP, 2006) with comparison to linear model fit.

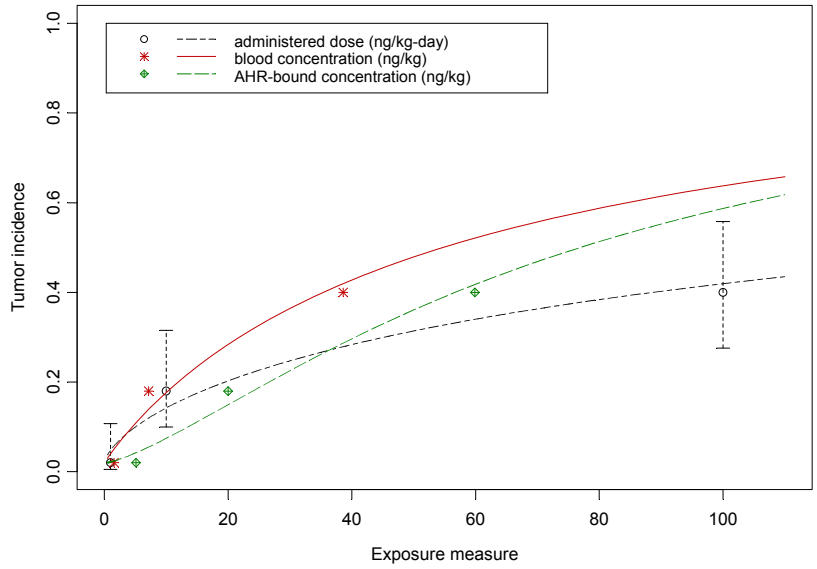


Figure 5-7. Weibull model fits to Kociba/G&S liver tumor response data with alternative dose metrics. Weibull powers are 0.68, 0.95, and 1.4 for the administered dose, blood concentration and AhR-bound concentration fits, respectively.

6. FEASIBILITY OF QUANTITATIVE UNCERTAINTY ANALYSIS FROM NAS EVALUATION OF THE 2003 REASSESSMENT

6.1. INTRODUCTION

This section focuses on the third area for improvement in the 2003 Reassessment that was identified by the National Academy of Sciences (NAS) review committee (NAS, 2006a), i.e., improving transparency, thoroughness, and clarity in *quantitative uncertainty analysis*. Although the NAS committee summarized the shortfalls in the 2003 Reassessment categorically, the elaborations within their report often contain the qualification "if possible" and do not take a position with regard to the feasibility of many of its suggestions. With appreciation for the extent of information available for dioxin, the goal of this section is to circumscribe the feasibility of a data-driven quantitative uncertainty analysis for 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) dose-response assessment. Following brief highlights of the evolution of quantitative uncertainty analysis for such applications, this section lays out definitions of key terms, reviews U.S. Environmental Protection Agency (EPA)'s position regarding cancer and noncancer endpoints, summarizes the NAS critique, and evaluates the feasibility of quantitative uncertainty analysis for TCDD within the framework of EPA's reference dose (RfD) and cancer dose-response methodologies.

6.1.1. Historical Context for Quantitative Uncertainty Analysis

The basic methods of probabilistic risk assessment were developed in the aerospace program in the 1960s, and they found their first full-scale application in the Nuclear Regulatory Commission's *Reactor Safety Study of 1975*—including accident consequence analysis and uncertainty analysis (U.S. NRC, 1975). This study, commonly referred to as the Rasmussen Report after its lead author, is considered to be the first modern probabilistic risk assessment. In the aftermath of the 1979 Three Mile Island accident, a new generation of probabilistic risk assessments (PRAs) appeared in which some of the methodological problems of the 1975 study were avoided. These advances were reflected in the Commission's *Fault Tree Handbook* (U.S. NRC, 1981) and PRA guide (U.S. NRC, 1983), which shored up and standardized much of the risk assessment methodology. An extensive chapter of the latter was devoted to uncertainty and sensitivity analysis. These documents formed the basis for standards and guidelines

established by other agencies, including the U.S. Department of Energy (U.S. DOE, 1992) and National Aeronautics and Space Administration (NASA, 2002).

In 1991, a set of U.S. Nuclear Regulatory Commission (U.S. NRC) studies known as NUREG 1150 used structured expert judgment to quantify uncertainty and set new standards for uncertainty analysis, in particular with regard to expert elicitation (U.S. NRC, 1991). This was followed by a joint U.S.-European Union (EU) program for quantifying uncertainty in accident consequence models. Expert judgment methods were further elaborated in those evaluations, as well as screening, dependence modeling and sensitivity analysis (EC, 2009). Studies building off of this work have performed a large-scale uncertainty analysis of European consequence models and provided extensive guidance on identifying important variables; selecting, interviewing and combining experts; propagating uncertainty; inferring distributions on model parameters; and communicating results, as documented by Goossens et al. (1996, 1997, 1998, 2001a-g) and others (Brown et al., 1997; Harper et al., 1995, 2002).

The National Research Council (NRC) has been a persistent voice in urging the government to enhance its risk assessment methodology since its report on risk assessment in the federal government (NRC, 1983). The Council's 1989 report, *Improving Risk Communication*, inveighed against minimizing the existence of uncertainty and noted the importance of considering the distribution of exposure and sensitivities in a population (NRC, 1989). The issue of uncertainty was a clear concern in subsequent reports, including those assessing human exposure to airborne pollutants (NRC, 1991). Building on its evaluation of *Issues in Risk Assessment* (NRC, 1993), the landmark study *Science and Judgment in Risk Assessment* (NRC, 1994) gathered many of these themes in a plea for quantitative uncertainty analysis as "the only way to combat the 'false sense of certainty,' which is *caused* by a refusal to acknowledge and (attempt to) quantify the uncertainty in risk predictions." A subsequent report *Estimating The Public Health Benefits of Proposed Air Pollution Regulations* (NRC, 2002) identified three barriers to the broad acceptance of recent EPA health benefit analyses: (1) the large amount of uncertainty inherent in these analyses, (2) the manner in which EPA deals with this uncertainty, and (3) "… projected health benefits are often reported as absolute numbers of avoided death or adverse health outcomes without a context of population size or total numbers of outcomes." The Council encouraged EPA to "explore alternative options for incorporating expert judgment into its probabilistic uncertainty analyses."

In 2006, the Office of Management and Budget released a draft bulletin proposing technical guidance for risk assessments produced by the federal government (OMB, 2006). An NRC (2007) review found many shortfalls in this proposal and recommended its retraction, and a revision was undertaken. In an early 2009 report, *Science and Decisions: Advancing Risk Assessment*, the NRC committee on improving risk analysis encouraged EPA to harmonize approaches for cancer and noncancer dose-response assessment (NRC, 2009), which involves uncertainty issues discussed in this section. Even more recently, EPA released a draft white paper, *Using Probabilistic Methods to Enhance the Role of Risk Analysis in Decision Making* (U.S. EPA, 2009d). Although not focused specifically on quantitative uncertainty analysis, there is overlap with the issues treated here, and relevant insights are anticipated from ongoing efforts in this area.

6.1.2. Definition of Terms

For purposes of this study, the following definitions are adopted:

Uncertainty Characterization: This consists of a *Structured Uncertainty Narrative* and, if the uncertainty is supported by quantitative models, *Quantitative Uncertainty Analysis*.

Structured Uncertainty Narrative: This identifies assumptions conditional on which uncertainty is to be characterized and delineates the type of arguments with supporting evidence that buttress these assumptions.

Joint Distribution/Marginal Distribution: For a set of uncertain quantities, a joint distribution is an assignment of probabilities (or probability densities) for each possible combination of values of these quantities. Each uncertain quantity has a marginal distribution, that is, an assignment of probabilities (or probability densities) to each possible value of that quantity. Assigning a marginal distribution to each quantity is not equivalent to assigning a joint distribution to the set of quantities, unless the quantities are independent; in this case the joint distribution is just the product of the margins.

Quantitative Uncertainty Analysis: This is a quantification of the uncertainty attending the use of quantitative models. It applies to a mathematical model of physical phenomena, some of whose parameter values are not known with certainty. A joint distribution is assigned to uncertain model parameters and propagated through the model to yield a joint distribution over the model output. Thus, a quantitative uncertainty analysis always has a joint distribution over model outputs as its result.

Qualitative/Informal Uncertainty Analysis: This assembles the arguments and evidence and provides an assessment of their plausibility in terms of verbal modifiers. The meaning of verbal modifiers such as "likely/unlikely" or "plausible/implausible" in the natural language is indeterminate and context dependent. The way in which these

qualifiers combine in the natural language requires critical attention from a quantitative viewpoint (if A is likely and B is likely and C is likely, is A and B and C likely?). Before the advent of personal computers, various shorthand techniques were developed for computing system risk. In control theory schemes of 'interval probabilities' were proposed which could be propagated through a system to yield bounds on system reliability. Whereas these bounds originally reflected accuracy of shorthand approximations of complex formulae, their offspring have been proposed as quantifications of uncertainty. Alternative notions of uncertainty are also proposed with the goal of simplifying the assessment and computational burden or capturing putative features of uncertainty which are overlooked in probability theory. These include possibility theory, fuzzy numbers, qualitative algebra, imprecise probabilities, belief functions, certainty factors, and the like. Nonmonotonic reasoning systems attempt to capture reasoning about knowledge, or reasoning from partial knowledge; they include default logic, defeasible logic, abductive logic, autoepistemic logic, to name a few. This is not the place to discuss foundational issues, except to remark that the practitioner should carefully explore the whole range of alternatives and critically examine the operational meaning of the primitive notions in each alternative.

Sensitivity Analysis: If a quantitative model uses "nominal values" for various input parameters, a sensitivity analysis is performed by choosing different values for these parameters and re-running the model to assess the impact of changes in these parameters on model output. Applicable methods include one- and two-at-a-time methods, design of experiments and Morris's method. They aim at estimating first-, and perhaps higher-order effects with a minimal number of model runs, by systematically varying the nominal values. In large uncertainty analyses, sensitivity analysis is used to screen variables for in-depth uncertainty quantification, and thus is part of a quantitative uncertainty analysis. As a note, the NAS committee report does not distinguish between uncertainty and sensitivity analysis. In fields which have not developed a tradition in uncertainty quantification, the spread of values generated by a sensitivity analysis is sometimes presented as a representation of uncertainty (Murphy et al., 2004). That may or may not be the case; the question is moot so long as the uncertainty on model input parameters is not quantified. Systematically varying input values is not the same as sampling input parameter values from their uncertainty distributions. In any event, a systematic approach to parameter variation is essential; simply changing a few values happenstance and generating different output does not serve a scientific purpose, and inevitably raises questions of selection bias. That said, if alternative values recommend themselves, then running these through the models can help sensitize users to parameter variations.

Cognitive Uncertainty: This concerns uncertainty regarding what is the case. This may be conceived as uncertainty over the set of possible worlds. Uncertainty over possible worlds may be represented as probability. Two interpretations or operationalizations of the probability formalism are current, the objective or frequentist interpretation and the subjective or Bayesian interpretation. These interpretations are not mutually exclusive, as subjective probabilities can and often do track relative frequencies.

Volitional Uncertainty: This concerns uncertainty regarding what to do. In the natural language, being unsure which course of action to choose is also called "uncertainty." Insofar as uncertainty on the best course of action can be translated into a claim about the state of the world, volitional uncertainty can be translated into cognitive uncertainty. A regulatory body charged with setting a speed limit is obliged to make a decision. The decision may be cautious or reckless, well or poorly motivated, wise or foolish; but it cannot be true or false. Since the decision makes no claim about the state of the world, it cannot be uncertain in the cognitive sense. The uncertainty cannot be analyzed by sampling from some distribution. However, if the decision is based on the claim that the chosen speed limit minimizes accidents while maintaining a prescribed traffic volume, that claim may be uncertain and may be subjected to quantitative uncertainty analysis. A discretionary decision of a regulatory body may entrain cognitive uncertainty, but it becomes amenable for quantitative uncertainty analysis only when it is linked to a claim about the state of the world.

Aleatoric/Epistemic Uncertainty: This distinction is also termed *Variability/Uncertainty*. A variable whose uncertainty is aleatoric for a given population takes different, uncertain, values for each member of the population. If its uncertainty is epistemic, it takes the same uncertain value for all members of the population. Issues involving uncertainty and variability or epistemic and aleatory uncertainty translate into issues of dependence, when conducting a quantitative uncertainty analysis (see Section 6.1.3.3). In its *Science and Judgment* report, NRC (1994) correctly remarks that "the amount of variability is generally itself an uncertain parameter." It is natural to ask whether a given uncertainty is aleatoric or epistemic, whereas it is awkward to ask whether this uncertainty is uncertain or variable—which explains the preference for the epistemic/aleatoric terminology.

6.1.3. Basic Requirements of a Quantitative Uncertainty Analysis

The uncertainty propagation can be performed by some rough estimation, as for example the delta method (Oehlert, 1992) or in rare cases it can be performed analytically, as in simple error propagation. Most often, however, it will be performed using Monte Carlo simulation. A joint distribution is assigned to the parameters of a quantitative model and then propagated through the model by sampling repeatedly from this joint distribution, computing model output and generating a distribution of model output. Every uncertainty analysis is conditional on initial assumptions. A "complete" uncertainty analysis is an unattainable goal; the best that can be done in practice is to identify and motivate the assumptions that are used. This section is not a how-to guide, but a to-do list of key tasks involved in any quantitative uncertainty analysis.

6.1.3.1. *Quantitative Model*

The starting point of any quantitative uncertainty analysis is a mathematical model or prescription for calculating quantities of interest. A structured narrative explains the choice of quantitative models. If some values of input parameters for this calculation are not known with certainty, then the question arises: "What is the uncertainty attending the use of this model?" This is the question a quantitative uncertainty analysis seeks to answer.

6.1.3.2. *Marginal Distributions over Model Parameter*

If the model parameters are directly measurable with sampling error, then the sampling distribution may itself be used in the quantitative uncertainty analysis. If the model parameters are fit to data that are sampled from a known or hypothesized distribution, then by resampling this distribution and refitting the model, distributions over the model parameters may be constructed. Physically-based simulation models, such as pharmacokinetic models or environmental transport models, may be solved analytically if equilibrium reaction rates (the transfer coefficients) are constant. If these rates are not constant, as when concentrations are near saturation levels, then simulation is indicated. Structured expert judgment has been applied for uncertainty quantification within the engineering community since the time of the Rasmussen Report. More recently, this approach has been "test-driven" by EPA in assessing health effects of fine particulates (Walker et al., 2009), and its potential application has been identified in the Agency's *Guidelines for Carcinogen Risk Assessment*, commonly referred to as the cancer guidelines (U.S. EPA, 2005).[34]

6.1.3.3. *Dependence Between Parameter Uncertainties: Aleatoric and Epistemic (Uncertainty and Variability)*

Two uncertain quantities are independent if knowledge about one of them does not alter our uncertainty regarding the other. The quantities are dependent if they are not independent. The role of dependence modeling in quantitative uncertainty analysis must be addressed. To illustrate, cigarette smoking and body fat are both thought to influence biomarkers for toxic

[34]The U.S. EPA (2005) cancer guidelines state: "In many of these scientific and engineering disciplines, researchers have used rigorous expert elicitation methods to overcome the lack of peer-reviewed methods and data…. These cancer guidelines are flexible enough to accommodate the use of expert elicitation to characterize cancer risks, as a complement to the methods presented in the cancer guidelines. According to NRC (2002), the rigorous use of expert elicitation for the analyses of risks is considered to be quality science."

response to dioxin exposure, such as ethoxyresorufin-*O*-deethylase (EROD) activity (Pereg et al., 2002). Both cigarette smoking and percent body fat in a random individual sampled from a target population are uncertain.[35] However, these uncertainties are not independent, inasmuch as smokers tend to have less body fat (Vanni et al., 2009).

Issues involving uncertainty and variability or epistemic and aleatory uncertainty translate into issues of dependence, when conducting a quantitative uncertainty analysis. For example, a constant used to estimate the biokinetic behavior of dioxin may be uncertain. If it is believed to be the same for every member of the population, the uncertainty is termed "epistemic." In a quantitative uncertainty analysis, this factor would be sampled from its uncertainty distribution on each Monte Carlo run and used for *all* members of the population. Body fat, in contrast, is aleatoric. We do not sample one value from the body fat distribution and use this value for *all* members of the population on each Monte Carlo run. Instead we sample a body fat value for each individual on each run. Because body fat is correlated with other relevant variables (e.g., smoking, sex, age, and socioeconomic status), all of these variables should be sampled in a manner that reflects their dependences. Kinetic constants whose uncertainty is epistemic are completely correlated across individuals: if the value is 0.5 for one individual, it is 0.5 for everyone. Body fat values vary from individual to individual, and they are correlated through a host of other variables.

6.1.3.4. *Model Uncertainty*

All models, being idealizations, are false; on this there is no uncertainty to quantify. However, the choice of model may constrain the ability to represent uncertainty in observable phenomena. Thus, in a linear low-dose model, uncertainty over a cancer slope factor may be quantified, but uncertainty regarding changes in slope at distinct low-dose regimes cannot be captured. When the model choice imposes severe and potentially unwelcome constraints on uncertainty quantification, this must be addressed. Distributions over model parameters may be

[35]Because dioxins generally distribute to body fat/lipid, the percent body fat is often used to estimate body burden; a default value of 25% is common (Connor and Aylward, 2006). However, body fat percentage varies widely between individuals, from a minimum essential level (e.g., 2% for men, 10% for women) to obesity (e.g., 38% or more for men, 42% for women). Considering that current estimates suggest 30% of the U.S. population are obese, an uncertainty analysis of dioxin risk in this population should sample individuals from their gender/body fat distribution and correlate this with other known or suspected covariates influencing toxic response (such as diet, smoking, natural and endogenous ligands, disease, and age).

selected and evaluated based on their ability to reflect uncertainty distributions over observable phenomena predicted by the models. In such cases, the uncertainty propagated through the quantitative model is not strongly model-dependent. In other cases, multiple model alternatives may be applied, whose "probability of being the true model" is known or assumed. Since different models can always be regarded as specializations of more general models, the distinction between parameter and model uncertainty is sometimes more apparent than real. The multistage and Weibull dose-response models both contain the model $Pr(x) = \gamma + (1 - \gamma)(1 - e^{-\beta 1 x})$ as a sub model, to which they collapse if other parameters are zero (multistage) or one (Weibull). Recalling that the function $1/(1 + x)$ is first order equivalent to $(1 - x)$ for small x, the same may be said for logistic models as well. In this case, these models could easily be parameterized within one family, rendering the choice between them a choice of parameter values. Similarly, the choice between sub-, supra-, and linear models is sometimes reduced to parameter estimation within a more general class of model (Hoel and Portier, 1994).

In other cases, the reduction of model uncertainty to parameter uncertainty is less natural. For example, according to the "chemoprotection model" of Greenlee et al. (2001), dioxin's binding to the aryl hydrocarbon receptor (AhR) inhibits proliferation in tumor cells and thus suppresses mammary tumors. Dose-dependent protection and cancer induction can both be true, each applying to different tissues. Although mathematical models exhibiting these twin features are under development (Kohn and Melnick, 2002), these models are not yet readily estimable from data, and the choice between them is referred to the structured narrative.

6.1.3.5. *Sampling Method*

All sampling on a computer is "pseudo random." Significant issues arise in choosing a method for sampling high-dimensional distributions with dependence. If evaluating the quantitative model is very time consuming, various 'quasi random' schemes may be applied, including Latin hypercube sampling, importance sampling, and Hammersley sampling. These methods involve trade-offs between economy and accuracy of the dependence modeling.

6.1.3.6. *Method for Extracting and Communicating Results*

When a large quantitative uncertainty analysis has been performed, the method for identifying important contributors and communicating this information to users is not

straightforward. Analysts have proposed many ways to quantify the uncertainty contribution of one variable, or set of variables, on others,[36] and his/her decision at this juncture may strongly impact the "take-home" message from the study. An importance measure that averages over an entire sample space may obscure the features of real interest. For example, the drivers of cancer induction at low doses might be different from the drivers at high doses. If the drivers of low-dose cancer induction are of interest, then importance measures that average over all doses should not be considered.

6.2. EPA APPROACHES FOR ORAL CANCER AND NONCANCER ASSESSMENT

EPA typically develops different types of toxicity information in its oral cancer and noncancer dose-response assessments, although efforts to harmonize these approaches continue to be made. Noncancer endpoints are usually assessed using the RfD methodology to derive a threshold below which there is likely no appreciable risk. (Note "risk" is used here simply as a general term indicating the potential for adverse effects.) In contrast, cancer endpoints are commonly assessed using a dose-response function with the probability of excess risk above background modeled as a linear function of dose, for doses down to zero. The RfD method relies on a point of departure (POD), and the dose-response method uses a POD if the linear model is chosen by default. From the U.S. EPA (2005) guidelines, cancer endpoints can also be assessed using the RfD methodology if the proof burden is satisfactorily met (previously described in Section 5.2.3.3).

Toxicity reference values have generally been derived for human noncancer endpoints based on a no-observed-adverse-effect level (NOAEL) or lowest-observed-adverse-effect level (LOAEL) from animal bioassay studies. This terminology suggests a biological population threshold beneath which no harm is anticipated. Reference values such as the RfD or reference

[36]A few examples may suffice. The standard Pearson correlation coefficient measures the linear dependence between two variables, positive or negative. The rank or Spearman correlation coefficient measures the monotone dependence. The correlation ratio measures the (unsigned) variance contribution of an explanatory variable on a target variable. The regression coefficient measures the expected change in standard (not natural!) units of a target variable, per standard unit change in an explanatory variable and assumes this expected change is independent of the values of the explanatory variables. Multiple correlation measures the correlation between a given variable and its best linear predictor based on another set of variables. The partial correlation of two variables given a set of other variables is their correlation after discounting the influence of the other variables. The correlation ratio, multiple correlation and the regression coefficient are not symmetric, the correlation multiple regression are always non-negative (Saltelli et al., 2000; Kurowicka and Cooke, 2006).

concentration are derived by applying uncertainty factors (UFs) to a POD. Depending on the nature of available data and modeling choice, a POD can be selected from values other than a NOAEL or LOAEL, such as an ED_x (effective dose eliciting x percent response), or a benchmark dose (BMD) or its lower confidence bound (BMDL). The BMD is the dose that induces a benchmark response (BMR), which is often chosen to represent a 5 or 10% increase in excess risk above background. The POD is divided by one or more uncertainty factors that represent knowledge gaps (see Section 6.4.1.2 for details on specific types of UFs).

An RfD is described as "likely to be without appreciable risk" but the probabilistic language has not as yet been operationalized. There is no quantitative definition of "appreciable" and no mechanism to compute risk as a function of dose, so as to ascertain that the risk is indeed not appreciable. In addition, different participants in discussions over threshold/nonthreshold models for dioxin may have different perspectives regarding how to define "appreciable risk." Under the current POD/UF framework, dose-response functions are not provided for calculating the actual risk at or above the RfD. Instead, to provide a "risk indicator" for use in screening for health hazards, a hazard quotient (HQ) is computed as the ratio of a given exposure to the RfD, or a margin of exposure (MOE) is estimated as the ratio of the POD to the human exposure level.

For the cancer endpoint, an oral cancer slope factor may be derived for human health risk assessment, typically based on tumor incidence data from an animal bioassay or on cancer incidence or deaths from an epidemiologic study. In U.S. EPA's Cancer Guidelines, cancer is predominantly thought to have no population biological threshold and a linear extrapolation to zero is applied from the POD based on extra risk above background, i.e., the probability of the endpoint decreases linearly in dose from the POD to zero or to a population background level. In the absence of sufficient evidence supporting low-dose cancer induction, the linear model is applied as a default. Cancer endpoints can also be evaluated using a "nonlinear" model. In this case, the proof burden clearly rests on the nonlinear model; it must have a preponderance of evidence to override the health-protective default choice, as described in U.S. EPA Cancer Guidelines. U.S. EPA (2005) cancer guidelines state, "When adequate data on mode of action provide sufficient evidence to support a nonlinear mode of action *for the general population* (emphasis added) and/or any subpopulations of concern, a different approach—a reference dose/reference concentration that assumes that nonlinearity—is used." In current terminology,

the RfD methodology applies to the cancer endpoint if there is sufficient evidence supporting a "zero slope at zero" model; otherwise, the linear nonthreshold model applies by default. (See Section 5.2.3.3 for a detailed discussion of linear vs. nonlinear extrapolations below the observed data, population vs. individual thresholds, and for the how U.S. EPA (2005) Cancer Guidelines are applied in choosing dose-response model forms for risk assessment.)

6.3. HIGHLIGHTS OF NAS REVIEW COMMENTS ON UNCERTAINTY QUANTIFICATION FOR THE 2003 REASSESSMENT

The NAS (2006a, b) identified a number of uncertainty characterization issues for the 2003 Reassessment; key sources of uncertainty for which quantification is suggested are highlighted in Table 6-1. The discussion in this section focuses on comments related to dose response.

There are several nuances in the NAS position relative to the need for substantial improvement in transparency, thoroughness, and clarity in quantitative uncertainty analysis for the 2003 Reassessment. These nuances concern whether the nonlinear model (note that the NAS committee uses "sublinear" and "nonlinear" interchangeably) is scientifically better supported than the linear model, and if the sublinear model is better supported, whether this is based on data or on apodictic knowledge (knowledge without uncertainty) of the mode of action (MOA). The NAS committee does not distinguish between individual and population dose-response models; however the criteria from U.S. EPA Cancer Guidelines clearly apply to population models. Assuming that the AhR-mediated MOA implies a threshold for each individual, the step to a population "zero slope at zero" model requires the following, as identified and discussed in detail in Section 5.2.3.3:

1. The distribution the individual thresholds induced by the MOA, and
2. The dose-response function for values above the thresholds.

This information can either come from data, or from apodictic knowledge of the MOA, but the burden of proof clearly rests on the nonlinear model. This section summarizes the NAS committee's overall positions. Responses to specific suggestions are given in Section 6.4 and

summarized in Section 6.5. Several excerpts of specific comments from NAS (2006a) illustrate key issues.

The NAS committee favors the nonlinear model with a threshold:

> …the committee concludes that, although it is not possible to scientifically prove the absence of linearity at low doses, the scientific evidence, based largely on mode of action, is adequate to favor the use of a nonlinear model that would include a threshold response over the use of the default linear assumption. *(p. 126)*

The committee does not state whether the threshold applies to the population, or whether each individual has his/her own threshold.

The NAS also comments on whether the nonlinear model should be used to compare with the health-protective linear default:

> Because the committee concludes that the data support the hypothesis that the dose-response relationship for dioxin and cancer is sublinear, it recommends that EPA include a nonlinear model for cancer risk estimates, but that EPA also use the current linear models for comparative purposes. *(p. 12)*

The committee does not suggest what the (population/individual) threshold might be, nor how it might be supported on the basis of data. Rather, the apodictic knowledge that there *is* a (population/individual) threshold places the dioxin risk assessment within the RfD framework, using a HQ or MOE as the basis for indicating the potential risks from exposure. The committee further asks for a quantitative characterization of the range of uncertainty:

> The committee determined that the available data support the use of a nonlinear model, which is consistent with receptor-mediated responses and a potential threshold, with subsequent calculations and interpretation of MOEs. EPA's sole use of the default assumption of linearity and selection of ED_{01} as the only POD to quantify cancer risk does not provide an adequate quantitative characterization of the overall range of uncertainties associated with the final estimates of cancer risk. *(p. 17)*

Regarding the preponderance of evidence, the committee indicates that quantitative evidence will not decide the linearity/nonlinearity (nonthreshold/threshold) issue, but knowledge (without uncertainty) of the MOA should be used:

> Quantitative evidence of nonlinearity below the POD, the ED_{01}, will never be available because the POD is chosen to be at the bottom end of the available dose-response data. ... EPA should give greater weight to knowledge about the mode of action and its impact on the shape of the dose-response relationship. *(p. 126)*

The comment continues, with the committee implicitly acknowledging that there is no evidence arguing against linearity, but that the lack of evidence should not justify using the linear model:

> The committee considers that the absence of evidence that argues against linearity is not sufficient justification for adopting linear extrapolation, even over a dose range of one or two orders of magnitude or to the assumption of linearity through zero, which would not normally be applied to receptor-mediated effects. *(p. 126)*

Whereas EPA has applied its own guidance on cancer risk assessment and adopted linearity as a health-protective default in the absence of sufficient evidence (volitional uncertainty), the NAS committee views the absence of evidence as imposing a burden of proof on the linear model (cognitive uncertainty). (As a note, terminology issues and conceptual inconsistencies within the NAS report illustrate the importance of clarity and are not unexpected, given the complexity of these issues and the nature of a committee process.)

The thrust of the NAS remarks regarding transparency, thoroughness and clarity in quantitative uncertainty analysis relevant to dose-response can be summarized as follows:

1. The uncertainty of cancer risks due to dioxin exposure should be quantified.
2. Dioxin cancer risk should be treated either as a threshold phenomenon, thus following the basic RfD methodology, or should be modeled using a sublinear dose-response function below the observed data, with the linear model used for comparison.
3. The POD should be subjected to quantitative uncertainty analysis.

A similar point of view has been indicated by others.[37] Detailed suggestions regarding specific improvements for quantitative uncertainty analysis in the 2003 Reassessment are outlined in the next section and summarized in Section 6.5.

6.4. FEASIBILITY OF CONDUCTING A QUANTITATIVE UNCERTAINTY ANALYSIS FOR TCDD

This section focuses on uncertainty analysis for TCDD dose response, which involves a range of issues as highlighted in Table 6-1.

6.4.1. Feasibility of Conducting a Quantitative Uncertainty Analysis under the RfD Methodology

This discussion applies to all noncancer endpoints of TCDD, and to cancer endpoints insofar as they fall under the RfD methodology. An RfD is obtained through the following steps

1. Choose a POD, then
2. Apply uncertainty factors (UFs) to account for knowledge shortfalls.

The method of uncertainty factors harkens back to the engineering practice of safety factors (Lehman and Fitzhugh, 1954). To illustrate, if the reference load for an engineered structure is X, then engineers might design the structure to withstand load 3X, using a safety factor of 3 to create a margin of safety. If the structure functions in a corrosive environment, another factor could be multiplied to guarantee safety for that condition, and another factor could be applied for heat, another for vibrations, and so on. The choice of values is simply based on good engineering practice, i.e., reflecting what has worked in the past. Although safety factors are still common in engineering, they are giving way to probabilistic design in many applications. The reason is that compounding safety factors quickly leads to overdesigning. Compounding safety margins for spaceflight systems may render them too heavy to fly. As our understanding of a system increases, it becomes possible to guarantee the requisite safety by leveraging our scientific understanding of the materials and processes. That of course requires

[37]For example, from Popp et al. (2006): "Overall, the evidence indicates that (1) TCDD causes cancer *via* a receptor-mediated process; (2) this dose-response is non-linear; and (3) a threshold region exists for TCDD-induced cancer below which adverse effects are unlikely to occur."

formulating clear probabilistic safety goals and developing the techniques to demonstrate compliance.

The engineering community has never sought to account for uncertainty by treating safety factors as random variables and assigning them (marginal) distributions; such an approach would not counteract the overdesigning inherent in safety factors. Many authors, including the recent national committee for *Science and Decisions* (NRC, 2009), have advocated just such a probabilistic approach to the apparent "overdesigning" of the RfD when multiple UFs are used in its derivation.

The NAS committee that evaluated the 2003 Reassessment does not discuss how to perform uncertainty analysis. But their call for substantial improvement in quantitative uncertainty analysis with TCDD falling under the RfD framework entails examining the *feasibility* of quantitative uncertainty analysis within this framework. (Note that the EPA Integrated Risk Information System database uses uncertainty factors without probabilistic interpretations; some context for this is offered in Section 6.4.1.2.)

6.4.1.1. *Feasibility of Conducting a Quantitative Uncertainty Analysis for the Point of Departure*

The POD plays a role in both the noncancer RfD methodology and the cancer dose-response methodology. The POD can be selected from various options, such as a NOAEL or LOAEL, a BMDL, or an ED_x. The feasibility of quantitative uncertainty analysis for each of these three options is considered below.

As described by Swartout et al. (1998), "The NOAEL is the highest of the tested doses in a toxicological experiment that is judged not to have caused an adverse effect" (with dose expressed as a dose rate, in mg/kg-day). A quantitative uncertainty analysis for a NOAEL or LOAEL encounters the following problem. In an experiment involving a small, positive response, the probability of seeing no response can be calculated using a binomial model with the number of exposed animals and the observed number of responses. However, in an experiment with no response, the probability of having observed a response cannot be calculated without assuming a response probability. Such an assumption could not be based on observed data. If the analysis is restricted to experiments showing a positive response, the results will be biased. The probability of a higher NOAEL or higher LOAEL can be computed, but not that of a

lower NOAEL or LOAEL. In other words, the probability that an experiment with a positive result may have yielded a null response can be estimated, but not the probability that an experiment with a null response might have yielded a positive response.

In addressing uncertainty quantification for a BMDL or ED_x, two questions must be distinguished regarding the response:

1. What is the distribution of possible doses that causes an x% increase over background?
2. What is the distribution for possible values of increase over background caused by a given dose?

The BMD is defined as the dose that realizes a BMR. It is an estimate from bioassay data that requires choosing a BMR and fitting a dose-response curve. The BMR, being a choice, is not amenable to quantitative uncertainty analysis, but the choice can be motivated in a structured narrative. The BMDL is the lower confidence limit (e.g., 5%) that can be found based on the uncertainty in the parameters of the dose-response relationship. Thus, the BMDL is addressed to the first question above, and represents in this case the 5% lower confidence band of the distribution of possible doses causing an x% increase over background. In the standard approach, the uncertainty captured by the BMDL is sampling uncertainty *conditional* on the truth of the dose-response model. Different models might fit the data equally well yet lead to different BMDLs.

The BMDL is also influenced by the constraints imposed on the parameter fitting. Suppose that the slope is expected to be greater than one, and that the maximum likelihood estimate of the slope is slightly greater than one. Since the constraint is not binding, the constrained and unconstrained model would have the same Akaike Information Criterion and would be equivalent in this sense. However, computing the BMDL with the slope constraint can lead to very different values than without this constraint. In the latter case, slope values less than one contribute to the uncertainty in the dose causing the BMR (see Cooke, 2009).

The ED_x can also be taken as a POD. It is similar in spirit to the BMD; however, as used here, the term ED_x applies when the dose causing an x% extra risk over background has actually

been observed, not estimated from a fitted dose-response model.[38] The observations are subject to sample fluctuations, and if the experiment on which the ED_x is based were repeated, different values might be found. It is helpful to consider a numerical example. Suppose a background response rate of 10% is established based on many observations of nonexposed individuals. In a given experiment, involving say 100 individuals given dose *d*, 14 individuals responded. The percent increase *x* over background (extra risk) is found by solving $14/100 = 10/100 + x \times 90/100$, or $x = 4.4\%$. We conclude that $d = ED_{4.4}$. We may assume that if the experiment were repeated with 100 new individuals sampled independently from the whole population, the response would be given by a binomial distribution with parameters (14, 100). The number of responses might be greater or smaller than four, there is a 16% chance of observing 10 or fewer responses. The response to dose *d* would not be distinguished from the background in that case, and a higher dose would be used for the POD.

The uncertainty analysis of ED_x as the POD involves addressing the second question above, without a quantitative dose-response model. A quantitative uncertainty analysis is hampered, however, by the possibility that dose *d* would produce a response less than or equal to the background, in which case the POD is indeterminate—another experiment with a different dose would be chosen as the POD. A true quantitative uncertainty analysis of ED_x as the POD would thus require a full bioassay experimental design, with binomial sampling of response rates at each dose level in the assay. Absent that, quantitative uncertainty analysis is not possible.

The interplay of choice and estimation ingredients in the POD depends on the type of POD. The main features of the above discussion are captured in Table 4-2. This table notes that the BMDL captures the uncertainty caused by sampling fluctuations *given* that the dose-response model is true. Other methods are available to compute the BMDL using (1) model-independent, observable uncertainty; (2) nonparametric Bayesian dose-response models; or (3) Bayesian model averaging (Cooke, 2009). Only the ED_x can be subject to a quantitative uncertainty analysis, and then only if a full bioassay data set is available.

[38]This definition of ED_x is adopted to distinguish the modeled response (BMD) and the observed response (ED_x), and it is more restrictive than usages common in the literature.

6.4.1.2. *Feasibility of Conducting a Quantitative Uncertainty Analysis with Uncertainty Factors*

Uncertainty factors are chosen based on a structured narrative characterizing knowledge shortfalls involving

1. Interspecies extrapolation, (UF_A: from animal data to humans).
2. Intraspecies extrapolation (UF_H: to account for human interindividual variability, considering sensitive subgroups).
3. LOAEL to NOAEL extrapolation (UF_L: to estimate the dose corresponding to no adverse effect from a reported LOAEL).
4. Subchronic to chronic extrapolation (UF_S: to estimate effects of chronic exposures from a subchronic study).
5. Database deficiency (UF_D: to extrapolate from an incomplete data set, e.g., in terms of endpoints assessed or study design, i.e., from a poor to a sufficient or rich data context).

The standard chronic RfD can represent a sensitive human (H) response to a toxic substance under chronic (C) exposure conditions. Suppose a BMDL POD were based on animal (A) data from a subchronic (S) study. The database for that chemical could be rich (R), e.g., involving multiple (and at least sensitive) species/strains, both genders, multiple life stages, with multiple endpoints observed under sound study designs. Or the data could be poor (P), with limited measurements from only a subchronic animal study (ASP) forming the basis for estimating a general reference value for humans (including sensitive subgroups) under chronic exposure conditions. For that case, the UF method would be applied as follows

$$RfD = \frac{ASP}{UF_A \times UF_S \times UF_D \times UF_H} \qquad \text{(Eq. 6-1)}$$

where UF_A, UF_S, UF_D, and UF_H are the uncertainty factors for extrapolating from animals to humans (UF_A), subchronic to chronic exposure conditions (UF_S), without adequate endpoint coverage (UF_D), and considering sensitive human subpopulations (UF_H). It is possible to assign distributions to the UFs in Eq. 6-1, and to perform a Monte Carlo analysis to produce a quantitative uncertainty distribution over the dose or value likely to be without appreciable risk

to humans for chronic exposures. Many authors have proposed such an approach,[39] and the recent NRC (2009) report on science and decisions emphatically counsels EPA to move in this direction.

The idea of using a Monte Carlo analysis to develop quantitative uncertainty distributions for the RfD is simple, but the data on which the UFs are based and the assumptions that would need to be made should be further explored. For example, it is assumed that the extrapolation from subchronic to chronic exposure (UF_S) is the same whether applied to animals or humans, and whether applied to sufficient (rich) or deficient (poor) data contexts. Swartout et al. (1998) write "Within the current RfD methodology, UF_S does not consider differences among species, endpoints, or severity of effects; the same factor is applied in all cases." In addition, due to the paucity of relevant human data, the same authors suggest the use of other endpoints as surrogates in estimating the extrapolation from animals to humans, UF_A. Further, few data exist in humans to accurately portray the interindividual variability in humans represented by UF_H. Much of the data gathered to date on distributions of UFs have aggregated across other extrapolations; that is, data from subchronic to chronic ratios are aggregated over different species and different data contexts. Finally, it may be noted that an important issue is the data on which empirical distributions of response ratios are based. Brand et al. (1999, 2001) studied the sampling behavior of response ratios and raise concerns with regard to their informativeness.

Detailed analyses of the data underlying a Monte Carlo uncertainty analysis of Eq. 6-1 would afford the possibility of verifying at least some of the assumptions and numerical estimations such an analysis must make. Even if the assumption that the same UF_S is applicable for all species, endpoints, and effect severities is thought to be biological plausible, the question of whether a given set of chemicals reflect this assumption, and hence are suitable for Monte

[39]There has been considerable work on giving a probabilistic interpretation of the UFs, including by Abdel-Rahman and Kadry (1995), Vermeire et al. (1999), Baird et al. (1996), Swartout et al. (1998), Slob and Pieters (1998), Evans and Baird (1998), Calabrese and Gilbert (1993), Calabrese and Baldwin (1995), Hattis et al. (2002), Kang et al. (2000), and Pekelis et al. (2003). These evaluations can be considered to frame what might be called a *random chemical* approach. Several authors adduce properties based on log normal distributions. Insightful studies by Kodell and Gaylor (1999; Gaylor and Kodell, 2000) suggest that uncertainty factors are independent log normal variables. Combining uncertainty factors involves multiplying the median values, and combining the "error factors" according to the formula $K_{S\times H} = exp[1.6449 \times \sqrt{(\sigma_S^2 + \sigma_H^2)}]$, where σ_S^2, σ_H^2 are the variances of $\ln(UF_S)$ and $\ln(UF_H)$. Thus $UF_S \times UF_H$ is a lognormal variable with Median($UF_S \times UF_H$) = Median(UF_S) × Median(UF_H), and 95th percentile given by Median($UF_S \times UF_H$) × $K_{S\times H}$. If U_S and U_H each have an error factor or 10, then the error factor of $UF_S \times UF_H$ is not 100 but 25.95. Several authors suggest that multiplying uncertainty factors might over-protect. Recent proposals from the National Research Council reflect the random chemical concept, and they inherit its problems (NRC, 2009).

Carlo analysis of Eq. 6-1, can only be decided by data evaluation. Data are the ultimate arbiter of whether quantitative uncertainty analysis with uncertainty factors, as currently envisioned, has sufficient evidentiary support.

6.4.1.3. *Conclusion on Feasibility of Quantitative Uncertainty Analysis with the RfD Approach*

A quantitative uncertainty analysis of the POD is not feasible for PODs based on NOAELs or LOAELs. For the BMDL, such an analysis is not appropriate because the BMDL is already a quantile from an uncertainty distribution of the BMD. However, this uncertainty distribution can be obtained in different ways that capture different aspects of uncertainty. Quantitative uncertainty analysis is feasible if the POD is based on the ED_x (as defined above) and is supported by a full set of bioassay data. A quantitative uncertainty analysis based on a probabilistic interpretation of uncertainty factors in their present form invokes strong assumptions. The data on which the distributions of uncertainty factors are based could be used to check at least some of these assumptions.

6.4.2. Feasibility of Conducting a Quantitative Uncertainty Analysis for TCDD Under the Dose-Response Methodology

Quantitative uncertainty analysis starts with a mathematical model and seeks to quantify the uncertainty attending the use of this model. Dose-response relations are mathematical models expressing the probability of response as a mathematical function of dose. For several decades, the uncertainty attending the use of dose-response models has been an abiding concern in many sectors, including the chemical and nuclear industries as well as the public health sector. Given a set of animal bioassay data, quantifying dose-response uncertainty may be approached in different ways. The differences reflect different types of uncertainty that are captured. A recent evaluation enumerates the following possible methodologies (Bussard et al., 2009):

> **Benchmark Dose Modeling (BMD)**: Choose the 'best' model, and assess uncertainty assuming this model is true. Supplemental results can compare estimates obtained with different models, and sensitivity analyses can investigate other modeling issues.

Probabilistic Inversion with Isotonic Regression (PI-IR): Define model-independent 'observational' uncertainty, and look for a model that captures this uncertainty by assuming the selected model is true and providing for a distribution over its parameters.

Non-Parametric Bayes (NPB): Choose a prior mean response (potency) curve (potentially a "non-informative prior") and a precision parameter to express prior uncertainty over all increasing dose-response relations, and update this prior distribution with the bioassay data.

Bayesian Model Averaging (BMA) (as considered here): Choose an initial set of models, and then estimate the parameters of each model with maximum likelihood. Use classical methods to estimate parameter uncertainty, given the truth of the model. Determine a probability weight for each model using the Bayes Information Criterion, and use these weights to average the model results.

The first of the above methods involves standard classical statistical methods and captures sampling uncertainty conditional on the truth of the model used. The other methods are "exotic" in the sense that they attempt to capture uncertainty that is not conditional on the truth of a given model. All have been subjected to peer review and published, but they do not enjoy the market penetration of the standard classical methods. The Bayesian models involve subjective choices of prior distributions. Insofar as the final result is largely independent of the choice of prior, these methods conform to the current starting point of focusing on data-driven methods and not appealing to structured expert judgment. (Structured expert judgment can also be considered an exotic method; an explanation of this approach falls outside the scope of this report.)

It should be emphasized that the four approaches above were illustrated as part of an integrated bench-testing exercise using simplified data for demonstration purposes. Application to actual, complicated data sets would provide additional insights into the feasibility of each per the nature of those data.

Many steps are involved in arriving at such a starting point, to frame the extrapolation of data from animal bioassays to human reference values together with consideration of epidemiological data from studies of workers (routine exposures) or the general public (including dietary exposures and those via discrete poisonings or accidental releases). Major issues are summarized below.

6.4.2.1. *Feasibility of Quantitatively Characterizing the Uncertainties Encountered when Determining Appropriate Types of Studies (Epidemiological, Animal, Both, and Other)*

The risk assessor must choose the data set(s) that will serve as a starting point for dose-response modeling. With respect to TCDD, a wealth of animal bioassay data exist in the scientific literature, across species ranging from rats, mice, guinea pigs, and hamsters to mink, dogs and monkeys, and a variety of tissues, organs, and systems. In addition, there are historical human data such as from clinical/case reports, accidental releases, and occupational exposures, including epidemiological data for several cohorts. Historical data are usually attended with large uncertainties regarding the doses actually received by individuals. Both clinical/case reports and epidemiological data pose issues with regard to possible confounders. Different TCDD studies have assessed various endpoints, ranging from chloracne to hepatic enzyme induction, abnormal fasting glucose, diabetes, lung and cardiovascular effects, neurological effects, developmental delay, dental defects, sex ratio, epididymal sperm count, endometriosis, non-Hodgkin's lymphoma, susceptibility to infection, depression, hostility, and anger, among others. Confounders that have been evaluated in the epidemiological studies include gender, body mass index, age, cigarette and alcohol consumption, and hair and eye color (Eskenazi et al., 2002a, b; Pereg et al., 2002; Baccarelli et al., 2005, 2006).

There is disagreement in the literature over the nature, scope, and quality of the historical and evolving epidemiological data for TCDD. Popp et al. (2006) put forth a position similar to that taken by the NAS:

> According to IARC [1997], the strongest overall evidence for the carcinogenicity of TCDD is for all cancers combined, rather than for any specific site. The relative risk for all cancers combined in the most highly exposed and longer-latency subcohorts is 1.4. Although IARC indicated that this relative risk does not appear likely to be explained by confounding, relative risks of this magnitude for studies of other substances, particularly with a lower confidence interval at or near one, are generally found to be the result of confounders. Few examples (perhaps only smoking and ionizing radiation in atomic bomb survivors) exist of agents known to cause an increase in cancers at many sites. This lack of precedent for a multi-site carcinogen without particular sites predominating, combined with the very small excess relative risks, means that the epidemiological findings must be treated with caution.

On the other hand, U.S. EPA Cancer Guidelines express a clear preference for epidemiological studies over animal data. The question here is whether quantitative uncertainty analyses based on bioassay data and different epidemiological data can be combined in some overall uncertainty assessment. Diverse studies are sometimes combined in meta-analyses. Different studies have different protocols, and ignoring these differences can distort results. To illustrate, placental EROD activity was 4-fold higher in a group of highly polychlorinated biphenyl (PCB)-exposed Inuit women compared to a group of women from southern Quebec with lower exposure. However the differences in EROD activity disappeared when smoking was taken into account. Had tobacco consumption been unknown in one of the groups, spurious conclusions would have been drawn (Pereg et al., 2002). Can uncertainty arising from combining studies with different protocols be taken into account in quantitative uncertainty analysis? The question is similar to that of accounting for uncertainty due to missing covariates in Cox regression (see Section 6.4.2.2).

Standard statistical tools will be of no avail, as they quantify uncertainty in model parameters estimated from data. However, exotic methods, such as Bayesian methods, probabilistic inversion, or structured expert judgment may be applicable. These methods can be applied when a quantitative model *predicts* other phenomena, even though these phenomena could not be used to estimate the model. The question of whether such methods could remain sufficiently tethered to data, or whether structured expert judgment is unavoidable, is a subject for future research.

6.4.2.2. *Uncertainty in TCDD Exposure/Dose in Epidemiological Studies*

There are many types of epidemiological studies including case controls, longitudinal and cross-sectional cohorts, and simple exposed/unexposed studies. The variety of mathematical models includes Cox proportional hazard, Poisson regression, and logistic regression. The outputs vary accordingly: relative risk ratios, odds ratios, and standardized mortality ratios (SMR, ratio of observed to expected deaths). Exposure uncertainties often concern back-casted exposures based on current serum lipid concentrations, estimated/self reported dietary habits, fish consumption, placenta lipid concentrations, and other measures.

Uncertainty in exposure is often compensated by coarsening the output, e.g., SMRs instead of dose-dependent risk factors. Packages computing the outputs routinely produce

confidence intervals that reflect sampling fluctuations, assuming truth of the model. Additional uncertainty could be factored in with exotic methods. A significant issue in epidemiological studies is the effect of omitted covariates. Omitted covariates in Cox regression will bias the estimates of effects of included covariates. If the omitted covariates are independent of the included covariates, the bias is toward zero in absolute value (Bretagnolle and Huber-Carol, 1988); if the omitted covariates are not independent, little can be said.

With regard to individual studies, it might be possible to identify specific opportunities for uncertainty quantification. This is illustrated with the study of Steenland et al. (2001) of more than 3,500 male workers exposed to TCDD-contaminated products at eight U.S. chemical plants. Each worker was assigned an exposure score based on estimated level of contact with TCDD, the degree of TCDD contamination of product at each plant over time, and the fraction of a workday in contact with the product. For 170 workers, the serum TCDD levels were also measured. The serum levels were back-extrapolated to the last time of exposure using a constant biological half life, and regressed on the exposure scores. This regression model was used to predict the dose in all workers, and predicted dose was correlated with cancer mortality. Figure 6-1 shows a scatter plot of back-casted versus predicted TCDD serum levels for the 170 workers on which the regression was based.

Given a predicted TCDD level, the uncertainty on the back-casted TCDD value could be inferred from such data by various techniques. A key question is whether the actual cancer mortalities among 170 back-casted workers are randomly placed in the conditional distribution given predicted TCDD. Imagine, in other words, that the mortalities among the 170 back-casts are colored red in the above picture. At any given level of TCDD prediction, are the red points evenly distributed, or are they shifted to the right? In principle, the correlation between mortality and back-casted TCDD level, given the predicted level, could be estimated. This amounts to estimating heteroscedasticity in the regression model. Then, for each of the 3,538 workers, given his predicted TCDD level, we could sample a back-casted TCDD level, appropriately correlating with mortality, and recompute the dose response analysis. Repeating this many times we could build up a distribution for excess lifetime cancer mortality risk.

It is instructive to step through similar issues with regard to biological half life, background and body fat. The Steenland et al. (2001) analysis assumed a constant TCDD biological half life (8.7 years). A distribution over this half life could plausibly be developed

from published sources. Assuming this half life is constant for all workers, but uncertain (epistemic uncertainty), this distribution could easily supplement the previous distribution: first sample a half life (to be applied to all workers), then estimate the regression model for this half life, and sample back-casted TCDD levels given each worker's exposure score, taking account of correlation with mortality. This works if the half life uncertainty is epistemic. However, since the half life is estimated from data, it is more reasonable to suppose that the half life varies from worker to worker (aleatoric uncertainty). Here again the correlation with mortality must be taken into account, indeed it seems reasonable to suppose that the 256 cancer deaths involved workers with longer half lives. However, there is no way ex post of determining the biological half life in the deceased workers.

Background exposure and body fat are similar to half life. The study authors held background level constant at the median level (6.1 ppt, range 2.0−19.7) for 79 nonexposed workers from whom blood was also drawn (see also Section 6.4.2.4). The full distribution of TCDD levels for these nonexposed workers could be used as well. Is it reasonable to suppose that responsive workers have background levels that are sampled randomly from this distribution, or might they not plausibly come from the high end of the distribution? The analysis also assumed a constant percentage of body fat (30%), whereas body fat percentage varies in the general population, e.g., for men this has been reported to range from 2 to 38% or more (see Footnote 32). The body fat distribution in the worker population could have been ascertained, but again the question arises, are the responsive workers sampled randomly from this distribution?

These three factors, variable half life, variable background, variable body fat percentage, might conspire to make the effective dose level among the responsive workers significantly higher than would appear in a study that assumes these factors to be constant. However, such concerns cannot be addressed in a quantitative uncertainty analysis, unless cancer mortality can be correlated with these variables. In an optimal study design, this information could be retrieved from the data, but optimality is a luxury that is seldom afforded. The optimal should not be the enemy of the good; it might be possible to estimate these correlations in some other defensible manner, in which case the effect of exposure uncertainty could be quantified and propagated. Such an analysis would involve substantial effort and should not be undertaken under assumptions that are themselves implausible. Experimental protocols do not currently

require such uncertainty quantification; perhaps they should and perhaps someday they will. In any event Steenland et al. (2001) must be applauded for conscientiously identifying these key issues.

6.4.2.3. *Uncertainty in Toxicity Equivalence (TEQ) Exposures in Epidemiological Studies*

Toxicity equivalence factors (TEFs) are used to infer the health effects of dioxin-like compounds based on their relative potencies compared to TCDD. These factors are not known with certainty, and the question arises whether uncertainty in TEFs can be incorporated into a quantitative uncertainty analysis. The process of deriving TEFs applied by the World Health Organization (WHO, 2005) is described in Van den Berg et al. (2006). Distributions of relative potencies (REPs) were developed from the scientific literature, with preference for in vivo studies, as supplemented by in vitro studies. An expert panel used a consensus process to select a TEF value for each congener, in half log steps "Thus, the TEF is a central value with a degree of uncertainty assumed to be at least ± half a log, which is one order of magnitude. However, it should be realized that TEF assignments are usually within the 50^{th} to 75^{th} percentile of the REP distribution, with a general inclination toward the 75^{th} percentile in order to be health protective" (Van den Berg et al., 2006) (see Figure 6-2).

The WHO considers the uncertainty in TEFs to span one order of magnitude (presumably log uniformly distributed). It would be tempting to use the distributions in Figure 6-2 to quantify uncertainty in the TEFs in a quantitative uncertainty analysis. However, the issue of dependence in this case is daunting. For example, should values of 1,2,3,7,8,-pentachlorodibenzofuran and 2,3,4,7,8-pentachlorodibenzofuran be sampled independently? The choice of dependence structure will have a large effect. As described by Van den Berg et al. (2006), the differences in REPs reflects differences in dosing regimens, species, endpoints, mechanisms, and calculation methods. In a quantitative uncertainty analysis one must insure that these are not double counted.

> Reasons for significant differences in REPs for the same congener can be caused by the use of different dosing regimens (acute vs. subchronic), different endpoints, species, and mechanisms (e.g., tumor promotion caused by at least two different mechanisms as for mono-*ortho*-substituted PCBs), as well as different methods used for calculating REPs. Thus, different methodological approaches used in different studies clearly provide uncertainties when deriving and comparing REPs.

If future study designs to derive REPs were more standardized and similar, the variation in REPs when using the same congener, endpoint, and species might be expected to be smaller (Van den Berg et al., 2006).

Although the TEFs themselves and the distributions underlying them are based on expert judgment, it is possible to incorporate these into a quantitative uncertainty analysis; however, it is not simply a matter of taking the distributions in Figure 6-2 to predict the results, with uncertainty, of exposure to dioxin-like compounds. The issues of dependence and double counting must first be addressed. Inasmuch as the distributions are the result of expert judgment, this would reasonably involve structured expert judgment as well. (Procedures for this type of assessment have been developed and applied, and it would entail a significant level of effort.)

6.4.2.4. *Uncertainty in Background Feed Exposures in Bioassays*

TCDD is not produced intentionally but rather is formed as a byproduct of volcano eruptions, forest fires, manufacturing of steel and certain chemicals (including some pesticides and paints), pulp and paper bleaching, exhaust emissions, and incineration. It enters the food supply primarily via aerial transport and deposition of emissions, and it bioaccumulates in animal fat. In general, food of animal origin contributes to about 80% of the overall human exposure. For example, Schecter et al. (1997) measured dioxins in pooled food samples collected in 1995 from supermarkets across the United States. Reported as parts per trillion (ppt) toxicity equivalences (TEQs), fresh water fish had the highest level (1.43); followed by butter (1.07); hotdog/bologna (0.54); ocean fish (0.47); cheese (0.40); beef (0.38); eggs (0.34); ice cream (0.33); chicken (0.32); pork (0.32); milk (0.12); and vegetables, fruits, grains, and legumes (0.07). (More recent exposure studies by Lorber et al. [2009] and others have updated these older estimates of dietary levels.)

These results illustrate that a person's dietary intake of dioxins depends on the relative intake of foods with high or low levels of contamination, and human background levels will vary accordingly. The same applies to experimental animals in bioassays, although in those cases the background intake can in principle be controlled. Some of the effects of TCDD and other AhR agonists in enhancing the early initiation stages of cancers are considered to occur as a result of prenatal exposures that are not included in the standard National Toxicology Program (NTP) bioassay protocol (Brown et al., 1998; Muto et al., 2001). Further, to enhance reproducibility

and keep statistical fluctuations to a minimum, the standard NTP assays are deliberately run on groups of animals that are relatively uniform genetically, fed uniform diets, and have the minimum possible exposures to toxicants other than the agent(s) being tested. This tends to reduce the potential for observing the consequences of potential interactive effects that might occur in the diverse human population with its variety of dietary and other exposures to a wide range of potentially interacting substances and conditions.

A critical question is the extent to which the background exposure influences the dose-response curve, and how this background should be taken into account.[40] One idea, articulated in the recent NRC (2009) report on science and decisions, involves an "interacting background." This can be implemented by computing a virtual dose B which, according to the selected dose-response model, would explain a chosen fraction of the background response. If the chosen model for dose δ is $f(\delta)$, the model can be adapted to account for an interacting background by writing $f^*(\delta) = f(\delta + B) - f(B)$. This can alter the model's behavior at zero dose.

For example, if $f(\delta) = \delta^n/(\delta^n + EC_{50}{}^n)$, the derivative $d(f)/d(\delta)$ is $n\delta^{n-1}EC_{50}{}^n/(\delta^n + EC_{50}{}^n)^2$, which goes to zero as $\delta \rightarrow 0$, if $n > 1$. However, replacing δ with $(\delta + B)$ evidently changes the derivative at zero to $nB^{n-1}EC_{50}{}^n/(B^n + EC_{50}{}^n)$. This model is not yet estimable from data, as we have no way of choosing from the available animal data the fraction of background response to be explained by the model when applied to humans (although judgments could be made if we had better information about the details of the processes that are involved in causing various human health effects). However, as a conceptual model, it serves to remind us that the manner of accounting for background exposures can influence a model's behavior in the low-dose region.

6.4.2.5. *Feasibility of Quantifying the Uncertainties Encountered When Choosing Specific Studies and Subsets of Data (e.g., Species and Gender)*

Species, strain, gender, life stage, and other characteristics of experimental animals are selected for a given study based on previous knowledge (e.g., of the species sensitivity,

[40]"Effects of exposures that add to background processes and background endogenous and exogenous exposures can lack a threshold if a baseline level of dysfunction occurs without the toxicant and the toxicant adds to or augments the background process. Thus, even small doses may have a relevant biologic effect. That may be difficult to measure because of background noise in the system but may be addressed through dose-response modeling procedures" (NRC, 2009).

availability of strains having little genetic variation for the endpoints in question, relevance of the MOA, and degree to which the endpoints are similar for humans). Many other decisions are made in designing a bioassay study; will the animals be sacrificed at the termination of the study (if not a lifetime study), or will they be allowed to live out their natural lives? What dosing regimen should be applied? How will the animals be fed and handled? Although such questions may engender uncertainty in the minds of the experimenters, and reviewers; such uncertainty is not amenable for quantitative uncertainty analysis unless and until there are quantitative models, with parameters estimable from data, that can predict the effect of these choices on the response function.

6.4.2.6. *Feasibility of Quantifying the Uncertainties Encountered when Choosing Specific Endpoints for Dose-Response Modeling*

Standard experimental protocols guide the selection of exposure/dosing conditions for a given bioassay, including the amount, delivery vehicle, route, timing, dosing frequency and duration, and dose spacing. The goal is to find the dose range where the experimental animals begin to respond adversely, to help anchor the lower end of the dose-response relationship, and to avoid multiple experiments in which all or none of the animals respond. A common recommendation is that the dose levels be chosen such that the increments in probability of response are roughly equal. Hence, the choice of endpoint, spacing, and number of animals should be made with these factors in mind. Of particular importance is the number of animals at each dose level in relation to the choice of endpoint and probability of response. Using more animals at the lower dose levels increases the probability of seeing some animals respond; on the other hand, this approach will give higher weight to the low-dose responses in model fitting and uncertainty quantification. Including many experiments at low dose with no expected response could provide deceptive graphs.

One of the most clearly nonlinear responses to high-dose dioxin exposure cited by the committee is cholangiocarcinoma (bile duct cancer) in rats (see Figure 6-3). The model $\text{Prob(dose)} = \text{dose}^n/(\text{dose}^n + EC_{50}^n)$ with Hill coefficient n was fit to these data, yielding a maximum likelihood estimate of $n = 2.79$ (Hattis, 2009). For $n > 1$, this is a zero slope at zero dose model (see Section 5.2.3.3). If we choose a linear extrapolation below the POD at dose 3.82 ng/kg-day with probability of response = 0.02, then at dose 1.7 ng/kg-day, the predicted

probability of response is 0.009; with 46 rats exposed to this dose we expect to see no responses, which is what we see. However, sublinear models will yield a higher probability of seeing nothing. A graph like Figure 6-3 would suggest that the linear extrapolation was missing the experiments at doses 1.7 and 0.52 ng/kg-day, indicating lack of fit, whereas the null response is what we should expect on the linear model.

The conclusion with regard to the feasibility of this quantitative uncertainty analysis echoes that of the previous paragraph: such uncertainty is not amenable for quantitative analysis unless and until there are quantitative models, with parameters estimable from data, that predict the effect of these choices on the response function, although sensitivity analyses can be done showing the consequences of assuming different amounts of interacting background within the context of a specific nonlinear model.

6.4.2.7. *Feasibility of Quantifying the Uncertainties Encountered when Choosing a Specific Dose Metric (Trade-Off Between Confidence in Estimated Dose and Relevance of MOA)*

The concept of dose is not straightforward. To review, U.S. EPA (2005) Cancer Guidelines provide the following taxonomy:

- *Exposure* is contact of an agent with the outer boundary of an organism.
- *Exposure concentration* is the concentration of a chemical in its transport or carrier medium at the point of contact.
- *Dose* is the amount of a substance available for interaction with metabolic processes or biologically significant receptors after crossing the outer boundary of an organism.
- *Potential dose* is the amount ingested, inhaled, or applied to the skin.
- *Applied dose* is the amount of a substance presented to an absorption barrier and available for absorption (although not necessarily having yet crossed the outer boundary of the organism).
- *Absorbed dose* is the amount crossing a specific absorption barrier (e.g., the exchange boundaries of skin, lung, and digestive tract) through uptake processes.
- *Internal dose* is a more general term, used without respect to specific absorption barriers or exchange boundaries. *Delivered dose* is the amount of the chemical available for interaction by any particular organ or cell.

Using the applied dose or absorbed dose would yield statistically more powerful results and enable more precise predictions. If it is not possible to measure these, then other available dose metrics such as potential dose or exposure are used. Due to toxicokinetic variability, different individuals receiving the same exposure may not have the same applied or absorbed dose. Hence, use of an exposure metric would add variability to the predicted results. The dose metric should be selected that (1) has the most proximate possible causal relation to the production of an adverse health endpoint, and (2) is most easily related to the units of (external) exposure that will be the basis for assessing and controlling human exposures.

6.4.2.8. *Feasibility of Quantifying the Uncertainties Encountered When Choosing Model Type and Form*

U.S. EPA (2009d) draft white paper on probabilistic methods notes: "There is no consensus on any one well-accepted general methodology for dealing with model uncertainty, although there are various examples of efforts to do so." Model uncertainty was introduced in Section 6.1.3.4. Many statistical techniques are available to evaluate model adequacy or to choose a "best" model. Although it is tempting to qualify such deliberations as "uncertainty that a model is true," one must remember that all models, being idealizations, are false. Ultimately, one is interested in uncertainty with regard to observable phenomena, not with regard to models. Models are merely tools for describing the phenomena. Nonetheless, the choice of a model constrains the ways in which uncertainty can be represented, and the question addressed in this paragraph is: How should one deal with these constraints? A recent study of uncertainty modeling in dose response (Cooke, 2009) addresses precisely this issue and provides ample technical detail to frame possible options.

Before elaborating on exotic approaches to model uncertainty, it is well advised to appreciate a feature in the standard statistical treatment of uncertainty. Consider a model based on experimental data, typically bioassay data, in which a certain number of study subjects are exposed to varying doses of a test substance, and in which the numbers of subjects exhibiting a response are tallied. Values for the parameters in the model are chosen by the principle of maximal likelihood: those values are chosen which render the data as likely as possible. According to standard practice, a model is chosen that best fits the data according to one of the

accepted criteria, such as reduced R^2, or the Akaike Information Criterion. There might be many incompatible models that are nearly as good.

One can ask the following: If the experiments on which the model is based were repeated, sampling the same number of experimental subjects from the distribution posited by the model, how much could our parameter estimates change? This is described by a joint distribution over the model's parameters, which captures sampling uncertainty under the assumption that the model is true. Now, all models are false, and as our sample sizes grow the lack of fit in the model becomes increasingly apparent. At the same time, the sample fluctuations in parameter estimates—*assuming the model is true*—become ever smaller. In very large epidemiological studies, standard statistical methods can produce razor-thin confidence bands in this way, which fail to capture experts' uncertainty regarding observable phenomena.[41]

The exotic methods sketched in the beginning of Section 6.4.2 may be viewed as attempts to deal with this feature. Probabilistic inversion methods were deployed on a large scale in the joint U.S. NRC-EU uncertainty analyses noted in Section 6.1. Distributions over model parameters are intended to capture an antecedently defined uncertainty over observable phenomena predicted by the model. This method was applied in dispersion and deposition modeling and further environmental transport models (including uptake) for radionuclides. In most cases, the observable uncertainty was based on structured expert judgment, but it has also been based on binomial uncertainty in bioassay studies. A potential drawback is that it may not prove possible to capture the observable uncertainty in this way with a classically best-fitting model, and new models may be required.

Nonparametric Bayesian methods arose in the biomedical and reliability fields. They start with a prior distribution over all nondecreasing dose-response functions, and update these with observations from a bioassay study. No further assumptions regarding parametric form are introduced, but the prior distribution remains important for doses outside the range of observations. Bayesian model averaging starts with a prior distribution over a set of candidate models, and updates this distribution with bioassay data. The method is flexible and intuitive, though attenuation of the effect of the prior on the posterior must be verified.

[41]See for example Tuomisto et al. (2008, Table 6) for a comparison of experts' uncertainty in health effects of fine particulates with uncertainties derived from sampling uncertainty from large epidemiological studies.

All these approaches represent attempts to capture "extramodel uncertainty," that is, uncertainty that is not conditional on the truth of the model. This is an active research area, and the last word on capturing extramodel uncertainty in quantitative uncertainty analysis has not been spoken. A major effort with regard to dioxin would be indicated when the strengths and weakness of the exotic methods are well understood.

6.4.2.9. *Threshold MOA for Cancer*

The NAS committee avers that knowledge of the AhR binding MOA entails that there is a response threshold for TCDD cancer induction. The differences between individual and population thresholds are not discussed, but the following two possibilities are distinguishable:

1. The threshold is the same for each individual; since human variability in AhR binding affinity is rather large (see Section 5.2.3.3), this entails that the threshold is not affected by the binding affinity.
2. The threshold varies across individuals and is related to the individual AhR binding affinity.

These two positions are different. As shown in Section 5.2.3.3, it is quite possible that each individual in a population has a threshold whereas the population dose-response relation is linear. Because the NAS committee does not distinguish which of these positions it holds, the feasibility of quantitative uncertainty analysis is examined here for both.

i. Quantitative uncertainty analysis concerns a mathematical model. In case (i), this model would show how the existence of the AhR binding would induce a threshold, independently of the strength of the binding. Assessing the feasibility of quantitative uncertainty analysis must await the elaboration of such a model.

ii. In this case, it must be shown that the distribution of thresholds, and the dose-response function above the threshold, is such as to induce a population "zero slope at zero dose" (ZS@Z) model. Recall, the burden of proof is on this (ZS@Z) model. Scoping the population variability with regard to AhR-mediated mechanisms in general, and dioxin sensitivity in particular, is an active area of research. It involves phenotyping human AhR-mediated responsiveness and relating this to polymorphisms in the human population. Harper et al. (2002) report that a 10-fold variation in binding affinity of AhR for TCDD in human placental tissue did not reveal any polymorphisms, suggesting that the relation between phenotypical and genotypical variation is tenuous at best. Tuomisto et al. (1999) demonstrate large variations in efficacy in two rat strains whose binding

affinity is similar (Long-Evans, K_d = 3.4, Han/Wistar, K_d = 3.9 [Connor and Alyward, 2006]), and they also show that this variation is endpoint-specific. The responses in both strains are similar for cytochrome P450 (CYP)1A1 induction, but very dissimilar for thymus atrophy, serum bilirubin, and mortality. Toide et al. (2003) suggest that common biochemical measures of EROD activity might be mediated by CYP1B1 and CYP1A2. The differences in serum bilirubin at doses around 10 μg/kg are about a factor of 30. Han/Wistar rats seldom die at this dose, while mortality of Long Evans rats is about 50%. The mechanisms are not understood.

Although the mass action dose-response model does not have a threshold, it is possible that certain enzymes block the receptor binding, and until these are overwhelmed, no response occurs. The availability of such enzymes may vary from individual to individual, and may or may not covary with the dissociation constant K_d. Pursuing these lines of research may result in a convincing demonstration of a population (ZS@Z) model. Such a model would express the individual threshold in terms of parameters that could be estimated with uncertainty from the data.

6.4.2.10. *Feasibility of Quantifying the Uncertainties Encountered when Selecting the BMR*

The NAS committee explicitly requested that the uncertainty attending the choice of a BMR be quantified. First of all, simply plugging other values in for the BMR does not constitute a quantitative analysis of uncertainty. The plugged-in values must be sampled from some uncertainty distribution. Since this concerns volitional uncertainty, there is no underlying distribution from which to sample, unless the choice of BMR is related to some claim about the state of the world.

6.5. CONCLUSIONS

The main conclusions regarding the feasibility of quantitative uncertainty analysis are first summarized in relation to specific suggestions made by the NAS committee. A suggested research agenda follows.

6.5.1. Summary of NAS Suggestions and Responses

On page 130 of their report (NAS, 2006a), NAS makes specific suggestions regarding uncertainty quantification. These are reformatted and presented in italics below. Following each

suggestion, a summary of the discussion in this section is given, with reference to the section in which it is addressed.

EPA should have addressed quantitatively the following sources of uncertainty:

- *Basis for risk quantification:*
    1. *bioassay data,*
    2. *occupational cohort data.*

Response: (a) Classical statistical methods yield distributions on model parameters which reflect sample fluctuations, assuming that the model is true. This type of uncertainty is taken into account in the BMDL. Exotic methods can account for uncertainty which is not conditional on the truth of a model, at least for bioassay data (see Section 6.4.2). (b) For epidemiological data, the dose reconstruction often involves assumptions which may support data driven uncertainty analysis, if sufficient data can be retrieved. Examples discussed above include back-casted TCDD level, biological half life, body fat and background (see Section 6.4.2.2). Uncertainty in the choice of bioassay data sets or choice of occupational cohort data sets is volitional, and is not quantified by sampling an input distribution. To be amenable for quantitative uncertainty analysis, the choice must be linked to a statement about the state of the world (see Section 6.1.1).

- *Epidemiology data to use:*
    1. *risk estimate developed with data aggregated from all suitable studies,*
    2. *risk estimate or estimates developed using each study individually.*
- *Factors affecting extrapolation from occupational to general population cohorts, including differences in baseline health status, age distribution, the healthy worker survivor effect, and background exposures.*

Response: (a) Quantitative uncertainty analysis based on meta-analysis data poses challenges owing to differences in study protocols. Exotic methods might take us further, the question is whether the restriction to data driven methods (as opposed to expert judgment or Bayesian methods) could be maintained (see Sections 6.4.2.2 and 6.4.2.3). (b) If the general population is characterized by distributions over known confounders whose coefficients are estimated from the epidemiological studies, then uncertainty over these coefficients can be extracted with the methods mentioned in Section 6.4.2.1. Uncertainty due to missing covariates is intractable for data driven uncertainty analysis (see Section 6.4.2.2).

- *Bioassay data to use:*
    1. *risk estimate developed with the single data set implying the greatest risk (that is, single study, tumor site, gender),*

2. *risk estimate developed with multiple data sets satisfying an a priori set of selection criteria.*

Response: (a) Echoing response 1.b, uncertainty in choice of data sets is volitional and is not quantified by sampling an input distribution. To be amenable for quantitative uncertainty analysis the choice must be linked to a statement about the state of the world (see Section 6.1.1). (b) The issue here is similar to the meta-analysis addressed in (2.a).

- *Dose-response model:*
  1. *linear dose response,*
  2. *nonlinear dose.*

Response: (a) Presumably the low dose extrapolation is meant. When the linear model is used by default, the uncertainty is volitional. To be amenable for quantitative uncertainty analysis the choice must be linked to a statement about the state of the world (see Section 6.1.1). The ED_x as POD for the linear extrapolation can be subjected to quantitative uncertainty analysis, if based on sufficient bioassay data. (b) presumably 'nonlinear dose response' is meant. It is possible that human thresholds exist, and that the distribution of thresholds can be characterized in the human population. In as much as the mechanisms for this are not yet understood, there is no quantitative model expressing threshold as a function of parameters which could be estimated, with uncertainty, from data. This currently limits the application of uncertainty quantification (see Section 6.4.2.9).

- *Dose metric:*
  1. *average daily intake,*
  2. *area under the blood concentration-time curve,*
  3. *lifetime average body burden,*
  4. *peak body burden,*
  5. *other.*

Response: The dose metric is chosen to maximize causal proximity to the endpoint, while maintaining the link to measured exposure (see Section 6.4.2.7). There may be uncertainty with regard to which metric is optimal. If an inappropriate metric is chosen in a bioassay study, this would be expressed in noisier responses which would tend to suppress the dependence of endpoint on dose. A data driven quantitative uncertainty analysis of dose metric would require a mathematical model expressing endpoints as a function, inter alia, of dose metric, with parameters estimated from data.

- *Dose metric—biological measure:*
  1. *free dioxin,*
  2. *bound dioxin.*

Response: The issue is whether all TCDD available for AhR binding, or only the bound TCDD, should be used as a dose metric. Binding affinity is determined by more factors than genetic polymorphisms and these other factors are poorly understood (see Section 6.4.2.9). A quantitative uncertainty analysis must await the formulation of a quantitative model expressing binding affinity in terms of parameters which can be estimated from data.

- *POD:*
    a. *ED_{10},*
    b. *ED_{05},*
    c. *ED_{01}*

Response: Uncertainty in choosing a POD is volitional. Uncertainty in the value of an ED_x can be quantified in a data driven manner if sufficient bioassay data is at hand (see Section 6.4.1.1).

- *Value from ED distribution to use:*
    1. *ED,*
    2. *lower confidence bound value for the ED (LED),*
    3. *upper confidence bound for the ED (UED).*

Response: Given that uncertainty on the POD is quantified, it is a trivial exercise to derive a distribution of the slopes of a linear low dose extrapolation, and hence a distribution of a risk specific dose. This would seem preferable to choosing between a lower, upper, or nominal value.

- *Where alternative assumptions or methodologies could not be ruled out as implausible or unreasonable, EPA could have estimated the corresponding risks and reported the impact of these alternatives on the risk assessment results. The potential impacts of four sources of uncertainty are discussed below.*
    - *The full range of plausible parameter values for the dose-response functions used to characterize the dose-response relationship for the three occupational cohort studies selected by EPA (Ott and Zober, 1996; Becher et al., 1998; Steenland et al., 2001).*
    - *Use of other points of departure, not just the ED_{01} (or LED_{01}), to develop a CSF.*
    - *Alternative dose-response functional forms as well as goodness of fit of all models, especially at low doses.*
    - *Uncertainty introduced by estimation of historical occupational exposures.*

Response: (a) The study of Steenland et al. (2001) was selected to illustrate the possibilities and limitations of quantitative uncertainty analysis for this type of study (see Section 6.4.2.2). (b) The possibilities for uncertainty quantification with regard to the POD are discussed in Section 6.4.1.1 and in the POD bullet above. (c) Goodness of fit at

any measured dose is evaluated in standard packages. There may be different models with comparable goodness of fit at observed doses which differ strongly at doses outside the measured range. Extra model uncertainty, that is, uncertainty which is not conditional on the truth of any given model, is addressed by the exotic methods (see Section 6.4.2). (d) The feasibility of quantifying uncertainty in occupational exposure is study specific. The example of Steenland et al. (2001) was discussed in some detail (see Section 6.4.2.2). In general, the problem is not so much quantifying the exposure uncertainty, but in quantifying the dependence between the endpoints and the exposure uncertainty.

6.5.2. How Forward? Beyond RfDs and Cancer Slope Factors to Development of Predictive Human Dose-Response Functions

Uncertainty quantification is an emerging area in science. There are many examples of highly vetted and peer-reviewed uncertainty analyses based on structured expert judgment. Under this process, experts in effect synthesize a wide diversity of information in generating their subjective probability distributions. Where considerable data exist for an environmental pollutant, such as for the well-studied TCDD, it is natural to ask whether these extensive data can be leveraged more directly in uncertainty quantification. This is an area where research could be focused. The requisite knowledge does not yet exist, but there are promising lines of attack. It is therefore not a question of convening blue-ribbon panels to reveal the proper approach; instead multiple approaches should be encouraged to try out new ideas and share experiences.

An important idea that has been pioneered in Europe is to organize bench-test exercises where different approaches are applied to a common problem. This focuses the discussion on real issues and builds a community of capable practitioners. Such initiatives have proven much more productive than simply supporting individual researchers to explore their ideas.

Areas for which bench-test exercises might be appropriate include:

- Testing "exotic methods" for capturing model uncertainty
- Combining bioassay and epidemiological data for uncertainty quantification
- Applicability of structured expert judgment, e.g., for low-dose extrapolation
- Conducting dependence modeling, dependence inference, and dependence elicitation (such as with regard to TEFs).

Looking beyond compounds for which considerable data exist, there will always be a need to evaluate new substances. The target will be a simple method that:

1. Can yield predictions of toxicological indicators with uncertainty via a valid probabilistic mechanism.
2. Could evolve from approaches based on similarities (such as a random chemical model) under which the new substance could be seen as a random sample from a reference distribution of chemicals considered sufficiently similar, e.g., in terms of structure, physicochemical properties, and biological activity (potency).
3. Is nondisruptive to the traditional process.

This last feature is very important because as a practical matter, any collective push forward in this area must extend from (not abruptly depart from) current methodology. For example, the discussion surrounding uncertainty factors suggests that a probabilistically valid inference system could substantially differ from the current system. Nonetheless, to meld with current practice, it must initialize on the current system and allow this system to evolve in a measured fashion.

Additional research topics relevant to dioxin that could further inform health assessments include population variability of biokinetic constants, threshold mechanisms for the mass action model, and polymorphism at frequencies lower than 1%. Further data and improved methodologies in these areas, combined with developments illustrated elsewhere in this report, will help reduce uncertainties and strengthen our understanding of potential health implications of environmental contaminants.

Table 6-1. Key sources of uncertainty

Selection of endpoint and of species/strain, gender, life stage, other subject characteristics - critical effect - sensitivity (e.g., species, life stage) - human relevance
Selection of key study(ies): human data and bioassays (strength, inclusion criteria) - epidemiological studies, clinical/case reports (exposure estimate) - adequacy of study design, statistical power (exposure term, histopathology) - human relevance of bioassays (TK, MOA, endpoint) - data uncertainty, confidence in data; database deficiencies
Use of TK, dosimetry; body burden; species differences, cross-species extrapolation - bioavailability, dose dependence - half life, life stage, body fat, other compartments, age, other factors - body burden (peak, steady state, lifetime average) - physiologically-based pharmacokinetic (PBPK) modeling - scaling (human equivalents), adjustments (default and non-default; with TD)
Selection of dose metric - intake (averaging time) - background (what place on the dose-response curve) - free vs. receptor-bound TCDD - tissue-specific concentration - lipid-normalized level
Selection of POD - selection (e.g., NOAEL/LOAEL, BMDL, $ED_{01, 05, 10}$) - derivation method (e.g., BMD) - choice of model form (e.g., Hill, Weibull) - statistical uncertainty at/confidence in POD
Selection of dose-response model (e.g., biologically based, multi-stage) and of BMR - biological plausibility, MOA - model type and form, alternative functional forms - range of plausible parameter values - goodness of fit, especially at low doses
Selection of low-dose extrapolation approach - linear/nonlinear - threshold/nonthreshold
Human population variability - subpopulations (e.g., occupational, general public, sensitive groups) - polymorphisms - life stage, other features - individual vs. population threshold
Characterization of risk/effect - adversity of effect (vs. in normal range of variation and adaptation) - uncertainty factors (TK; TD; chemical-specific vs. default; justification) - consistency of methods for endpoints with common MOA - back-extrapolation from occupational data - MOE, RfD; beyond a point estimate for SF

PBPK = physiologically-based pharmacokinetic; SF = slope factor; TD = toxicodynamic; TK = toxicokinetic.

Table 6-2. PODs and amenability for uncertainty quantification

POD	Data profile	Choice	Uncertainty quantification
LOAEL	Experimental dose level from set of exposure-response data	Choose set of exposure-response measurements	No
NOAEL	Experimental dose level from set of exposure-response data	Choose set of exposure-response measurements	No
BMDL	Estimate from bioassay data	Choose BMR, choose dose-response relation	No, the BMDL is a quantile of an uncertainty distribution assuming that the dose-response model is true
ED_x	Estimate from set of exposure-response data	Choose bioassay experiments to estimate ED_x	Yes, if full bioassay data are available

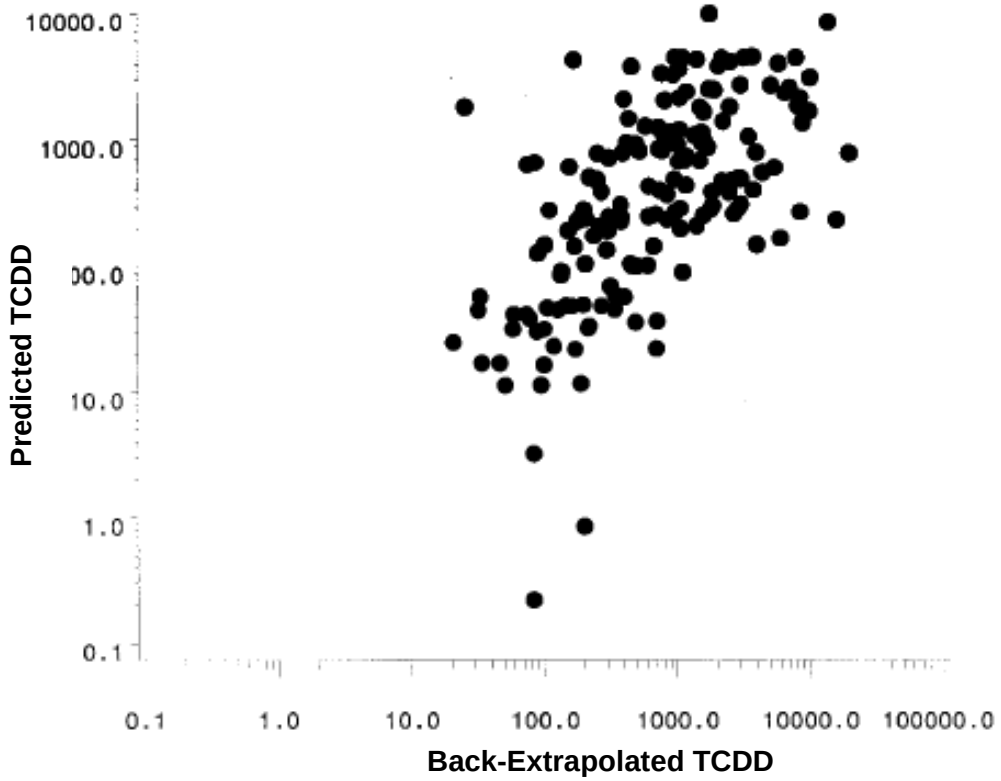


Figure 6-1. Back-casted vs. predicted TCDD serum levels for a worker subset.

Source: Steenland et al. (2001).

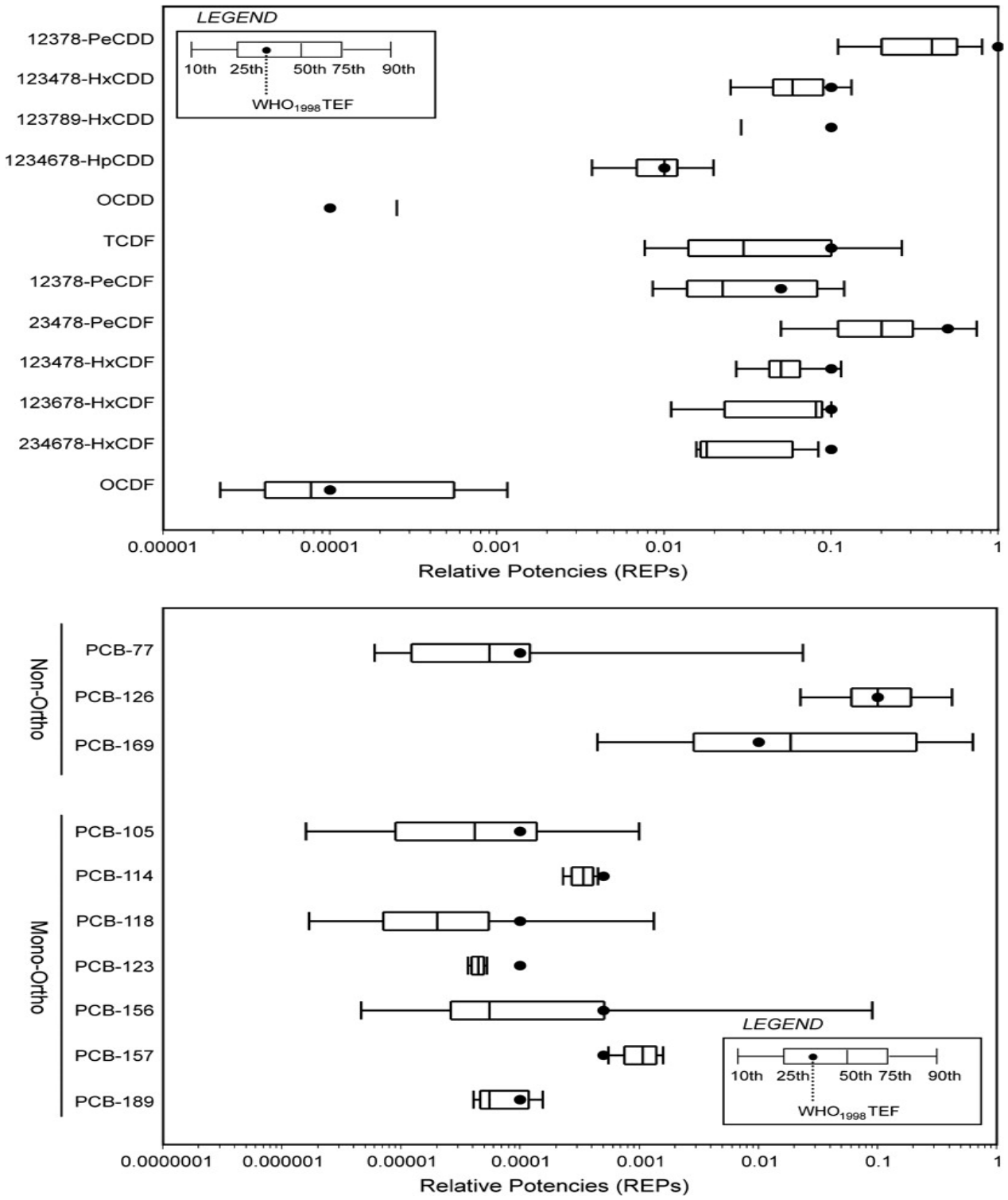


Figure 6-2. Distribution of in vivo unweighted REP values in the 2004 database.

Source: Van den Berg et al. (2006), reprinted with permission from Haws et al. (2006).

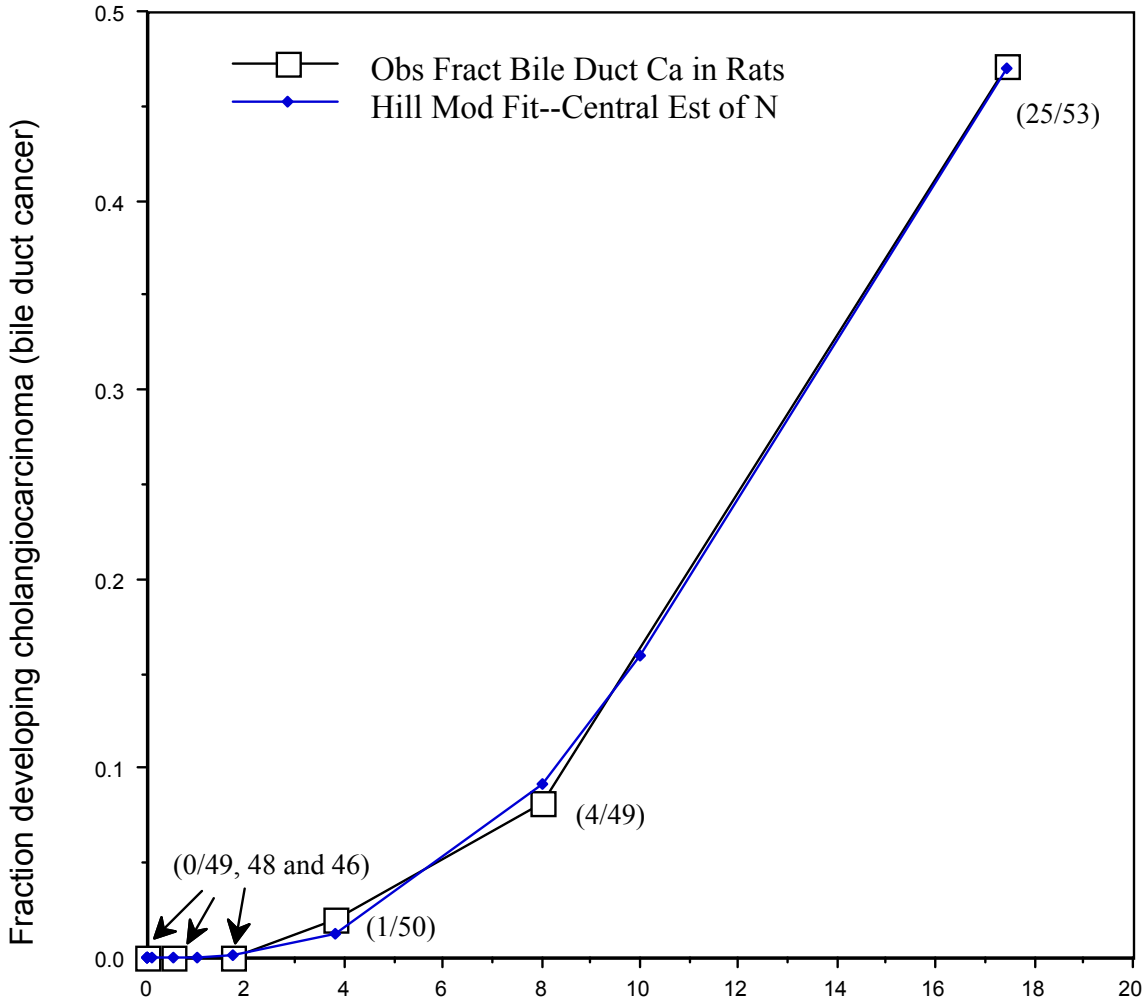


Figure 6-3. Plot of observed rat cholagiocarcinoma incidence and central estimate of Hill Model fit to the data vs. human equivalent dose.

Source: Hattis (2009).

REFERENCES

Abbott, BD; Birnbaum, LS; Diliberto, JJ. (1996). Rapid distribution of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) to embryonic tissues in C57BL/6N mice and correlation with palatal uptake in vitro. Toxicol. Appl. Pharmacol 141(1): 256–263.

Abraham, K; Krowke, R; Neubert, D. (1988) Pharmacokinetics and biological activity of 2,3,7,8-tetrachlorodibenzo-p-dioxin. 1. Dose-dependent tissue distribution and induction of hepatic ethoxyresorufin O-deethylase in rats following a single injection. Arch Toxicol 62(5):359−368.

Abraham, K; Knoll, A; Ende, M; et al. (1996) Intake, fecal excretion, and body burden of polychlorinated dibenzo-p-dioxins and dibenzofurans in breast-fed and formula-fed infants. Pediatr Res 40(5):671–679.

Abraham, K; Geusau, A; Tosun, Y; et al. (2002) Severe 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) intoxication: Insights into the measurement of hepatic cytochrome P450 1A2 induction. Clin Pharm Therapeut 72(2):163−174.

Aittomaki, A; Lahelma, E; Roos, E; et al. (2005) Gender differences in the association of age with physical workload and functioning. Occup Environ Med 62(2):95−100.

Akhmedkhanov, A; Revich, B; Adibi, JJ; et al. (2002) Characterization of dioxin exposure in residents of Chapaevsk, Russia. J Expo Anal Environ Epidemiol 12(6):409−417.

Akhtar, FZ; Garabrant, DH; Ketchum, NS; et al. (2004) Cancer in US Air Force veterans of the Vietnam War. J Occup Environ Med 46(2):123−136.

Alaluusua, S; Calderara, P; Gerthoux, PM; et al. (2004) Developmental dental aberrations after the dioxin accident in Seveso. Environ Health Perspect 112(13):1313−1318.

Alvarez-Pedrerol, M; Ribas-Fito, N; Torrent, M; et al. (2008) Thyroid disruption at birth due to prenatal exposure to beta-hexachlorocyclohexane. Environ Intl 34:737−740.

Amin, S; Moore, RW; Peterson, RE; et al. (2000) Gestational and lactational exposure to TCDD or coplanar PCBs alters adult expression of saccharin preference behavior in female rats. Neurotoxicol Teratol 22(5):675–682.

Andersen, ME; Mills, JJ; Gargas, ML; et al. (1993) Modeling receptor-mediated processes with dioxin: implications for pharmacokinetics and risk assessment. Risk Anal 13(1):25−36.

Andersen, ME; Birnbaum, LS; Barton, HA; et al. (1997) Regional hepatic CYP1A1 and CYP1A2 induction with 2,3,7,8-tetrachlorodibenzo-p-dioxin evaluated with a multicompartment geometric model of hepatic zonation. Toxicol Appl Pharmacol 144(1):145−155.

Andersson, P; McGuire, J; Rubio, C; et al. (2002) A constitutively active dioxin/aryl hydrocarbon receptor induces stomach tumors. Proc Natl Acad Sci 99(15): 9990–9995.

Ariens, EJ; Van Rossum, JM; Koopman, PC. (1960) Receptor reserve and threshold phenomena. I. Theory and experiments with autonomic drugs tested on isolated organs. Arch Int Pharmacodyn 127:459−478.

Armstrong, BG. (1995) Comparing standardized mortality ratios. Ann Epidemiol 5(1):60−64.

ATSDR (Agency for Toxic Substances and Disease Registry), Public Health Service, U.S. Department of Health and Human Services. (1998) Toxicological profile for chlorinated dibenzo-p-dioxins (CDDs). ATSDR, Atlanta, GA. Available at: http://www.atsdr.cdc.gov/toxprofiles/tp104.pdf.

Aylward, LL; Hays, SM; Karch, NJ; et al. (1996) Relative susceptibility of animals and humans to the cancer hazard posed by 2,3,7,8-tetrachlorodibenzo-p-dioxin using internal measures of dose. Environ Sci Technol 30(12):3534−3543.

Aylward, L; Kirman, C; Cher, D; et al. (2003) Reanalysis of dioxin cancer threshold. Environ Health Perspect 111(10):A510. Available at: http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=1241631&blobtype=pdf.

Aylward, LL; Brunet, RC; Carrier, G; et al. (2005a) Concentration dependent TCDD elimination kinetics in humans: toxicokinetic modeling for moderately to highly exposed adults from Seveso, Italy, and Vienna, Austria, and impact on dose estimates for the NIOSH cohort. J Expo Anal Environ Epidemiol 15:51−65.

Aylward, LL; Brunet, RC; Starr, TB; et al. (2005b) Exposure reconstruction for the TCDD-exposed NIOSH cohort using a concentration- and age-dependent model of elimination. Risk Anal 25(4):945−956.

Aylward, LL; Bodner, KM; Collins, JJ. (2007) Exposure reconstruction for a dioxin-exposed cohort: integration of serum sampling data and work histories. Organohal Comp 69:2063−2066.

Aylward, LL; Goodman, JE; Charnley, G; et al. (2008) A margin-of-error approach to assessment of noncancer risks of dioxins based on human exposure and response data. Environ Health Perspect 116(10):1344−1351.

Aylward, LL; Bodner, KM; Collins, JJ; et al. (2009) TCDD exposure estimation for workers at a New Zealand 2,4,5-T manufacturing facility based on serum sampling data. J Expo Sci Environ Epidemiol (epub doi: 10.1038/jes.2009.31). Available at http://www.nature.com/jes/journal/vaop/ncurrent/full/jes200931a.html.

Baccarelli, A; Mocarelli, P; Patterson, DG, Jr; et al. (2002) Immunologic effects of dioxin: new results from Seveso and comparison with other studies. Environ Health Perspect 110(12):1169−1173.

Baccarelli, A; Pesatori, AC; Masten, SA; et al. (2004) Aryl-hydrocarbon receptor-dependent pathway and toxic effects of TCDD in humans: a population-based study in Seveso, Italy. Toxicol Lett 149(1−3):287−293.

Baccarelli, A; Pesatori, AC; Consonni, D; et al. (2005) Health status and plasma dioxin levels in chloracne cases 20 years after the Seveso, Italy accident. Br J Dermatol 152(3):459−465.

Baccarelli, A; Hirt, C; Pesatori, AC; et al. (2006) T(14;18) translocations in lymphocytes of healthy dioxin-exposed individuals from Seveso, Italy. Carcinogenesis. 27(10):2001−2007. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16543249.

Baccarelli, A; Giacomini, SM; Corbetta, C; et al. (2008) Neonatal thyroid function in Seveso 25 years after maternal exposure to dioxin. PLos Medicine 5(7):1133−1142. Available at: http://www.plosmedicine.org/article/info:doi/10.1371/journal.pmed.0050161.

Bang, KM; Kim, JH. (2001) Prevalence of cigarette smoking by occupation and industry in the United States. Am J Ind Med 40(3):233−239.

Banks, YB; Birnbaum, LS. (1991) Absorption of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) after low dose dermal exposure. Toxicol Appl Pharmacol 107(2):302−310.

Banks, YB; Brewster, DW; Birnbaum, LS. (1990) Age-related changes in dermal absorption of 2,3,7,8-tetrachlorodibenzo-p-dioxin and 2,3,4,7,8-pentachlorodibenzofuran. Fundam Appl Toxicol 15(1):163−173.

Baron, JM; Zwadio-Klarwasser, G; Jugert, F; et al. (1998) Cytochrome P450 1B1: a major P450 isoenzyme in human blood monocytes and macrophage subsets. Biochem Pharmacol 56:1105−1110.

Barouki, R; Coumoul, X; Fernandez-Salguero, PM. (2007) The aryl hydrocarbon receptor, more than a xenobiotic-interacting protein. FEBS Lett. 581(19):3608–3615.

Bastomsky, CH. (1977) Enhanced thyroxine metabolism and high uptake goiters in rats after a single dose of 2,3,7,8-tetrachlorodibenzo-p-dioxin. Endocrinology 101:292–296.

Bates, MN; Buckland, SJ; Garrett, N; et al. (2004) Persistent organochlorines in the serum of the non-occupationally exposed New Zealand population. Chemosphere 54(10):1431−1443.

Becher, H; Flesch-Janys, D; Kauppinen, T; et al. (1996) Cancer mortality in German male workers exposed to phenoxy herbicides and dioxins. Cancer Causes Control 7:312−321.

Becher, H; Steindorf, K; Flesch-Janys, D. (1998) Quantitative cancer risk assessment for dioxins using an occupational cohort. Environ Health Perspect 106(Suppl 2):663−670.

Beebe, LE; Anver, MR; Riggs, CW; et al. (1995) Promotion of N-nitrosodimethylamine-initiated mouse lung tumors following single or multiple low dose exposure to 2,3,7,8- tetrachlorodibenzo-p-dioxin. Carcinogenesis 16:1345−1349.

Bell, DR; Clode, S; Fan, MQ; et al. (2007a) Toxicity of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin in the developing male Wistar(Han) rat. II: Chronic dosing causes developmental delay. Toxicol Sci 99(1):224−233.

Bell, DR; Clode, S; Fan, MQ; et al. (2007b) Relationships between tissue levels of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD), mRNAs, and toxicity in the developing male Wistar (Han) rat. Toxicol Sci 99(2):591–604.

Bernert, JT; Turner, WE; Patterson, DG; et al. (2007) Calculation of serum total lipid concentrations for the adjustment of persistent organohalogen toxicant measurements in human samples. Chemosphere 68:824−831.

Bertazzi, PA; Zocchetti, C; Pesatori, AC; et al. (1989) Ten-year mortality study of the population involved in the Seveso incident in 1976. Am J Epidemiol 129(6):1187−1200.

Bertazzi, A; Pesatori, AC; Consonni, D; et al. (1993) Cancer incidence in a population accidentally exposed to 2,3,7,8-tetrachlorodibenzo-para-dioxin. Epidemiology 4(5):398−406.

Bertazzi, PA; Zocchetti, C; Guercilena, S; et al. (1997) Dioxin exposure and cancer risk: a 15-year mortality study after the "Seveso accident". Epidemiology 8(6):646−652.

Bertazzi, PA; Consonni, D; Bachetti, S; et al. (2001) Health effects of dioxin exposure: a 20-year mortality study. Am J Epidemiol 153(11):1031−1044.

Birnbaum, LS. (1986) Distribution and excretion of 2,3,7,8-tetrachlorodibenzo-p-dioxin in congenic strains of mice which differ at the Ah locus. Drug Metab Dispos 14(1):34−40.

Black, JW: Leff, P. (1983) Operational models of pharmacological agonism. Proc. R. Soc. Lond. B 220:141-162.

Blankenship, A; Matsumura, F. (1997) 2,3,7,8-Tetrachlorodibenzo-p-dioxin-induced activation of a protein tyrosine kinase pp60src in murine hepatic cytosol using a cell-free system. Mol Pharmacol 52: 667−675.

Bock, KW. (1994) Aryl hydrocarbon or dioxin receptor: biologic and toxic responses. Rev Physiol Biochem Parmacol 125:1−42.

Bock, KW; Gschaidmeier, H; Heel, H; et al. (1998) AH receptor-controlled transcriptional regulation and function of rat and human UDP-glucuronosyltransferase isoforms. Adv Enzyme Regul 38:207−222.

Bodner, KM; Collins, JJ; Bloemen, LJ; et al. (2003) Cancer risk for chemical workers exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Occup Environ Med 60(9):672−675.

Bond, GG; Wetterstroem, NH; Roush, GJ; et al. (1988) Cause specific mortality among employees engaged in the manufacture, formulation, or packaging of 2,4-dichlorophenoxyacetic acid and related salts. Br J Ind Med 45(2):98−105.

Bond, GG; McLaren, EA; Brenner, FE; et al. (1989) Incidence of chloracne among chemical workers potentially exposed to chlorinated dioxins. J Occup Med 31(9):771−774.

Boverhoff, DR; Burgoon, LD; Tashiro, C; et al. (2005) Temporal and dose-dependent hepatic gene expression patterns in mice provide new insights into TCDD-mediated hepatotoxicity. Toxicol Sci 85:1048−1063.

Bowman, RE; Schantz, SL; Weerasinghe, NCA; et al. (1989a) Chronic dietary intake of 2,3,7,8-tetrachlorobibenzo-p-dioxin (TCDD) at 5 or 25 parts per trillion in the monkey: TCDD kinetics and dose-effect estimate of reproductive toxicity. Chemosphere 18:243−252.

Bowman, RE; Schantz, SL; Gross, ML; et al. (1989b) Behavioral effects in monkeys exposed to 2,3,7,8-TCDD transmitted maternally during gestation and for four months of nursing. Chemosphere 18:235−242.

Brand, KP; Rhomberg, L; Evans, JS. (1999) Estimating noncancer uncertainty factors: are ratios NOAELS informative? Risk Anal 19(2):295−308. Available at: http://www.springerlink.com/content/l376403247u14513/fulltext.pdf.

Brand, KP; Catalano, PJ; Hammitt, JK; et al. (2001) Limitations to empirical extrapolation studies: the case of BMD ratios. Risk Anal 21(4):625-640.

Bretagnolle, J; Huber-Carol, C. (1988) Effects of omitting covariates in Cox's model of survival data. Scand J Stat 15:125−138.

Brouwer, A; Morse, DC; Lans, MC; et al. (1998) Interactions of persistent environmental organohalogens with the thyroid hormone system: mechanisms and possible consequences for animal and human health. Toxicol Indust Health 14(1−2):59−84.

Brown, J; Goossens, LH; Kraan, BCP. (1997) Probabilistic accident consequence uncertainty study: food chain uncertainty assessment. Report Number: NUREG/CR-6523, EUR 16771, SAND97-0335. Nuclear Regulatory Commission, Washington, DC; Commission of the European Communities, Brussels; 74 pp.

Brown, NM; Manzolillo, PA; Zhang, JX; et al. (1998) Prenatal TCDD and predisposition to mammary cancer in the rat. Carcinogenesis 19(9):1623−1629.

Budinsky, RA; Paustenbach, D; Fontaine, D; et al. (2006) Recommended relative potency factors for 2,3,4,7,8 pentachlorodibenzofuran: the impact of different dose metrics. Toxicol Sci 91(1):275−285.

Buelke-Sam, J; Holson, JF; Nelson, CJ. (1982a) Blood flow during pregnancy in the rat: II. Dynamics of and litter variability in uterine flow. Teratology 26: 279–288.

Buelke-Sam, J; Nelson, CJ; Byrd, RA; et al. (1982b) Blood flow during pregnancy in the rat: I. Flow patterns to maternal organs. Teratology 26: 269–277.

Bueno de Mesquita, HB; Doornbos, G; van der Kuip, DAM; et al. (1993) Occupational exposure to phenoxy herbicides and chlorophenols and cancer mortality in The Netherlands. Am J Ind Med 23:289−300.

Burleson, GR; Lebrec, H; Yang, YG; et al. (1996) Effect of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) on influenza virus host resistance in mice. Fund Appl Toxicol 29:40−47.

Bussard, D; Preuss, P; White, P. (2009) Conclusions. In: Cooke, RM; ed. Uncertainty modeling in dose response: bench testing environmental toxicity. Statistics in Practice. New York, NY: Wiley.

Cantoni, L; Salmona, M; Rizzardini, M. (1981) Porphyrogenic effect of chronic treatment with 2,3,7,8-tetrachlorodibenzo-p-dioxin in female rates. Dose-effect relationship following urinary excretion of porphyrins. Toxicol Appl Pharmacol 57:156−157.

Carrier, G; Brunet, RC; Brodeur, J. (1995a) Modeling of the toxicokinetics of polychlorinated dibenzo-p-dioxins and dibenzofurans in mammalians, including humans. I. Nonlinear distribution of PCDD/PCDF body burden between liver and adipose tissues. Toxicol Appl Pharmacol 131(2):253–266.

Carrier, G; Brunet, RC; Brodeur, J. (1995b) Modeling of the toxicokinetics of polychlorinated dibenzo-p-dioxins and dibenzofurans in mammalians, including humans. II. Kinetics of absorption and disposition of PCDDs/PCDFs. Toxicol Appl Pharmacol 131(2):267–276.

Cesana, GC; de Vito, G; Ferrario, M; et al. (1995) Trends of smoking habits in northern Italy (1986-1990). Euro J Epidemiol 11(3):251–258.

Checkoway, H; Pearce, N; Crawford-Brown, DJ. (1989) Research methods in occupational epidemiology. New York, NY: Oxford University Press; 366 pp.

Cheng, H; Aylward, L; Beall, C; et al. (2006) TCDD exposure-response analysis and risk assessment. Risk Anal 26(4):1059–1071.

Chevrier, J; Eskenazi, B; Bradman, A; et al. (2007) Associations between prenatal exposure to polychlorinated biphenyls and neonatal thyroid-stimulating hormone levels in a Mexican-American population, Salinas Valley, California. Environ Health Perspect 115:1490–1496.

Chiaro, C; Morales, JL; Prabhu, KS; et al. (2008) Leukotriene A4 metabolites are endogenous ligands for the AH receptor. Biochemistry 47(32):8445–8455.

Choi, BC. (1992) Definition, sources, magnitude, effect modifiers, and strategies of reduction of the healthy worker effect. J Occup Med 34(10):979–988.

Chu, I; Lecavalier, P; Håkansson, H; et al. (2001) Mixture effects of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin and polychlorinated biphenyl congeners in rats. Chemosphere 43:807–814.

Clark, GC; Tritscher, A; Maronpot, R; et al. (1991) Tumor promotion by TCDD in female rats. In: Gallo, M; Scheuplein, R; van der Heijden, K; eds. Banbury Report 35: biological basis for risk assessment of dioxin and related compounds. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory; pp. 389–404.

Clegg, LX; Li, FP; Hankey, BF; et al. (2002) Cancer survival among US whites and minorities: a SEER (Surveillance, Epidemiology, and End Results) program population-based study. Arch Intern Med 162:1985–1993.

Clewell, HJ; Gentry, PR; Covington, TR; et al. (2004) Evaluation of the potential impact of age- and gender-specific pharmacokinetic differences on tissue dosimetry. Toxicol Sci 79(2):381–93.

Cole, P; Trichopoulos, D; Pastides, H; et al. (2003) Dioxin and cancer: a critical review. Regul Toxicol Pharmacol 38(3):378–388.

Collins, JJ; Bodner, K; Aylward, LL; et al. (2009) Mortality rates among trichlorophenol workers with exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Am J Epid 170(4):501–506.

Connor, KT; Aylward, L. (2006) Human response to dioxin: aryl hydrocarbon receptor (AhR) molecular structure, function, and dose-response data for enzyme induction indicate an impaired human AhR. J Toxicol Environ Health B 9(2):147−171.

Consonni, D; Pesatori, AC; Zocchetti, C; et al. (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):847–858.

Cooke, RM; ed. (2009) Uncertainty modeling in dose response: bench testing environmental toxicity. Statistics in Practice. New York, NY: Wiley, John & Sons, Inc; 248 pp.

Cooper, GS; Klebanoff, MA; Promislow, J; et al. (2005) Polychlorinated biphenyls and menstrual cycle characteristics. Epidemiology 16(2):191–200.

Cox, DR. (2006) Combination of data. In: Encyclopedia of Statistical Sciences, Kotz, S; Read, CB; Balakrishnan, N; et al., editors. Hoboken: Wiley. pp. 1074–1081.

Crofton, KM; Craft, ES; Hedge, JM; et al. (2005) Thyroid-hormone-disrupting chemicals: evidence for dose-dependent additivity or synergism. Environ Health Perspect 113(11):1549–1554.

Croutch, CR; Lebofsky, M; Schramm, K-W; et al. (2005) 2,3,7,8-tetrachlorordibenzo-*p*-dioxin (TCDD) and 1,2,3,4,7,8-hexachlorodibenzo-*p*-dioxin (HxCDD) alters body weight by decreasing insulin-like growth factor I (IGF-I) signaling. Toxicol Sci 85:560–571.

Crump, K. 2002. Critical issues in benchmark calculations from continuous data. Crit Rev. Toxicol. 21:133–153.

Crump, KS; Hoel, DG; Langley, CH; et al. (1976) Fundamental carcinogenic processes and their implications for low dose risk assessment. Cancer Research 36:2973–2979.

Crump, KS; Canady, R; Kogevinas, M. (2003) Meta-analysis of dioxin cancer dose response for three occupational cohorts. Environ Health Perspect 111(5):681–687.

D'Amico, M; Agozzino, E; Biagino, A; et al. (1999) Ill-defined and multiple causes on death certificates--a study of misclassification in mortality statistics. Eur J Epidemiol 15(2):141–148.

DeCaprio, AP; McMartin, DN; O'Keefe, PE; et al. (1986) Subchronic oral toxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in the guinea pig: comparisons with a PCB-containing transformer fluid pyrolysate. Fund Appl Toxicol 6:454–463.

DeKoning, EP; Karmaus, W. (2000) PCB exposure in utero and via breast milk. A review. J Expo Anal Environ Epidemiol 10:285–293.

Devito, MJ; Ma, X; Babish, JG; et al. (1994) Dose-response relationships in mice following subchronic exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin: CYP1A1, CYP1A2, estrogen receptor, and protein tyrosine phosphoylation. Toxicol Appl Pharmacol 124:82–90.

Diliberto, JJ; Akubue, PI; Luebke, RW; et al. (1995) Dose-response relationships of tissue distribution and induction of CYP1A1 and CYP1A2 enzymatic activities following acute exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in mice. Toxicol Appl Pharmacol 130(2):197–208.

Diliberto, JJ; Jackson, JA; Birnbaum, LS. (1996) Comparison of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) disposition following pulmonary, oral, dermal and parenteral exposures to rats. Toxicol Appl Pharmacol 138:158–168.

Diliberto, JJ; Burgin, DE; Birnbaum, LS. (1997) Role of CYP1A2 in hepatic sequestration of dioxin: studies using CYP1A2 knock-out mice. Biochem Biophys Res Commun 236:431–433.

Diliberto, JJ; Burgin, DE; Birnbaum, LS. (1999) Effects of CYP1A2 on disposition of 2,3,7,8-tetrachlorodibenzo-p-dioxin, 2,3,4,7,8-pentachlorodibenzofuran, and 2,2',4,4',5,5'-hexachlorobiphenyl in CYP1A2 knockout and parental (C57BL/6N and 129/Sv) strains of mice. Toxicol Appl Pharmacol 159(1):52–64.

Diliberto, JJ; DeVito, MJ; Ross, DG; et al. (2001) Subchronic exposure of [^{3}H]–2,3,7,8–tetrachlorodibenzo-p-dioxin (TCDD) in female B6C3F1 mice: relationship of steady–state levels to disposition and metabolism. Toxicol Sci 61:241–255.

Dolwick, KM; Schmidt, JV; Carver, LA; et al. (1993) Cloning and expression of a human Ah receptor cDNA. Mol Pharmacol 44(5):911–917.

Dunson, DB; Baird, DD. (2001) A flexible parametric model for combining current status and age at first diagnosis data. Biometrics 57(2):396–403.

EC (European Commission). (2009) CORDIS, Community Research and Development Information Service. Nuclear Energy Library, Archives (archived July 17). Available at: http://cordis.europa.eu/fp5-euratom/src/lib_docs.htm.

Ema, M; Ohe, N; Suzuki, M; et al. (1994) Dioxin binding activities of polymorphic forms of mouse and human arylhydrocarbon receptors. J Biol Chem 269(44):27337–27343.

Emond, C; Birnbaum, LS; DeVito, M. (2004) Physiologically based pharmacokinetic model for developmental exposures to TCDD in the rat. Toxicol Sci 80(1):115–133.

Emond, C; Michalek, JE; Birnbaum, L; et al. (2005) Comparison of the use of a physiologically-based pharmacokinetic model and a classical pharmacokinetic models for dioxin exposure assessments. Environ. Health Perspect 113(12):1666–1668.

Emond, C; Birnbaum, LS; Devito, MJ. (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. Environ. Health Perspect 114(9):1394–1400.

Eskenazi, B; Mocarelli, P; Warner, M; et al. (2000) Seveso women's health study: a study of the effects of 2,3,7,8-tetrachlorodebenzo-p-dioxin on reproductive health. Chemosphere 40:1247–1253.

Eskenazi, B; Warner, M; Mocarelli, P; et al. (2002a) Serum dioxin concentrations and menstrual cycle characteristics. Am J Epidemiol 156(4):383–392. Available at: http://aje.oxfordjournals.org/cgi/reprint/156/4/383.

Eskenazi, B; Mocarelli, P; Warner, M; et al. (2002b) Serum dioxin concentrations and endometriosis: a cohort study in Seveso, Italy. Environ Health Perspect 110(7):629–634.

Eskenazi, B; Mocarelli, P; Warner, M; et al. (2003) Maternal serum dioxin levels and birth outcomes in women of Seveso, Italy. Environ Health Perspect 111(7):947–953.

Eskenazi, B; Mocarelli, P; Warner, M; et al. (2004) Relationship of serum TCDD concentrations and age at exposure of female residents of Seveso, Italy. Environ Health Perspect 112(1):22–27.

Eskenazi, B; Warner, M; Marks, AR; et al. (2005) Serum dioxin concentrations and age at menopause. Environ Health Perspect 113(7):858–862.

Eskenazi, B; Warner, M; Samuels, S; et al. (2007) Serum dioxin concentrations and risk of uterine leiomyoma in the Seveso Women's Health Study. Am J Epidemiol 166(1):79–87.

Fattore, E; Trossvik, C; Håkansson, H. (2000) Relative potency values derived from hepatic vitamin A reduction in male and female Sprague-Dawley rats following subchronic dietary exposure to individual polychlorinated dibenzo-*p*-dioxin and dibenzofuran congeners and a mixture thereof. Toxicol and Appl Pharmacol 165:184–194.

Fernandez-Salguero, PM; Hilbert, DM; Rudikoff, S; et al. (1996) Aryl-hydrocarbon receptor-deficient mice are resistant to 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced toxicity. Toxicol Appl Pharmacol 140:173–179.

Ferriby, LL; Knutsen, JS; Harris, M; et al. (2007) Evaluation of PCDD/F and dioxin-like PCB serum concentration data from the 2001-2002 National Health and Nutrition Examination Survey of the United States population. J Expo Sci Environ Epidemiol 17(4):358–371.

Fingerhut, MA; Halperin, WE; Marlow, DA; et al. (1991) Cancer mortality in workers exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin. N Engl J Med 324(4):212–218.

Fisher, J. W., Whittaker, T. A., Taylor, D. H., Clewell, H. J., 3rd, and Andersen, M. E. (1989). Physiologically based pharmacokinetic modeling of the pregnant rat: a multiroute exposure model for trichloroethylene and its metabolite, trichloroacetic acid. Toxicol Appl Pharmacol 99:395–414.

Flesch-Janys, D. (1997) Analyses of exposure to polychlorinated dibenzo-p-dioxins, furans, and hexachlorocyclohexane and different health outcomes in a cohort of former herbicide-producing workers in Hamburg, Germany. Teratog Carcinog Mutagen 17(4−5):257−264.

Flesch-Janys, D; Gurn, P; Jung, D. (1994) First results of an investigation of the elimination of polychlorinated dibenzo-p-dioxins and dibenzofurans (PCDD/F) in occupationally exposed persons. Organohal Comp (21):93–99.

Flesch-Janys, D; Berger, J; Gurn, P; et al. (1995) Exposure to polychlorinated dioxins and furans (PCDD/F) and mortality in a cohort of workers from a herbicide-producing plant in Hamburg, Federal Republic of Germany. Am J Epidemiol 142(11):1165–1175.

Flesch-Janys, D; Becher, H; Gurn, P; et al. (1996) Elimination of polychlorinated dibenzo-p-dioxins and dibenzofurans in occupationally exposed persons. J Tox Environ Health 47(4):363–378.

Flesch-Janys, D; Steindorf, K; Gurn, P; et al. (1998) Estimation of the cumulated exposure to polychlorinated dibenzo-p-dioxins/furans and standardized mortality ratio analysis of cancer mortality by dose in an occupationally exposed cohort. Environ Health Perspect 106(Suppl 2):655–662.

Flodstrom, S; Ahlborg, UG. (1991) Promotion of hepatocarcinogenesis in rats by PCDDs and PCDFs. In: Gallo, MA; Scheuplein, RJ; van der Heijden, KA; eds. Banbury Report 35: biological basis for risk assessment of dioxin and related compounds. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory; pp. 405−414.

Fox, TR; Best, LL; Goldsworthy, SM; et al. (1993) gene expression and cell proliferation in rate liver after 2,3,7,8-tetrachlorodibenzo-p-dioxin exposure. Cancer Research 53:2265−2271.

Franczak, A; Nynca, A; Valdez, KE; et al. (2006) Effects of acute and chronic exposure to the aryl hydrocarbon receptor agonist 2,3,7,8-tetrachlorodibenzo-*p*-dioxin on the transition to reproductive senescence in female Sprague-Dawley rats. Biol Reprod 74:125−130.

Fretland, AJ; Safe, S; Hankinson, O. (2004) Lack of antagonism of 2,3,7,8-tetrachlorodibenzo-p-dioxin's (TCDDs) induction of cytochrome P4501A1 (CYP1A1) by the putative selective aryl hydrocarbon receptor modulator 6-alkyl-1,3,8-trichlorodibenzofuran (6-MCDF) in the mouse hepatoma cell line Hepa-1c1c7. Chemico-Biol Interac 150:161–170.

Fritz, W; Lin, TM; Safe, S; et al. (2009) The selective aryl hydrocarbon receptor modulator 6-methyl-1,3,8-trichlorodibenxofuran inhibits prostate tumor metastasis in TRMP mice. Biochem Pharmacol 77(7):1151−1160.

Fujii-Kuriyama, Y; Ema, M; Mimura, J; et al. (1995) Polymorphic forms of the Ah receptor and induction of the CYP1A1 gene. Pharmacogenetics 5(S):149–153.

Gasiewicz, TA; Henry, EC; Collins, LL. (2008) Expression and activity of aryl hydrocarbon receptors in development and cancer. Crit Rev Eukaryot Gene Expr 18: 279–321.

Gaylor, DW; Kodell, RL. (2000) Percentiles of the product of uncertainty factors for establishing probabilistic risk doses. Risk Anal 20(2):245−250.

Ge, N-L; Elferink, CJ. (1998) A direct interation between the aryl hydrocarbon receptor and retinoblastoma protein. J Biol Chem 273(35):22708−22713.

Geusau, A; Schmaldienst, S; Derfler, K; et al. (2002) Severe 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) intoxication: kinetics and trials to enhance elimination in two patients. Arch Toxicol 76:316−325.

Geyer, H; Scheunert, I; Korte, F. (1986) Bioconcentration potential of organic environmental chemicals in humans. Regul Toxicol Pharmacol 6(4):313−347.

Geyer, HJ; Scheuntert, I; Rapp, K; et al. (1990) Correlation between acute toxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and total body fat content in mammals. Toxicology 65(1−2):97−107.

Geyer, HJ; Schramm, K-W; Scheunert, I; et al. (1997) Considerations on genetic and environmental factors that contribute to resistance or sensitivity of mammals including humans to toxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and related compounds. Ecotoxicol Environ Safety 36:213–230.

Gielen, JE; Nebert, DW. (1971) Aryl hydrocarbon hydroxylase induction in mammalian liver cell culture. I. Stimulation of enzyme activity in nonhepatic cells and in hepatic cells by phenobarbital, polycyclic hydrocarbons, and 2,2-bis(p-chlorophenyl)-1,1,1-trichloroethane. J Biol Chem 246:5189–5198.

Goodman, DG; Sauer, RM. (1992) Hepatotoxicity and carcinogenicity in female Sprague-Dawley rats treated with 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD): a pathology working group reevaluation. Regulat Toxicol Pharmacol 15:245–252.

Goossens, LH; Cooke, RM; Kraan, BCP. (1996) Evaluation of weighting schemes for expert judgment studies. Commission of European Communities, Directorate-General for Science, Research and Development, XII-F-6, Delft University of Technology, Delft, The Netherlands.

Goossens, LH; Boardman, J; Harper, FT; et al. (1997) Probabilistic accident consequence uncertainty study: uncertainty assessment for deposited material and external doses. Washington, DC, Brussels-Luxembourg; NUREG/CR-6526, EUR 16772, SAND97-2323.

Goossens, LH; Harrison, JD; Harper, FT; et al. (1998) Probabilistic accident consequence uncertainty study: uncertainty assessment for internal dosimetry. Washington, DC, Brussels-Luxembourg; NUREG/CR-6571, EUR 16773, SAND98-0119.

Goossens, LH; Kraan, BCP; Cooke, RM; et al. (2001a) Methodology and processing techniques. EUR 18827EN, European Commission, Luxembourg.

Goossens, LH; Kraan, BCP; Cooke, RM; et al. (2001b) Overall uncertainty analysis. European Commission, Luxembourg; Project Report, EUR 18826EN.

Goossens, LH; Kraan, BCP; Cooke, RM; et al. (2001c) Nuclear science and technology countermeasures uncertainty assessment. European Commission, Luxembourg; EUR 18821EN.

Goossens, LH; Kraan, BCP; Cooke, RM; et al. (2001d) Uncertainty from the atmospheric dispersion and deposition module. European Commission, Luxembourg; EUR 18822EN.

Goossens, LH; Kraan, BCP; Cooke, RM; et al. (2001e) Uncertainty from the early and late health effects module. European Commission, Luxembourg; EUR 18824EN.

Goossens, LH; Kraan, BCP; Cooke, RM; et al. (2001f) Uncertainty from the food chain module. European Commission, Luxembourg; EUR 18823EN.

Goossens, LH; Kraan, BCP; Cooke, RM; et al. (2001g) Uncertainty from the dose module. European Commission, Luxembourg; Project Report, EUR 18825EN.

Graham, MJ; Lucier, GW; Linko, P; et al. (1988) Increases in cytochrome P-450 mediated 17ß-estradiol 2-hydroxylase activity in rat liver microsomes after both acute administration and subchronic administration of 2,3,7,8-tetrachlorodibenzo-p-dioxin in a two-stage hepatocarcinogenises model. Carcinogenesis 9(11):1935–1941.

Grassman, JA; Needham, LL; Masten, SA; et al. (2000) Evidence of hepatic sequestration of dioxin in humans? An examination of tissue levels and CYP1A2 expression. Organohal Comp 48:87–90**.**

Greenlee, WF; Hushka, LJ; Hushka, DR. (2001) Molecular basis of dioxin actions: evidence supporting chemoprotection. Toxicol Path 29(1):6–7. Available at: http://tpx.sagepub.com/cgi/reprint/29/1/6.

Guess, HA; Hoel, DG. (1977) The effect of dose on cancer latency period. J Environ Pathol Tox 1(2):279–286**.**

Haarmann-Stemmann, T; Bothe, H; Abel, J. (2009) Growth factors, cytokines and their receptors as downstream targets of arylhydrocarbon receptor (AhR) signaling pathways. Biochem Pharmacol 77:508-520.

Hahn, ME; Allan, LL; Sherr, DH. (2009) Regulation of constitutive and inducible AHR signaling: complex interactions involving the AHR repressor. Biochem Pharmacol 77(4):485–497.

Hakk, H; Diliberto, JJ; Birnbaum, LS. (2009) The effect of dose on 2,3,7,8-TCDD tissue distribution, metabolism and elimination in CYP1A2 (-/-) knockout and C57BL/6N parental strains of mice. Toxicol Appl Pharmacol Nov 15; 241(1):119–126. Epub 2009 Aug 18.

Harper, FT; Goossens, LH; Cooke, RM; et al. (1995) Probabilistic accident consequence uncertainty study: dispersion and deposition uncertainty assessment. Washington, DC, Brussels-Luxembourg; NUREG/CR-6244, EUR 15855 EN, SAND94-1453.

Harper, PA; Wong, JMY; Lam, MSM; et al. (2002) Polymorphisms in the human AH receptor. Chemico-Biol Interac 141(1–2):161–187.

Harrad, S; Wang, Y; Sandaradura, S; et al. (2003) Human dietary intake and excretion of dioxin-like compounds. J Environ Monit 5:224−228.

Hassoun, EA; Wilt, SC; DeVito, MJ; et al. (1998) Induction of oxidative stress in brain tissues of mice after subchronic exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Toxicol Sci 42:23−27.

Hassoun, EA; Li, F; Abushaban, A; et al. (2000) The relative abilities of TCDD and its congeners to induce oxidative stress in the hepatic and brain tissues of rats after subchronic exposure. Toxicology 145:103−113.

Hassoun, EA; Wang, H; Abushaban, A. (2002) Induction of oxidative stress following chronic exposure to TCDD, 2,3,4,7,8-pentachlorodibenzofuran, and 2,3',4,4',5-pentachlorobiphenyl. J Toxicol Environ Health A 65:825−842.

Hassoun, EA; Al-Ghafri, M; Abushaban, A. (2003) The role of antioxidant enzymes in TCDD-induced oxidative stress in various brain regions of rats after subchronic exposure. Free Rad Biol Medicine 35(9):1028−1036.

Hattis, D. (1996) Human interindividual variability in susceptibility to toxic effects--from annoying detail to a central determinant of risk. Toxicology 111:5–14.

Hattis, D. (2009) High-throughput testing – the NRC vision, the challenge of modeling dynamic changes in biological systems, and the reality of low-throughput environmental health decision making. Risk Anal 29(4):483−484.

Hattis, D; Burmaster, DE. (1994) Assessment of variability and uncertainty distributions for practical risk analyses. Risk Anal 14:713–730.

Hattis, D; Banati, P; Goble, R. (1999) Distributions of individual susceptibility among humans for toxic effects--for what fraction of which kinds of chemicals and effects does the traditional 10-fold factor provide how much protection?" Ann NY Acad Sci 895:286–316.

Hattis, D; Baird, S; Goble, R. (2002) A straw man proposal for a quantitative definition of the RfD. Drug Chem Toxicol 25(4):403−436.

Hattis, D; Ginsberg, G; Sonawane, B; et al. (2003) Differences in pharmacokinetics between children and adults−II. Childrens variability in drug elimination half-lives and in some parameters needed for physiologically-based pharmacokinetic modeling. Risk Anal 23:117−142.

Haws, LC; Su, SH; Harris, M; et al. (2006) Development of a refined database of mammalian relative potency estimates for dioxin-like compounds. Toxicol Sci 89(1):4−30.

Henck, JM; New, MA; Kociba, RJ; et al. (1981) 2,3,7,8-Tetrachlorodibenzo-p-dioxin: acute oral toxicity in hamsters. Toxicol Appl Pharmacol 59:405−407.

Henrikson, GL; Ketchum, NS; Michalek, JE; et al. (1997) Serum dioxin and diabetes mellitus in veterans of operation ranch hand. Epidemiology 8(3):252−258.

Hertz-Picciotto, I. (1995) Epidemiology and quantitative risk assessment: a bridge from science to policy. Am J Public Health 85(4):484−491.

Higgins, JPT; Thompson, SG; Spiegelhalter, DJ. (2009) Re-evaluation of random-effects meta analysis. J R Stat Soc Series A. 172(1): 137–159.

Hochstein, JR; Render, JA; Bursian, SJ; Aulerich, RJ. (2001) Chronic toxicity of dietary 2,3,7,8-tetrachlorodibenzo-p-dioxin to mink. Vet Human Toxicol 43(3):134−139.

Hoel, DG; Portier, CJ. (1994) Nonlinearity of dose-response functions for carcinogenicity. Environ Health Perspec 102(Suppl 1):109−113. Available at: http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=1566901&blobtype=pdf.

Hoglund, M; Sehn, L; Connors, JM; et al. (2004) Identification of cytogenetic subgroups and karyotypic pathways of clonal evolution in follicular lymphomas. Genes Chromos Cancer 39(3):195−204.

Hojo, R; Stern, S; Zareba, G; et al. (2002) Sexually dimorphic behavioral responses to prenatal dioxin exposure. Environ Health Perspect 110(3):247−254.

Hooiveld, M; Heederik, DJ; Kogevinas, M; et al. (1998) Second follow-up of a Dutch cohort occupationally exposed to phenoxy herbicides, chlorophenols, and contaminants. Am J Epidemiol 147(9):891−901.

Huff, JE. (1992) 2,3,7,8-TCDD: a potent and complete carcinogen in experimental animals. Chemosphere 25:173−176.

Huff, JE; Salmon, AG; Hooper, NK; et al. (1991) Long-term carcinogenesis studies on 2,3,7,8-tetrachlorodibenzo-p-dioxin and hexachlorodibenzo-p-dioxins. Cell Biol Toxicol 7(1):67−93.

Hurst, CH; Abbott, BD; DeVito, MJ; et al. (1998) 2,3,7,8-tetrachlorodibenzo-p-dioxin in pregnant Long Evans rats: disposition to maternal and embryo/fetal tissues. Toxicol Sci 45:129−136.

Hurst, CH; DeVito, MJ; Birnbaum, LS. (2000a) Tissue disposition of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in maternal and developing Long-Evans rats following subchronic exposure. Toxicol Sci 57(2):275−283.

Hurst, CH; DeVito, MJ; Setzer, RW; et al. (2000b) Acute administration of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in pregnant Long Evans rats: association of measured tissue concentrations with developmental effects. Toxicol Sci 3(2):411−420.

IARC (International Agency for Research on Cancer). (1997) IARC Monographs on the Evaluation of Carcinogenic Risks to Humans. Vol. 69. Polychlorinated dibenzo-para-dioxins and polychlorinated dibenzofurans. Lyon, France: International Agency for Research on Cancer; 666 pp.

Ikeda, M; Tamura, M; Yamashita, J; et al. (2005) Repeated in utero and lactational 2,3,7,8-tetrachlorodibenzo-*p*-dioxin exposure affects male gonads in offspring, leading to sex ratio changes in F2 progeny. Toxicol Appl Pharmacol 206:351−355.

ILSI (International Life Sciences Institute). (1994) Physiological parameter values for PBPK models. Washington, DC: Risk Science Institute.

Institute of Medicine. (1994) Veterans and Agent Orange. Washington, DC: National Academies Press; 832 pp.

Institute of Medicine. (2006) Veterans and Agent Orange: update 2000. 7th edition. Washington, DC: National Academies Press; 896 pp.

Ishihara, K; Warita, K; Tanida, T; et al. (2007) Does paternal exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) affect sex ratio of offspring? J Vet Med Sci 69(4):347−352.

James, WH. (1995) What stabilizes the sex ratio? Ann Hum Genet 59(2):243−249.

Jorgensen, N; Andersen, AG; Eustache, F; et al. (2001) Regional differences in semen quality in Europe. Hum Reprod 16:1012−1019.

Kang, S-H; Kodell, RL; Chen, JJ. (2000) Incorporating model uncertainties along with data uncertainties in microbial risk assessment. Regul Toxicol Pharmacol 32:68−72. Available at: http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6WPT-45BCPMP-14&_user=1722207&_rdoc=1&_fmt=&_orig=search&_sort=d&_docanchor=&view=c&_searchStrId=1043774060&_rerunOrigin=google&_acct=C000053990&_version=1&_urlVersion=0&_userid=1722207&md5=29e6d3709cdba2152a90d24cea739a68.

Kattainen, H; Tuukanan, J; Simanainen, U; et al. (2001) In utero/lactational 2,3,7,8-tetrachlorodibenzo-*p*-dioxin exposure impairs molar tooth development in rats. Toxicol Appl Pharmacol 17:216−224.

Kauppinen, TP; Partanen, TJ; Hernberg, SG; et al. (1993) Chemical exposures in manufacture of phenoxy herbicides and chlorophenols and in spraying of phenoxy herbicides. Amer J Indust Med 23(6):903−920.

Keller, JM; Huet-Hudson, YM; Leamy, LJ. (2007) Qualitative effects of dioxin on molars vary among inbred mouse strains. Arch Oral Biol 52:450−454.

Keller, JM; Zelditch, ML; Huet, YM; et al. (2008a) Genetic differences in sensitivity to alterations of mandible structure caused by the teratogen 2,3,7,8-tetrachlorodibenzo-*p*-dioxin. Toxicol Pathol 36:1006–1013.

Keller, JM; Huet-Hudson, Y; Leamy, LJ. (2008b) Effects of 2,3,7,8-tetrachlorodibenzo-p-dioxin on molar development among non-resistant inbred strains of mice: a geometric morphometric analysis. Growth Devel Aging 71:3−16.

Kerger, BD; Leung, HW; Scott, P; et al. (2006) Age- and concentration-dependent elimination half-life of 2,3,7,8-tetrachlorodibenzo-p-dioxin in Seveso children. Environ Health Perspect 114(10):1596−1602.

Kerger, BD; Leung, HW; Scott, PK; et al. (2007) Refinements on the age-dependent half-life model for estimating child body burdens of polychlorodibenzodioxins and dibenzofurans. Chemosphere 67(9):S272−278.

Ketchum, NS; Michalek, JE; Burton JE. (1999) Serum dioxin and cancer in veterans of Operation Ranch Hand. Am J Epidemiol 149(7):630−639.

Kim, AH; Kohn, MC; Nyska, A; et al. (2003) Area under the curve as a dose metric for promotional responses following 2,3,7,8-tetrachlorodibenzo-p-dioxin exposure. Toxicol Appl Pharmacol 191:12−21.

Kitchin, KT; Woods, JS. (1979) 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) effects on hepatic microsomal cytochrome P-448-mediated enzyme activities. Toxicol Appl Pharmacol 47:537−546.

Kociba, RJ; Keeler, PA; Park, GN; et al. (1976) 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD): results of a 13 week oral toxicity study in rats. Toxicol Appl Pharmacol 35:553−574.

Kociba, RJ; Keyes, DG; Beyer, JE; et al. (1978) Results of a two-year chronic toxicity and oncogenicity study of 2,3,7,8-tetrachlorodibenzo-p-dioxin in rats. Toxicol Appl Pharmacol 46(2):279−303.

Kogevinas, M; Becher, H; Benn, T; et al. (1997) Cancer mortality in workers exposed to phenoxy herbicides, chlorophenols, and dioxins. An expanded and updated international cohort study. Am J Epidemiol 145:1061−1075.

Kohn, MC; Melnick, RL. (2002) Biochemical origins of the non-monotonic receptor-mediated dose-response. J Mol Endocrinol 29:113−123. Available at: http://jme.endocrinology-journals.org/cgi/reprint/29/1/113.

Kohn, MC; Lucier, GW; Clark, GC; et al. (1993) A mechanistic model of effects of dioxin on gene expression in the rat liver. Toxicol Appl Pharmacol 120(1):138−154.

Kohn, MC; Sewall, CH; Lucier, GW; et al. (1996) A mechanistic model of effects of dioxin on thyroid hormones in the rat. Toxicol Appl Pharmacol 136(1):29−48.

Kohn, MC; Walker, NJ; Kim, AH; et al. (2001) Physiological modeling of a proposed mechanism of enzyme induction by TCDD. Toxicology 162:193–208.

Kolluri, SK; Weiss, C; Koff, A; et al. (1999) p27(Kip1) induction and inhibition of proliferation by the intracellular Ah receptor in developing thymus and hepatoma cells. Genes Dev. 13:1742–1753.

Kopylev L; Chen C; White P. (2007) Towards quantitative uncertainty assessment for cancer risks: central estimates and probability distributions of risk in dose-response modeling. Regul Toxicol Pharmacol 49(3):203–207.

Kopylev, L; Fox, J; Chen, C. (2009) Combining risks from several tumors using Markov chain Monte Carlo. In: Cooke, R; ed. Uncertainty modeling in environmental toxicity dose response: bench testing the models. New York: Wiley; 230 pp.

Kreuzer, PE; Csanády, GA; Baur, C; et al. (1997) 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) and congeners in infants. A toxicokinetic model of human lifetime body burden by TCDD with special emphasis on its uptake by nutrition. Arch Toxicol 71(6):383−400.

Krishnan, K; Andersen, ME. (1991) Interspecies scaling in pharmacokinetics. In: Rescingo, A; Thakkur, A; eds. New trends in pharmacokinetics. New York, NY: Plenum Press; pp. 203−226.

Krowke, R; Chahoud, I; Baumann-Wilschke, I.; et al. (1989) Pharmacokinetics and biological activity of 2,3,7,8-tetrachlorodibenzo-p-dioxin. 2. Pharmacokinetics in rats using a loading-dose/maintenance-dose regime with high doses. Arch Toxicol 63:356−360.

Kuchiiwa, S; Cheng, S-B; Nagatomo, I; et al. (2002) In utero and lactational exposure to 2,3,7,8-tetrachlorodibenso-*p*-dioxin decreases serotonin-immunoreactive neurons in raphe nuclei of male mouse offspring. Neurosci Lett 317:73−76.

Kurowicka, D; Cooke, RM. (2006) Uncertainty analysis with high dimensional dependence modelling. West Sussex, England: John Wiley & Sons; 302 pp.

LaKind, JS; Berlin, CM; Park, CN; et al. (2000) Methodology for characterizing distributions of incremental body burdens of 2,3,7,8-TCDD and DDE from breast milk in North American nursing infants. J Toxicol Environ Health A 59(8):605−639.

Lakshmanan, MR; Campbell, BS; Chirtel, SJ; et al. (1986) Studies on the mechanism of absorption and distribution of 2,3,7,8-tetrachlorodibenzo-p-dioxin in the rat. J Pharmacol Exp Ther 239:673−677.

Landi, MT; Consonni, D; Patterson, DG, Jr; et al. (1998) 2,3,7,8-Tetrachlorodibenzo-p-dioxin plasma levels in Seveso 20 years after the accident. Environ Health Perspect 106(5):273−277.

Larsen, JC. (2006) Risk assessments of polychlorinated dibenzo-p-dioxins, polychloriniated dibenzofurans, and dioxin-like polychlorinated biphenyls in food. Mol Nutr Food Res 50(10):885−896.

Latchoumycandane, C; Mathur, PP. (2002) Effects of vitamin E on reactive oxygen species-mediated 2,3,7,8-tetrachlorodi-benzo-p-dioxin toxicity in rat testis. J Appl Toxicol 22(5):345–351.

Latchoumycandane, C; Chitra, KC; Mathur, PP. (2002a) The effect of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin on the antioxidant system in mitochondrial and microsomal fractions in rat testis. Toxicology 171:127−135.

Latchoumycandane, C; Chitra, KC; Mathur, PP. (2002b) Induction of oxidative stress in rat epididymal sperm after exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Arch Toxicol 76:113−118.

Latchoumycandane, C; Chitra, KC; Mathur, PP. (2003) 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) induces oxidative stress in the epididymis and epididymal sperm of adult rats. Arch Toxicol 77:280−284.

Lawrence, GS; Gobas, FA. (1997) A pharmacokinetic analysis of interspecies extrapolation in dioxin risk assessment. Chemosphere 35(3):427−452.

Lean, MEJ; Han, TS; Deurenberg, P. (1996) Predicting body composition by densitometry from simple anthropometric measurements. Am J Clin Nut 63:4−14.

Lee, DJ; Fleming, LE; Arheart, KL; et al. (2007) Smoking rate trends in U.S. occupational groups: the 1987 to 2004 National Health Interview Survey. J Occup Environ Med 49(1):75−81.

Lehman, AJ; Fitzhugh, OG. (1954) 100-Fold margin of safety. Assoc Food Drug Off USQ Bull 18:33−35.

Leo, A; Hansch, C; Elkins, D. (1971) Partition coefficients and their uses. Chem Rev 71:525−616.

Leung, HW; Ku, RH; Paustenbach, DJ; et al. (1988) A physiologically based pharmacokinetic model for 2,3,7,8-tetrachlorodibenzo-p-dioxin in C57BL/6J and DBA/2J mice. Toxicol Lett 42(1):15−28.

Leung, HW; Poland, A; Paustenbach, D; et al. (1990) Pharmacokinetics of (125-I)-2-lodo-3,7,8-trichlorodibenzo-p-dioxin in mice: analysis with a physiological modeling approach. Toxicol Appl Pharmacol 103:411−419.

Leung, HW; Kerger, BD; Paustenbach, DJ. (2006) Elimination half-lives of selected polychlorinated dibenzodioxins and dibenzofurans in breast-fed human infants. J Toxicol Environ Health A 69(6):437−443.

Li, B; Liu, H-Y; Dai, L-J; et al. (2006) The early embryo loss caused by 2,3,7,8-tetrachlorodibenzo-*p*-dioxin may be related to the accumulation of this compound in the uterus. Reprod Toxicol 21:301−306.

Li, CY; Sung, FC. (1999) A review of the healthy worker effect in occupational epidemiology. Occup Med 49(4):225−229.

Li, X; Johnson, DC; Rozman, KK. (1997) 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) increases release of luteinizing hormone and follicle-stimulating hormone from the pituitary of immature female rats in vivo and in vitro. Toxicol Appl Pharmacol 142(2): 264-−269.

Limbird, LE. (1996) Cell surface receptors: a short course on theory and method. USA: Springer; 258 pp.

Longnecker, MP; Gladen, BC; Patterson, DG, Jr; et al. (2000) Polychlorinated biphenyl (PCB) exposure in relation to thyroid hormone levels in neonates. Epidemiology 11:249−254.

Lorber, M; Patterson, D; Huwe, J; et al. (2009) Evaluation of background exposures of Americans to dioxin-like compounds in the 1990s and the 2000s. Chemosphere 77:640–651.

Lorenzen, A; Okey, AB. (1991) Detection and characterization of Ah receptor in tissue and cells from human tonsils. Toxicol Appl Pharmacol 107(2):203–214.

Lucier, GW; Rumbaugh, RC; McCoy, Z; et al. (1986) Ingestion of soil contaminated with 2,3,7,8-tetrachloro-dibenzo-p-dioxin (TCDD) alters hepatic enzyme activities in rats. Fund Appl Toxicol 6:364–371.

Lucier, GW; Tritscher, A; Goldsworthy, T; et al. (1991) Ovarian hormones enhance 2,3,7,8-tetrachlorodibenzo-p-dioxin-mediated increases in cell proliferation and preneoplastic foci in a two-stage model for rat hepatocarcinogenesis. Cancer Res 51(5):1391–1397.

Lutz, WK. (1990) Dose-response relationship and low dose extrapolation in chemical carcinogenesis. Carcinogenesis 11:1243–1247.

Lutz, WK. (1999) Dose-response relationships in chemical carcinogenesis reflect diffeences in individual susceptibility. Consequences for canceer risk assessment, extrapolation, and prevention. Human Exp Toxicol 18:707-712.

Lutz, WK. (2001) Susceptibility differences in chemical carcinogenesis linearize the dose-response relationship: threshold doses can be defined only for individuals. DNA Repair 482:71–76.

Lutz, WK; Gaylor, DW. (2008) Letter to the editor. Dose-response relationships for cancer incidence reflect susceptibility distributions. Chem. Res. Toxicol. 21:971–973.

Mackie, D; Liu, J; Loh, YS; et al. (2002) No evidence of dioxin cancer threshold. Environ Health Perspect 111(9):1145−1147. Available at: http://www.ehponline.org/members/2003/5730/5730.pdf.

Mally, A; Chipman, JK. (2002) Non-genotoxic carcinogens: early effects on gap junctions, cell proliferation and apoptosis in the rat. Toxicology 180:233–248.

Manchester, DK; Gordon, SK; Golas, C; et al. (1987) Ah receptor in human placenta: stabilization by molybdate and characterization of binding of 2,3,7,8-tetrachlorodibenzo-p-dioxin, 3-methylcholanthrene, and benzo(a)pyrene. Cancer Res 47(18):4861–4868.

Manz, A; Berger, J; Dwyer, JH; et al. (1991) Cancer mortality among workers in chemical plant contaminated with dioxin. Lancet 338(8773):959−964.

Markowski, VP; Zareba, G; Stern, S; et al. (2001) Altered operant responding for motor reinforcement and the determination of benchmark doses following perinatal exposure to low-level 2,3,7,8-tetrachlorodibenzo-p-dioxin. Environ Health Perspect 109(6):621−627.

Maronpot, RR; Montgomery, CA; Boorman, GA; et al. (1986) National Toxicology Program nomenclature for hepatoproliferative lesions of rats. Toxicol Pathol 14(2):263−273.

Maronpot, RR; Pitot, HC; Peraino, C. (1989) Use of rat liver altered focus models for testing chemicals that have completed two-year carcinogenicity studies. Toxicol Pathol 17(4):651−662.

Maronpot, RR; Foley, JF; Takahashi, K; et al. (1993) Dose-response for TCDD promotion of hepatocarcinogenesis in rats initiated with DEN: histologic, biochemical, and cell proliferation endpoints. Environ Health Perspect 101:634−642.

Maruyama, W; Yshida, K; Tanaka, T; et al. (2002) Determination of tissue-blood partition coefficients for a physiological model for humans, and estimation of dioxin concentration in tissues. Chemosphere 46:975−985.

Matsumoto, Y; Ide, F; Kishi, R; et al. (2007) Aryl hydrocarbon receptor plays a significant role in mediating airborne particulate-induced carcinogenesis in mice. Environ Sci Technol 41(10):3775−3780.

McBride, DI; Collins, JJ; Humphry, NF; et al. (2009a) Mortality in workers exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin at a trichlorophenol plant in New Zealand. J Occup Environ Med 51(9):1049–1056.

McBride, DI; Burns, CJ; Herbison, GP; et al. (2009b) Mortality in employees at a New Zealand agrochemical manufacturing site. Occup Med 59(4):255−263.

McEwen, LN; Kim, C; Haan, M; et al. (2006) Diabetes reporting as a cause of death. Diabetes Care 29(2):247−253.

McMichael, AJ. (1976) Standardized mortality ratios and the "healthy worker effect": scratching beneath the surface. J Occup Med 18:128−131.

McMillan, BJ; Bradfield, CA. (2007) The aryl hydrocarbon receptor sans xenobiotics: endogenous function in genetic model systems. Mol Pharmacol 72:487–498.

McNulty, WP; Nielsen-Smith, KA; Lay, JO, Jr; et al. (1982) Persistence of TCDD in monkey adipose tissue. Food Chem Toxicol. 20(6):985−986.

Michalek, JE; Pavuk, M. (2008) Diabetes and cancer in veterans of Operation Ranch Hand after adjustment for calendar period, days of spraying, and time spent in Southeast Asia. J Occup Environ Med 50(3):330−340.

Michalek, JE; Pirkle, JL; Caudill, SP; et al. (1996) Pharmacokinetics of TCDD in veterans of Operation Ranch Hand: 10-year follow-up. J Tox Environ Health 47(3):209−220.

Michalek, JE; Pirkle, JL; Needham, LL; et al. (2002) Pharmacokinetics of 2,3,7,8-tetrachlorodibenzo-p-dioxin in Seveso adults and veterans of operation Ranch Hand. J Expo Anal Environ Epidemiol 12:44−53.

Micka, J; Milatovich, A; Menon, A; et al. (1997) Human Ah receptor (AHR) gene: localization to 7p15 and suggestive correlation of polymorphism with CYP1A1 inducibility. Pharmacogenetics 7:95−101.

Miettinen, HM; Sorvari, R; Alaluusua, S; et al. (2006) The Effect of perinatal TCDD exposure on caries susceptibility in rats. Toxicol Sci 91(2):568−575.

Milbrath, MO; Wenger, Y; Chang, CW; et al. (2009) Apparent half-lives of dioxins, furans, and polychlorinated biphenyls as a function of age, body fat, smoking status, and breast-feeding. Environ Health Perspect 117(3):417−425.

Mocarelli P; Brambilla P; Gerthoux PM; et al. (1996) Change in sex ratio with exposure to dioxin [letter]. Lancet 348:409.

Mocarelli, P; Gerthoux, PM; Ferrari, E; et al. (2000) Paternal concentrations of dioxin and sex ratio of offspring. Lancet 355:1858−1863.

Mocarelli, P; Gerthoux, PM; Patterson, DG, Jr; et al. (2008) Dioxin exposure, from infancy through puberty, produces endocrine disruption and affects human semen quality. Environ Health Perspect 116(1):70−77.

Monson, RR. (1986) Observations on the healthy worker effect. J Occup Med 28(6):425−433.

Moser, GA; McLaghlan, MS. (2001) The influence of dietary concentration on the absorption and excretion of persistent lipophilic organic pollutants in the human intestinal tract. Chemosphere 45:201−211.

Motulsky, HJ; Christopoulos, A. (2003) Fitting models to biological data using linear and nonlinear regression. A practical guide to curve fitting. San Diego: GraphPad Software, Inc., www.graphpad.com

Muller, A; De La Rochebrochard, E; Labbe-Decleves, C; et al. (2004) Selection bias in semen studies due to self-selection of volunteers. Human Reprod 19:2838−2844.

Murdoch, DJ; Krewski, D. (1988) Carcinogenic risk assessment with time-dependent exposure patterns. Risk Anal Dec;8(4):521–530.

Murdoch, DJ; Krewski, D; Wargo, J. (1992) Cancer risk assessment with intermittent exposure. Risk Anal Dec;12(4):569–577.

Murphy, JM; Sexton, DMH; Barnett, DN; et al. (2004) Quantification of modeling uncertainties in a large ensemble of climate change simulations. Nature 430:768−772. Available at: http://www.seas.harvard.edu/climate/pdf/Uncertainty1.pdf.

Murray, FJ; Smith, FA; Nitschke, KD; et al.(1979) Three-generation reproduction study of rats given 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in the diet. Toxicol Appl Pharmacol 50:241−252.

Muto, T; Wakui, S; Imano, N; et al. (2001) In-utero and lactational exposure of 3,3',4,4',5-pentachlorobiphenyl modulate dimethlbenz[a]anthracene-induced rat mammary carcinogenesis. J Toxicol Pathol 14(3):213−224. Available at: http://www.jstage.jst.go.jp/article/tox/14/3/213/_pdf.

Myers, HE; Thompson, ML. (1998) Meta-analysis and occupational epidemiology. Occup Medic 48:99−101.

Nagel, S; Berger, J; Flesch-Janys, D; et al. (1994) Mortality and cancer mortality in a cohort of female workers of a herbicide producing plant exposed to polychlorinated dibenzo-p-dioxins and furans. Inform Biomet Epidemiol Med Biol 25:32–38.

NAS (National Academy of Sciences). (2006a) Health risks from dioxin and related compounds: evaluation of the EPA reassessment. Washington, DC: National Academies Press. Available at http://www.nap.edu/catalog.php?record_id=11688.

NAS (National Academy of Sciences). (2006b) Health risks from dioxin and related compounds: listen to the briefing (audio recording of public briefing led by D. Eaton, chair, NAS Dioxin Committee), Washington, DC. Available at: http://www.nap.edu/webcast/webcast_detail.php?webcast_id=328 (July 11).

NAS (National Academy of Sciences). (2009) Toward a unified approach to dose-response assessment: the need for an improved dose-response framework. Science and Decisions: Advancing Risk Assessment (pp. 127–187). National Resource Council. Washington: National Academies Press.

NASA (National Aeronautics and Space Administration). (2002) Probabilistic risk assessment procedures guide for NASA managers and practitioners. Version 1.1. Office of Safety and Mission Assurance, NASA Headquarters Washington, DC (Aug.). Available at: http://www.hq.nasa.gov/office/codeq/doctree/praguide.pdf.

Nebert, DW; Petersen, DD; Fornace, AJ, Jr. (1990) Cellular responses to oxidative stress: the [Ah] gene battery as a paradigm. Environ Health Perspect 88:13–25.

Nebert, DW; Peterson, DD; Puga, A. (1991) Human Ah locus polymorphism and cancer: Inducibility of CYPIA1 and other genes by combustion products and dioxin. Pharmacogenetics 1:68–78.

Needham, LL; Gerthoux, PM; Patterson, DG; et al. (1994) Half-life of 2,3,7,8-tetrachlorodibenzo-p-dioxin in serum of Seveso adults: Interim report. Organohal Comp 21:81−85.

Needham, LL; Gerthoux, PM; Patterson, DG, Jr; et al. (1997) Serum dioxin levels in Seveso, Italy, population in 1976. Teratog Carcinog Mutagen 17:225−240.

Needham, LL; Barr, DB; Caudill, SP; et al. (2005) Concentrations of environmental chemicals associated with neurodevelopmental effects in the US population. NeuroToxicology 26(4):531–545.

Nessel, CS; Amoruso, MA; Umbreit, TH; et al. (1992) Transpulmonary uptake and bioavailability of 2,3,7,8-TCDD from respirable soil particles. Chemosphere 25(1–2):29–32.

Nilsson, CB; Hakansson, H. (2002) The retinoid signaling system – a target in dioxin toxicity. Crit Rev Toxicol 32(3):211–232.

Nishimura, N; Yonemoto, J; Miyabara, Y; et al. (2005) Altered thyroxin and retinoid metabolic response to 2,3,7,8-tetrachlorodibenzo-p-dioxin in aryl hydrocarbon receptor-null mice. Arch Toxicol 79:260–267.

Niskar, A; Needham, LL; Rubin, C; et al. (2009) Serum dioxin, polychlorinated biphenyls, and endometriosis: a case-control study in Atlanta. Chemosphere 74(7):944–949.

Nohara, K; Fujimaki, N; Tsukumo, S; et al. (2000) The effects of perinatal exposure to low doses of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin on immune organs in rats. Toxicology 154:123–133.

Nohara, K; Izumi, H; Tamura, S; et al. (2002) Effect of low-dose 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) on influenza A virus-induced mortality in mice. Toxicology 170:131–138.

Nolan, KJ; Smith, FA; Hefner, JG. (1979) Elimination and tissue distribution of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in female guinea pigs following a single oral dose. Toxicol Appl Pharmacol 48:162.

NRC (National Research Council). (1983) Risk assessment in the federal government: managing the process. Washington, DC: National Academies Press. Available at: http://www.nap.edu/openbook.php?isbn=0309033497.

NRC (National Research Council). (1989) Improving risk communication. Washington, DC: National Academies Press. Available at: http://www.nap.edu/openbook.php?isbn=0309039436.

NRC (National Research Council). (1991) Human exposure assessment for airborne pollutants: advances and opportunities. Washington, DC: National Academies Press. Available at: http://books.nap.edu/openbook.php?record_id=1544.

NRC (National Research Council). (1993) Issues in risk assessment. Washington, DC: National Academies Press. Available at: http://www.nap.edu/catalog.php?record_id=2078.

NRC (National Research Council). (1994) Science and judgment in risk assessment. Washington, DC: National Academies Press. Available at: http://www.nap.edu/openbook.php?record_id=2125&page=629.

NRC (National Research Council). (2002) Estimating the public health benefits of proposed air pollution regulations. Washington, DC: National Academies Press. Available at: http://www.nap.edu/catalog.php?record_id=10511. **)**

NRC (National Research Council). (2006) Health risks from dioxin and related compounds: evaluation of the EPA reassessment. Washington, DC: National Academies Press. Available at: http://www.nap.edu/catalog.php?record_id=11688.

NRC (National Research Council). (2007) Scientific review of the proposed risk assessment bulletin from the Office of Management and Budget. Washington, DC: National Academies Press. Available at: http://www.nap.edu/catalog.php?record_id=11811.

NRC (National Research Council). (2009) Science and decisions: advancing risk assessment. Committee on Improving Risk Analysis Approaches Used by the U.S. EPA. Washington, DC: National Academies Press. Available at: http://www.nap.edu/catalog.php?record_id=12209**.**

NTP (National Toxicology Program). (1982) NTP Technical Report on carcinogenesis bioassay of 2,3,7,8-tetrachlorodibenzo-p-dioxin in Osborne-Mendel rats and B6C3F1 mice (gavage study). Public Health Service, U.S. Department of Health and Human Services; NTP TR 209. Available from the National Institute of Environmental Health Sciences, Research Triangle Park, NC.

NTP (National Toxicology Program). (2006) 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 3,3',4,4',5-pentachlorobiphenyl (PCB 126) (CAS No. 57465-28-8) in female Harlan Sprague-Dawley Rats (gavage studies). Public Health Service, U.S. Department of Health and Human Services; NTP TR 526. Available from the National Institute of Environmental Health Sciences, Research Triangle Park, NC. Available at: http://ntp.niehs.nih.gov/index.cfm?objectid=070B7300-0E62-BF12-F4C3E3B5B645A92B.

Oehlert, GW. (1992) A note on the delta method. Am Stat 46(1):27–29.

Ohsako, S; Miyabara, Y; Nishimura, N; et al. (2001) Maternal exposure to a low dose of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) suppressed the development of reproductive organs of male rats: dose-dependent increase of mRNA levels of 5α-reductase type 2 in contrast to decrease of androgren receptor in the pubertal ventral prostate. Toxicol Sci 60:132–143.

Okey, AB; Riddick, DS; Harper; PA. (1994) The Ah receptor: mediator of the toxicity of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) and related compounds. Toxicol Lett 70:1–22.

Olson, JR; Holscher, MA; Neal, RA. (1980) Toxicity of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin in the Golden Syrian Hamster. Toxicol Appl Pharmacol 55:67–78.

Olson, JR; McGarrigle, BP; Gigliotti, PJ; et al. (1994) Hepatic uptake and metabolism of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin and 2,3,7,8-tetrachlorodibenzofuran. Fundam Appl Toxicol 22:631–640.

OMB (Office of Management and Budget). (2006) Proposed risk assessment bulletin (Draft for public review). Washington, DC (Jan.). Available at: http://www.whitehouse.gov/omb/inforeg/proposed_risk_assessment_bulletin_010906.pdf.

Ott, MG; Zober A. (1996) Cause specific mortality and cancer incidence among employees exposed to 2,3,7,8-TCDD after a 1953 reactor accident. Occup Environ Med 53:606–612.

Ott, MG; Olson, RA; Cook, RR; et al. (1987) Cohort mortality study of chemical workers with potential exposure to the higher chlorinated dioxins. J Occup Med 29(5):422–429.

Ott, MG; Messerer, P; Zober, A. (1993) Assessment of past occupational exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin using blood lipid analyses. Int Arch Occup Environ Health 65:1–8.

Papke, O; Ball, M; Lis, A. (1994) PCDD/PCDF in humans - an update of background data. Chemosphere 29:2355−2360.

Pekelis, M; Nicolich, MJ; Gauthier, JS. (2003) Probabilistic framework for the estimation of the adult and child toxicokinetic intraspecies uncertainty factors. Risk Anal 23(6):1239–1255. Available at: http://www3.interscience.wiley.com/cgi-bin/fulltext/118881007/PDFSTART.

Percy, C; Stanek, E, III; Gloeckler, L. (1981) Accuracy of cancer death certificates and its effect on cancer mortality statistics. Am J Public Health 71(3):242−250.

Pereg, D; Dewailly, E; Poirier, GG; et al. (2002) Environmental exposure to polychlorinated biphenyls and placental CYP1A1 activity in Inuit women from northern Québec. Environ Health Perspect 110(6):607–613 (June). Available at: http://www.pubmedcentral.nih.gov/picrender.fcgi?artid=1240878&blobtype=pdf.

Pesatori, AC; Zocchetti, C; Guercilena, S; et al. (1998) Dioxin exposure and non-malignant health effects: A mortality study. Occup Environ Med 55:126–131.

Pesatori, AC; Consonni, D; Bachetti, S; et al. (2003) Short- and long-term morbidity and mortality in the population exposed to dioxin after the "Seveso Accident." Industrial Health 41:127−138.

Piacitelli, L; Sweeney, MH; Fingerhut, M. (1992) Serum levels of 2,3,7,8-substituted PCDDs among workers exposed to 2,3,7,8-TCDD contaminated chemicals. Chemosphere 25:251−254.

Pipe, NG; Smith, T; Halliday, D; et al. (1979) Changes in fat, fat-free mass and body water in human normal pregnancy. Br J Obstet Gynaecol 86(12):929−940.

Pirkle, JL; Wolfe, WH; Patterson, DG; et al. (1989) Estimates of the half-life of 2,3,7,8-tetrachlorodibenzo-p-dioxin in Vietnam Veterans of Operation Ranch Hand. J Tox Environ Health 27(2):165−171.

Pitot, HC; Goldsworthy, T; Campbell, HA; et al. (1980) Quantitative evaluation of the promotion by 2,3,7,8-tetrachlorodibenzo-p-dioxin of hepatocarcinogenesis from diethylnitrosamine. Cancer Res 40:3616−3620.

Pohjanvirta, R; Tuomisto, J. (1994) Short-term toxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in laboratory animals: Effects, mechanisms, and animal models. Pharmacol Rev 46:483–549.

Pohjanvirta, R; Tuomisto, L; Tuomisto, J. (1989) The central nervous system may be involved in TCDD toxicity. Toxicology 58:167−174.

Poiger, M; Schlatter, C. (1986) Pharmacokinetics of 2,3,7,8-TCDD in man. Chemosphere 15:9−12.

Poland, A;Glover, E. (1980) 2,3,7,8-Tetrachlorodibenzo-p-dioxin: Segregation of toxicity with the Ah locus. Mol. Pharmacol. 17:86–94.

Poland, A; Glover, E. (1990) Characterization and strain distribution pattern of the murine Ah receptor specified by the Ahd and Ahb-3 alleles. Mol Pharmacol 38(3):306−312.

Poland, A; Palen, D; Glover, E. (1982) Tumor promotion by TCDD in skin of HRS/J mice. Nature 300(5889):271−273.

Poland, A; Palen, D; Glover, E. (1994) Analysis of the four alleles of the murine aryl hydrocarbon receptor. Mol Pharmacol 46:915−921.

Popp, JA; Crouch, E; McConnell, EE. (2006) A weight-of-evidence analysis of the cancer dose-response characteristics of 2,3,7,8-tetrachlorodibenzodioxin (TCDD). Toxicol Sci 89(2): 361−369.

Potter, CL; Sipes, IG; Russell, DH. (1983) Hypothyroxinemia and hypothermia in rats in response to 2,3,7,8-tetrachlorodibenzo-p-dioxin administration. Toxicol Appl Pharmacol 69(1):89−95.

Potter, CL; Moore, RW; Inhorn, SL; et al. (1986) Thyroid status and thermogenesis in rats treated with 2,3,7,8-tetrachlorodibenzo-p-dioxin. Toxicol Appl Pharmacol 84(1):45−55.

Poulin P; Theil FP. (2001) Prediction of pharmicokinetics prior to in vivo studies. 1. mechanism-based prediction of volume of distribution. J Pharm Sci 91(1):129−156.

Puga, A; Nebert, DW; Carier, F. (1992) Dioxin induces expression of c-fos and c-jun proto-oncogenes and a large increases in transcription factor AP-1. DNA Cell Biol 11:269−281.

Ramadoss, P; Perdew, GH. (2004) Use of 2-Azido-3-[125I]iodo-7,8-dibromodibenzo-*p*-dioxin as a probe to determine the relative ligand affinity of human versus mouse aryl hydrocarbon receptor in cultured cells. Molec Pharmacol 66:129−136.

Ramsey, JC; Hefner, JG; Karbowski, RJ; et al. (1982) The in vivo biotransformation of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in the rat. Toxicol Appl Pharmacol 65(1):180−184.

Rao, MS; Subbarao, V; Prasad, JD; et al. (1988) Carcinogenicity of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin in the Syrian golden hamster. Carcinogenesis 9(9):1677–1679.

Reddy, M; Yang, R; Clewell, HJ; et al. (2005) Physiologically based pharmacokinetic modeling: science and applications. Hoboken, New Jersey: John Wiley & Sons.

Revich, B; Aksel, E; Ushakova, T; et al. (2001) Dioxin exposure and public health in Chapaevsk, Russia. Chemosphere 43(4−7):951−966.

Revich, B; Sergeyev, O; Zeilert, V; et al. (2005) Chapaevsk, Russia: 40 years of dioxins exposure on the human health and 10 years of Russian – USA epidemiological studies.

Rier, SE; Martin, DC; Bowman, RE; et al. (1993) Endometriosis in rhesus monkeys (*Macaca mulata*) following chronic exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Fundam Appl Toxicol 21:433−441.

Rier, SE; Martin, DC; Bowman, RE; et al. (1995) Immunoresponsiveness in endometriosis: implications of estrogenic toxicants. Environ Health Perspect 103 (Suppl 7):151−156.

Rier, SE; Turner, WE; Martin, DC; et al. (2001a) Serum levels of TCDD and dioxin-like chemicals in rhesus monkeys chronically exposed to dioxin: correlation of increased serum PCB levels with endometriosis. Toxicol Sci 59:147–159.

Rier, SE; Coe, CL; Lemieux, AM; et al. (2001b) Increased tumor necrosis factor-α production by peripheral blood leukocytes from TCDD-exposed rhesus monkeys. Toxicol Sci 60:327–337.

Roberts, EA; Shear, NH; Okey, AB; et al. (1985) The Ah receptor and dioxin toxicity: from rodent to human tissues. Chemosphere 14:661–674.

Roberts, EA; Golas, CL; Okey, AB. (1986) Ah receptor mediating induction of aryl hydrocarbon hydroxylase: detection in human lung by binding of 2,3,7,8-[3H]tetrachlorodibenzo-p-dioxin. Cancer Res 46:3739–3743.

Rohde, S; Moser, GA; Päpke,O; et al. (1999) Clearance of PCDD/Fs via the gastrointestinal tract in occupationally exposed persons. Chemosphere 38(14):3397–3410.

Roth, WL; Ernst, S; Weber, LW; et al. (1994) A pharmacodynamically responsive model of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) transfer between liver and fat at low and high doses. Toxicol Appl Pharmacol 127:151–162.

Rothman, KJ. (1986) Modern epidemiology. Toronto: Little, Brown, and Company.

Roy, T; Hammerstrom, K; Schaum, J. (2008) Percutaneous absorption of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) from soil. J Toxicol Environ Health A 71(23):1509–1515.

Rozman, KK. (2000) The role of time in toxicology or Haber's c × t product. Toxicol 149:35–42.

Ryan, JJ; Schecter, A. (2000) Exposure of Russian phenoxy herbicide producers to dioxin. J Occup Environ Med 42:861–870.

Ryan, JJ; Amirova, Z; Carrier, G. (2002) Sex ratios of children of Russian pesticide producers exposed to dioxin. Environ Health Perspect 110(11):A699−701.

Saltelli, A; Chan, K; Scott, EM; eds. (2000) Sensitivity analysis. Wiley Series in Probability and Statistics, John Wiley & Sons Ltd., England (Oct.) (reprinted Nov. 2001, John Wiley & Sons, Inc., New York, NY).

Sandau, CD; Ayotte, P; Dewailly, E; et al. (2002) Pentachlorophenol and hydroxylated polychlorinated biphenyl metabolites in umbilical cord plasma of neonates from coastal populations in Quebec. Environ Health Perspect 110(4):411–417.

Santostefano, MJ; Johnson, KL; Whisnant, NA; et al. (1996) Subcellular localization of TCDD differs between the liver, lungs, and kidneys after acute and subchronic exposure: species/dose comparison and possible mechanism. Fundam Appl Toxicol 34(2):265–275.

Santostefano, MJ; Wang, X; Richardson, VM; et al. (1998) A pharmacodynamic analysis of TCDD-induced cytochrome P450 gene expression in multiple tissues: dose- and time-dependent effects. Toxicol Appl Pharmacol 151:294–310.

Saracci, R; Kogevinas, M; Bertazzi, PA; et al. (1991) Cancer mortality in workers exposed to chlorophenoxy herbicides and chlorophenols. Lancet 338(8774):1027−1032.

Sauer, RM. (1990) 2,3,7,8-tetrachlorodibenzo-p-dioxin in Sprague-Dawley rats. Submitted to the Maine Scientific Advisory Panel by Pathco, Inc., Ijamsville, MD. March 13, 1990.

Schantz, SL; Bowman, RE. (1989) Learning in monkeys exposed perinatally to 2,3,7,8-tetrachloridibenzo-p-dioxin (TCDD). Neurotoxicol Teratol 11:13–19.

Schantz, SL; Laughlin, NK; Van Valkenberg, HC; et al. (1986) Maternal care by rhesus monkeys exposed to either lead or 2,3,7,8-tetrachlorodibenzo-p-dioxin. Neurotoxicology 7(2):637−650.

Schantz, SL; Seo, BW; Moshtaghian, J; et al. (1996) Effects of gestational and lactational exposure to TCDD or coplanar PCBs on spatial learning. Neurotoxicol Teratol 18(3):305–313.

Schecter, A; Cramer P; Boggess K; et al. (1997) Levels of dioxins, dibenzofurans, PCB and DDE congeners in pooled food samples collected in 1995 at supermarkets across the United States. Chemosphere 34(5–7):1437–1447.

Schwartz, M; Appel, KE. (2005) Carcinogenic risks of dioxin: mechanistic considerations. Regul Toxicol Pharmacol 43(1):19–34.

Seidel, SD; Winters, GM; Rogers, WJ; et al. (2001) Activation of the Ah receptor signaling pathway by prostaglandins. J Biochem Mol Toxicol 15:187–196.

Self, SG; Liang, KY. (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J Am Stat Assoc 82: 605–610.

Seo, B-W; Li, M-H; Hansen, LG; et al. (1995) Effects of gestational and lactational exposure to coplanar polychlorinated biphenyl (PCB) congeners or 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) on thyroid hormone concentrations in weanling rats. Toxicology Letters 78(3):253-−262.

Sewall, CH; Lucier, GW; Tritscher, AM; et al. (1993) TCDD-mediated changes in hepatic epidermal growth factor receptor may be a critical event in the hepatocarcinogenic action of TCDD. Carcinogenesis 14:1885–1893.

Sewall, CH; Flagler, N; Vanden Heuvel, JP; et al. (1995) Alterations in thyroid function in female Sprague-Dawley rats following chronic treatment with 2,3,7,8-tetrachlorodibenzo-p-dioxin. Toxicol Appl Pharmacol 132:237–244.

Shi, Z; Valdez, KE; Ying, AY; et al. (2007) Ovarian endocrine disruption underlies premature reproduction senescence following environmentally relevant chronic exposure to aryl hydrocarbon receptor agonist 2,3,7,8-tetrachlorodibenzo-*p*-dioxin. Biol Reprod 30(4):293−342.

Shu, H; Teitelbaum, P; Ebb, AS; et al. (1988) Bioavailability of soil-bound TCDD: dermal bioavailability in the rat. Fundam Appl Toxicol (2):335−343.

Siemiatycki, J; Wacholder, S; Dewar, R; et al. (1988) Degree of confounding bias related to smoking, ethnic group, and socioeconomic status in estimates of the associations between occupation and cancer. J Occup Med 30(8):617−625.

Sikov, M. (1970) Radiation biology of the fetal and juvenile mammal. Science 167(3925):1640−1641.

Simanainen, U; Tuomisto, JT; Tuomisto, J; Viluksela, M. (2002) Structure-activity relationships and dose responses of polychlorinated dibenzo-p-dioxins for short-term effects in 2,3,7,8-tetrachlorodibenzo-p-dioxin-resistant and –sensitve rat strains. Toxicol Appl Pharmacol 181:38-47.

Simanainen, U; Haavisto, T; Tuomisto, J.T; et al. (2004) Pattern of male reproductive system effects after in utero and lactational 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) exposure in three differentially TCDD-sensitive rat lines. Toxicol Sci 80(1):101−108.

Simon, T; Aylward, LL; Kirman, CR; et al. (2009) Estimates of cancer potency of 2,3,7,8-tetrachlorodibenzo(p)dioxin using linear and non-linear dose-response modeling and toxicokinetics. Toxicol Sci Advance access published online on September 23, 2009. doi:10.1093/toxsci/kfp232.

Slezak, BP; Hatch, GE; DeVito, MJ; et al. (2000) Oxidative stress in female B6C3F1 mice following acute and subchronic exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD). Toxicol Sci 54:390−398.

Smart, J; Daly, A. (2000) Variation in induced CYP1A1 levels: relationship to CYP1A1, Ah receptor, and GSTM1 polymorphisms. Pharmacogenetics 10(1):11−24.

Smialowicz, RJ; Burgin, DE; Williams, WC; et al. (2004) CYP1A2 is not required for 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced immunosuppression. Toxicology 197(1):15-22.

Smialowicz, RJ; DeVito, MJ; Williams, WC; et al. (2008) Relative potency based on hepatic enzyme induction predicts immunosuppressive effects of a mixture of PCDDS/PCDFS and PCBS. Toxicol Appl Pharmacol 227:477–484.

Smith, AH; Lopipero, P. (2001) Invited commentary: How do the Seveso findings affect conclusions concerning TCDD as a human carcinogen? Am J Epidemiol 153(11):1045−1047.

Smith, AH; Fisher, DO; Pearce, N; et al. (1982) Congenital defects and miscarriages among New Zealand 2, 4, 5-T sprayers. Arch Environ Health 37(4):197−200.

Spiegelhalter, D; Thomas, A; Best, N; et al. (2003) BUGS 0.5 Bayesian inference using Gibbs sampling manual, version ii. MRC Biostatistics Units, Institute of Public Health, Cambridge. Available at: http://193.60.86.19/bugs/documentation/Download/manual05.pdf

Squire, RA. (1980) Pathologic evaluations of selected tissues from the Dow Chemical TCDD and 2,4,5-T rat studies. Submitted to Carcinogen Assessment Group, U.S. Environmental Protection Agency on August 15 under Contract No. 68-01-5092.

Starr, TB. (2003) Significant issues raised by meta-analyses of cancer mortality and dioxin exposure. Environ Health Perspect 111(12):1443−1147.

Staskal, DF; Diliberto, JJ; Devito, MJ; et al. (2005) Inhibition of human and rat CYP1A2 by TCDD and dioxin-like chemicals. Toxicol Sci 84:225–231.

Stayner, L; Bailer, AJ; Smith, R; et al. (1999) Sources of uncertainty in dose-response modeling of epidemiological data for cancer risk assessment. Ann NY Acad Sci 895:212−222.

Stayner, L; Steenland, K; Dosemeci, M; et al. (2003) Attenuation of exposure-response curves in occupational cohort studies at high exposure levels. Scand J Work Environ Health 29(4):317−324.

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

Steenland, K; Piacitelli, L; Deddens, J; et al. (1999) Cancer, heart disease, and diabetes in workers exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin. J Natl Cancer Inst 91(9):779−786.

Steenland, K; Deddens, J; Piacitelli, L. (2001) Risk assessment for 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) based on an epidemiologic study. Am J Epidemiol 154(5):451−458.

Stellman, SD; Stellman, JM. (1986) Estimation of exposure to Agent Orange and other defoliants among American troops in Vietnam: a methodological approach. Am J Ind Med 9:305−321.

Stephenson, RP. (1956) A modification of receptor theory. Br J Pharmacol 11:379.

Sugita-Konishi, Y; Kobayashi, K; Naito, H; et al. (2003) Effect of lactational exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin on the susceptibility to *Listeria* infection. Biosci Biotechnol Biochem 67(1):89−93.

Swartout, JC; Price, PS; Dourson. ML; et al. (1998) A probabilistic framework for the reference dose (probabilistic RfD). Risk Anal 18(3):271−282.

Takemoto, K; Nakajima, M; Fujiki, Y; et al. (2004) Role of the aryl hydrocarbon receptor and Cyp1b1 in the anti-estrogenic activity of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin. Arch Toxicol 78:309−315.

Thiess, AM; Frentzel-Beyme, R. (1977) Mortality study of persons exposed to dioxin following an accident which occurred in the BASF on 17 November 1953. In: Proceedings of the 5th International Conference Medichem, 1977. San Francisco, CA. pp. 228−236.

Thiess, AM; Frentzel-Beyme, R; Link, R. (1982) Mortality study of persons exposed to dioxin in a trichlorophenol-process accident that occurred in the BASF AG on November 17, 1953. Am J Ind Med 3(2):179−189.

Tian, Y; Ke, S; Denison, MS; et al. (1999) Ah receptor and NF-kB interactions, a potential mechanism for dioxin toxicity. J Biol Chem 274: 510−515.

t'Mannetje, A; McLean, D; Cheng, S; et al. (2005) Mortality in New Zealand workers exposed to phenoxy herbicides and dioxins. Occup Environ Med 62(1):34−40.

Toide, K; Yamazaki, JH; Nagashima, R; et al. (2003) Aryl hydrocarbon hydroxylase represents CYP1B1 and not CYP1A1, in human freshly isolated white cells: Trimodal distribution of Japanese population according to induction of CYP1B1 mRNA by environmental dioxins. Cancer Epidemiol Biomark Prev 12:219−222.

Toth, L; Somfai-Relle, S; Sugár, J; et al. (1979) Carcinogenicity testing of herbicide 2,4,5-trichlorophenoxyethanol containing dioxin and of pure dioxin in Swiss mice. Nature 278:548−549.

Tritscher, AM; Mahler, J; Portier, CJ; et al. (2000) Induction of lung lesions in female rats following chronic exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Toxicol Pathol 28:761−769.

Tuomisto, JT; Viluksela, M; Pohjanvirta, R; et al. (1999) The AH receptor and a novel gene determine acute toxic responses to TCDD: segregation of the resistant alleles to different rat lines. Toxicol Appl Pharm 155(1):71−81.

Tuomisto, JT; Wilson, AM; Evans, JS; et al. (2008) Uncertainty in mortality response to airborne fine particulate matter: combining European air pollution experts. Reliab Eng Syst Saf 93(5):775−777.

Umemura, T; Kai, S; Hasegawa1, R; et al. (1999) Pentachlorophenol (PCP) produces liver oxidative stress and promotes but does not initiate hepatocarcinogenesis in B6C3F1 mice. Carcinogenesis 20(6):1115–1120.

U.S. DOE (Department of Energy). (1992) DOE standard, hazard categorization, and accident analysis techniques for compliance with DOE Order 5480.23, nuclear safety analysis reports. DOE-STD-1027-92 (Change Notice No.1, Sept. 1997), Washington, DC. Available at: http://www.hss.doe.gov/nuclearsafety/ns/techstds/standard/std1027/s1027cn1.pdf.

U.S. EPA. (1996a) Proposed guidelines for carcinogen risk assessment. Risk Assessment Forum, Washington, DC. Federal Register 61:17960–18011.

U.S. EPA. (1998b) Guidelines for neurotoxicity risk assessment. Federal Register 63(93):26926-26954. Available from: http://cfpub.epa.gov/ncea/raf/recordisplay.cfm?deid=12479.

U.S. EPA. (2000) Benchmark dose technical guidance document. October. External review draft. Washington, DC. EPA/630/R-00/001.

U.S. EPA. (2003) Exposure and human health reassessment of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) and related compounds [NAS review draft]. Volumes 1−3. National Center for Environmental Assessment, Washington, DC; EPA/600/P-00/001 Cb. Available at: http://www.epa.gov/nceawww1/pdfs/dioxin/nas-review/.

U.S. EPA. (2005) Guidelines for carcinogen risk assessment. Risk Assessment Forum, Washington, DC. EPA/630/P-03/001F. Available at http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=116283.

U.S. EPA. (2008a) Framework for application of the toxicity equivalence methodology for polychlorinated dioxins, furans, and biphenyls in ecological risk assessment. Risk Assessment Forum, Washington, DC; EPA/100/R-08/004

U.S. EPA. (2008b) 2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD) dose-response studies: preliminary literature search results and request for additional studies. National Center for Environmental Assessment, Cincinnati, OH; EPA/600/R-08/119. Available at: http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=199923.

U.S. EPA. (2008c) Benchmark dose technical guidance document. Risk Assessment Forum, Washington, DC; EPA/630/R-00/0001F.

U.S. EPA. (2009a) Integrated Risk Information System (IRIS). National Center for Environmental Assessment, Washington, DC. Available at: http://www.epa.gov/ncea/iris.

U.S. EPA. (2009b) Summary of U.S. EPA dioxin workshop: February 18−20, 2009. National Center for Environmental Assessment, Cincinnati, OH; EPA/600/R-09/027.

U.S. EPA. (2009c) Using probabilistic methods to enhance the role of risk analysis in decision-making with case study examples. Risk Assessment Forum, Washington, DC. EPA/100/R-09/001.

U.S. EPA. (2010) Recommended toxicity equivalency factors (TEFs) for human health risk assessments of dioxin and dioxin-like compounds [external review draft]. Risk Assessment Forum, Washington, DC; NCEA-C-2100.

U.S. NRC (Nuclear Regulatory Commission). (1975) Reactor safety study – an assessment of accident risks in U.S. commercial nuclear power plants. NUREG-75/014 (WASH-1400). Prepared by N. Rasmussen (MIT) et al., for the U.S. Nuclear Regulatory Commission, Rockville, MD.

U.S. NRC (Nuclear Regulatory Commission). (1981) Fault tree handbook. NUREG-0492, Systems and Reliability Research, Office of Nuclear Regulatory Research, Rockville, MD. Available at: http://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr0492/sr0492.pdf.

U.S. NRC (Nuclear Regulatory Commission). (1983) A guide to the performance of probabilistic risk assessments for nuclear power plants. Final Report. NUREG/CR-2300. Rockville, MD. Available at: http://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr2300/vol2/cr2300v2-a.pdf.

U.S. NRC (Nuclear Regulatory Commission). (1991) Severe accident risks: an assessment for five U.S. nuclear power plants. NUREG-1150. Final Summary Report, Final Report. Rockville, MD. Selected portions available at: http://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr1150/; http://www.nrc.gov/reading-rm/doc-collections/nuregs/staff/sr1150/v3/sr1150v3.pdf.

Van Birgelen, AP; Van den Berg, M. (2000) Toxicokinetics. Food Addit Contam 17:267−273.

Van Birgelen, AP; Van der Kolk, J; Fase, KM; et al. (1995) Subchronic dose-response study of 2,3,7,8-tetrachlorodibenzo-p-dioxin in female Sprague-Dawley rats. Toxicol Appl Pharmacol 132:1-13.

Van den Berg, M; de Vroom, E; Olie, K; et al. (1986) Bioavailability of PCDDs and PCDFs of fly ash after semi-chronic oral ingestion by guinea pig and Syrian golden hamster. Chemosphere 15(4):519–533.

Van den Berg, M; Birnbaum, L; Bosveld, AT; et al. (1998) Toxic equivalency factors (TEFs) for PCBs, PCDDs, PCDFs for humans and wildlife. Environ Health Perspect 106(12):775−792.

Van den Berg, M; Birnbaum, LS; Denison, M; et al. (2006) The 2005 World Health Organization reevaluation of human and mammalian toxic equivalency factors for dioxins and dioxin-like compounds. Tox Sci 93(2):223–241. Available at: http://toxsci.oxfordjournals.org/cgi/reprint/93/2/223.

Van der Molen, G; Kooijman, A; Slob, W. (1996) A generic toxicokinetic model for persistent lipophilic compounds in humans: an application to TCDD. Fund Appl Toxicol 31:83–94.

Van der Molen, GW; Kooijman, SALM; Michalek, JE; et al. (1998) The estimation of elimination rates of persistent compounds: a re-analysis of 2,3,7,8-tetrachlorodibenzo-p-dioxin levels in Vietnam veterans. Chemosphere 37(9–12):1833–1844.

Van der Molen, GW; Kooijman, BA; Wittsiepe, J; et al. (2000) Estimation of dioxin and furan elimination rates with a pharmacokinetic model. J Expo Anal Environ Epidemiol 10:579–585.

Vanden Heuvel, JP; Clark, GC; Tritscher, A; et al. (1994) Accumulation of polychlorinated dibenzo-p-dioxins and dibenzofurans in liver of control laboratory rats. Fundam Appl Toxicol 23:465–469.

Vanni, H; Kazeros, A; Wang, R; et al. (2009) Cigarette smoking induces overexpression of a fat-depleting gene AZGP1 in the human airway epithelium. Chest 135(5):1197–1208. Available at: http://chestjournal.chestpubs.org/content/135/5/1197.full.pdf+html.

Viluksela, M; Bager, Y; Tuomisto**,** JT. (2000) Liver tumor-promoting activity of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in TCDD-sensitive and TCDD-resistant rat strains. Cancer Res 60:6911–6920.

Vos, JG; Moore, JA; Zinkl, JG. (1973) Effect of 2,3,7,8-tetrachlorodibenzo-p-dioxin on the immune system of laboratory animals. Environ Health Perspect 5:149–162.

Walker, NJ; Portier, CJ; Lax, SF; et al. (1999) Characterization of the dose-response of CYP1B1, CYP1A1, and CYP1A2 in the liver of female Sprague-Dawley rats following chronic exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. Toxicol Appl Pharmacol 1;154(3):279–286.

Wang, X; Santostefano, MJ; Evans, MV; et al. (1997) Determination of parameters responsible for pharmacokinetic behavior of TCDD in female sprague-dawley rats. Toxicol Appl Pharmacol147:151–168.

Wang, X; Santostefano, MJ; Devito, MJ; et al. (2000). Extrapolation of a PBPK model for dioxins across dosage regimen, gender, strain, and species. Toxicol Sci 56(1):49–60.

Wang, SL; Su, PH; Jong, SB; et al. (2005) In utero exposure to dioxins and polychlorinated biphenyls and its relations to thyroid function and growth hormone in newborns. Environ Health Perspect 113:1645−1650.

Warner, M; Eskenazi, B; Mocarelli, P; et al. (2002) Serum dioxin concentrations and breast cancer risk in the Seveso Women's Health Study. Environ Health Perspect 110(7):625−628.

Warner, M; Samuels, S; Mocarelli, P; et al. (2004) Serum dioxin concentrations and age at menarche. Environ Health Perspect 112(13):1289−1292.

Warner, M; Eskenazi, B; Olive, DL; et al. (2007) Serum dioxin concentrations and quality of ovarian function in women of Seveso. Environ Health Perspect 115(3):336−340.

Weber, R; Schmitz, H-J; Schrenk, D; et al. (1997) Metabolic degradation, inducing potency, and metabolites of fluorinated and chlorinated-fluorinated dibenzodioxins and dibenzofurans. Chemosphere 34(1):29–40.

Wendling, JM; Orth, RG; Poiger, H. (1990) Determination of [3H]-2,3,7,8-tetrachlorodibenzo-p-dioxin in human feces to ascertain its relative metabolism in man. Anal Chem 62(8):796–800.

White, KL, Jr; Lysy, HH; McCay, JA; et al. (1986) Modulation of serum complement levels following exposure to polychlorinated dibenzo-p-dioxins. Toxicol Appl Pharmacol 84:209–219.

WHO (World Health Organization). (1978) International Classification of Diseases: Ninth Revision. ISBN-13 9789241541336.

WHO (World Health Organization). (1988) Assessment of the health risk of dioxins: re-evaluation of the tolerable daily intake (TDI). WHO Consultation May 25−29, 1998. Geneva, Switzerland: WHO European Centre for Environmental Health and International Programme on Chemical Safety.

WHO (World Health Organization). (1998) Executive summary, assessment of health risk of dioxins: re-evaluation of the tolerable daily intake (TDI), WHO Consultation, May 25–29.

WHO (World Health Organization). (2005) Chemical-specific adjustment factors for interspecies differences and human variability: guidance document for use of data in dose/concentration–response assessment. Harmonization Project Document 2, IPCS project on Harmonization of Approaches to the Assessment of Risk from Exposure to Chemicals. Geneva.

Whysner, J; Williams, GM. (1996) 2,3,7,&tetrachlorodibenzo+hdioxin mechanistic data and risk assessment: gene regulation, cytotoxicity, enhanced cell proliferation, and tumor promotion. Pharmacol Ther 71(l/2):193–223.

Wittsiepe, J; Erlenkämper, B; Welge, P; et al. (2007) Bioavailability of PCDD/F from contaminated soil in young Goettingen minipigs. Chemosphere 67(9):S355−S364.

Wong, T K; Domin, BA; Bent, PE; et al. (1986) Correlation of placental microsomal activities with protein detected by antibodies to rabbit cytochrome P-450 isozyme 6 in preparations from humans exposed to polychlorinated biphenyls, quaterphenyls, and dibenzofurans. Cancer Res 46:999–1004

Woods, CG; Burns, A; Bradford, BU; et al. (2007) WY-14,643-induced cell proliferation and oxidative stress in mouse liver are independent of NADPH oxidase. Toxicol Sci 98(2):366–374.

Wyde, M; Wong, VA; Kim, AH; et al. (2001) induction of hepatic 8-oxo-deoxyguanosine adducts by 2,3,7,8-tetrachlorodibenzo-p-dioxin in Sprague-Dawley rats is female-specific and estrogen-dependent. Chem Res Toxicol 14:849–855.

Wyde, M; Cambre, T; Lebetkin, M; et al. (2002) Promotion of altered hepatic foci by 2,3,7,8-tetrachlorodibenzo-p-dioxin and 17b-estradiol in male Sprague-Dawley rats. Toxicol Sci 68:295–303.

Yang, JZ; Agarwal, SK; Foster, WG. (2000) Subchronic exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin modulates the pathophysiology of endometriosis in the cynomolgus monkey. Toxicol Sci 56:374–381.

Youakim, S. (2006) Risk of cancer among firefighters: a quantitative review of selected malignancies. Int Arch Occup Environ Health 61(5):223−231.

Zack, JA; Gaffey, WR. (1983) A mortality study of workers employed at the Monsanto Company plant in Nitro, West Virginia. Environ Sci Res 26:575−591.

Zack, JA; Suskind, R. (1980) The mortality experience of workers exposed to tetrachlorodibenzodioxin in a trichlorophenol process accident. J Occup Med 22:11–14.

Zareba, G; Hojo,R; Zareba, GM; et al. (2002) Sexually dimorphic alterations of brain cortical dominance in rats prenatally exposed to TCDD. J Appl Toxicol 22:129–137.

Zober, A; Papke, O. (1993) Concentrations of PCDDs and PCDFs in human tissue 36 years after accidental dioxin exposure. Chemosphere 27:413−418.

Zober, A; Messerer, P; Huber, P. (1990) Thirty-four-year mortality follow-up of BASF employees exposed to 2,3,7,8-TCDD after the 1953 accident. Int Arch Occup Environ Health 62(2):139−157.

Zober, A; Ott, MG; Messerer, P. (1994) Morbidity follow-up study of BASF employees exposed to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) after a 1953 chemical reactor incident. J Occup Environ Med 51:479–486.

Zober, A; Schilling, D; Ott, M; et al. (1998) *Helicobacter pylori* infection: prevalence and clinical relevance in a large company. Occup Environ Med 40(7):586–594.