EPA/635/R-22/277Fb
www.epa.gov/iris

IRIS Toxicological Review of Perfluorobutanoic Acid (PFBA, CASRN 375-22-4) and Related Salts

Supplemental Information—Appendices A though F

December 2022

Integrated Risk Information System
Center for Public Health and Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Washington, DC

DISCLAIMER

This document has been reviewed by the U.S. Environmental Protection Agency, Office of Research and Development and approved for publication. Any mention of trade names, products, or services does not imply an endorsement by the U.S. government or the U.S. Environmental Protection Agency. EPA does not endorse any commercial products, services, or enterprises.

CONTENTS

TABLES

FIGURES

ABBREVIATIONS AND ACRONYMS

ACO	acyl-CoA oxidase
ADME	absorption, distribution, metabolism, and excretion
AFFF	aqueous film-forming foam
AIC	Akaike's information criterion
ALP	alkaline phosphatase
ALT	alanine aminotransferase
AST	aspartate aminotransferase
atm	atmosphere
ATSDR	Agency for Toxic Substances and Disease Registry
AUC	area-under-the-concentration curve
BMD	benchmark dose
BMDL	benchmark dose lower confidence limit
BMDS	Benchmark Dose Software
BMR	benchmark response
BW	body weight
C_{AVG}	average concentration
C_{MAX}	maximum concentration
CA	Cochran-Armitage
CAR	constitutive androstane receptor
CASRN	Chemical Abstracts Service registry number
CDR	Chemical Data Reporting
CI	confidence interval
CL	clearance
CL_A	clearance in animals
CL_H	clearance in humans
CPAD	Chemical and Pollutant Assessment Division
CPHEA	Center for Public Health and Environmental Assessment
CV	constant variance
CYP450	cytochrome P450 superfamily
DAF	dosimetric adjustment factor
DNA	deoxyribonucleic acid
DNT	developmental neurotoxicity
DOD	Department of Defense
EPA	Environmental Protection Agency
EOP	Executive Office of the President
ER	extra risk
FLR	full-litter resorption
FTOH	fluorotelomer alcohol
GD	gestation day
GFR	glomerular filtration rate
GGT	γ-glutamyl transferase
GRADE	Grading of Recommendations Assessment, Development, and Evaluation
GSH	glutathione
HAWC	Health Assessment Workspace Collaborative
HED	human equivalent dose
HERO	Health and Environmental Research Online
HISA	highly influential scientific information
HPT	hypothalamic-pituitary-thyroid
IRIS	Integrated Risk Information System
i.v.	intravenous
IQ	intelligence quotient
IQR	interquartile range
ISI	influential scientific information
IUR	inhalation unit risk
LLOQ	lower limit of quantitation
LN	log-normal
LOAEL	lowest-observed-adverse-effect level
MBq	megabecquerel
MOA	mode of action
NCEA	National Center for Environmental Assessment
NCV	nonconstant variance
NIOSH	National Institute for Occupational Safety and Health
NIS	sodium-iodide symporter
NOAEL	no-observed-adverse-effect level
NPL	National Priority List
NTP	National Toxicology Program
OAT	organic anion transporter
OECD	Organisation for Economic Co-operation and Development
OMB	Office of Management and Budget
ORD	Office of Research and Development
OSF	oral slope factor
PC	partition coefficient
PBPK	physiologically based pharmacokinetic
PBTK	physiologically based toxicokinetic
PECO	Populations, Exposures, Comparators, Outcomes
PFAA	perfluoroalkyl acid
PFAS	per- and polyfluoroalkyl substances
PFBA	perfluorobutanoic acid
PFBS	perfluorobutane sulfonate
PFCA	perfluoroalkyl carboxylic acid
PFDA	perfluorodecanoic acid
PFHxA	perfluorohexanoic acid
PFHxS	perfluorohexane sulfonate
PFNA	perfluorononanoic acid
PFOA	perfluorooctanoic acid
PFOS	perfluorooctane sulfonate
PK	pharmacokinetic
PND	postnatal day

POD	point of departure
POD_{HED}	human equivalent dose POD
PPAR	peroxisome proliferator-activated receptor
PQAPP	Programmatic Quality Assurance Project Plan
PT	prothrombin time
PXR	pregnane X receptor
QA	quality assurance
QAPP	Quality Assurance Project Plan
QMP	Quality Management Plan
RBC	red blood cell
RD	relative deviation
RfC	inhalation reference concentration
RfD	oral reference dose
RS	Rao-Scott
SD	standard deviation
S-D	Sprague-Dawley
SE	standard error
TD	toxicodynamic
TH	thyroid hormone
TK	toxicokinetic
TPO	thyroid peroxidase
TRI	Toxic Release Inventory
TSCA	Toxic Substances Control Act
TSCATS	Toxic Substances Control Act Test Submissions
TSH	thyroid-stimulating hormone
TSHR	thyroid-stimulating hormone receptor
UCMR	Unregulated Contaminant Monitoring Rule
UDP-GT	uridine 5′-diphospho-glucuronosyltransferase
UF	uncertainty factor
UF_A	animal-to-human uncertainty factor
UF_C	composite uncertainty factor
UF_D	database deficiencies uncertainty factor
UF_H	human variation uncertainty factor
UF_L	LOAEL-to-NOAEL uncertainty factor
UF_S	subchronic-to-chronic uncertainty factor
V_d	volume of distribution
VOC	volatile organic compound
WOS	Web of Science

APPENDIX A. SYSTEMATIC REVIEW PROTOCOL FOR THE PFAS IRIS ASSESSMENTS

A single systematic review protocol was used to guide the development of five, separate IRIS PFAS assessments (i.e., PFBA, PFHxA, PFHxS, PFNA, and PFDA). This "systematic review protocol for the PFAS IRIS assessments" was released for public comment and subsequently amended. The amended protocol and prior revisions can be found at the following location: http://cfpub.epa.gov/ncea/iris_drafts/recordisplay.cfm?deid=345065.

When the assessment references a particular section or page number in Appendix A, please refer to that section in the systematic review protocol linked above.

APPENDIX B. ADDITIONAL DETAILS OF SYSTEMATIC REVIEW METHODS AND RESULTS

Table B-1. Perfluorobutanoic acid (PFBA) database search strategy

Search	Search strategy	Dates of search[a]
PubMed		
Search terms	375-22-4[rn] OR "Heptafluoro-1-butanoic acid"[tw] OR "Heptafluorobutanoic acid"[tw] OR "Heptafluorobutyric acid"[tw] OR "Kyselina heptafluormaselna"[tw] OR "Perfluorobutanoic acid"[tw] OR "Perfluorobutyric acid"[tw] OR "Perfluoropropanecarboxylic acid"[tw] OR "2,2,3,3,4,4,4-heptafluoro-Butanoic acid"[tw] OR "Butanoic acid, 2,2,3,3,4,4,4-heptafluoro-"[tw] OR "Butanoic acid, heptafluoro-"[tw] OR "Perfluoro-n-butanoic acid"[tw] OR "Perfluorobutanoate"[tw] OR "2,2,3,3,4,4,4-Heptafluorobutanoic acid"[tw] OR "Butyric acid, heptafluoro-"[tw] OR "Fluorad FC 23"[tw] OR "H 0024"[tw] OR "NSC 820"[tw] OR ((PFBA[tw] OR "FC 23"[tw] OR HFBA[tw]) AND (fluorocarbon*[tw] OR fluorotelomer*[tw] OR polyfluoro*[tw] OR perfluoro-*[tw] OR perfluoroa*[tw] OR perfluorob*[tw] OR perfluoroc*[tw] OR perfluorod*[tw] OR perfluoroe*[tw] OR perfluoroh*[tw] OR perfluoron*[tw] OR perfluoroo*[tw] OR perfluorop*[tw] OR perfluoros*[tw] OR perfluorou*[tw] OR perfluorinated[tw] OR fluorinated[tw] OR PFAS[tw] OR PFOS[tw] OR PFOA[tw]))	No date limit–7/19/2017
Literature update search terms	(((375-22-4[rn] OR "Heptafluoro-1-butanoic acid"[tw] OR "Heptafluorobutanoic acid"[tw] OR "Heptafluorobutyric acid"[tw] OR "Kyselina heptafluormaselna"[tw] OR "Perfluorobutanoic acid"[tw] OR "Perfluorobutyric acid"[tw] OR "Perfluoropropanecarboxylic acid"[tw] OR "2,2,3,3,4,4,4-heptafluoro-Butanoic acid"[tw] OR "Butanoic acid, 2,2,3,3,4,4,4-heptafluoro-"[tw] OR "Butanoic acid, heptafluoro-"[tw] OR "Perfluoro-n-butanoic acid"[tw] OR "Perfluorobutanoate"[tw] OR "2,2,3,3,4,4,4-Heptafluorobutanoic acid"[tw] OR "Butyric acid, heptafluoro-"[tw] OR "Fluorad FC 23"[tw] OR "H 0024"[tw] OR "NSC 820"[tw] OR ((PFBA[tw] OR "FC 23"[tw] OR HFBA[tw]) AND (fluorocarbon*[tw] OR fluorotelomer*[tw] OR polyfluoro*[tw] OR perfluoro-*[tw] OR perfluoroa*[tw] OR perfluorob*[tw] OR perfluoroc*[tw] OR perfluorod*[tw] OR perfluoroe*[tw] OR perfluoroh*[tw] OR perfluoron*[tw] OR perfluoroo*[tw] OR perfluorop*[tw] OR perfluoros*[tw] OR perfluorou*[tw] OR perfluorinated[tw] OR fluorinated[tw] OR PFAS[tw] OR PFOS[tw] OR PFOA[tw])) AND ("2017/08/01"[PDAT] : "2018/02/14"[PDAT])	8/1/2017–2/14/2018

Search	Search strategy	Dates of search[a]
Web of Science		
Search terms	TS="Heptafluoro-1-butanoic acid" OR TS="Heptafluorobutanoic acid" OR TS="Heptafluorobutyric acid" OR TS="Kyselina heptafluormaselna" OR TS="Perfluorobutanoic acid" OR TS="Perfluorobutyric acid" OR TS="Perfluoropropanecarboxylic acid" OR TS="2,2,3,3,4,4,4-heptafluoro-Butanoic acid" OR TS="Butanoic acid, 2,2,3,3,4,4,4-heptafluoro-" OR TS="Butanoic acid, heptafluoro-" OR TS="Perfluoro-n-butanoic acid" OR TS="Perfluorobutanoate" OR TS="2,2,3,3,4,4,4-Heptafluorobutanoic acid" OR TS="Butyric acid, heptafluoro-" OR TS="Fluorad FC 23" OR TS="H 0024" OR TS="NSC 820" OR (TS=(PFBA OR "FC 23" OR HFBA) AND TS=(fluorocarbon* OR fluorotelomer* OR polyfluoro* OR perfluoro-* OR perfluoroa* OR perfluorob* OR perfluoroc* OR perfluorod* OR perfluoroe* OR perfluoroh* OR perfluoron* OR perfluoroo* OR perfluorop* OR perfluoros* OR perfluorou* OR perfluorinated OR fluorinated OR PFAS OR PFOS OR PFOA))	No date limit–7/20/2017
Literature update search terms	((TS="Heptafluoro-1-butanoic acid" OR TS="Heptafluorobutanoic acid" OR TS="Heptafluorobutyric acid" OR TS="Kyselina heptafluormaselna" OR TS="Perfluorobutanoic acid" OR TS="Perfluorobutyric acid" OR TS="Perfluoropropanecarboxylic acid" OR TS="2,2,3,3,4,4,4-heptafluoro-Butanoic acid" OR TS="Butanoic acid, 2,2,3,3,4,4,4-heptafluoro-" OR TS="Butanoic acid, heptafluoro-" OR TS="Perfluoro-n-butanoic acid" OR TS="Perfluorobutanoate" OR TS="2,2,3,3,4,4,4-Heptafluorobutanoic acid" OR TS="Butyric acid, heptafluoro-" OR TS="Fluorad FC 23" OR TS="H 0024" OR TS="NSC 820") OR TS=(PFBA OR "FC 23" OR HFBA) AND TS=(fluorocarbon* OR fluorotelomer* OR polyfluoro* OR perfluoro-* OR perfluoroa* OR perfluorob* OR perfluoroc* OR perfluorod* OR perfluoroe* OR perfluoroh* OR perfluoron* OR perfluoroo* OR perfluorop* OR perfluoros* OR perfluorou* OR perfluorinated OR fluorinated OR PFAS OR PFOS OR PFOA)) AND PY=2017-2018	2017–2018
Toxline		
Search terms	(375-22-4 [rn] OR "heptafluoro-1-butanoic acid" OR "heptafluorobutanoic acid" OR "heptafluorobutyric acid" OR "kyselina heptafluormaselna" OR "perfluorobutanoic acid" OR "perfluorobutyric acid" OR "perfluoropropanecarboxylic acid" OR "2,2,3,3,4,4,4-heptafluoro-butanoic acid" OR "butanoic acid 2 2 3 3 4 4 4-heptafluoro-" OR "butanoic acid heptafluoro-" OR "perfluoro-n-butanoic acid" OR "perfluorobutanoate" OR "2,2,3,3,4,4,4-heptafluorobutanoic acid" OR "butyric acid heptafluoro-" OR "fluorad fc 23" OR "h 0024" OR "nsc 820" OR ((pfba OR "fc 23" OR hfba) AND (fluorocarbon* OR fluorotelomer* OR polyfluoro* OR perfluoro* OR perfluorinated OR fluorinated OR pfas OR pfos OR pfoa))) AND (ANEUPL [org] OR BIOSIS [org] OR CIS [org] OR DART [org] OR EMIC [org] OR EPIDEM [org] OR HEEP [org] OR HMTC [org] OR IPA [org] OR RISKLINE [org] OR MTGABS [org] OR NIOSH [org] OR NTIS [org] OR PESTAB [org] OR PPBIB [org]) AND NOT PubMed [org] AND NOT pubdart [org]	No date limit–7/20/2017

Search	Search strategy	Dates of search[a]
Literature update search terms	@AND+@OR+("heptafluoro-1-butanoic acid"+"heptafluorobutanoic+acid"+"heptafluorobutyric+acid"+"kyselina+heptafluormaselna"+"perfluorobutanoic+acid"+"perfluorobutyric+acid"+"perfluoropropanecarboxylic +acid"+"2 2 3 3 4 4 4-heptafluoro-butanoic+acid"+"butanoic+acid+2 2 3 3 4 4 4-heptafluoro-"+"butanoic+acid+heptafluoro-"+"perfluoro-n-butanoic acid"+"perfluorobutanoate"+"2 2 3 3 4 4 4-heptafluorobutanoic+acid"+"butyric+acid+heptafluoro-"+"fluorad+fc+23"+"h0024"+"nsc+820"+@TERM+@rn+375-22-4("pfba"+"fc+23"+"hfba"))+(fluorocarbon*+ fluorotelomer*+polyfluoro*+perfluoro*+perfluorinated+fluorinated+pfas+pfos+pfoa)+@RANGE+yr+2017+2018	2017–2018
TSCATS		
Search terms	375-22-4[rn] AND tscats[org]	No date limit–7/20/2017

[a] Yearly spring updates are conducted following release of the draft for public comment; see also the docket ("EPA-HQ-ORD-2020-0675-0022") for studies identified after the last formal update preceding public release.

Table B-2. Title/abstract-level screening criteria for the initial literature searches

	Inclusion criteria	Exclusion criteria
Populations	Humans Standard mammalian animal models, including rat, mouse, rabbit, guinea pig, hamster, monkey, dog Alternative animal models in standard laboratory conditions (e.g., *Xenopus*, zebrafish, minipig) Human or animal cells, tissues, or organs (not whole animals); bacteria, nonmammalian eukaryotes; other nonmammalian laboratory species	Ecological species
Exposures	Exposure is to PFBA Exposure via oral, inhalation, dermal, intraperitoneal, or intravenous injection routes Exposure is measured in air, dust, drinking water, diet, gavage, injection or via a biomarker of exposure (PFBA levels in whole blood, serum, plasma, or breastmilk)	Study population is not exposed to PFBA Exposure is to a mixture only
Outcomes	Studies that include a measure of one or more health effect endpoints, including but not limited to, effects on reproduction, development, developmental neurotoxicity, liver, thyroid, immune system, nervous system, genotoxicity, and cancer In vivo or in vitro studies related to toxicity mechanisms, physiological effects/adverse outcomes, and studies useful for elucidating toxic modes of action (MOAs) Qualitative or quantitative description of absorption, distribution, metabolism, excretion, toxicokinetic or toxicodynamic models (e.g., PBPK, PBTK, PBTK/TD) Studies addressing risks to infants, children, pregnant women, occupational workers, the elderly, and any other susceptible or differentially exposed populations	

	Inclusion criteria	Exclusion criteria
Other	Structure and physiochemical properties Reviews and regulatory documents	Not on topic, including: Abstract only, inadequately reported abstract, or no abstract and not considered further because study was not potentially relevant Bioremediation, biodegradation, or chemical or physical treatment of PFBA, including evaluation of wastewater treatment technologies and methods for remediation of contaminated water and soil Ecosystem effects Studies of environmental fate and transport of PFBA in environmental media Analytical methods for detecting/measuring PFAS compounds in environmental media and use in sample preparations and assays Studies describing the manufacture and use of PFBA Not chemical specific (studies that do not involve testing of PFBA) Studies that describe measures of exposure to PFBA without data on associated health effects

MOA = mode of action; PBPK = physiologically based pharmacokinetic; PBTK = physiologically based toxicokinetic; TD = toxicodynamic.

Table B-3. Example DistillerSR form questions to be used for title/abstract-level and full text-level screening for literature search updates from 2019

	Used in title/abstract and full-text screening					Used in full text screening only	
Question	**Source of study if not identified from database search?**	**Does the article meet PECO criteria?**	**If meets PECO, what type of evidence?**	**If supplemental, what type of information?**	**Which PFAS did the study report?**	**If meets PECO, which health outcome(s) apply?**	**If meets PECO and endocrine outcome, which endocrine tags apply?**
Answer options (can select multiple options)	Source other than HERO database search	Yes No Unclear Tag as potentially relevant supplemental information	Human Animal (mammalian models) In vitro or in silico genotoxicity PBPK or PK model	In vivo mechanistic or MOA studies, including non-PECO routes of exposure (e.g., injection) and populations (e.g., nonmammalian) In vitro or in silico studies (nongenotoxicity) ADME/toxicokinetic (excluding models) Exposure assessment or characterization (no health outcome) PFAS mixture study (no individual PFAS comparisons) Human case reports or case series Ecotoxicity studies	PFBA PFHxA PFHxS PFNA PFDA	General toxicity, including body weight, mortality, and survival Cancer Cardiovascular, including serum lipids Endocrine (hormone) Gastrointestinal Genotoxicity Growth (early life) and development Hematological, including nonimmune/hepatic/ renal clinical chemistry measures Hepatic, including liver measures and serum markers (e.g., ALT; AST) Immune/ inflammation	Adrenal Sex hormones (e.g., androgen; estrogen; progesterone) Neuroendocrine Pituitary Steroidogenesis Thyroid

	Used in title/abstract and full-text screening					Used in full text screening only	
Question	**Source of study if not identified from database search?**	**Does the article meet PECO criteria?**	**If meets PECO, what type of evidence?**	**If supplemental, what type of information?**	**Which PFAS did the study report?**	**If meets PECO, which health outcome(s) apply?**	**If meets PECO and endocrine outcome, which endocrine tags apply?**
				Environmental fate or occurrence (including food) Manufacture, engineering, use, treatment, remediation, or laboratory methods Other assessments or records with no original data (e.g., reviews, editorials, commentaries)		Musculoskeletal Nervous system, including behavior and sensory function Nutrition and metabolic Ocular PBPK or PK model Renal, including urinary measures (e.g., protein) Reproductive Respiratory Skin and connective tissue effects	

ADME = absorption, distribution, metabolism, and excretion; ALT = alanine aminotransferase; AST = aspartate aminotransferase; HERO = Health and Environmental Research Online; MOA = mode of action; PBPK = physiologically based pharmacokinetic; PECO = Populations, Exposures, Comparators, and Outcomes; PFAS = per- and polyfluoroalkyl substance; PFBA = perfluorobutanoic acid; PFDA = perfluorodecanoic acid; PFHxA = perfluorohexanoic acid; PFHxS = perfluorohexanesulfonate; PFNA = perfluorononanoic acid; PK = pharmacokinetic.

APPENDIX C. ADDITIONAL TOXICOKINETIC INFORMATION IN SUPPORT OF DOSE-RESPONSE ANALYSIS

C.1. USE OF HALF-LIVES OF EXCRETION FOR DOSIMETRIC ADJUSTMENTS

The pharmacokinetics of PFBA have only been measured after direct administration of PFBA in single-exposure/single-day studies in animals (Chang et al., 2008). For the mouse, Chang et al. (2008) performed 24-hour toxicokinetic studies after 10, 30, and 100 mg/kg oral doses. Based on the area-under-the-concentration-curve (AUC) and maximum concentration (C_{max}), the data also appear approximately linear below 30 mg/kg but show some saturation above that dose rate (see Figure C-1, Figure C-2).

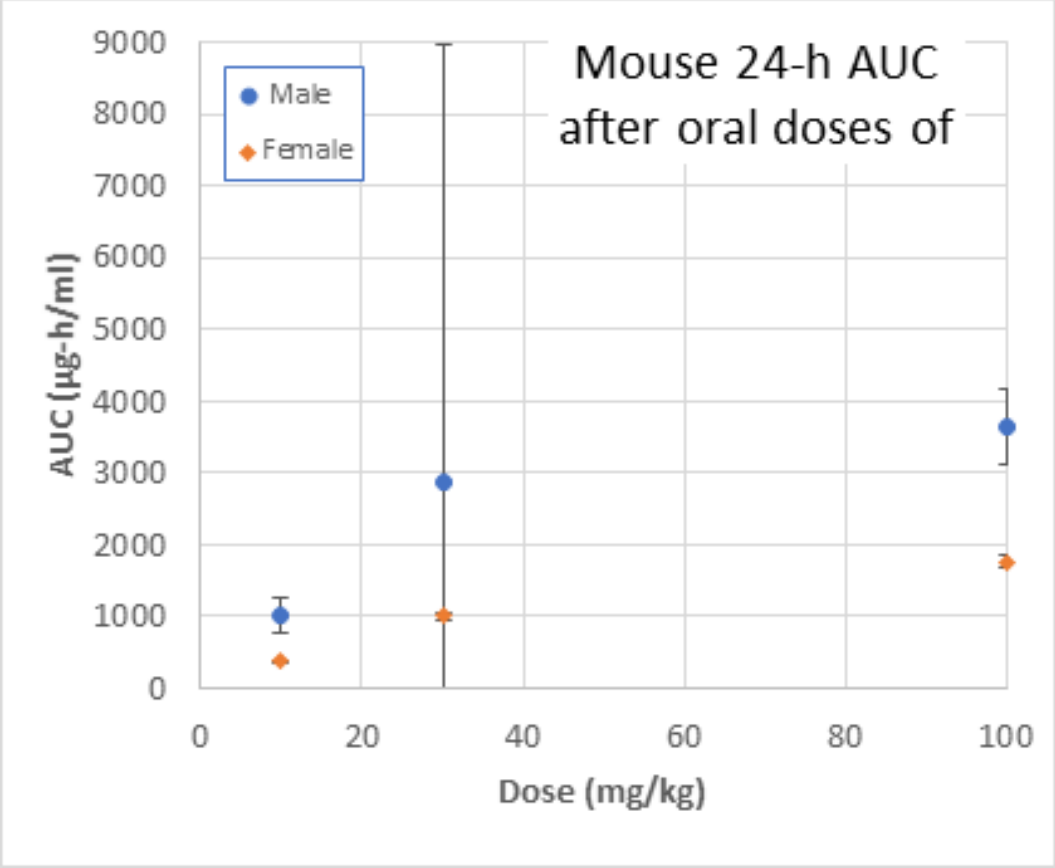

Figure C-1. Mouse AUC after oral doses of PFBA.

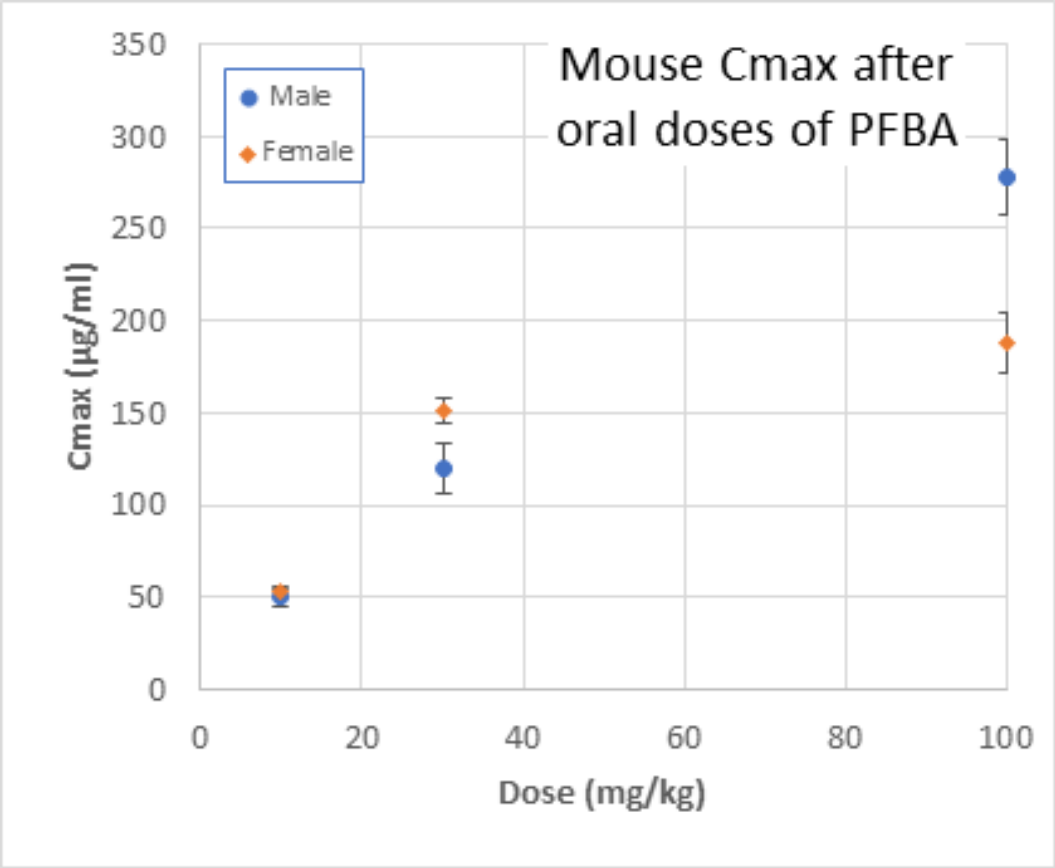

Figure C-2. Mouse C_{max} after oral doses of PFBA.

Chang et al. (2008) reported serum and liver concentrations in male rats and serum concentrations in female rats given a 3–300 mg/kg oral dose of PFBA at 24 hours after dosing. Although the time point for these measurements is not ideal given the short half-life of PFBA, the data indicate that the dosimetry is approximately linear up to 100 mg/kg in male rats and up to 30 mg/kg in female rats (see Figure C-3, Figure C-4). Tissue levels then appear to saturate or decline; this might be due to incomplete absorption at higher doses, saturable renal resorption, or both, whereby excretion is more rapid for concentrations above the level of saturable resorption in the kidney. With the half-life in female rats being ~3 hours, the female serum 24-hour data are particularly subject to experimental noise, but at least provide an indication that use of the half-life measured using a 30 mg/kg dose is applicable to BMD levels from bioassays at or below this dose rate.

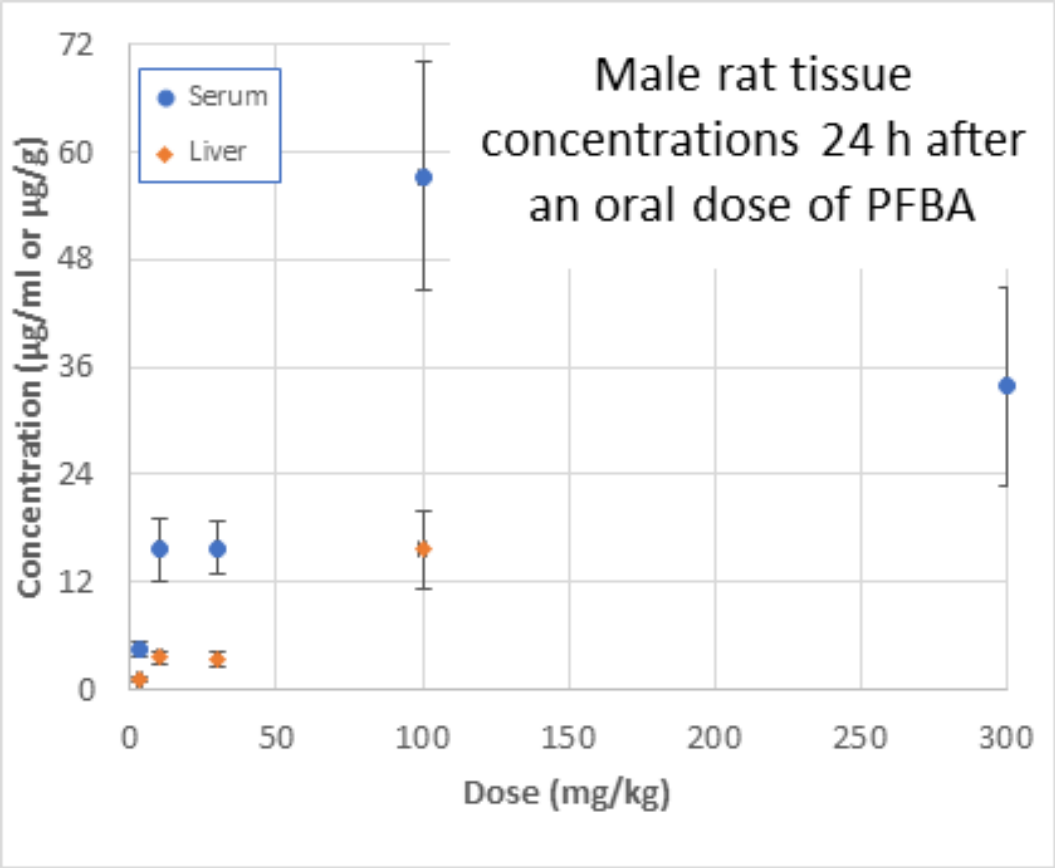

Figure C-3. Rat AUC after oral doses of PFBA.

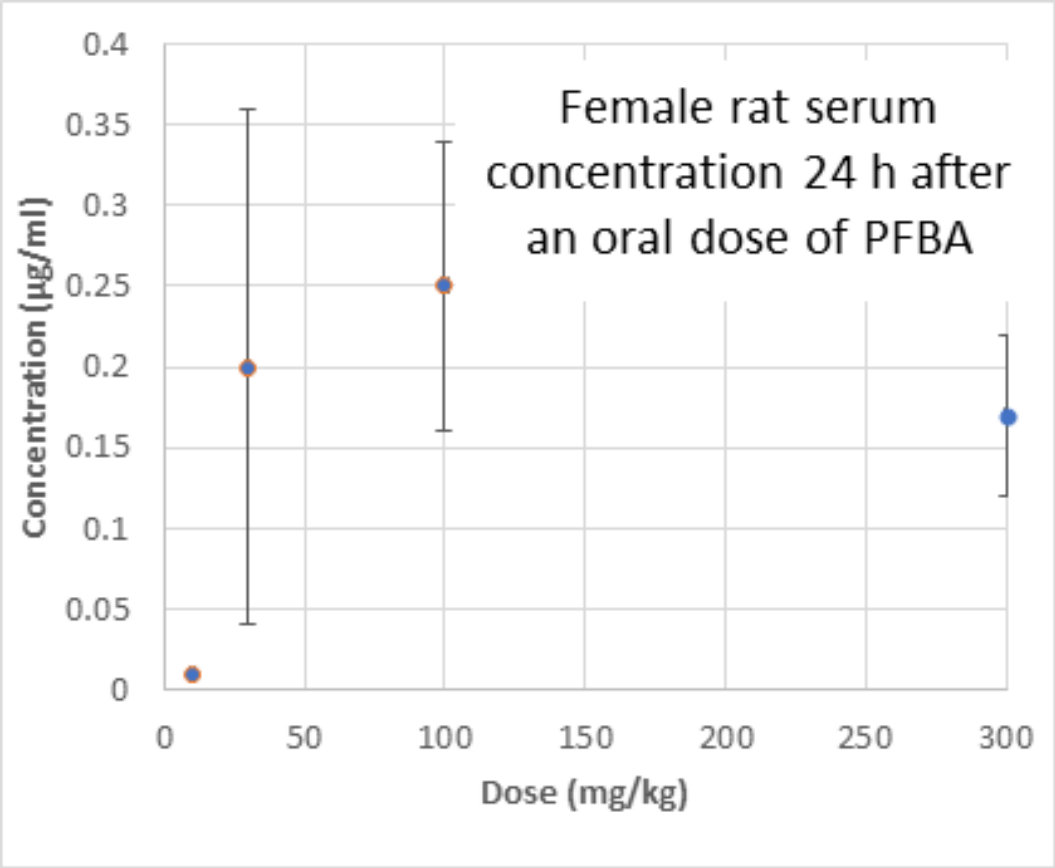

Figure C-4. Rat C_{max} after oral doses of PFBA.

For the human data analyzed by Chang et al. (2008), detailed toxicokinetic parameters are not available, but one can evaluate the relationship between the initial concentration and $t_{1/2}$. Here only data for subjects in which the final concentration is greater than the limit of quantification is considered to avoid statistical artifacts due to limited observational data. Although the lower half-life of the subject with the highest initial concentration indicates a possible negative trend, the half-life is in the range of subjects with lower initial concentrations. Thus, these data do not show a clear dose dependence for half-life and are interpreted as only showing interindividual variation (see Figure C-5). The human data appear consistent with first-order clearance across the range of concentrations observed.

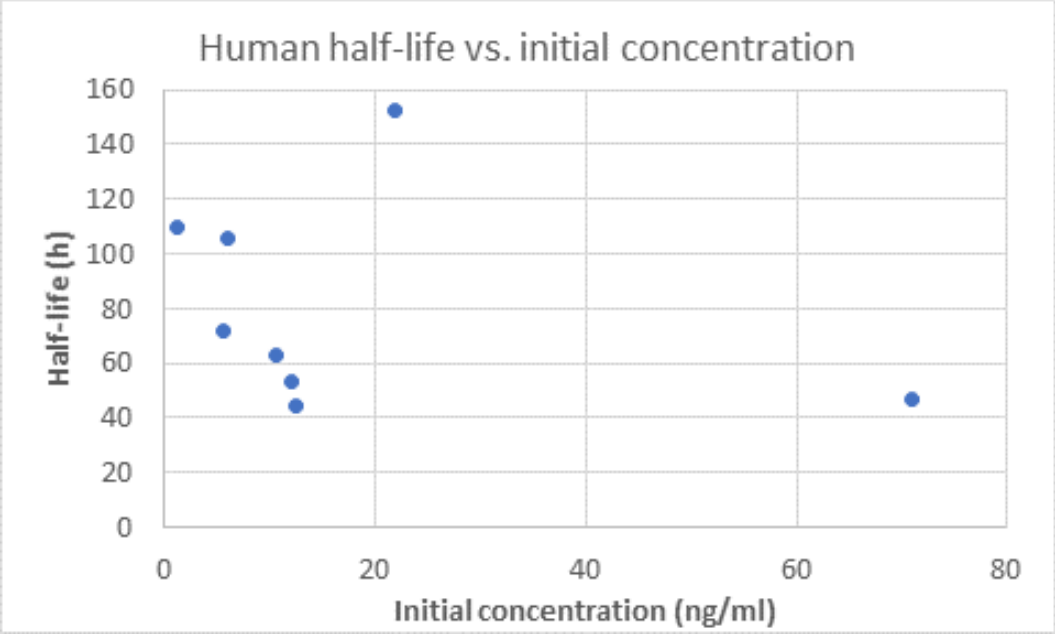

Figure C-5. Estimated human half-lives versus initial serum concentrations.

Chang et al. (2008) only evaluated one PFBA dose in monkeys, so determining whether the biphasic clearance pattern is due to the classical distinction between distribution and excretion phases or a nonlinearity in clearance is not possible. The data show linear clearance from 1–7 or 10 days after the i.v. dose was given, however, when serum concentrations were below 100 ng/mL. Thus, interpreting these data as showing linear kinetics for serum concentrations below 100 ng/mL under long-term exposure conditions seem reasonable. Because the highest initial condition of the human subjects in Chang et al. (2008) was 72 ng/mL, to the extent that kinetics in monkeys can be extrapolated to humans, the results for monkeys confirm the conclusion that human kinetics are also reasonably assumed linear below ~100 ng/mL. This is approximately 1,000-fold below the range of linearity in mice and rats, however, so uncertainty exists as to whether the range of linear kinetics in humans and monkeys extends into the range of rodent-based points of departure.

Russell et al. (2015) attempted to evaluate the kinetics of PFBA as a metabolite of 6:2 fluorotelomer alcohol (FTOH) during a 1-day inhalation study (6-hour exposure, 24-hour observation) and at the end of 23 days of exposure. The half-life of PFBA, however, could not be estimated from the single-day data for male rats and could be estimated only for the high-level exposure in female rats, with yields of PFBA 0.2% in males and not detectable or 0.02% in females. Also, three metabolic intermediates occur between 6:2 FTOH and PFBA, but the model appears to have assumed direct, instantaneous transformation through the first two steps. Assumptions about the volume of distribution were made by (Russell et al., 2015). These simplifications in the model likely explain the large discrepancy between the PFBA half-life determined from the single-day exposure 6:2 FTOH for female rats (19 hours) and the half-life obtained for direct exposure to PFBA (1.4-hour average) by (Chang et al., 2008). Russell et al. (2015) used only male rats in the 23-day

6:2 FTOH inhalation study, from which they estimated a half-life of 27.7 hours, over three times higher than the average obtained by (Chang et al., 2008). The discrepancy also could be due to an underestimation of the metabolic yield from the 1-day experiments. In summary, whereas Russell et al. (2015) described measurements of PFBA in male rats from 23 days of exposure to 6:2 FTOH, the results for female rats after a single exposure are completely inconsistent with the results of (Chang et al., 2008). Therefore, the conclusions from the multiday study are considered too unreliable to be used.

The other long-term data available on internal dosimetry are from the bioassays (Butenhoff et al., 2012; Das et al., 2008; van Otterdijk, 2007b). Serum concentrations in nonpregnant female mice after 17 days of exposure (24 hours after the last dose) are 2.0 ± 1.0 and 2.4 ± 1.7 μg/mL, and for pregnant mice are 3.8 ± 1.0 and 4.4 ± 0.7 μg/mL, for the 35- and 175-mg/kg dose groups, respectively (Das et al., 2008). For female mice dosed with 30- and 100-mg/kg PFBA, Chang et al. (2008) reported 4.1 ± 1.7 and 6.4 ± 3.9 μg/mL in serum 24 hours after the dose; using linear extrapolation based on the difference in dose, one might expect 4.8 and 11.2 μg/mL at 24 hours after doses of 35 and 175 mg/kg, given these data. Although the concentrations in the Das et al. (2008) study are somewhat lower than these projections, the difference, especially at the low dose, is within the range of uncertainty and precision expected for PK analysis.

Of note is that, given an average clearance of 28 mL/kg-hour obtained by Chang et al. (2008) after 10- and 30-mg/kg doses, the predicted average serum concentrations for a 35-mg/kg dose is 52 μg/mL. This average concentration reflects the much higher concentrations expected in the first few hours after each dose.

For male rats, Butenhoff et al. (2012) measured end-of-treatment serum levels of 38 ± 23 and 52 ± 25 μg/mL after 28 and 90 days, respectively, at 30 mg/kg-day; EPA presumes these measurements were made 24 hours after the last dose. The corresponding values reported by Chang et al. (2008) for a 30-mg/kg oral dose in the dose-range and time-course studies are 16 ± 3 and 29 ± 13 μg/mL, respectively. Although again, some discrepancy is found between the short-term PK data and the bioassay measurements, the difference is that it is roughly within a factor of 2, which is acceptable for PK analysis and does not indicate a strong time dependence in the PK. One should keep in mind that the estimated clearance and half-life values are based on multiple time points at which the serum concentration is measured, while the comparisons above use only a single time point, 24 hours after dosing, when the result will be sensitive to experimental variation.

Given these data and results, the half-life or clearance of PFBA measured in single-day exposures by Chang et al. (2008) will be assumed to predict dosimetry after repeated exposures that occur in bioassays. This is a common assumption for chemicals with relatively short half-lives because pharmacokinetic studies are typically confined to a single day or less. Clearance in rats and mice might include a slower beta phase, like that observed in monkeys. If a slow clearance phase exists, internal dose from long-term exposure will be higher than is effectively estimated using the

clearance rate determined from single-day exposures, which would increase the HED compared with the current prediction. Using an animal-human ratio of clearance values to estimate the HED relies only on the assumptions that the average serum concentration (C_{AVG}) is predictive of systemic effects in adults and that the relationship between C_{AVG} and dose rate is linear with the proportionality determined by the clearance values estimated here (i.e., the clearance from single-day experiments is predictive of bioassay conditions).

The human half-life estimates were from subjects who had been occupationally exposed to PFBA, with the duration of the PK observation 7–10 days. Thus, those results are reasonably expected to represent clearance under (subsequent to) chronic exposure conditions. The primary uncertainty in predicting human clearance comes from assuming a volume of distribution equal to that estimated for monkeys, which is thought modest given the physiological similarity between monkeys and humans. Thus, the overall uncertainty from using the animal-human clearance ratio to predict the HED for systemic effects in adults appears modest, especially compared to the case where PK data such as used here are not available.

Because developmental effects are usually presumed to depend on peak concentration rather than average concentration, it must be noted that use of the clearance ratio to estimate HEDs for those endpoints also involves an assumption that the absorption rate in humans is similar to that of animals. For PFBA, the absorption rate in mice and rats is fairly rapid, with the peak concentration occurring 0.6–4 hours after bolus oral doses (Chang et al., 2008). That absorption in humans would be faster than in rodents seems unlikely, and exposures are more likely spread out over the day than in the animal bioassays. Therefore, the most likely case is that the peak concentration in humans exposed at the HED will be lower than the peak concentration in mice or rats at the corresponding dose rate. Thus, although this assumption creates uncertainty in the dose extrapolation, the result is not expected to underpredict human health risks.

C.2. MIXED MODELING TO ESTIMATE HALF-LIFE IN HUMANS

A linear mixed-effects model was additionally used to estimate a $t_{1/2}$ for PFBA according to methods described in (Li et al., 2018). Briefly, linear mixed-effect models are extensions of simple linear models that use the best linear unbiased prediction estimator to estimate random and fixed effects for clustered data. One important consequence of clustering is that measurements of serum PFBA units within the same person (cluster) are more similar than measurements on serum PFBA in different people (i.e., other clusters). Failure to account for the intracluster correlation would result in misleading inferences. Each individual in Chang et al. (2008) was assumed to have been selected randomly from a larger population. Below is the mixed model formula used for estimating the half-life of serum PFBA:

$$\ln(\text{PFBA}_{ij}) = (\alpha_{\text{pop}} + \alpha_i) + (k_{\text{pop}} + k_i) \times t_{ij} + \varepsilon_{ij} \quad \text{(C-1)}$$

where ln(PFBA$_{ij}$) is the natural logarithm of the serum PFBA concentrations measured at the *jth* time point for the *ith* subject, α_{pop} is the population mean (also known as the fixed intercept for the population); $\alpha_i \sim N(0, \sigma^2_\alpha)$ is a random intercept for the *ith* subject; k_{pop} is the fixed slope for the population (also known as the average excretion rate constant for serum PFBA for the whole population); $k_i \sim N(0, \sigma^2_k)$ is the random slope for the *ith* subject that allows the excretion rate to vary by individuals; t_{ij} represents the observation time for the *jth* measurement of serum PFBA for *ith* subject; and $\varepsilon_{ij} \sim N(0, \sigma^2_\varepsilon)$ is the random-error effect (residual) for *jth* measurement of *ith* subject. Of note, the small sample sizes (due to the exclusion of the only two subjects identified as females) limited our ability to draw clear conclusions in gender-stratified comparisons.

The subjects from Chang et al. (2008) used in this analysis and the half-lives estimated by Chang et al. (2008) are listed in the following table. As explained in section 5.2.1, Approach for Animal-Human Extrapolation of Perfluorobutanoic Acid (PFBA) Dosimetry, subjects whose second concentration measurement was below the lower limit of quantitation (LLOQ) were excluded from analysis because the half-life for these subjects is highly uncertain. This choice is expected to bias the analysis towards higher half-lives but given the small number of human subjects for which data are available, and that variability in clearance among the human population is expected, this is considered a reasonably health-protective choice.

Subject	Sex	Reported half-life (h)
Cottage Grove Subject 1	NS	105.3
Cottage Grove Subject 1	NS	109.7
Cordova Subject 2	M	53
Cordova Subject 3	M	72
Cordova Subject 4	M	44
Cordova Subject 6	M	152
Cordova Subject 8	M	63
Cordova Subject 9	M	47

The half-life of serum PFBA for the study population ($t_{1/2,\,pop}$) then was estimated as:

$$t_{1/2,pop} = \left|\frac{\ln(2)}{k_{pop}}\right| \qquad \text{(C-2)}$$

The mixed-effects model estimated k_{pop} to be −0.010, therefore resulting in an estimated $t_{1/2}$ of 67.9 hours. Of the estimated half-lives reported by Chang et al. (2008), including those excluded from analysis by EPA, five values (42% of the population) were greater than 67.9 hours and seven (58% of the population) were below. This value also matches very closely to the median value calculated when not taking clustering into account, and therefore was considered a reasonable estimate of the population mean and was used in estimation of clearance in humans.

APPENDIX D. BENCHMARK DOSE MODELING RESULTS

D.1. BENCHMARK DOSE MODELING APPROACHES

As discussed in Section 5 of the body of the Toxicological Review, the endpoints selected for benchmark dose (BMD) modeling were relative liver weight, liver hypertrophy, total T4, and thyroid follicular hypertrophy incidence from Butenhoff et al. (2012) and relative liver weight, full litter resorption, delayed eye opening, delayed vaginal opening, and delayed preputial separation from (Das et al., 2008). The animal doses in the study were used in the BMD modeling and then converted to human equivalent doses (HEDs) using the ratio of animal-to-human clearance values; the modeling results are presented in this appendix.

D.1.1. Modeling Procedure for Dichotomous Noncancer Data

BMD modeling of dichotomous noncancer data was conducted using EPA's Benchmark Dose Software (BMDS, version 3.1.2). For these data, the Gamma, Logistic, Log-Logistic, Log-Probit, Multistage, Probit, Weibull, and Dichotomous Hill models available within the software were fit using a benchmark response (BMR) of 10% extra risk (see Toxicological Review, Section 4.2.1 for justification of selected BMRs). The Multistage model is run for all polynomial degrees up to $n - 2$, where n is the number of dose groups including control. Adequacy of model fit was judged on the basis of χ^2 goodness-of-fit p-value ($p > 0.1$), scaled residuals at the data point (except the control) closest to the predefined benchmark response (absolute value <2.0), and visual inspection of the model fit. In the cases where no best model was found to fit to the data, a reduced data set without the high-dose group was further attempted for modeling and the result presented with that of the full data set. In cases where a model with several parameters equal to the number of dose groups was fit to the data set, all parameters were estimated, and no p-value was calculated, that model was not considered for estimating a point of departure (POD) *unless* no other model provided adequate fit. Among all models providing adequate fit, the benchmark dose lower confidence limit (BMDL) from the model with the lowest Akaike's information criterion (AIC) was selected as a potential POD when BMDL values were sufficiently close (within threefold). Otherwise, the lowest BMDL was selected as a potential POD.

D.1.2. Modeling Procedure for Continuous Noncancer Data

BMD modeling of continuous noncancer data was conducted using EPA's Benchmark Dose Software (BMDS, version 3.1.2). For these data, the Exponential, Hill, Polynomial, and Power models available within the software are fit using a BMR of 1 standard deviation (SD) when no toxicological

information was available to determine an adverse level of response. When toxicological information was available, the BMR was based on relative deviation, as outlined in the Benchmark Dose Technical Guidance (U.S. EPA, 2012) (see Toxicological Review, Section 5.2.1 for justification for using BMRs); when a BMR based on relative deviation was used, modeling results using BMRs based on SD are included for reference. An adequate fit is judged on the basis of χ^2 goodness-of-fit *p*-value ($p > 0.1$), scaled residuals at the data point (except the control) closest to the predefined benchmark response (absolute value <2.0), and visual inspection of the model fit. In addition to these three criteria for judging adequacy of model fit, a determination is made on whether the variance across dose groups is homogeneous. If a homogeneous variance model is deemed appropriate based on the statistical test provided by BMDS (i.e., Test 2), the final BMD results are estimated from a homogeneous variance model. If the test for homogeneity of variance is rejected ($p < 0.05$), the model is run again while modeling the variance as a power function of the mean to account for this nonhomogeneous variance. If this nonhomogeneous variance model does not adequately fit the data (i.e., Test 3; $p < 0.05$), alternative approaches were assessed on a case-by-case basis. For example, in cases where neither variance model fit, or constant variance did not fit (with adequate Test-4 *p*-value) and nonconstant variance did fit (with inadequate Test-4 *p*-value), the log-normal distribution was attempted.

In cases where a model with several parameters equal to the number of dose groups was fit to the data set, all parameters were estimated, and no *p*-value was calculated, that model was not considered for estimating a POD *unless* no other model provided adequate fit. Among all models providing adequate fit, the BMDL from the model with the lowest AIC was selected as a potential POD when BMDL estimates differed by less than threefold. When BMDL estimates differed by greater than threefold, the model with the lowest BMDL was selected to account for model uncertainty.

In situations where there are multiple, related continuous endpoints in an organ system, or endpoints that inform a generalized effect to the exposed organism (e.g., developmental delays in multiple organ systems), modeling a combined endpoint of "total affected" animals was not pursued. Such a systematic multi-endpoint modeling approach is not currently available in BMDS other than the MS-Combo model that requires an assumption that the endpoints are independent; such an assumption is likely not valid with respect to continuous non-cancer endpoints in the PFBA toxicity database. Further, with respect to combining multiple continuous and dichotomous endpoints outside of a specific multi-endpoint model, first a cut-off value would need to be established for continuous endpoints to determine when an animal has "responded" (i.e., the continuous data would need to be "dichotomized"). Such dichotomization of continuous data results in a loss of precision and is recommended against in the BMD Technical Guidance (U.S. EPA, 2012).

D.1.3. Modeling Procedure for Continuous Noncancer Developmental Toxicity Data

For continuous developmental toxicity data, data for individual animals were requested from the study authors when possible. The use of individual animal data allows for the correct measure of variance to be calculated. When a biological rationale for selecting a benchmark response level is lacking, a BMR equal to 0.5 SD was used. The use of 1 SD for the BMR for continuous endpoints is based on the observation that shifting the distribution of the control group by 1 SD results in ~10% of the animal data points falling beyond an adversity cutoff defined at the ~1.5 percentile (Crump, 1995). This approximates the 10% extra risk commonly used as the BMR for dichotomous endpoints. Thus, the use of 0.5 SD for continuous developmental toxicity endpoints approximates the extra risk commonly used for dichotomous developmental toxicity endpoints.

D.1.4. Modeling Procedure for Dichotomous Noncancer Developmental Toxicity Data

For dichotomous developmental toxicity data, data for individual animals were requested from the study authors when possible. This allowed the use of the nested logistic model, which statistically accounts for intralitter similarity (the propensity of littermates to respond more like one another than pups from another litter) by estimating intralitter correlation and using litter-specific covariates. Other models (Rai and van Ryzin, NCTR) that also account for intralitter similarity were not considered in modeling dichotomous developmental toxicity data as they are not currently implemented in BMDS 3.2. Judging model fit for this model is identical to the procedure used for regular dichotomous models.

D.1.5. Data Used for Modeling

The source of the data used for modeling is provided in Table D-1. For endpoints from the Das et al. (2008) study, the study authors kindly provided individual dam-level data to facilitate modeling and to provide corrected data where needed. These data also are included in full in the tables below.

Table D-1. Sources of data used in benchmark dose modeling of PFBA endpoints

Endpoint/Reference	Reference	Location	HAWC link
Relative liver weight	Butenhoff et al. (2012)	Appendix 1, page 37 (van Otterdijk, 2007b)	https://hawcprd.epa.gov/ani/endpoint/100507453/
Relative liver weight	Das et al. (2008)	Figure 2, page 175	https://hawcprd.epa.gov/ani/endpoint/100507508/
Liver hypertrophy	Butenhoff et al. (2012)	Table 9, page 523	https://hawcprd.epa.gov/ani/endpoint/100507383/
Total T4	Butenhoff et al. (2012)	Table 8, page 522	https://hawcprd.epa.gov/ani/endpoint/100507375/
Full litter resorption	Das et al. (2008)	Table D-2	

Endpoint/Reference	Reference	Location	HAWC link
Fetal mortality (full litter resorptions combined with fetal death from litters without full litter resorptions)	Das et al. (2008)	Table D-3	
Eyes opening	Das et al. (2008)	Table D-4	
Vaginal opening	Das et al. (2008)	Table D-5	
Preputial separation	Das et al. (2008)	Table D-6	

Table D-2. Data received from study authors for Das et al. (2008) on full litter resorptions (FLR)

Dose (mg/kg-d)	Number of implants FLR
0	8
0	18
35	2
175	2
175	2
175	9
175	5
350	3
350	2
350	13
350	13
350	3
350	14
350	13

Table D-3. Data received from study authors for Das et al. (2008) on fetal death (litters without full litter resorptions) combined with full litter resorptions

Dose (mg/kg-d)	Number of implants	Number of dead	Dam weight on GD1 (litter-specific covariate)
0	16	1	30
0	16	1	28.2
0	11	2	27.7

Dose (mg/kg-d)	Number of implants	Number of dead	Dam weight on GD1 (litter-specific covariate)
0	11	0	27.4
0	12	3	25.9
0	11	0	24.1
0	15	0	29.2
0	14	1	28
0	12	3	27.1
0	14	0	26.8
0	16	1	26.6
0	13	2	25.1
0	17	3	30.1
0	14	0	29
0	6	0	27.5
0	9	2	28.1
0	6	0	26.9
0	13	1	26.7
0	11	0	23.3
0	8	8	25.8
0	18	18	31.4
35	15	3	28.1
35	13	0	29.3
35	13	0	27.4
35	14	1	27
35	15	2	26.9
35	13	2	25.7
35	12	4	31.6
35	13	0	29.2
35	14	1	27.7
35	16	0	27.5
35	13	2	28.1
35	7	3	25.5
35	15	1	30.3
35	13	0	27.5
35	14	1	28.1

Dose (mg/kg-d)	Number of implants	Number of dead	Dam weight on GD1 (litter-specific covariate)
35	13	1	27.9
35	11	0	26.4
35	10	1	27.4
35	13	1	27.9
35	13	0	26.1
35	13	1	24.8
35	12	1	24.8
35	2	2	23.1
175	14	1	28.1
175	15	0	27.5
175	14	0	27.4
175	14	1	27.5
175	15	2	29.4
175	14	1	27.5
175	15	0	26
175	16	2	26.2
175	11	0	23.4
175	16	3	29.1
175	11	0	28.2
175	13	0	25.8
175	11	2	26.8
175	15	1	26.9
175	14	1	25
175	13	1	26.7
175	2	2	25.5
175	2	2	25.4
175	9	9	29
175	5	5	25
350	7	2	29.2
350	12	1	26.3
350	16	3	27.4
350	11	0	25.1
350	14	2	25.3

Dose (mg/kg-d)	Number of implants	Number of dead	Dam weight on GD1 (litter-specific covariate)
350	12	1	29.5
350	16	2	28.8
350	17	2	26.2
350	12	2	26.2
350	16	0	27.3
350	9	3	27.6
350	13	0	27.7
350	13	0	27.4
350	13	1	26.4
350	7	1	24.6
350	3	3	21.5
350	2	2	23
350	13	13	25.8
350	13	13	24.6
350	3	3	25.1
350	14	14	28.2
350	13	13	29.2
350	1	1	25.4

Table D-4. Data received from study authors for Das et al. (2008)on delayed eye opening

Dose (mg/kg-d)	Average day of eye opening
0	16.27
0	15.57
0	15.22
0	15.27
0	14.55
0	14.91
0	17.64
0	15.69
0	15.00
0	17.57
0	17.71

Dose (mg/kg-d)	Average day of eye opening
0	14.91
0	16.50
0	17.58
0	16.50
0	16.25
0	15.20
0	17.25
0	18.00
0	18.00
35	16.00
35	17.31
35	18.00
35	17.23
35	17.23
35	16.82
35	18.78
35	17.31
35	17.57
35	17.53
35	18.00
35	15.25
35	17.00
35	17.82
35	18.09
35	17.70
35	16.11
35	18.29
35	17.50
35	17.55
35	17.60
35	17.78
175	17.69
175	17.67

Dose (mg/kg-d)	Average day of eye opening
175	15.71
175	17.77
175	16.91
175	18.00
175	17.69
175	17.27
175	17.17
175	17.64
175	18.00
175	18.00
175	18.09
175	18.88
175	18.00
175	18.00
175	18.20
350	15.00
350	18.64
350	17.85
350	17.64
350	18.00
350	17.36
350	17.85
350	17.93
350	18.00
350	18.00
350	18.00
350	18.60
350	18.00
350	18.09
350	18.00

Table D-5. Data received from study authors for Das et al. (2008) on delayed vaginal opening

Dose (mg/kg-d)	Average day of vaginal opening
0	32.40
0	27.00
0	30.80
0	30.20
0	34.17
0	33.67
0	30.33
0	28.00
0	30.14
0	33.67
0	28.00
0	31.90
0	32.50
0	34.00
0	29.25
0	28.00
0	29.33
0	35.57
0	34.83
35	28.20
35	34.00
35	37.25
35	34.00
35	31.00
35	31.20
35	35.67
35	34.25
35	35.38
35	30.00
35	31.50
35	31.20

Dose (mg/kg-d)	Average day of vaginal opening
35	33.50
35	32.50
35	37.67
35	35.00
35	35.20
35	33.00
35	34.50
35	38.50
35	34.30
175	31.60
175	29.40
175	33.67
175	31.67
175	34.20
175	34.50
175	37.00
175	32.22
175	38.00
175	34.50
175	34.33
175	34.67
175	37.86
175	33.00
175	36.50
175	35.33
175	39.25
350	35.00
350	36.00
350	33.80
350	33.00
350	32.00
350	31.17
350	33.57

Dose (mg/kg-d)	Average day of vaginal opening
350	34.10
350	33.33
350	38.70
350	36.33
350	36.00
350	37.25
350	35.00
350	38.50

Table D-6. Data received from study authors for Das et al. (2008) on delayed preputial separation

Dose (mg/kg-d)	Average day of preputial separation
0	29.00
0	28.20
0	28.20
0	28.00
0	31.80
0	29.20
0	28.71
0	30.00
0	31.00
0	28.29
0	30.00
0	29.80
0	31.00
0	29.50
0	29.00
0	31.00
0	29.67
35	27.40
35	33.40
35	28.20
35	31.80

Dose (mg/kg-d)	Average day of preputial separation
35	30.00
35	31.33
35	35.50
35	30.22
35	33.17
35	30.00
35	29.00
35	30.14
35	30.29
35	29.80
35	30.43
35	30.00
35	27.50
35	28.20
35	28.57
35	29.25
35	30.17
175	26.60
175	28.80
175	30.50
175	31.71
175	31.11
175	32.33
175	28.00
175	31.00
175	35.00
175	30.60
175	30.13
175	29.50
175	30.00
175	31.60
175	31.00
175	30.17

Dose (mg/kg-d)	Average day of preputial separation
175	31.50
350	28.00
350	31.80
350	31.50
350	32.40
350	31.83
350	30.80
350	31.17
350	33.80
350	34.00
350	30.33
350	30.00
350	33.17
350	32.00
350	32.80

D.2. RELATIVE LIVER WEIGHT—MALE RATS EXPOSED 90 DAYS (BUTENHOFF ET AL., 2012; VAN OTTERDIJK, 2007B)[1]

Table D-7. Dose-response data for relative liver weight in male rats following 90 day exposure (Butenhoff et al., 2012; van Otterdijk, 2007b)

Dose (mg/kg-d)	*n*	Mean	SD
0	10	2.11	0.13
1.2	10	2.29	0.14
6	10	2.26	0.16
30	10	2.8	0.32

[1]Throughout this document, if a model was selected as appropriately fitting the modeled data, that model's entries in the tables are in green shaded cells and the text is bolded.

Table D-8. Benchmark dose results for relative liver weight in male rats exposed 90 days —constant variance, BMR = 10% relative deviation (Butenhoff et al., 2012; van Otterdijk, 2007b)

Models	Restriction[a]	10% Relative deviation		*p*-Value	AIC	BMDS classification[b]	BMDS notes
		BMD	BMDL				
Constant variance							
Exponential 2 (CV—normal)	Restricted	11.3634	9.4685	0.1720	−8.8244	Questionable	Constant variance test failed (Test 2 *p*-value < 0.05) Modeled control response SD >|1.5| actual response SD
Exponential 3 (CV—normal)	Restricted	11.3634	9.4572	0.1720	−8.8244	Questionable	Constant variance test failed (Test 2 *p*-value < 0.05) Modeled control response SD >|1.5| actual response SD
Exponential 4 (CV—normal)	Restricted	10.4110	4.8569	0.0584	−6.7628	Questionable	Constant variance test failed (Test 2 *p*-value < 0.05) Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Exponential 5 (CV—normal)	Restricted	10.4033	4.8563	0.0584	−6.7621	Questionable	Constant variance test failed (Test 2 *p*-value < 0.05) Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Hill (CV—normal)	Restricted	6.6152	6.0656	NA	−4.1913	Questionable	Constant variance test failed (Test 2 *p*-value < 0.05) Modeled control response SD >|1.5| actual response SD df = 0, saturated model (goodness-of-fit *p*-value cannot be calculated)

Models	Restriction[a]	10% Relative deviation		*p*-Value	AIC	BMDS classification[b]	BMDS notes
		BMD	BMDL				
Constant variance							
Polynomial (3 degree) (CV—normal)	Restricted	12.8952	8.4671	0.0624	−6.8714	Questionable	Constant variance test failed (Test 2 *p*-value < 0.05) Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Polynomial (2 degree) (CV—normal)	Restricted	12.1463	8.4560	0.0611	−6.8370	Questionable	Constant variance test failed (Test 2 *p*-value < 0.05) Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Power (CV—normal)	Restricted	10.4151	8.4328	0.1668	−8.7631	Questionable	Constant variance test failed (Test 2 *p*-value < 0.05) Modeled control response SD >|1.5| actual response SD
Linear (CV—normal)	Unrestricted	10.4151	8.4328	0.1668	−8.7631	Questionable	Constant variance test failed (Test 2 *p*-value < 0.05) Modeled control response SD >|1.5| actual response SD

[a]"Restriction" column denotes the restriction status of applied models.

[b]"Classification" column denotes whether a model can be considered for model selection purposes. See BMDS User Guide: https://www.epa.gov/bmds.

Table D-9. Benchmark dose results for relative liver weight in male rats exposed 90 days—nonconstant variance, BMR = 10% relative deviation (Butenhoff et al., 2012; van Otterdijk, 2007b)

Models	Restriction	10% Relative deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Nonconstant variance							
Exponential 2 (NCV—normal)	Restricted	11.3982	9.0908	0.0362	−15.2001	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 3 (NCV—normal)	Restricted	11.3962	9.0911	0.0362	−15.2001	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 4 (NCV—normal)	Restricted	10.5179	5.2058	0.0096	−13.1325	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 5 (NCV—normal)	Restricted	10.5091	5.2055	0.0096	−13.1313	Questionable	Goodness-of-fit *p*-value < 0.1
Hill (NCV—normal)	Restricted	11.1854	7.9783	0.0090	−13.0126	Questionable	Goodness-of-fit *p*-value < 0.1
Polynomial (3 degree) (NCV—normal)	Restricted	12.7313	8.1751	0.0104	−13.2674	Questionable	Goodness-of-fit *p*-value < 0.1
Polynomial (2 degree) (NCV—normal)	Restricted	11.9089	8.1513	0.0100	−13.2065	Questionable	Goodness-of-fit *p*-value < 0.1
Power (NCV—normal)	Restricted	10.5174	8.1228	0.0350	−15.1326	Questionable	Goodness-of-fit *p*-value < 0.1
Linear (NCV—normal)	Unrestricted	10.5179	8.1236	0.0350	−15.1326	Questionable	Goodness-of-fit *p*-value < 0.1

Table D-10. Benchmark dose results for relative liver weight in male rats exposed 90 days—log-normal distribution, constant variance, BMR = 10% relative deviation (Butenhoff et al., 2012; van Otterdijk, 2007b)

Models[a]	Restriction	10% Relative deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Log-normal distribution, constant variance							
Exponential 2 (CV—log-normal)	Restricted	11.5672	9.5455	0.1004	−14.1752	Viable—Alternate	Modeled control response SD >|1.5| actual response SD
Exponential 3 (CV—log-normal)	**Restricted**	**11.5672**	**9.6019**	**0.1004**	**−14.1752**	**Viable—Recommended**	**Lowest AIC Modeled control response SD >|1.5| actual response SD**
Exponential 4 (CV—log-normal)	Restricted	10.6449	5.1404	0.0311	−12.1242	Questionable	Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Exponential 5 (CV—log-normal)	Restricted	10.6419	5.1401	0.0311	−12.1239	Questionable	Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Hill (CV—log-normal)	Restricted	10.5728	4.9799	0.0976	−14.1178	Questionable	Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Polynomial (3 degree) (CV—log-normal)	Restricted	12.6948	8.5635	0.0328	−12.2144	Questionable	Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Polynomial (2 degree) (CV—log-normal)	Restricted	11.9903	8.5515	0.0321	−12.1783	Questionable	Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD

Models[a]	Restriction	10% Relative deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Log-normal distribution, constant variance							
Power (CV—log-normal)	Restricted	10.6452	8.5334	0.0979	−14.1242	Questionable	Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Linear (CV—log-normal)	Unrestricted	10.6452	8.5334	0.0979	−14.1242	Questionable	Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD

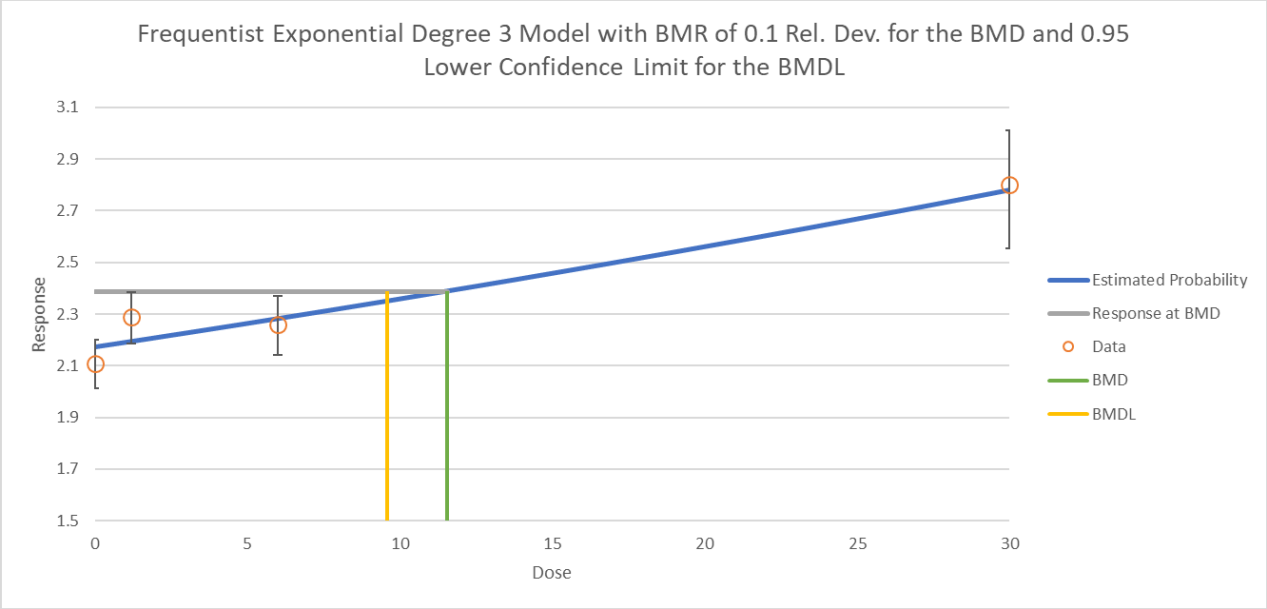

Figure D-1. Dose-response curve for the Exponential M3 model fit to relative liver weight in male rats exposed 90 days (Butenhoff et al., 2012; van Otterdijk, 2007b).

User Input

Info	
Model	frequentist Exponential degree 3 v1.1
Dataset Name	Butenhoff_90_Lweight_rel
User notes	[Add user notes here]
Dose-Response Model	M[dose] = a * exp(±1 * (b * dose)^d)
Variance Model	Var[i] = alpha

Model Options	
BMR Type	Rel. Dev.
BMRF	0.1
Tail Probability	-
Confidence Level	0.95
Distribution Type	Log-normal
Variance Type	Constant

Model Data	
Dependent Variable	[Dose]
Independent Variable	[Mean]
Total # of Observations	4
Adverse Direction	Automatic

Model Results

Benchmark Dose	
BMD	11.56718731
BMDL	9.60187006
BMDU	14.67526197
AIC	-14.17517344
Test 4 P-value	0.100441772
D.O.F.	2

Model Parameters	
# of Parameters	4
Variable	Estimate
a	2.171112769
b	0.0082397
d	Bounded
log-alpha	-5.045994496

Goodness of Fit

Dose	Size	Estimated Median	Calc'd Median	Observed Mean	Estimated GSD	Calc'd GSD	Observed SD	Scaled Residual
0	10	2.171112769	2.10600663	2.11	1.08352413	1.063487	0.13	-0.17835832
1.2	10	2.192686432	2.28573248	2.29	1.08352413	1.062982	0.14	0.284010771
6	10	2.281146197	2.25435749	2.26	1.08352413	1.073268	0.16	-0.061715421
30	10	2.779944166	2.78189148	2.8	1.08352413	1.120657	0.32	0.058533184

Likelihoods of Interest

Model	Log Likelihood*	# of Parameters	AIC
A1	12.38576382	5	-14.771528
A2	15.32442666	8	-14.648853
A3	12.38576382	5	-14.771528
fitted	10.08758672	3	-14.175173
R	-8.71328445	2	21.4265689

* Includes additive constant of -70.8323. This constant was not included in the LL derivation prior to BMDS 3.0.

Tests of Interest

Test	-2*Log(Likelihood Ratio)	Test df	p-value
1	48.07542222	6	<0.0001
2	5.877325671	3	0.11773355
3	5.877325671	3	0.11773355
4	4.596354207	2	0.10044177

Table D-11. Benchmark dose results for relative liver weight in male rats exposed 90 days—log-normal distribution, constant variance, BMR = 1 standard deviation (Butenhoff et al., 2012; van Otterdijk, 2007b)

Models	Restriction	1 Standard deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Log-normal distribution, constant variance							
Exponential 2 (CV—log-normal)	Restricted	9.7357	7.6047	0.1004	−14.1752	Viable—Alternate	Modeled control response SD >|1.5| actual response SD
Exponential 3 (CV—log-normal)	**Restricted**	**9.7356**	**7.6049**	**0.1004**	**−14.1752**	**Viable—Recommended**	**Lowest AIC Modeled control response SD >|1.5| actual response SD**
Exponential 4 (CV—log-normal)	Restricted	8.8962	0.0000	0.0311	−12.1242	Questionable	Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Exponential 5 (CV—log-normal)	Restricted	8.8943	6.9746	0.0311	−12.1239	Questionable	Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Hill (CV—log-normal)	Restricted	8.8323	4.0523	0.0976	−14.1178	Questionable	Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Polynomial (3 degree) (CV—log-normal)	Restricted	10.7197	6.8148	0.0328	−12.2144	Questionable	Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Polynomial (2 degree) (CV—log-normal)	Restricted	10.1369	6.8036	0.0321	−12.1783	Questionable	Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Power (CV—log-normal)	Restricted	8.8972	6.7871	0.0979	−14.1242	Questionable	Goodness-of-fit *p*-value < 0.1 Modeled control response SD >|1.5| actual response SD
Linear (CV—log-normal)	Unrestricted	8.8972	6.7871	0.0979	−14.1242	Questionable	Goodness-of-fit *p*-value < 0.1

Models	Restriction	1 Standard deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Log-normal distribution, constant variance							
							Modeled control response SD >|1.5| actual response SD

D.3. RELATIVE LIVER WEIGHT—P_0 MICE (DAS ET AL., 2008)

Table D-12. Dose-response data for relative liver weight in pregnant mice (Das et al., 2008)

Dose (mg/kg-d)	*n*	Mean	SD
0	6	8.04	0.66
35	6	8.76	1.37
175	7	10.28	0.75
350	6	10.65	0.62

Table D-13. Benchmark dose results for relative liver weight in pregnant mice—constant variance, BMR = 10% relative deviation (Das et al., 2008)

Models	Restriction	10% Relative deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Constant variance							
Exponential 2 (CV—normal)	Restricted	130.2877	98.9543	0.0486	73.1479	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 3 (CV—normal)	Restricted	130.2877	99.1362	0.0486	73.1479	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 4 (CV—normal)	**Restricted**	**36.1911**	**15.1545**	**0.8612**	**69.1285**	**Viable—recommended**	**Lowest AIC**
Exponential 5 (CV—normal)	Restricted	39.4346	15.2398	NA	71.0979	Questionable	df = 0, saturated model (goodness-of-fit *p*-value cannot be calculated)
Hill (CV—normal)	Restricted	38.7873	12.3846	NA	71.0979	Questionable	df = 0, saturated model (goodness-of-fit *p*-value cannot be calculated)
Polynomial (3 degree) (CV—normal)	Restricted	115.5880	84.4884	0.0736	72.3159	Questionable	Goodness-of-fit *p*-value < 0.1
Polynomial (2 degree) (CV—normal)	Restricted	115.5878	84.4883	0.0736	72.3159	Questionable	Goodness-of-fit *p*-value < 0.1
Power (CV—normal)	Restricted	115.5870	84.4876	0.0736	72.3159	Questionable	Goodness-of-fit *p*-value < 0.1
Linear (CV—normal)	Unrestricted	115.5882	84.4875	0.0736	72.3159	Questionable	Goodness-of-fit *p*-value < 0.1

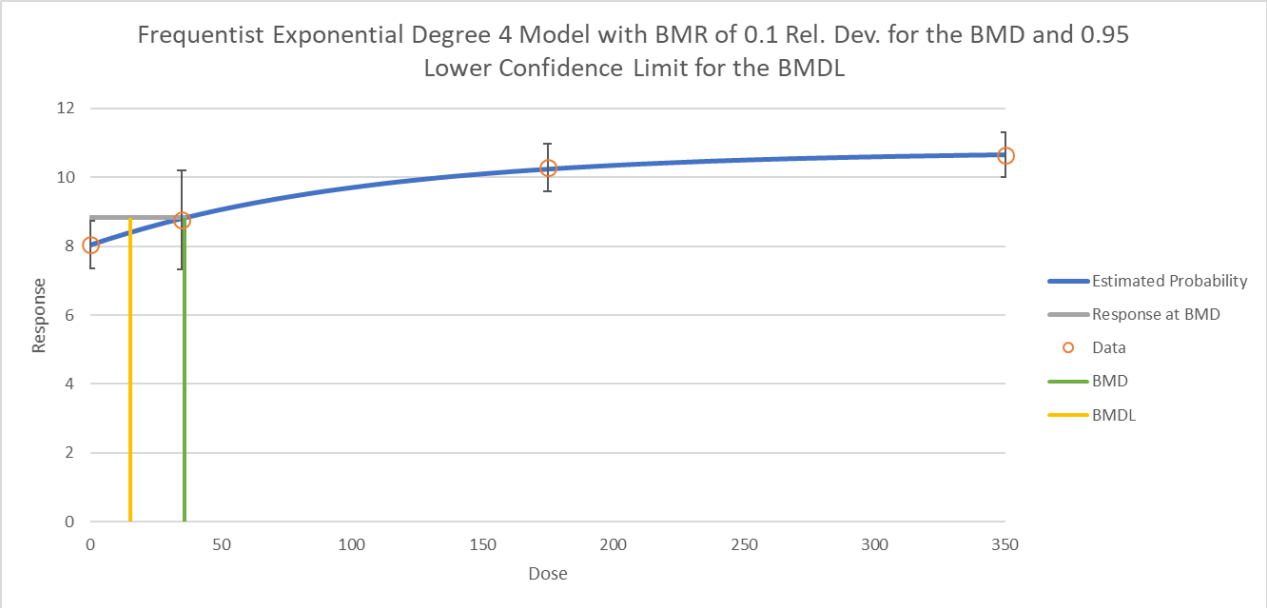

Figure D-2. Dose-response curve for the Exponential M4 model fit to relative liver weight in pregnant mice (Das et al., 2008).

User Input

Info	
Model	frequentist Exponential degree 4 v1.1
Dataset Name	Das_p_Lweight_rel
User notes	[Add user notes here]
Dose-Response Model	M[dose] = a * [c-(c-1) * exp(-b * dose)]
Variance Model	Var[i] = alpha

Model Options	
BMR Type	Rel. Dev.
BMRF	0.1
Tail Probability	-
Confidence Level	0.95
Distribution Type	Normal
Variance Type	Constant

Model Data	
Dependent Variable	[Dose]
Independent Variable	[Mean]
Total # of Observations	4
Adverse Direction	Automatic

#NAME?

Benchmark Dose	
BMD	36.19110286
BMDL	15.15446485
BMDU	87.70968183
AIC	69.12846157
Test 4 P-value	0.861196136
D.O.F.	1

Model Parameters	
# of Parameters	4
Variable	Estimate
a	8.018710905
b	0.009531749
c	1.342753894
log-alpha	-0.39273843

Goodness of Fit

Dose	Size	Estimated Median	Calc'd Median	Observed Mean	Estimated SD	Calc'd SD	Observed SD	Scaled Residual
0	6	8.018710905	8.04	8.04	0.82170879	0.66	0.66	0.063462168
35	6	8.798356028	8.76	8.76	0.82170879	1.37	1.37	-0.114338192
175	7	10.24876199	10.28	10.28	0.82170879	0.75	0.75	0.100580637
350	6	10.66937939	10.65	10.65	0.82170879	0.62	0.62	-0.057769406

Likelihoods of Interest

Model	Log Likelihood*	# of Parameters	AIC
A1	-30.54894422	5	71.0978884
A2	-27.8068244	8	71.6136488
A3	-30.54894422	5	71.0978884
fitted	-30.56423079	4	69.1284616
R	-42.8486201	2	89.6972402

* Includes additive constant of -22.97346. This constant was not included in the LL derivation prior to BMDS 3.0.

Tests of Interest

Test	-2*Log(Likelihood Ratio)	Test df	p-value
1	30.08359139	6	<0.0001
2	5.484239634	3	0.13958431
3	5.484239634	3	0.13958431
4	0.030573129	1	0.86119614

Table D-14. Benchmark dose results for relative liver weight in pregnant mice—constant variance, BMR = 1 standard deviation (Das et al., 2008)

Models	Restriction	1 Standard deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Constant variance							
Exponential 2 (CV—normal)	Restricted	141.5518	104.9937	0.0524	73.6332	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 3 (CV—normal)	Restricted	141.5511	104.9942	0.0524	73.6331	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 4 (CV—normal)	**Restricted**	**37.2658**	**16.6945**	**0.5517**	**70.0879**	**Viable—recommended**	**Lowest AIC**
Exponential 5 (CV—normal)	Restricted	40.3641	16.7699	NA	71.7337	Questionable	df = 0, saturated model (goodness-of-fit *p*-value cannot be calculated)
Hill (CV—normal)	Restricted	39.5789	13.8731	NA	71.7337	Questionable	df = 0, saturated model (goodness-of-fit *p*-value cannot be calculated)
Polynomial (3 degree) (CV—normal)	Restricted	124.9178	90.1236	0.0725	72.9822	Questionable	Goodness-of-fit *p*-value < 0.1
Polynomial (2 degree) (CV—normal)	Restricted	124.9176	90.1235	0.0725	72.9822	Questionable	Goodness-of-fit *p*-value < 0.1
Power (CV—normal)	Restricted	124.9169	90.1256	0.0725	72.9822	Questionable	Goodness-of-fit *p*-value < 0.1
Linear (CV—normal)	Unrestricted	124.9180	90.1238	0.0725	72.9822	Questionable	Goodness-of-fit *p*-value < 0.1

D.4. LIVER HYPERTROPHY—MALE RAT (BUTENHOFF ET AL., 2012; VAN OTTERDIJK, 2007B)

Table D-15. Dose-response data liver hypertrophy in male rats (Butenhoff et al., 2012; van Otterdijk, 2007b)

Dose (mg/kg-d)	*n*	Incidence
0	10	0
1.2	10	0
6	10	0
30	10	9

Table D-16. Benchmark dose results for liver hypertrophy in rats—BMR = 10% extra risk (Butenhoff et al., 2012; van Otterdijk, 2007b)

Models	Restriction	10% Extra risk		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Gamma	Restricted	16.2946	5.3859	1.0000	8.5017	Viable—alternate	
Log-logistic	Restricted	23.5001	5.4486	1.0000	10.5017	Viable—alternate	
Multistage 3rd	Restricted	10.8404	5.0184	0.9796	8.8673	Viable—alternate	
Multistage 2nd	Restricted	6.8934	3.6966	0.8078	10.2814	Viable—alternate	
Multistage 1st	Restricted	2.4428	1.4091	0.0817	18.5672	Questionable	Goodness-of-fit *p*-value < 0.1
Weibull	**Restricted**	**25.2757**	**5.3801**	**1.0000**	**8.5017**	**Viable—recommended**	**Lowest AIC**
Dichotomous Hill	Unrestricted	23.4994	5.8336	0.9995	12.5017	Viable—alternate	
Logistic	Unrestricted	23.4727	8.4278	1.0000	8.5017	Viable—alternate	
Log-probit	Unrestricted	20.1374	5.4722	1.0000	10.5017	Viable—alternate	
Probit	Unrestricted	21.2661	7.6123	1.0000	10.5017	Viable—alternate	

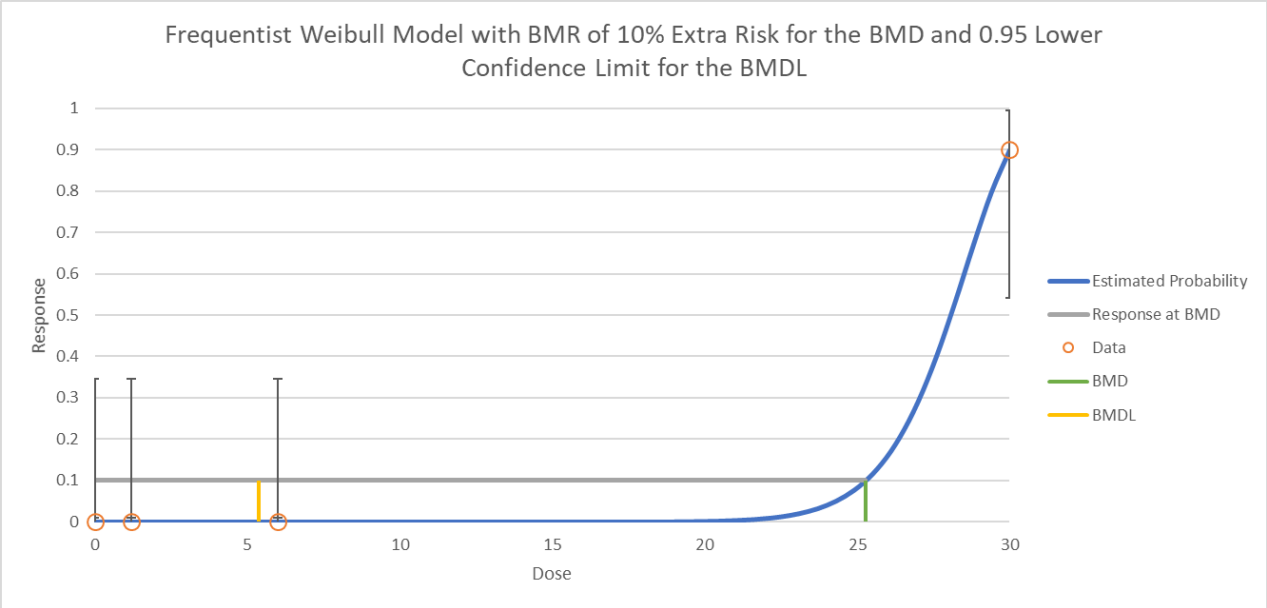

Figure D-3. Dose-response curve for the Weibull model fit to liver hypertrophy in male rats (Butenhoff et al., 2012; van Otterdijk, 2007b).

User Input

Info	
Model	frequentist Weibull v1.1
Dataset Name	Butenhoff_90_Lhypertrophy
User notes	[Add user notes here]
Dose-Response Model	P[dose] = g + (1-g)*[1-exp(-b*dose^a)]

Model Options	
Risk Type	Extra Risk
BMR	0.1
Confidence Level	0.95
Background	Estimated

Model Data	
Dependent Variable	Dose
Independent Variable	Incidence
Total # of Observations	4

Model Results

Benchmark Dose	
BMD	25.27565904
BMDL	5.380065202
BMDU	26.31774355
AIC	8.501660382
P-value	1
D.O.F.	3
Chi^2	4.56905E-07

Model Parameters	
# of Parameters	3
Variable	Estimate
g	Bounded
a	Bounded
b	5.94337E-27

Goodness of Fit

Dose	Estimated Probability	Expected	Observed	Size	Scaled Residual
0	1.523E-08	1.523E-07	0	10	-0.00039
1.2	1.523E-08	1.523E-07	0	10	-0.00039
6	1.52306E-08	1.52306E-07	0	10	-0.00039
30	0.899999999	8.999999992	9	10	8.003E-09

Analysis of Deviance

Model	Log Likelihood	# of Parameters	Deviance	Test d.f.	P Value
Full Model	-3.250829734	4	-	-	-
Fitted Model	-3.250830191	1	9.1381E-07	3	1
Reduced Model	-21.32655363	1	36.1514478	3	<0.0001

Table D-17. Dose-response data for liver hypertrophy (slight severity lesions) in male rats (Butenhoff et al., 2012; van Otterdijk, 2007b)

Dose (mg/kg-d)	*n*	Incidence
0	10	0
1.2	10	0
6	10	0
30	10	4

Table D-18. Benchmark dose results for liver hypertrophy (slight severity lesions) in male rats—BMR = 10% extra risk (Butenhoff et al., 2012; van Otterdijk, 2007b)

Models	Restriction	10% Extra risk		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Gamma	Restricted	23.1357	5.6717	1.0000	15.4602	Viable—alternate	
Log-logistic	Restricted	27.1575	5.5461	1.0000	17.4602	Viable—alternate	
Multistage 3rd	Restricted	17.7871	5.5407	0.9978	15.5422	Viable—alternate	
Multistage 2nd	Restricted	13.9892	5.1121	0.8984	17.8741	Viable—alternate	
Multistage 1st	**Restricted**	**8.1158**	**3.9098**	**0.5376**	**19.5942**	**Viable—recommended**	**Lowest BMDL**
Weibull	Restricted	27.4811	5.6718	1.0000	17.4602	Viable—alternate	
Dichotomous Hill	Unrestricted	27.1562	5.2830	0.9995	19.4602	Viable—alternate	BMD:BMDL ratio > 5
Logistic	Unrestricted	26.9449	13.6106	1.0000	15.4602	Viable—alternate	
Log-Probit	Unrestricted	24.8237	5.3131	1.0000	17.4602	Viable—alternate	
Probit	Unrestricted	25.5166	12.1561	1.0000	17.4602	Viable—alternate	

D.5. TOTAL T4—MALE RAT (BUTENHOFF ET AL., 2012; VAN OTTERDIJK, 2007B)

Table D-19. Dose-response data for total T4 levels in male rats (Butenhoff et al., 2012; van Otterdijk, 2007b)

Dose (mg/kg-d)	*n*	Mean	SD
0	10	5.27	0.71
1.2	10	5.97	1.08
6	9	4.46	0.88
30	9	3.23	0.55

Table D-20. Benchmark dose results for total T4 levels in male rats—constant variance, BMR = 1 standard deviation (Butenhoff et al., 2012; van Otterdijk, 2007b)

Models	Restriction	1 Standard deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Constant variance							
Exponential 2 (CV—normal)	Restricted	9.2322	6.5166	0.0138	104.3816	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 3 (CV—normal)	Restricted	9.2324	6.5166	0.0138	104.3816	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 4 (CV—normal)	Restricted	4.9496	2.5239	0.0075	104.9572	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 5 (CV—normal)	Restricted	5.7655	3.5138	NA	103.5642	Questionable	df = 0, saturated model (goodness-of-fit *p*-value cannot be calculated)
Hill (CV—normal)	Restricted	5.5394	3.2999	NA	103.5644	Questionable	df = 0, saturated model (goodness-of-fit *p*-value cannot be calculated)
Polynomial (3 degree) (CV—normal)	Restricted	11.5906	8.7704	0.0090	105.2374	Questionable	Goodness-of-fit *p*-value < 0.1
Polynomial (2 degree) (CV—normal)	Restricted	11.5906	8.7704	0.0090	105.2374	Questionable	Goodness-of-fit *p*-value < 0.1
Power (CV—normal)	Restricted	11.5906	8.7706	0.0090	105.2374	Questionable	Goodness-of-fit *p*-value < 0.1
Linear (CV—normal)	Unrestricted	11.5906	8.7704	0.0090	105.2374	Questionable	Goodness-of-fit *p*-value < 0.1

Table D-21. Benchmark dose results for total T4 levels in male rats—nonconstant variance, BMR = 1 standard deviation (Butenhoff et al., 2012; van Otterdijk, 2007b)

Models	Restriction	1 Standard deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Nonconstant variance							
Exponential 2 (NCV—normal)	Restricted	11.3786	7.8978	0.0182	102.5921	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 3 (NCV—normal)	Restricted	11.3789	7.8977	0.0182	102.5921	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 4 (NCV—normal)	Restricted	5.8707	2.9606	0.0104	103.1558	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 5 (NCV—normal)	Restricted	5.8297	3.9098	NA	102.1810	Questionable	df = 0, saturated model (goodness-of-fit *p*-value cannot be calculated)
Hill (NCV—normal)	Restricted	5.8562	3.7033	NA	102.1809	Questionable	df = 0, saturated model (goodness-of-fit *p*-value cannot be calculated)
Polynomial (3 degree) (NCV—normal)	Restricted	13.7327	10.1890	0.0130	103.2666	Questionable	Goodness-of-fit *p*-value < 0.1
Polynomial (2 degree) (NCV—normal)	Restricted	13.7329	10.1889	0.0130	103.2666	Questionable	Goodness-of-fit *p*-value < 0.1
Power (NCV—normal)	Restricted	13.7325	10.1890	0.0130	103.2666	Questionable	Goodness-of-fit *p*-value < 0.1
Linear (NCV—normal)	Unrestricted	13.7332	10.1889	0.0130	103.2666	Questionable	Goodness-of-fit *p*-value < 0.1

Table D-22. Benchmark dose results for total T4 levels in male rats—log-normal distribution, constant variance, BMR = 1 standard deviation (Butenhoff et al., 2012; van Otterdijk, 2007b)

Models	Restriction	1 Standard deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Log-normal distribution, constant variance							
Exponential 2 (CV—log-normal)	Restricted	12.0074	7.6347	0.0223	98.5676	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 3 (CV—log-normal)	Restricted	12.0074	7.6347	0.0223	98.5676	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 4 (CV—log-normal)	Restricted	5.7060	2.5325	0.0200	98.3698	Questionable	Goodness-of-fit *p*-value < 0.1
Exponential 5 (CV—log-normal)	Restricted	5.9263	3.4425	NA	97.5382	Questionable	df = 0, saturated model (goodness-of-fit *p*-value cannot be calculated)
Hill (CV—log-normal)	Restricted	-	-	-	-	Questionable	df = 0, saturated model (goodness-of-fit *p*-value cannot be calculated)
Polynomial (3 degree) (CV—log-normal)	Restricted	-	-	-	-	Questionable	Goodness-of-fit *p*-value < 0.1
Polynomial (2 degree) (CV—log-normal)	Restricted	-	-	-	-	Questionable	Goodness-of-fit *p*-value < 0.1
Power (CV—log-normal)	Restricted	-	-	-	-	Questionable	Goodness-of-fit *p*-value < 0.1
Linear (CV—log-normal)	Unrestricted	-	-	-	-	Questionable	Goodness-of-fit *p*-value < 0.1

D.6. INCREASED FETAL MORTALITY – MALE AND FEMALE F_1 MICE (DAS ET AL., 2008)

Table D-23. Dose-response data for increased fetal mortality (Das et al., 2008)

Dose (mg/kg-d)	*n* (No. of implants)	No. of dead fetuses/neonates by PND 21	Litter-specific covariate (Maternal weight on GD1)
0	16	1	30
0	16	1	28.2
0	11	2	27.7
0	11	0	27.4
0	12	3	25.9
0	11	0	24.1
0	15	0	29.2
0	14	1	28
0	12	3	27.1
0	14	0	26.8
0	16	1	26.6
0	13	2	25.1
0	17	3	30.1
0	14	0	29
0	6	0	27.5
0	9	2	28.1
0	6	0	26.9
0	13	1	26.7
0	11	0	23.3
0	8	8	25.8
0	18	18	31.4
35	15	3	28.1
35	13	0	29.3
35	13	0	27.4
35	14	1	27
35	15	2	26.9
35	13	2	25.7
35	12	4	31.6
35	13	0	29.2
35	14	1	27.7

Dose (mg/kg-d)	*n* (No. of implants)	No. of dead fetuses/neonates by PND 21	Litter-specific covariate (Maternal weight on GD1)
35	16	0	27.5
35	13	2	28.1
35	7	3	25.5
35	15	1	30.3
35	13	0	27.5
35	14	1	28.1
35	13	1	27.9
35	11	0	26.4
35	10	1	27.4
35	13	1	27.9
35	13	0	26.1
35	13	1	24.8
35	12	1	24.8
35	2	2	23.1
175	14	1	28.1
175	15	0	27.5
175	14	0	27.4
175	14	1	27.5
175	15	2	29.4
175	14	1	27.5
175	15	0	26
175	16	2	26.2
175	11	0	23.4
175	16	3	29.1
175	11	0	28.2
175	13	0	25.8
175	11	2	26.8
175	15	1	26.9
175	14	1	25
175	13	1	26.7
175	2	2	25.5
175	2	2	25.4
175	9	9	29
175	5	5	25

Dose (mg/kg-d)	*n* (No. of implants)	No. of dead fetuses/neonates by PND 21	Litter-specific covariate (Maternal weight on GD1)
350	7	2	29.2
350	12	1	26.3
350	16	3	27.4
350	11	0	25.1
350	14	2	25.3
350	12	1	29.5
350	16	2	28.8
350	17	2	26.2
350	12	2	26.2
350	16	0	27.3
350	9	3	27.6
350	13	0	27.7
350	13	0	27.4
350	13	1	26.4
350	7	1	24.6
350	3	3	21.5
350	2	2	23
350	13	13	25.8
350	13	13	24.6
350	3	3	25.1
350	14	14	28.2
350	13	13	29.2

Table D-24. Benchmark dose results for increased fetal mortality (male and female mice)—BMR = 1% extra risk (Das et al., 2008)

Models	Restriction	1% Extra risk		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Nested logistic (lsc+ilc+)	**Restricted**	**19.5989**	**5.7383**	**Infinity**	**0.2633**	**Viable—Recommended**	**Lowest AIC BMDL 3× lower than lowest non-zero dose**
Nested logistic (lsc+ilc−)	Restricted	326.9633	170.7455	Infinity	<0.0001	Questionable	Goodness of fit *p*-value < 0.1

Models	Restriction	1% Extra risk		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Nested logistic (lsc−ilc+)	Restricted	50.4014	10.1822	Infinity	0.0833	Questionable	Goodness of fit *p*-value < 0.1 BMDL 3× lower than lowest non-zero dose
Nested logistic (lsc−ilc−)	Restricted	191.2272	81.9934	Infinity	<0.0001	Questionable	Goodness of fit *p*-value < 0.1

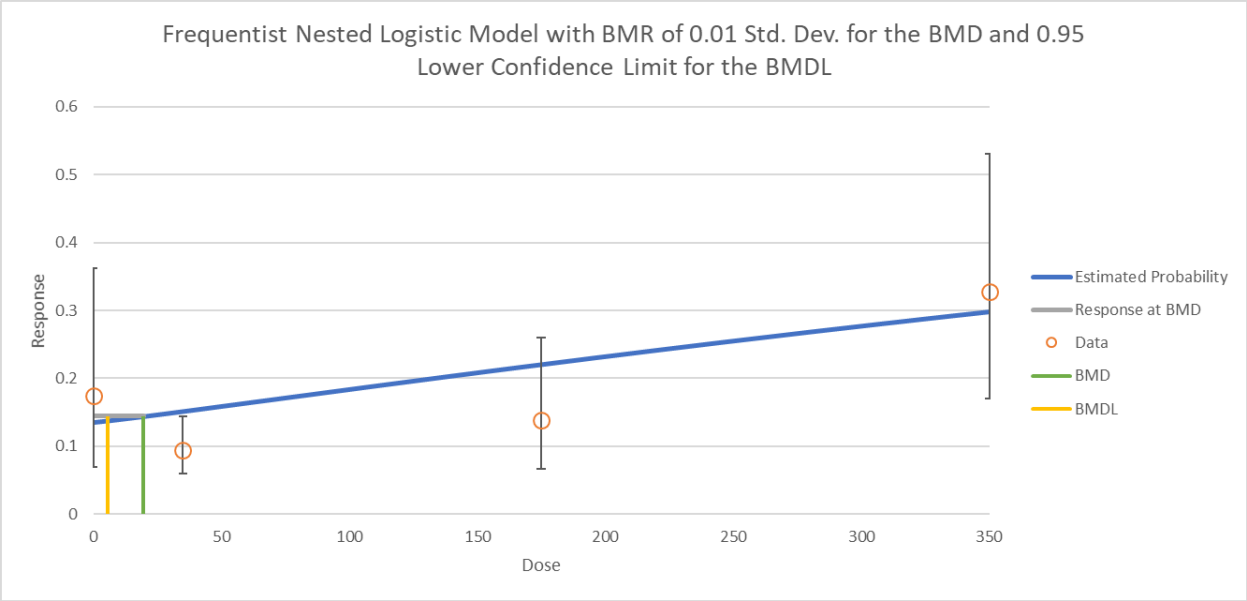

Figure D-4. Dose-response curve for the Nested-Logistic model fit to increased fetal mortality in male and female mice (Das et al., 2008).

User Input

Info	
Model	frequentist Nested Logistic_lsc+ilc+_ v2.2
Dataset Name	Das_FLR_Fetal_Death
User notes	[Add user notes here]
Dose-Response Model	P[dose] = alpha + theta1*Rij + [1 - alpha - theta1*Rij]/[1+exp(-beta-theta2*Rij-rho*log(dose))]

Model Options	
Risk Type	Extra Risk
BMR	0.01
Confidence Level	0.95
Litter Specific Covariate	Overall Mean
Intralitter Correlation	Estimate
Background	Estimate

Model Data	
Dependent Variable	Dose
Independent Variable	Incidence
Total # of Observations	87

Model Results

Benchmark Dose	
BMD	19.59891366
BMDL	5.738265629
BMDU	-
AIC	688.92042
P-value	0.263333333
D.O.F.	78
Chi^2	96.74138773

Model Parameters	
# of Parameters	9
Variable	Estimate
alpha	-0.5312932
beta	16.8290783
theta1	0.024711312
theta2	-0.913645475
rho	1.08467654
phi1	0.368252856
phi2	0.135465621
phi3	0.509745798
phi4	0.576861839

Bootstrap Results	
# Iterations	1000
Bootstrap Seed	1599045577
Log-likelihood	-335.46021
Observed Chi-square	96.74138773
Combined P-value	0.263333333

Bootstrap Runs

Run	Bootstrap Chi-square Percentiles				
	P-Value	50th	90th	95th	99th
1	0.285	85.65851617	109.848694	117.6	134.18995
2	0.258	85.12942257	110.722914	119.0851	131.33939
3	0.247	85.05473338	108.751327	115.9296	137.48443
Combined	0.263333333	85.30000651	109.644757	117.9646	135.17128

Scaled Residuals	
Minimum scaled residual for dose group nearest the BMD	-0.50395
Minimum ABS(scaled residual) for dose group nearest the BMD	0.503952
Average Scaled residual for dose group nearest the BMD	-0.50395
Average ABS(scaled residual) for dose group nearest the BMD	0.503952
Maximum scaled residual for dose group nearest the BMD	-0.50395
Maximum ABS(scaled residual) for dose group nearest the BMD	0.503952

Litter Data						
Dose	Lit. Spec. Cov.	Est. Prob.	Litter Size	Expected	Observed	Scaled Residual
0	23.3	0.044480361	11	0.489284	0	-0.330689547
0	24.1	0.06424941	11	0.706744	0	-0.401615664
0	25.1	0.088960722	13	1.156489	2	0.353012617
0	25.8	0.10625864	8	0.850069	8	4.33673202
0	25.9	0.108729771	12	1.304757	3	0.699492477
0	26.6	0.126027689	16	2.016443	1	-0.299772055
0	26.7	0.12849882	13	1.670485	1	-0.238711127
0	26.8	0.130969952	14	1.833579	0	-0.603802545
0	26.9	0.133441083	6	0.800646	0	-0.570250323
0	27.1	0.138383345	12	1.6606	3	0.498243205
0	27.4	0.145796738	11	1.603764	0	-0.633212408
0	27.5	0.14826787	6	0.889607	0	-0.606305933
0	27.7	0.153210132	11	1.685311	2	0.121734257
0	28	0.160623525	14	2.248729	1	-0.377819018
0	28.1	0.163094657	9	1.467852	2	0.241698202
0	28.2	0.165565788	16	2.649053	1	-0.434253259
0	29	0.185334837	14	2.594688	0	-0.741848906
0	29.2	0.190277099	15	2.854156	0	-0.756724162
0	30	0.210046149	16	3.360738	1	-0.567256611
0	30.1	0.21251728	17	3.612794	3	-0.138387912
0	31.4	0.244641985	18	4.403556	18	2.76675012
35	23.1	0.420439493	2	0.840879	2	1.558208765
35	24.8	0.193667312	12	2.324008	1	-0.612920728
35	24.8	0.193667312	13	2.517675	1	-0.657368032
35	25.5	0.160429819	7	1.123009	3	1.435714172
35	25.7	0.15539473	13	2.020131	2	-0.009511434
35	26.1	0.149721705	13	1.946382	0	-0.933727752
35	26.4	0.148530963	11	1.633841	0	-0.902727468
35	26.9	0.150776303	15	2.261645	2	-0.110930327
35	27	0.151716296	14	2.124028	1	-0.503951928
35	27.4	0.156698447	13	2.03708	0	-0.959178264
35	27.4	0.156698447	10	1.566984	1	-0.331093052
35	27.5	0.158199979	16	2.5312	0	-0.995853056
35	27.5	0.158199979	13	2.0566	0	-0.964622031
35	27.7	0.16145156	14	2.260322	1	-0.550928027
35	27.9	0.164988496	13	2.14485	1	-0.527947549
35	27.9	0.164988496	13	2.14485	1	-0.527947549
35	28.1	0.168763344	15	2.53145	3	0.189788804
35	28.1	0.168763344	14	2.362687	1	-0.585185094
35	28.1	0.168763344	13	2.193923	2	-0.088622513
35	29.2	0.192293335	13	2.499813	0	-1.08570969
35	29.3	0.194583201	13	2.529582	0	-1.093706422
35	30.3	0.218173919	15	3.272609	1	-0.834803748
35	31.6	0.249793258	12	2.997519	4	0.423637452
175	23.4	0.753292803	11	8.286221	0	-2.346999161
175	25	0.450913899	14	6.312795	1	-1.033293673
175	25	0.450913899	5	2.254569	5	1.415449135
175	25.4	0.381299381	2	0.762599	2	1.466122666
175	25.5	0.365523168	2	0.731046	2	1.516399081
175	25.8	0.322690467	13	4.194976	0	-0.932876612
175	26	0.298046899	15	4.470703	0	-0.884741991
175	26.2	0.276550941	16	4.424815	2	-0.460908554
175	26.7	0.235805287	13	3.065469	1	-0.505849379
175	26.8	0.229690589	11	2.526596	2	-0.152863219
175	26.9	0.2241858	15	3.362787	1	-0.512836334
175	27.4	0.204707536	14	2.865906	0	-0.687383811
175	27.5	0.202204115	14	2.830858	1	-0.441145101
175	27.5	0.202204115	14	2.830858	1	-0.441145101
175	27.5	0.202204115	15	3.033062	0	-0.683561326
175	28.1	0.194568014	14	2.723952	1	-0.421446879
175	28.2	0.194300585	11	2.137306	0	-0.659587572
175	29	0.199087513	9	1.791788	9	2.670218904
175	29.1	0.200338227	16	3.205412	3	-0.043633258
175	29.4	0.204695235	15	3.070429	2	-0.240145327
350	21.5	0.971795484	3	2.915386	3	0.201065485
350	23	0.901250165	2	1.8025	2	0.372789514
350	24.6	0.695111669	13	9.036452	13	0.848375256
350	24.6	0.695111669	7	4.865782	1	-1.502678856
350	25.1	0.601246578	3	1.80374	3	0.961150815
350	25.1	0.601246578	11	6.613712	0	-1.565382323
350	25.3	0.562306321	14	7.872288	2	-1.085132125
350	25.4	0.542866281	1	0.542866	1	0.91764604
350	25.8	0.467290016	13	6.07477	13	1.367719528
350	26.2	0.398842995	12	4.786116	2	-0.606043119
350	26.2	0.398842995	17	6.780331	2	-0.740295677
350	26.3	0.383300844	12	4.59961	1	-0.788584398
350	26.4	0.368470315	13	4.790114	1	-0.774204472
350	27.3	0.269132881	16	4.306126	0	-0.781258285
350	27.4	0.261775084	13	3.403076	0	-0.762807027
350	27.4	0.261775084	16	4.188401	3	-0.217527974
350	27.6	0.249017755	9	2.24116	3	0.246847242
350	27.7	0.243559594	13	3.166275	0	-0.726875687
350	28.2	0.224244443	14	3.139422	14	2.387132337
350	28.8	0.214724836	16	3.435597	2	-0.281313821
350	29.2	0.214117934	13	2.783533	13	2.454127084
350	29.2	0.214117934	7	1.498826	2	0.218630258
350	29.5	0.215777472	12	2.58933	1	-0.411519634

D.7. DELAYED EYE OPENING—F_1 MALE AND FEMALE MICE (DAS ET AL., 2008)

Table D-25. Dose-response data for delayed eye opening in male and female mice (Das et al., 2008)

Dose (mg/kg-d)	*n*	Mean	SD
0	20	16.28	1.19
35	22	17.38	0.79
175	17	17.69	0.68
350	15	17.8	0.83

Table D-26. Benchmark dose results for delayed eye opening in male and female mice—constant variance, BMR = 5% relative deviation (Das et al., 2008)

Models	Restriction	5% Relative deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Constant variance							
Exponential 2 (CV—normal)	Restricted	252.3387	178.6688	0.0008	211.1176	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Exponential 3 (CV—normal)	Restricted	252.3380	178.7347	0.0008	211.1176	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Exponential 4 (CV—normal)	Restricted	20.4436	0.0000	0.7270	198.8811	Unusable	BMD computation failed; lower limit includes zero BMDL not estimated
Exponential 5 (CV—normal)	Restricted	175.5239	0.0000	NA	215.6060	Unusable	BMD computation failed; lower limit includes zero BMDL not estimated |Residual at control| >2 df = 0, saturated model (goodness-of-fit *p*-value cannot be calculated)
Hill (CV—normal)	**Restricted**	**16.1508**	**4.8878**	**0.8659**	**198.7878**	**Viable—recommended**	**Lowest AIC BMDL 3× lower than lowest nonzero dose**
Polynomial (3 degree) (CV—normal)	Restricted	247.2477	172.9292	0.0008	210.9441	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Polynomial (2 degree) (CV—normal)	Restricted	247.2476	172.9292	0.0008	210.9441	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Power (CV—normal)	Restricted	247.2483	172.9366	0.0008	210.9441	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Linear (CV—normal)	Unrestricted	247.2471	172.9288	0.0008	210.9441	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2

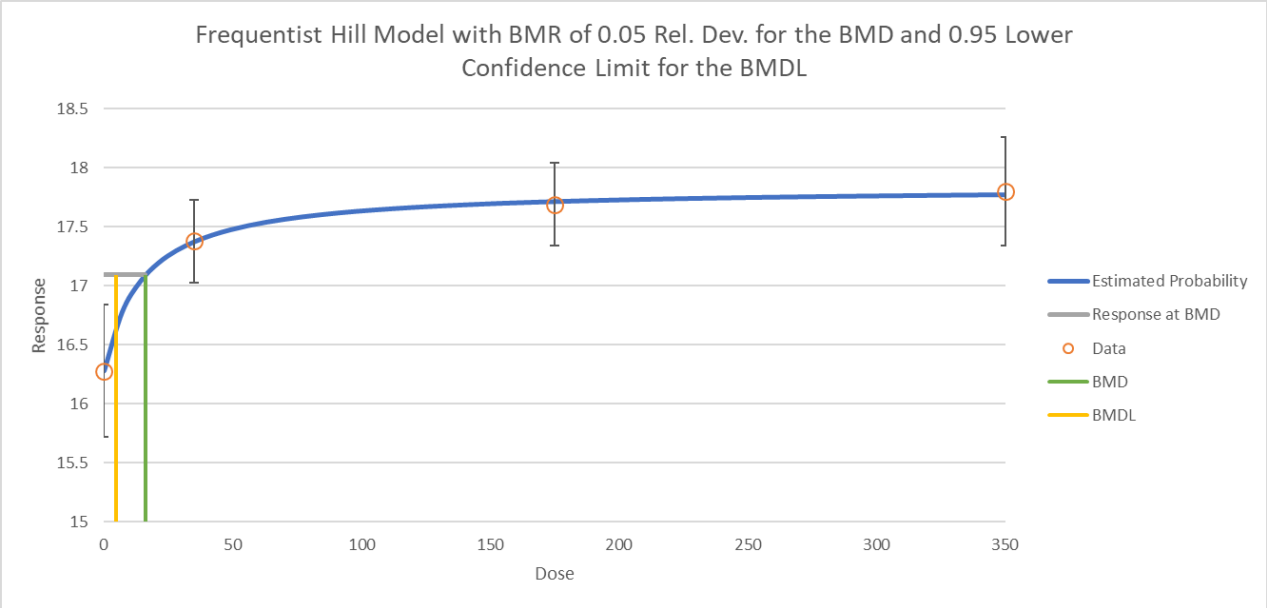

Figure D-5. Dose response curve for the Hill model fit to delayed eye opening in male and female mice (Das et al., 2008).

User Input

Info	
Model	frequentist Hill v1.1
Dataset Name	Das_EO_litter_SDs
User notes	[Add user notes here]
Dose-Response Model	M[dose] = g + v*dose^n/(k^n + dose^n)
Variance Model	Var[i] = alpha

Model Options	
BMR Type	Rel. Dev.
BMRF	0.05
Tail Probability	-
Confidence Level	0.95
Distribution Type	Normal
Variance Type	Constant

Model Data	
Dependent Variable	[Dose]
Independent Variable	[Mean]
Total # of Observations	4
Adverse Direction	Automatic

Model Results

Benchmark Dose	
BMD	16.15084927
BMDL	4.88775303
BMDU	58.67497527
AIC	198.7877861
Test 4 P-value	0.865852068
D.O.F.	1

Model Parameters	
# of Parameters	5
Variable	Estimate
g	16.28027637
v	1.557732828
k	14.75612987
n	Bounded
alpha	0.771309051

Goodness of Fit

Dose	Size	Estimated Median	Calc'd Median	Observed Mean	Estimated SD	Calc'd SD	Observed SD	Scaled Residual
0	20	16.28027637	16.28	16.28	0.87824202	1.19	1.19	-0.001407337
35	22	17.3760338	17.38	17.38	0.87824202	0.79	0.79	0.021182211
175	17	17.71687421	17.69	17.69	0.87824202	0.68	0.68	-0.126167037
350	15	17.77499146	17.8	17.8	0.87824202	0.83	0.83	0.110285841

Likelihoods of Interest

Model	Log Likelihood*	# of Parameters	AIC
A1	-95.37962446	5	200.759249
A2	-91.88601151	8	199.772023
A3	-95.37962446	5	200.759249
fitted	-95.39389305	4	198.787786
R	-109.7197233	2	223.439447

* Includes additive constant of -68.00145. This constant was not included in the LL derivation prior to BMDS 3.0.

Tests of Interest

Test	-2*Log(Likelihood Ratio)	Test df	p-value
1	35.6674235	6	<0.0001
2	6.987225901	3	0.07230604
3	6.987225901	3	0.07230604
4	0.028537187	1	0.86585207

Table D-27. Benchmark dose results for delayed eye opening in male and female mice—constant variance, BMR = 1 standard deviation (Das et al., 2008)

Models	Restriction	1 Standard deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Constant variance							
Exponential 2 (CV—normal)	Restricted	289.0417	204.0632	0.0008	211.1176	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Exponential 3 (CV—normal)	Restricted	289.0397	204.0631	0.0008	211.1176	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Exponential 4 (CV—normal)	**Restricted**	**23.0895**	**12.5328**	**0.7270**	**198.8811**	**Viable—recommended**	**Lowest AIC**
Exponential 5 (CV—normal)	Restricted	−9,999.0000	0.0000	NA	215.6060	Unusable	BMD computation failed BMD not estimated BMDL not estimated |Residual at control| >2 df = 0, saturated model (goodness-of-fit *p*-value cannot be calculated)
Hill (CV—normal)	Restricted	19.0723	0.0000	0.8659	198.7878	Unusable	BMD computation failed; lower limit includes zero BMDL not estimated
Polynomial (3 degree) (CV—normal)	Restricted	284.0211	198.2059	0.0008	210.9441	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Polynomial (2 degree) (CV—normal)	Restricted	284.0211	198.2059	0.0008	210.9441	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Power (CV—normal)	Restricted	284.0218	198.2009	0.0008	210.9441	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Linear (CV—normal)	Unrestricted	284.0204	198.2054	0.0008	210.9441	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2

D.8. VAGINAL OPENING—F_1 FEMALE MICE (DAS ET AL., 2008)

Table D-28. Dose response data for delayed vaginal opening in female mice (Das et al., 2008)

Dose (mg/kg-d)	*n*	Mean	SD
0	83	31.59	5.386
35	97	33.598	5.715
175	89	34.292	5.714
350	87	35.023	5.188

Table D-29. Benchmark dose results for delayed vaginal opening in female mice—constant variance, 5% relative deviation (Das et al., 2008)

Models	Restriction	5% Relative deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Constant variance							
Exponential 2 (CV—normal)	Restricted	199.6149	137.1410	0.0106	348.8761	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Exponential 3 (CV—normal)	Restricted	199.6216	137.1431	0.0106	348.8761	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Exponential 4 (CV—normal)	Restricted	17.1139	0.0000	0.6944	341.9320	Unusable	BMD computation failed; lower limit includes zero BMDL not estimated
Exponential 5 (CV—normal)	Restricted	30.5201	0.0000	NA	343.9392	Unusable	BMD computation failed; lower limit includes zero BMDL not estimated df = 0, saturated model (goodness-of-fit p-value cannot be calculated)
Hill (CV—normal)	**Restricted**	**13.5161**	**3.7929**	**0.8401**	**341.8184**	**Viable—recommended**	**Lowest AIC BMDL 3× lower than lowest nonzero dose**
Polynomial (3 degree) (CV—normal)	Restricted	193.4400	130.5619	0.0115	348.7113	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2

Models	Restriction	5% Relative deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Polynomial (2 degree) (CV—normal)	Restricted	193.4443	130.5615	0.0115	348.7113	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Power (CV—normal)	Restricted	193.4434	130.5626	0.0115	348.7113	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Linear (CV—normal)	Unrestricted	193.4436	130.5610	0.0115	348.7113	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2

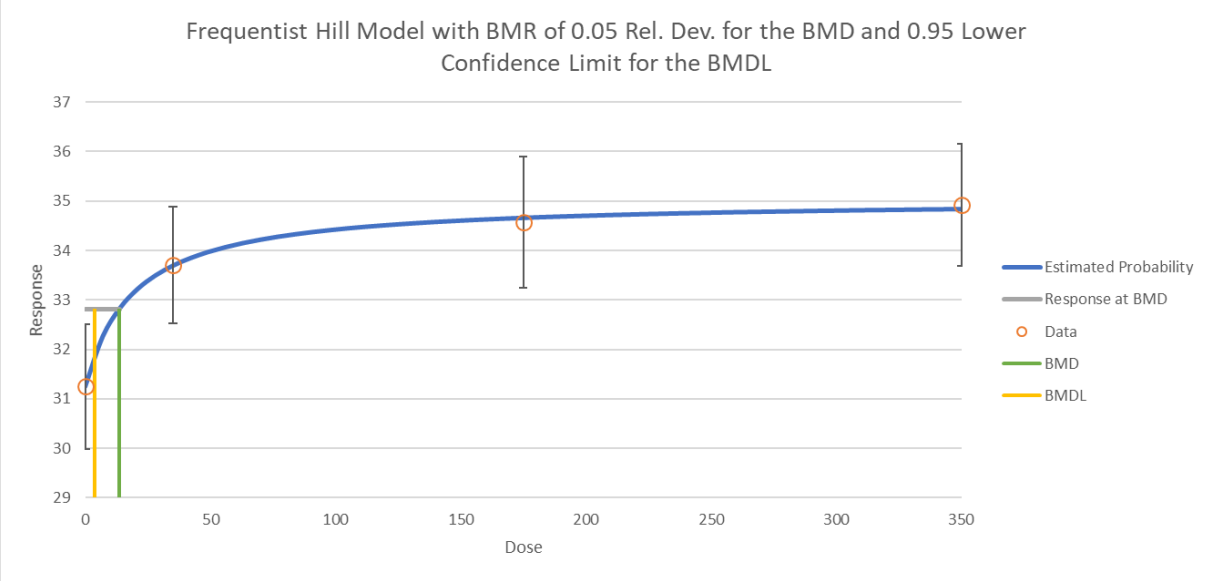

Figure D-6. Dose response curve for the Hill model fit to delayed vaginal opening in female mice (Das et al., 2008).

User Input

Info	
Model	frequentist Hill v1.1
Dataset Name	Das_VO_litter_SDs
User notes	[Add user notes here]
Dose-Response Model	M[dose] = g + v*dose^n/(k^n + dose^n)
Variance Model	Var[i] = alpha

Model Options	
BMR Type	Rel. Dev.
BMRF	0.05
Tail Probability	-
Confidence Level	0.95
Distribution Type	Normal
Variance Type	Constant

Model Data	
Dependent Variable	[Dose]
Independent Variable	[Mean]
Total # of Observations	4
Adverse Direction	Automatic

Model Results

Benchmark Dose	
BMD	13.51609885
BMDL	3.792905489
BMDU	58.81907947
AIC	341.8183924
Test 4 P-value	0.840124836
D.O.F.	1

Model Parameters	
# of Parameters	5
Variable	Estimate
g	31.25160173
v	3.782877454
k	19.2052612
n	Bounded
alpha	6.040525655

Goodness of Fit

Dose	Size	Estimated Median	Calc'd Median	Observed Mean	Estimated SD	Calc'd SD	Observed SD	Scaled Residual
0	19	31.25160173	31.25	31.25	2.45774809	2.62	2.62	-0.002840717
35	21	33.69418217	33.71	33.71	2.45774809	2.59	2.59	0.029493016
175	17	34.66038453	34.57	34.57	2.45774809	2.59	2.59	-0.151628625
350	15	34.83770206	34.92	34.92	2.45774809	2.23	2.23	0.129687238

Likelihoods of Interest

Model	Log Likelihood*	# of Parameters	AIC
A1	-166.8888479	5	343.777696
A2	-166.5982185	8	349.196437
A3	-166.8888479	5	343.777696
fitted	-166.9091962	4	341.818392
R	-177.364099	2	358.728198

* Includes additive constant of -66.16357. This constant was not included in the LL derivation prior to BMDS 3.0.

Tests of Interest

Test	-2*Log(Likelihood Ratio)	Test df	p-value
1	21.53176107	6	0.00147157
2	0.581258883	3	0.900709
3	0.581258883	3	0.900709
4	0.040696527	1	0.84012484

Table D-30. Benchmark dose results for delayed vaginal opening in female mice—constant variance, 1 standard deviation (Das et al., 2008)

Models	Restriction	1 Standard deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Constant variance							
Exponential 2 (CV—normal)	Restricted	316.9350	218.4320	0.0106	348.8761	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Exponential 3 (CV—normal)	Restricted	316.9457	218.4320	0.0106	348.8761	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Exponential 4 (CV—normal)	**Restricted**	**35.1705**	**15.4720**	**0.6944**	**341.9320**	**Viable—recommended**	**Lowest AIC**
Exponential 5 (CV—normal)	Restricted	34.9991	15.4632	NA	343.9392	Questionable	df = 0, saturated model (goodness-of-fit *p*-value cannot be calculated)
Hill (CV—normal)	Restricted	35.6204	0.0000	0.8401	341.8184	Unusable	BMD computation failed; lower limit includes zero BMDL not estimated
Polynomial (3 degree) (CV—normal)	Restricted	311.4806	211.1287	0.0115	348.7113	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Polynomial (2 degree) (CV—normal)	Restricted	311.4877	211.1313	0.0115	348.7113	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Power (CV—normal)	Restricted	311.4864	211.1303	0.0115	348.7113	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2
Linear (CV—normal)	Unrestricted	311.4866	211.1307	0.0115	348.7113	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2

D.9. PREPUTIAL SEPARATION—F_1 MALE MICE (DAS ET AL., 2008)

Table D-31. Dose-response data for delayed preputial separation in male mice (Das et al., 2008)

Dose (mg/kg-d)	*n*	Mean	SD
0	17	29.55	1.14
35	21	30.21	1.99
175	17	30.56	1.84
350	15	31.88	1.72

Table D-32. Benchmark dose results for delayed preputial separation in male mice—constant variance, BMR = 5% relative deviation (Das et al., 2008)

Models	Restriction	5% Relative deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Constant variance							
Exponential 2 (CV—normal)	Restricted	254.8183	179.1436	0.6004	277.5960	Viable—alternate	
Exponential 3 (CV—normal)	**Restricted**	**254.8005**	**179.1431**	**0.6004**	**277.5960**	**Viable—recommended**	**Lowest AIC**
Exponential 4 (CV—normal)	Restricted	252.8480	102.0115	0.3080	279.6149	Viable—alternate	
Exponential 5 (CV—normal)	Restricted	252.5410	101.9527	0.3076	279.6166	Viable—alternate	
Hill (CV—normal)	Restricted	194.2094	175.4639	0.2286	280.0252	Viable—alternate	
Polynomial (3 degree) (CV—normal)	Restricted	276.4524	176.5648	0.3427	279.4759	Viable—alternate	
Polynomial (2 degree) (CV—normal)	Restricted	269.5337	175.9153	0.3268	279.5372	Viable—alternate	
Power (CV—normal)	Restricted	252.7648	175.1179	0.5950	277.6140	Viable—alternate	
Linear (CV—normal)	Unrestricted	252.7653	175.1182	0.5950	277.6140	Viable—alternate	

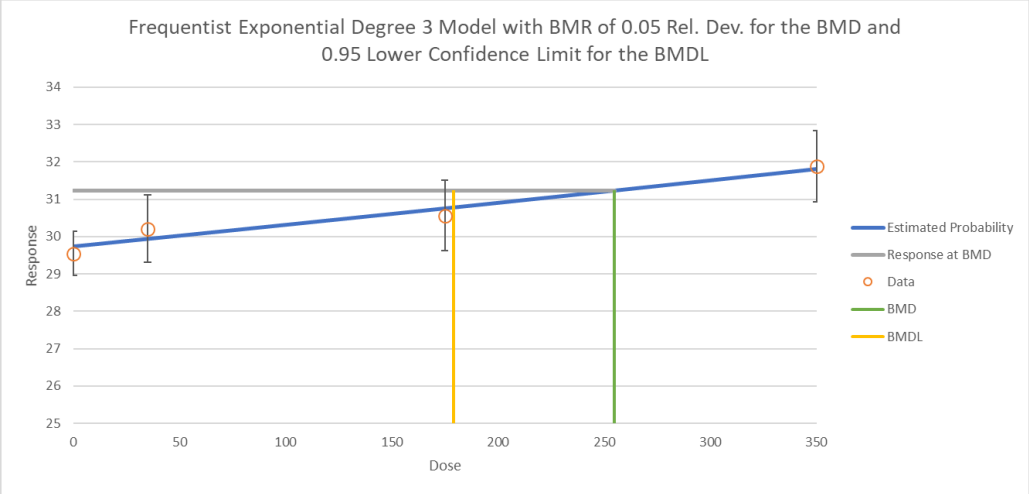

Figure D-7. Dose response curve for the Exponential 3 model fit to delayed preputial separation in male mice (Das et al., 2008).

User Input

Info	
Model	frequentist Exponential degree 3 v1.1
Dataset Name	Das_PS_litter_SDs
User notes	[Add user notes here]
Dose-Response Model	M[dose] = a * exp(±1 * (b * dose)^d)
Variance Model	Var[i] = alpha

Model Options	
BMR Type	Rel. Dev.
BMRF	0.05
Tail Probability	-
Confidence Level	0.95
Distribution Type	Normal
Variance Type	Constant

Model Data	
Dependent Variable	[Dose]
Independent Variable	[Mean]
Total # of Observations	4
Adverse Direction	Automatic

Model Results

Benchmark Dose	
BMD	254.8005164
BMDL	179.1431485
BMDU	443.2041287
AIC	277.5960319
Test 4 P-value	0.600364435
D.O.F.	2

Model Parameters	
# of Parameters	4
Variable	Estimate
a	29.74458616
b	0.000191484
d	Bounded
log-alpha	1.042066246

Goodness of Fit

Dose	Size	Estimated Median	Calc'd Median	Observed Mean	Estimated SD	Calc'd SD	Observed SD	Scaled Residual
0	17	29.74458616	29.55	29.55	1.68376629	1.14	1.14	-0.47649088
35	21	29.94460185	30.21	30.21	1.68376629	1.99	1.99	0.722313504
175	17	30.75820529	30.56	30.56	1.68376629	1.84	1.84	-0.485353184
350	15	31.80636595	31.88	31.88	1.68376629	1.72	1.72	0.169372344

Likelihoods of Interest

Model	Log Likelihood*	# of Parameters	AIC
A1	-135.2877975	5	280.575595
A2	-132.4445224	8	280.889045
A3	-135.2877975	5	280.575595
fitted	-135.7980159	3	277.596032
R	-142.6419354	2	289.283871

* Includes additive constant of -64.3257. This constant was not included in the LL derivation prior to BMDS 3.0.

Tests of Interest

Test	-2*Log(Likelihood Ratio)	Test df	p-value
1	20.39482594	6	0.00235492
2	5.686550161	3	0.12789698
3	5.686550161	3	0.12789698
4	1.020436835	2	0.60036443

Table D-33. Benchmark dose results for delayed preputial separation in male mice—constant variance, BMR = 1 standard deviation (Das et al., 2008)

Models	Restriction	1 Standard deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Constant variance							
Exponential 2 (CV—normal)	Restricted	287.5467	201.6707	0.6004	277.5960	Viable—alternate	
Exponential 3 (CV—normal)	**Restricted**	**287.5612**	**201.6697**	**0.6004**	**277.5960**	**Viable—recommended**	**Lowest AIC**
Exponential 4 (CV—normal)	Restricted	286.3951	198.7931	0.3080	279.6149	Viable—alternate	
Exponential 5 (CV—normal)	Restricted	286.1679	197.6553	0.3076	279.6166	Viable—alternate	
Hill (CV—normal)	Restricted	201.3711	94.7311	0.2286	280.0252	Viable—alternate	
Polynomial (3 degree) (CV—normal)	Restricted	302.3780	199.5688	0.3427	279.4759	Viable—alternate	
Polynomial (2 degree) (CV—normal)	Restricted	297.6581	198.8516	0.3268	279.5372	Viable—alternate	
Power (CV—normal)	Restricted	286.2526	197.9759	0.5950	277.6140	Viable—alternate	
Linear (CV—normal)	Unrestricted	286.2531	197.9763	0.5950	277.6140	Viable—alternate	

D.10. RELATIVE LIVER WEIGHT—MALE HUMANIZED PPARA MICE (FOREMAN ET AL., 2009)

Table D-34. Dose-response data for relative liver weight in male humanized PPARα mice (Foreman et al., 2009)

Dose (mg/kg-d)	*n*	Mean	SD
0	10	4.07	0.261
35	10	5.62	0.719
175	10	6.65	0.784
350	10	7.38	0.719

Table D-35. Benchmark dose results for relative liver weight in male humanized PPARα mice —nonconstant variance, BMR = 10% relative deviation (Foreman et al., 2009)

Models	Restriction	10% Relative deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Nonconstant variance							
Exponential 2 (NCV—normal)	Restricted	77.3820	62.7400	<0.0001	107.4138	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2 Modeled control response SD >|1.5| actual response SD
Exponential 3 (NCV—normal)	Restricted	77.3912	62.7399	<0.0001	107.4138	Questionable	Goodness-of-fit *p*-value < 0.1 |Residual at control| >2 Modeled control response SD >|1.5| actual response SD
Exponential 4 (NCV—normal)	Restricted	6.7656	4.8076	0.0951	80.0462	Questionable	Goodness-of-fit *p*-value < 0.1 BMD 3× lower than lowest nonzero dose BMDL 3× lower than lowest nonzero dose
Exponential 5 (NCV—normal)	Restricted	6.7678	4.8076	0.0951	80.0462	Questionable	Goodness-of-fit *p*-value < 0.1 BMD 3× lower than lowest nonzero dose BMDL 3× lower than lowest nonzero dose
Hill (NCV—normal)	**Restricted**	**5.4945**	**4.4070**	**0.2883**	**78.3878**	**Viable—recommended**	**Lowest AIC BMD 3× lower than lowest nonzero dose BMDL 3× lower than lowest nonzero dose**
Polynomial (3 degree) (NCV—normal)	Restricted	59.5695	46.0032	<0.0001	104.4698	Questionable	Goodness-of-fit *p*-value < 0.1 |residual for dose group near BMD| >2 |residual at control| >2 Modeled control response SD >|1.5| actual response SD

Models	Restriction	10% Relative deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Nonconstant variance							
Polynomial (2 degree) (NCV—normal)	Restricted	59.5723	46.0033	<0.0001	104.4698	Questionable	Goodness-of-fit *p*-value < 0.1 |residual for dose group near BMD| >2 |residual at control| >2 Modeled control response SD >|1.5| actual response SD
Power (NCV—normal)	Restricted	59.5691	46.0034	<0.0001	104.4698	Questionable	Goodness-of-fit *p*-value < 0.1 |residual for dose group near BMD| >2 |residual at control| >2 Modeled control response SD >|1.5| actual response SD
Linear (NCV—normal)	Unrestricted	59.5725	46.0031	<0.0001	104.4698	Questionable	Goodness-of-fit *p*-value < 0.1 |residual for dose group near BMD| >2 |residual at control| >2 Modeled control response SD >|1.5| actual response SD

D.11. RELATIVE LIVER WEIGHT—MALE RATS EXPOSED 28 DAYS (BUTENHOFF ET AL., 2012; VAN OTTERDIJK, 2007B)

Table D-36. Dose-response data for relative liver weight in male rats following 28 day exposure (Butenhoff et al., 2012; van Otterdijk, 2007b)

Dose (mg/kg-d)	*n*	Mean	SD
0	10	2.42	0.17
1.2	10	2.55	0.25
6	10	3	0.33
30	10	3.59	0.46

Table D-37. Benchmark dose results for relative liver weight in male rats exposed 28 days—nonconstant variance, BMR = 10% relative deviation (Butenhoff et al., 2012; van Otterdijk, 2007b)

(Das et al., 2008) Models	Restriction	10% Relative deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Nonconstant variance							
Exponential 2 (NCV—normal)	Restricted	39.0522	30.9899	0.0010	30.9052	Questionable	Goodness of fit *p*-value < 0.1 |Residual for Dose Group Near BMD| > 2 Modeled control response std. dev. >|1.5| actual response std. dev.
Exponential 3 (NCV—normal)	Restricted	39.0519	30.9899	0.0010	30.9052	Questionable	Goodness of fit *p*-value < 0.1 |Residual for Dose Group Near BMD| > 2 Modeled control response std. dev. >|1.5| actual response std. dev.
Exponential 4 (NCV—normal)	**Restricted**	**9.9467**	**6.3433**	**0.9596**	**19.1475**	**Viable - Recommended**	**Lowest AIC**
Exponential 5 (NCV—normal)	Restricted	10.1350	6.3447	NA	21.1450	Questionable	d.f.=0, saturated model (Goodness of fit test cannot be calculated)

(Das et al., 2008) Models	Restriction	10% Relative deviation		*p*-Value	AIC	BMDS classification	BMDS notes
		BMD	BMDL				
Nonconstant variance							
Hill (NCV—normal)	Restricted	9.9219	5.3433	NA	21.1450	Questionable	d.f.=0, saturated model (Goodness of fit test cannot be calculated)
Polynomial (3 degree) (NCV—normal)	Restricted	31.8784	23.5467	0.0028	28.8760	Questionable	Goodness of fit *p*-value < 0.1 |Residual for Dose Group Near BMD| > 2
Polynomial (2 degree) (NCV—normal)	Restricted	31.8784	23.5468	0.0028	28.8760	Questionable	Goodness of fit *p*-value < 0.1 |Residual for Dose Group Near BMD| > 2
Power (NCV—normal)	Restricted	31.8784	23.5470	0.0028	28.8760	Questionable	Goodness of fit *p*-value < 0.1 |Residual for Dose Group Near BMD| > 2
Linear (NCV—normal)	Unrestricted	31.8784	23.5468	0.0028	28.8760	Questionable	Goodness of fit *p*-value < 0.1 |Residual for Dose Group Near BMD| > 2

APPENDIX E. SUMMARY OF PUBLIC AND EXTERNAL PEER REVIEW COMMENTS AND EPA'S DISPOSITION

The Toxicological Review of Perfluorobutanoic Acid and Related Salts was released for public comment in August 2021. Public comments on the assessment were submitted to the U.S. Environmental Protection Agency (EPA) by November 8, 2021. The Toxicological Review has also undergone a formal external peer review in accordance with U.S. Environmental Protection Agency (EPA) guidance on peer review (U.S. EPA, 2015). A public, external peer-review meeting was held February 22 and 23, 2022, which included another opportunity for public comment. The external peer reviewers were tasked with providing written answers to general questions on the overall assessment approach, key conclusions, and areas of scientific controversy or uncertainty. A summary of comments made by the external peer reviewers and public commenters, as well as EPA's responses to these comments, are arranged by charge question as follows. In many cases, the comments of the individual external reviewers have been synthesized and paraphrased for brevity (please consult the final peer review report for the full text of the panel's comments: Peer Review Report). External Peer Reviewers were asked to prioritize their comments to indicate their relative importance. The prioritization instructions are duplicated below from the IRIS PFBA charge questions to the peer reviewers, which can be found in the public EPA docket: EPA-HQ-ORD-2020-0675:

> Tier 1: *Recommended Revisions* – Key major recommendations necessary for strengthening the scientific basis for the Toxicological Review of PFBA. The implication of such key Tier 1 recommendations is that the assessment conclusions are not adequately supported without addressing the recommendations and need to be reconsidered or better substantiated. For Tier 1 recommendations, please describe the specific revisions necessary to modify or better substantiate the most scientifically appropriate assessment conclusions.
>
> Tier 2: *Suggestions* – Recommendations that are encouraged to strengthen the scientific analyses and conclusions in the Toxicological Review of PFBA. That other factor (e.g., timeliness) also may also be considered before deciding to address or incorporate Tier 2 suggestions is understood. For Tier 2 recommendations, please provide specific suggestions to strengthen the scientific basis for assessment conclusions or improve the clarity of the analyses and presentation.
>
> Tier 3: *Future Considerations* – Scientific exploration that might inform future work. These recommendations are outside the immediate scope or needs of the current document under review but could inform future toxicological reviews or research efforts

Appendix E lists all Tier 1 recommendations and Tier 2 Suggestions from the external peer reviewers organized by charge question. For Tier 3 Considerations, please refer to the external peer review report linked above. Where public comments were made on topics raised by the external peer reviewers, they are noted alongside the external peer review comments. All Tier 1 recommendations were implemented in this revised assessment, either through revision or addition to the peer reviewed analyses or text. Tier 2 suggestions were considered in light of the extent to which those suggestions would impact the conclusions or quantitative analyses of the assessment, consistency across panelists in raising the suggestion, and the level of effort to implement. For this assessment, all Tier 2 suggestions deemed to be impactful to the toxicity value conclusions were implemented in this revised assessment. Additional public comments not raised by the peer reviewers are included in a separate section at the end of each charge question section. Where possible, the public comments have been reproduced in this Appendix as they were submitted, but in some cases have synthesized and paraphrased for brevity. A summary document collating all public comments was provided as a courtesy to the external peer review panel. Please see docket EPA-HQ-ORD-2020-0675 for both this summary document and the full text of the submitted public comments.

External peer reviewer and public comments regarding requests for additions of clarifying text or editorial or grammatical corrections have been made throughout the assessment as appropriate; these comments and responses have not been tracked in this Appendix.

E.1. CHARGE QUESTION 1 – SYSTEMATIC REVIEW

The Toxicological Review describes and applies a systematic review process for identifying and screening pertinent studies that is described in detail in Section 1.2.1 (Literature Search and Screening) and Appendix A (Systematic Review Protocol for the PFBA, PFHxA, PFHxS, PFNA, and PFDA IRIS Assessments). Please comment on whether the search strategy and screening criteria for PFBA are appropriate and clearly described. Please identify additional peer-reviewed studies of PFBA that the assessment should incorporate[2]*.*

E.1.1. Overarching External Peer Reviewer Comments on Systematic Review

"All reviewers agreed that the literature search was well done, noting that it was comprehensive and that the methods used were appropriate and clearly described. They also stressed how challenging it is to conduct a thorough literature search in such a rapidly evolving field, where information may be out of date in a matter of months or even weeks."

[2] Newly identified studies (i.e., studies identified by EPA or the public that meet PECO criteria but were not addressed in the external review draft, for example due to recent publication) will be characterized by EPA and presented to the peer review panel. This characterization will focus on EPA's judgment of whether the studies would have a material impact on the conclusions (i.e., identified hazards or toxicity values) in the external review draft. The peer review panel is asked to review EPA's characterization and provide tiered recommendations to EPA regarding which studies, if any, to incorporate into the assessment before finalizing.

E.1.2. Tier 1 Recommendations

Comment: EPA should clarify when and how papers identified from the related systematic reviews for the other PFAS compounds were included in the PFBA toxicological review. EPA could provide this clarification by adding a small section and/or a table describing how the health effects text in the PFBA report was similar or was supported by the application of information from the review of related PFAS compounds.

EPA Response: Section 1.2.1 "Literature Search and Screening" describes the identification of studies during problem formulation, scoping and title/abstract screening for other PFAS that are relevant to the PFBA toxicological review. Specifically, some studies relevant to PFBA were identified by searches focused on the other four PFAS currently being assessed by the Integrated Risk Information System (IRIS) Program (i.e., PFHxA, PFHxS, PFNA, and PFDA) or from other authoritative reviews (e.g., final EPA reviews). This mostly applied to epidemiological studies as animal and mechanistic studies on specific PFAS are better indexed by specific PFAS. In addition, Table 4-2 has been added to Section 4.1 to demonstrate similarities and differences (contingent on the availability of data) in the health effects observed across the EPA PFAS human health assessments published at the time of finalization of the PFBA assessment. As EPA finalizes more PFAS assessments, this table will be expanded in subsequent IRIS assessments.

Comment: Multiple reviewers recommended EPA update the literature search to include the most up-to-date set of studies. Specifically, one reviewer recommended that EPA incorporate the Weatherly et al. (2021) study before finalizing the Toxicological Review. Multiple public comments were also received recommending that EPA update its literature search and incorporate relevant studies (including Weatherly et al., 2021). Public comments also recommended that EPA explicitly state the date of the last literature search used for the Toxicological Review.

EPA Response: The date of the last literature search used for the Toxicological Review (April 2021) was added to Section 2.1. Updates to the literature incorporated into the public comment draft (after the last literature update) are reflected in a separate document posted to the docket ("EPA-HQ-ORD-2020-0675-0022") and provided to the peer reviewers. This document describes the consideration of the studies deemed relevant based on the methods laid out in the protocol and documents the justification for the subset of those incorporated into the revised assessment. A specific charge question was posed to the peer reviewers on these decisions and no disagreements were noted in the panel's final comments. The Weatherly et al. (2021) study was added to the assessment given it was specifically identified by the peer review panel (see Sections 3.2.2, 3.2.5, and 5.2.1).

Comment: To improve access to the studies identified, one reviewer recommended that EPA (1) develop a simple table explicitly listing all the in vitro, in silico, or nonmammalian model "supplemental material" studies that were considered and selected, and (2) develop a simple table listing all the studies that were considered but not selected, that also briefly identifies the reasons for rejecting each of these studies. Public comments were also received requesting EPA make available the lists of included and supplemental studies and to ensure that the list of studies in HAWC are accurate.

EPA Response: The included, excluded and supplemental studies can be found in HERO (https://hero.epa.gov/hero/index.cfm/project/page/search/true/isws/false/project_id/2632/). With respect to inclusion or exclusion, studies are excluded if they do not meet all PECO criteria. During screening, most studies are excluded because they do not meet any or only meet a few of the PECO criteria. Thus, a single screened out study typically has multiple reasons for exclusion which is unwieldy to document, especially at the title and abstract level when screening may be needed for thousands of studies. The annotation used in the assessment is consistent with the convention in systematic review (Page et al., 2021). Note also that multiple tags may be applied to a single study (e.g., tagged "supplemental" during title/abstract and "in vitro" during full-text screening) which results in potential discrepancies when cross referencing numbers between HERO and the literature flow diagrams or HAWC study evaluation heatmaps.

E.1.3. Tier 2 Suggestions

Comment: EPA should add an introductory preview on how it is approaching PFBA in relation to other forms of PFAS that have been much more extensively studied.

EPA Response: Text was added to Section 1.2.1 "Literature Search and Screening" explaining that relevant literature on PFBA was identified by searches focused on the other four PFAS currently being assessed by the Integrated Risk Information System (IRIS) Program (i.e., PFHxA, PFHxS, PFNA, and PFDA) or from other authoritative reviews (e.g., final EPA reviews). Table 4-2 has been added to Section 4.1 to demonstrate similarities and differences (contingent on the availability of data) in the health effects observed across the EPA PFAS human health risk assessments published at the time of finalization of the PFBA assessment. As EPA finalizes more PFAS assessments, this table will be expanded in subsequent IRIS assessments.

Comment: One reviewer suggested, as an enhancement to the added table (listing all the in vitro, in silico, or nonmammalian model "supplemental material" studies) recommended in the Tier 1 recommendation above, EPA should incorporate columns (a) summarizing qualitatively the confidence (low, medium, high) associated with the information presented

in each study, and (b) listing potential outcomes associated with the "supplemental material" studies.

EPA Response: Confidence in studies meeting PECO criteria is documented in HAWC (https://hawc.epa.gov/study/assessment/100500073/), and several additions were made in this revised assessment to ensure that the health outcome-specific confidence for these studies is conveyed in the figures or text of the relevant synthesis sections. Studies that are identified as potentially relevant supplemental material (e.g., mechanistic, pharmacokinetic) can be found in HERO.

Comment: Numerical inconsistencies in the number of studies listed in Figures 2-1, 2-2, and 2-3, the number of studies discussed in the text of the Toxicological Review, and the number of studies listed in HAWC should be corrected (a Tier 1 recommendation on this topic was also provided under Charge Question 2 below). Public comments were also received recommending that EPA ensure that all numbers of studies are properly reported within the document, figures, tables, and associated meta-data.

EPA Response: Eight epidemiology studies and seven animal studies were identified as meeting the PECO criteria following full-text review. Nine animal studies are listed in Figure 2-1 which includes the Butenhoff et al. (2012) study which reported the findings of two unpublished industry reports: a 28-day and 90-day gavage study fully reported in (van Otterdijk, 2007a, b). These industry reports were conducted at the same facility and largely by the same staff but independently of one another and at different times: July 26, 2006, through September 15, 2006, for the 28-day study and April 5, 2007, through August 6, 2007, for the 90-day study. Throughout the Toxicological Review, both (Butenhoff et al., 2012) and the relevant industry report are cited when discussing effects observed in these reports. However, only one study evaluation was performed for this group of citations in HAWC (see Figure 2-2), the overall confidence level of high applies to both the 28-day and 90-day reports and grouping of these studies accounts for the discrepancy between the number of animal studies in Figures 2-1 and 2-2.

Comment: EPA could add a statement about what kind of information would be required to change the overall analysis/conclusions, with a clearer description of when updates will be made.

EPA Response: With respect to information to change overall analysis or conclusions, this is implicit in the evidence synthesis and integration analysis. For example, conclusions of "evidence inadequate" are reached after describing specific limitations to the evidence base. These limitations can be translated by the research community into information gaps that, if filled, could potentially change an overall analysis or conclusion. Presentation of uncertainty

factors is another area in the IRIS assessment that provides an indication of information that could be impactful to change an overall analysis/conclusion. Once the IRIS Toxicological Review of PFBA is finalized, the IRIS program has no immediate plans on updating the assessment. Given the finite resources of the IRIS Program, IRIS assessment activities are based on the priority needs of EPA National Program and Regional Offices identified through a structured internal (to EPA) nomination process.

Comment: EPA could consider providing a brief discussion of what is known (and not known) to help inform animal-to-human extrapolation. For example, the relevance of PPAR as a mode of action is an important point and the degree to which it is or is not relevant to humans could be mentioned at the outset. Similarly, the dramatic sex differences in some rodents are clearly not applicable to humans for other forms of PFAS and presumably not for PFBA either.

EPA Response: The human relevance of the animal data is explicitly addressed within the context of the evidence available to inform each individual hazard. However, without specific evidence to the contrary, effects in animals are presumed relevant to humans (U.S. EPA, 2005, 1998, 1991). Once a determination is made that an effect is considered relevant to humans using the currently available evidence, the quantitative implications of the remaining uncertainties in extrapolation are addressed through dosimetric adjustment and application of the UF_A during dose-response analysis.

Comment: EPA could consider adding a section that discusses available information on PFBA's potential immunomodulation (immunosuppression) effects. Existing studies most probably cannot support derivation of relevant reference values, but compilation and evaluation of the available information can provide an initial framework for addressing this challenge in future revisions.

EPA Response: A discussion of potential immunomodulation was added to Section 3.2.5 "Other Non-Health Effects."

Comment: PFAS information submitted to IRIS should be available to all EPA programs and vice versa.

EPA Response: IRIS assessments rely on publicly available information in the published literature and can potentially include information submitted to EPA programs (e.g., TSCA), if those data can be made publicly available.

Comment: EPA should correct the PECO element Exposures statement that incorrectly suggests that the 6:2fluorotelomer is metabolized into two analytes; it is in fact multiple analytes.

EPA Response: The IRIS PFAS protocol states (text of interest underlined) the following: "[Note: although while these PFAS are not metabolized or transformed in the body, there are precursor compounds known to be biotransformed to a PFAS of interest; for example, 6:2 fluorotelomer alcohol is metabolized to PFHxA and PFBA (Russell et al., 2015). Thus, studies of precursor PFAS that identify and quantify a PFAS of interest will be tracked as potential supplemental material (e.g., for ADME analyses or interpretations)]". This text does not preclude that this compound can be metabolized to other analytes; it is simply emphasizing those analytes of interest to the protocol, namely PFBA, PFHxA, PFHxS, PFNA, or PFDA. However, a small editorial change to the text was made for clarity.

E.1.4. Public Comments

Comment: EPA improperly excluded its own relevant studies (i.e., Das et al., 2008) in developing the draft IRIS review. In the Draft IRIS Review, EPA made numerous comparisons between PFBA, a four carbon perfluoroalkyl carboxylate, and perfluorobutanesulfonate (PFBS), a four-carbon perfluoroalkyl sulfonate congener. While both PFBA and PFBS have generally been considered short-chain PFAS compounds given the relatively short serum elimination half-lives in the species evaluated (rodents, non-human primates, and humans), EPA leaned too heavily on this similarity and ignored relevant data available for PFBA itself.

EPA Response: This comment implies that the draft assessment did not include the Das et al. (2008) mouse developmental toxicity study in the PBFA assessment. This is incorrect. Section 5 indicates that multiple effects from Das et al. (2008) were modeled and considered for RfD derivation.

Additionally, the consideration of data regarding the toxicological effects of PFBS in the PFBA Toxicological Review is consistent with methods described in Appendix A (Protocol) supporting the consideration of data on similar chemicals to inform PFBA-specific data gaps. Thus, PFBS data is described in the PFBA assessment in cases where data was lacking for PFBA or when drawing parallels between chemicals was useful in discussing potentially consistent toxicological effects across PFAS. That said, PFBS data were not used in the PFBA assessment for quantitative purposes or toxicity value derivation, nor were they necessary to draw the evidence integration conclusions regarding PFBA, which were sufficiently supported by PFBA-specific data.

Comment: In analyzing the potential health effects of PFBA, the draft assessment makes several comparisons to data available for other PFAS. While such a "read-across" approach can be a useful in qualitatively assessing the potential for a compound to impact health endpoints, it is important that the comparison be made to compounds that are as structurally similar to the compound of interest as possible and, when appropriate, to indicate how structural differences may impact ability to compare the toxicokinetics, target organs, critical effects, potential molecular targets, and shapes of the dose-response curves. In the case of PFBA, other short-chain (<8 carbons) carboxylates are the most appropriate for comparison. Consequently, comparison to perfluorohexanoic acid (PFHxA) would be more appropriate than to a sulfonic acid like perfluorobutyl sulfonate (PFBS). In this regard, the lack of observed significant developmental effects associated with PFHxA is noted.

EPA Response: As noted above, the IRIS PFBA assessment considers the data available on other PFAS to inform PFBA-specific data gaps, with an emphasis on and increased use of data and judgments on PFAS for which EPA has available a final assessment at the time of developing the PFBA assessment. It is important to emphasize that EPA has not reached a final conclusion on whether the evidence supports a "lack of observed significant developmental effects associated with PFHxA." EPA's IRIS PFHxA assessment is still under development.

Comment: EPA should provide further clarification and better reporting when multiple publications of the same data are included. For example, the studies reported as van Otterdijk 2007c and van Otterdijk 2007d are industry documents available in EPA's HERO database but have also been published in the peer reviewed literature in the study by Butenhoff et al. 2012. That these studies contain overlapping and duplicative data, should be more clearly noted in the literature flow diagram (Figure 2-1) and the discussion of Study Evaluation Results in Section 2.2.

EPA Response: This is noted in Section 3.2.1 "Non-Cancer Evidence Synthesis and Integration" (footnote 11). Additional clarification was also added to Study Evaluation Results in Section 2.2: "The studies meeting PECO criteria at the full-text level included eight epidemiological studies, nine animal studies (including one published study [Butenhoff et al. (2012)] that reported on the same data in two unpublished industry reports [(van Otterdijk, 2007a) and (van Otterdijk, 2007b)]."

E.2. CHARGE QUESTION 2 – STUDY EVALUATION

The Toxicological Review describes the results of the evaluations of individual studies in Section 2.2 (Study Evaluation Results) and presents and analyzes the findings from those studies deemed informative in the relevant health effect-specific synthesis sections.

a. *Please comment on whether the study confidence conclusions for the PFBA studies are scientifically justified, giving appropriate consideration to important methodological features of the assessed outcomes. Please specify any study confidence conclusions that are not justified and explain any alternative study evaluation decisions.*

b. *Results from individual PFBA studies are presented and synthesized in the health system-specific sections. Please comment on whether the presentation and analysis of study results is clear, appropriate, and effective to allow for scientifically supported syntheses of the findings across sets of studies.*

E.2.1. Overarching External Peer Reviewer Comments on Study Evaluation

"Reviewers agreed that the confidence conclusions were scientifically justified, and that Section 2.2 was well done. Reviewers noted that the scientific justification presented was clear and effective, and found that the interactive visualizations provided a convenient overview. [One reviewer] also found it very beneficial that EPA presented the logic for giving more or less attention to particular studies and specific outcomes."

E.2.2. Tier 1 Recommendations

Comment: EPA should correct numerical inconsistencies in the number of studies listed in the Figures 2-1, 2-2, and 2-3, the number of studies discussed in the text of the Toxicological Review, and the number of studies listed in HAWC. (a Tier 2 suggestion on this topic was also provided under Charge Question 1 above).

EPA Response: See EPA Response to Charge Question 1 above and referenced in footnote 11 of the Toxicological Review which states "Eight epidemiology studies and seven animal studies were identified as PECO relevant following full-text review. Nine animal studies are listed in Figures 2-1 which includes the Butenhoff et al. (2012) study which reported the findings of two unpublished industry reports: a 28-day and 90-day gavage study fully reported in (van Otterdijk, 2007a, b). These industry reports were conducted at the same facility and largely by the same staff but independently of one another and at different times: July 26, 2006, through September 15, 2006, for the 28-day study and April 5, 2007, through August 6, 2007, for the 90-day study. Throughout the Toxicological Review, both Butenhoff et al. (2012) and the relevant industry report are cited when discussing effects observed in these reports. However, only one study evaluation was performed for this group of citations in HAWC (see Figure 2-2), the overall confidence level of *high* applies to both the 28-day and 90-day reports and accounts for the discrepancy between nine animal studies in Figures 2-1 and 2-2."

E.2.3. Tier 2 Suggestions

Comment: Comparison of effects in males and females is a common theme in the document and sometimes tentatively related to the pharmacokinetic differences of PFBA in females

(more rapid clearance) than males (slower clearance). EPA could select an administered dose in males that is nearly equivalent to females based on pharmacokinetic dose metrics and compare toxicity outcomes, as it may provide a clearer picture of the role that pharmacokinetics plays in toxicity in male and female lab animals.

EPA Response: A brief comparison and discussion of pharmacokinetic parameters and relative responses for liver and thyroid endpoints at a constant dose has been added to Section 4.3 (Conclusions Regarding Susceptible Populations and Lifestages).

Comment: While the information on the available epidemiologic studies is provided in the HAWC table, EPA could provide text that briefly notes the nature of the studies in general, and since these studies are so few in number, a sentence or two about each. It could be as little as 2-3 sentences that gives a synopsis of what was done (the PECO attributes), key methodologic limitations, and overall assignment regarding its quality. This is editorial in nature but would help make this a more transparent, user-friendly document. Those who want the full details could look to HAWC and appendices as needed.

EPA Response: Confidence in studies meeting PECO criteria is documented in HAWC (https://hawc.epa.gov/assessment/100500073/), and several additions were made in this revised assessment to ensure that the health outcome-specific confidence for these studies is conveyed in the figures or text of the relevant synthesis sections. Because study evaluations in IRIS are outcome-specific, the revised draft assessment does not include text summarizing the studies and their strengths/limitations in an outcome-nonspecific manner. Further, the health outcome-specific syntheses attempt to distill the available (epidemiological or other) evidence to those aspects of the studies (e.g., design; confidence) most pertinent to drawing hazard judgments; they intentionally try to avoid study-by-study summaries. This focus on developing concise IRIS assessments is based on feedback from reviewers over many years.

Comment: The Toxicological Review states that most of the animal studies evaluated for study confidence were adequate, but the specific rationale for the analyses was unclear because the interactive HAWC link did not work. Ultimately it was unclear how the study confidence conclusions were determined, other than professional judgement. This chapter could be improved by writing more about the overall confidence conclusions in the chapter, even if the information is found in links.

EPA Response: The HAWC site is now public (https://hawc.epa.gov/assessment/100500073/). Links were checked and fixed throughout the revised assessment, and the study confidence conclusions are detailed there. Please see responses above about text additions relating to study confidence ratings.

E.3. CHARGE QUESTIONS 3 AND 4 – HEPATIC EFFECTS

For each health effect considered in the assessment and outlined below, please comment on whether the available data have been clearly and appropriately synthesized to describe the strengths and limitations. For each, please also comment on whether the weight-of-evidence decisions for hazard identification have been clearly described and scientifically justified.

- *For hepatic effects, the Toxicological Review concludes that the available evidence indicates PFBA exposure is likely to cause hepatic effects in humans given relevant exposure circumstances, on the basis of a series of short-term, subchronic, and developmental studies in rats and mice demonstrating consistent and coherent effects with a clear biological gradient. Although the available mechanistic information indicates the effects in rodents are relevant to humans, some uncertainty remains regarding potential differences in sensitivity across species due to evidence for the involvement of both PPARα-dependent and PPARα-independent pathways in these effects (see Charge Question 4 requesting input specific to this latter uncertainty).*

Appendix A (Systematic Review Protocol for the PFBA PFHxA, PFHxS, PFNA, and PFDA IRIS Assessments) identifies the human relevance of hepatic effects in animals that involve peroxisome proliferator-activated receptor alpha (PPARα) receptors as a key science issue. To the extent supported by the PFBA literature (and to a lesser extent, literature for other PFAS), the Toxicological Review evaluates the evidence relevant to the potential involvement of PPARα and non-PPARα pathways with respect to the reported hepatic effects. The Toxicological Review ultimately concludes evidence from in vivo and in vitro studies support that multiple modes of action (MOA) are operant in the induction of hepatic effects by PFBA exposure and the relative contribution of these different MOAs cannot be concluded with confidence from the available data. Please comment on whether the available animal and mechanistic studies support this conclusion and whether the analysis presented in the Toxicological Review is clearly documented.

E.3.1. Overarching External Peer Reviewer Comments on Hepatic Effects

"All reviewers agreed that the PFBA document clearly and appropriately synthesizes available data to describe the strengths and limitations of hepatic effects. [One reviewer] commented that "The evidence integration section is well done and supported by a great summary in Table 3-8." All reviewers also found the document supported the conclusion that MOAs are operant in the induction of hepatic effects by PFBA exposures, and the relative contribution of these different MOAs cannot be concluded with confidence. [The same reviewer] commented that "EPA did a great job in describing the complications" of the data, and another stated that "the available animal and mechanistic studies are clearly documented, and the conclusion is supported.""

E.3.2. Tier 1 Recommendations

Reviewers had no Tier 1 recommendations.

E.3.3. Tier 2 Suggestions

Comment: EPA could consider adding tables and/or figures that would help readers visualize important EPA conclusions, such as coherence of liver histopathology with liver

weight effects (cited in Table 3-8), since these results are only presented in separate tables within the document (i.e., Table 3-7/Figure 3-6 for histopathology and Table 3-6/Figure 3-5 for liver weight changes).

EPA Response: A data pivot chart has been created in HAWC displaying liver histopathology and liver weight effects together for ease of assessing the consistency and coherence of liver effects across species, sexes, durations of exposure, and study types. A link to this chart has been added to Section 3.2.2 of the assessment.

Comment: The EPA should resolve an apparent discrepancy where statements regarding serum biomarker data are incoherent. Observations for serum biomarkers of altered liver function or injury appear as a factor that decreases certainty in Table 3-8 (p. 3-40) which contradicts statements in Section 3 that state that the inconsistent serum biomarker results did not influence the evidence integration judgements.

EPA Response: The statement in Table 3-8 those incoherent observations across serum biomarkers (e.g., increased ALP but decreased bilirubin) decreased certainty is in reference to the certainty in the evidence for the serum biomarkers specifically. For the overall evidence integration judgment about hepatic effects, the noted incoherence across some of the biomarkers findings was not influential, but rather this judgment was based on the strong evidence of consistent and coherent effects on liver weights and liver histopathology. The text in Section 3 has been revised to provide more clarity on exactly how the serum biomarker evidence was used in drawing the overall evidence integration judgment.

Comment: EPA should explicitly state the meaning of "consistent effects" in the sentence: "The available animal evidence for effects on the liver includes multiple high and medium confidence studies with consistent effects across multiple species, sexes, exposure durations, and study designs…". This phrase could have several meanings (e.g., all the same effects occurred at the same or similar doses across multiple species, sexes, exposure durations, and study designs; one or more hepatic effects were consistently found at he some dose across studies though the specific effects observed may vary across studies) and different readers can/will interpret it differently.

EPA Response: As described in the protocol, 'consistency' generally relates to findings for a given outcome (e.g., liver weight), while 'coherence' reflects the observed findings across related (e.g., through biological understanding or MOA) outcomes, such as effect on liver weight, and separately, histopathology, within individual studies or across multiple studies. Further, consistency is judged based on the pattern of findings across studies and comparisons and can range from consistency in the direction of the response to something more specific such as consistency in the magnitude of change in response at a given

exposure level (while the latter can provide stronger evidence of consistency, it is neither expected nor required). In the context noted in the comment, the summary statement about consistency in the evidence integration section reflects the analyses presented in detail within the preceding synthesis sections on individual liver outcomes. An edit was made to clarify consistency refers to the findings of increased liver weight and, separately, increased liver histopathology incidence across the available studies.

Comment: EPA should consider differences in metabolic pathways between species when comparing rodent to human exposures. The data on exposure obtained from the mouse used a humanized mouse model. Caution is always required when using data from a "humanized model" when the humanized model is limited to a single gene replacement. As noted in the public comments, only a single nuclear receptor was humanized (PPARα) and that there are other nuclear receptors exist in the rat that can be induced and may lead to hepatocellular hypertrophy. This would not necessarily lead to a similar effect in humans at similar doses. This Tier 2 comment was also provided under Charge Question 6; responses to both instances of this Tier 2 comment are provided here for brevity's sake.

EPA Response: Due to reported cross-species differences in PPARα signaling potency and dynamics, the potential human relevance of some hepatic effects has been questioned. Thus, the Foreman et al. (2009) study is informative in providing evidence on the relative contribution that PPARα has on PFBA-induced liver effects. While true that only PPARα was humanized in the Foreman et al. (2009) study, given the response in the humanized mice, including mouse PPARα–independent increases in hepatocellular hypertrophy and vacuolation, Foreman et al. (2009) provides evidence that rodent PPARα is not necessary for PFBA to induce some liver effects. The assessment indicates the data from Foreman et al. (2009) is largely used in qualitative analyses; thus, no claim is made that similar effects, whether in type or magnitude, would be observed in humans as in rats at similar doses. The mechanistic section discusses the activation and human relevance of other nuclear receptors (PXR, CAR) that might also contribute to the hepatic effects of PFBA.

Comment: The one issue that may call for more comment concerns the absence of evidence that PFBA affects liver enzymes (ALT, etc.) because that is the one human health endpoint that has consistently been found to be associated with other forms of PFAS. EPA should consider whether that contrast between an absence of an effect for PFBA in rodents and an apparent effect for PFOA and PFOS for humans has relevance to the final judgment.

EPA Response: The IRIS assessment does not conclude there is no evidence that PFBA exposure affects liver enzymes. The currently available data for PBFA on this endpoint are inconclusive. Specifically, only one epidemiology study was available and, although this study did not find any associations between serum biomarker levels and PFBA exposure, it

was considered *deficient* with respect to sensitivity and likely biased towards the null. Additionally, while the two animal studies were also ultimately inconclusive with regard to ALT and other enzymes, some effects on related biomarkers were observed. Specifically, while no effects were observed after 28 days of exposure, changes in several serum biomarkers potentially indicative of liver injury (e.g., increased ALP) were observed after 90 days of PFBA exposure, along with some incoherent findings (e.g., changes in the opposite direction as expected for bilirubin). Although uncertain, the serum biomarker data specifically were inconclusive and thus not necessarily in contrast to the findings for PFOA and PFOS. Regardless, in the case of PFBA, there is sufficient evidence on liver weight increases and increased histopathological lesions, as well as data informing mode-of-action, that the serum biomarker data, including if one were to more explicitly consider the serum biomarker data for other PFAS, would not change the overall evidence integration judgment of "evidence indicates (likely)" for PFBA.

Comment: Because the entire Toxicological Report rests on animal-to-human extrapolation, to the extent that there are mechanisms in animals *known* not to apply to humans, this should be explained and factored into the report.

EPA Response: A primary purpose of the ***Mechanistic Evidence and Supplemental Information*** section for hepatic effects is to evaluate the currently available mechanistic evidence informing whether the hepatic effects observed in rodents are relevant to human health (and thus suitable as the basis for reference value derivations). It is unclear what mechanisms in animals the commenter is referring to as "known" not to apply to humans, but given the focus on PPARα activation for PFAS, this is the mechanism of interest this response assumes. However, to clarify, activation of PPARα by other compounds can contribute to toxicity in humans and furthermore PFBA appears to be a ligand for PPARα in humans, given the data from the humanized PPARα mice and the observed interaction between PFBA and the human receptor in vitro. The concern is rather one of a differential response magnitude based on the presumption that hepatic responses to PPARα activation are exaggerated in rodents as compared to humans and thus may generally be more difficult to interpret as relating to a significant change in hepatic effects in most human exposure scenarios. Focusing on this concern, a key science issue addressed by the section noted above is whether the available data are sufficient to support dependence on PPARα activation for the observed hepatic effects of PFBA, or whether it is possible that the hepatic effects can also be mediated through non-PPARα modes-of-action. The other potential mechanisms discussed in this section are also considered in light of the available evidence on their human relevance, with a general assumption that animal mechanisms are relevant to humans without data to the contrary. The final conclusion, which is also supported in the peer review comments, is that PFBA's liver effects appear to be mediated through both

PPARα-dependent and -independent pathways and that the observed liver effects are relevant to human health. As this is already extensively discussed and explained (similarly to the explanations in the above response) in the peer reviewed draft, additional discussion on this topic was not added to this revised assessment.

E.3.4. Public Comments

Comment: EPA improperly characterizes liver hypertrophy in rats as an adverse effect of PFBA exposure. This conclusion is not consistent with EPA's own guidance, which states that reported liver hypertrophy is not an adverse effect. As noted in the EPA Office of Pesticide Programs Health Effects Division Guidance Document # G2002.01 on Hepatocellular Hypertrophy (U.S.EPA, 2002), liver hypertrophy does not necessarily represent liver toxicity, nor is it necessarily a precursor to a particular outcome of toxicity. In fact, Guidance Document # G2002.01 recommends a weight-of-evidence approach for determining whether a liver effect should be characterized as adverse. This approach includes evaluation of other findings, such as: (1) type and severity of observed effects; (2) onset, duration, and progression of effects; (3) study method and design; and (4) other relevant effects and data. This guidance states that liver size or weight changes may be "indicative of adaptation which, by itself, is not necessarily adverse."

EPA Response: The HED guidance document does not state that "liver hypertrophy is not an adverse effect" but rather, as the comment itself points out "[h]epatocellular hypertrophy (and its corresponding increased liver size/weight) may be indicative of adaptation which, by itself, is not necessarily adverse. *However, it might be associated with other more severe changes* [emphasis added]. These changes are usually accompanied by alterations in relevant clinical chemistry parameters and/or histopathology." The HED guidance document recommends a weight-of-evidence approach to determine whether hypertrophy and related liver weight changes are considered to reflect an adverse hepatic response. Consistent with the protocol, to judge the adversity of the observed liver effects, the PFBA IRIS assessment considered the panel recommendations outlined by Hall et al. (2012) and the HED 2002 Guidance document. Ultimately, using this paradigm, the assessment concluded "application of the recommendations from Hall et al. (2012) supports the conclusion that the multiple and interconnected effects observed in the livers of exposed animals meet the criteria for adversity" (i.e., the conclusion of adversity is not based on liver hypertrophy alone). This conclusion was unanimously supported by the external peer review panel as noted in the final Peer Review Report: "[a]ll reviewers agreed that the PFBA document clearly and appropriately synthesizes available data to describe the strengths and limitations of hepatic effects."

Comment: The reported liver hypertrophy in rats with PFBA exposure is an adaptive response occurring through increased activation of PPARα and CAR nuclear receptors. The activation of PPARα and CAR can lead to expansion of the smooth endoplasmic reticulum in the liver cell, and this added intracellular mass is reflected in the overall liver weight macroscopically and hypertrophy microscopically. Another consequence of PPARα and CAR/PXR activation in rodents is the potential stimulation of cell division (hyperplasia) and a decrease in the normal process of removal of worn-out cells (apoptosis). These processes also increase liver mass and can potentially lead to tumor formation in rodents. However, it is important to note that the hyperplastic processes are not present in mice with humanized PPARα expression (Foreman et al., 2009). Therefore, the processes that lead to tumor formation in rats as a result of PFBA exposure are not applicable to humans. In addition, Foreman et al. also demonstrated that activation of the human form of PPARα with PFBA does not produce frank liver toxicity. Moreover, Bjork and Wallace (2009) demonstrated that human hepatocytes did not respond to PFBA-induced PPARα activation at concentrations up to 200 μM (42,600 ppb); rat hepatocytes responded at PFBA concentration of 25 μM (5,325 ppb) and above.

EPA Response: Please see responses above relating to the adversity and human relevance of the hepatic changes observed in rodents following PFBA exposure, noting that the assessment does not draw a conclusion regarding the potential for PFBA exposure to lead to tumor formation (i.e., the available carcinogenicity evidence was judged as inadequate).

E.4. CHARGE QUESTION 3 – THYROID EFFECTS

For each health effect considered in the assessment and outlined below, please comment on whether the available data have been clearly and appropriately synthesized to describe the strengths and limitations. For each, please also comment on whether the weight-of-evidence decisions for hazard identification have been clearly described and scientifically justified.

- *For thyroid effects, the Toxicological Review concludes that the available evidence indicates PFBA exposure is likely to cause thyroid toxicity in humans given relevant exposure circumstances, primarily on the basis of short-term and subchronic studies in male rats reporting a consistent and coherent pattern of thyroid effects following PFBA exposure, but also drawing from the consistency of effects when considering evidence from structurally related PFAS. The Toxicological Review concludes the thyroid effects are considered relevant to humans in the absence of evidence to suggest otherwise.*

E.4.1. Overarching External Peer Reviewer Comments on Thyroid Effects

"Six of the seven reviewers found the conclusions for thyroid effects to be clearly described and scientifically justified. For example, [One reviewer] wrote: "The data leading to the conclusion

that the thyroid effects are considered relevant to humans were appropriately synthesized and their strengths and limitations were adequately described; the weight-of-evidence decisions for hazard identification have been adequately described; the weight-of-evidence decisions for hazard identification have been adequately described and justified." The seventh reviewer commented that there was a "lack of clarity on human exposures.""

E.4.2. Tier 1 Recommendations

Comment: One reviewer recommended that EPA provide a table presenting lab animal and human citations and data that show connections between a percent decrease in serum T4 and adverse outcome, relating percent change with adversity in humans especially. This reviewer noted that, "This argument needs to be solid and presenting data will help" to relate thyroid effects in rats and humans quantitatively.

EPA Response: To our knowledge, there is only limited information from human studies that demonstrate what percent decrease in T4 leads to adverse outcomes such as neurodevelopmental outcomes (note: these outcomes are the focus of this response and discussions in the assessment since these associations are the best studied). This is mainly due to the nature of epidemiological studies, typically with representative samples analyzed post hoc; many also bin data by "hypothyroid, euthyroid, hypothyroxinemic" based on reference ranges, and then correlate to adverse outcomes. There are a few human studies Jansen et al. (2019); Levie et al. (2018); Korevaar et al. (2016) where the sample sizes are large enough to capture a wide range of TSH and/or T4 values, which were then correlated to various neurodevelopmental outcomes that could be quantified. However, these studies still do not make direct comparisons from a percent decrease in hormones that would lead to an adverse effect; rather, they stratify their hormone samples by standard deviation to the mean/median, quartiles, etc. Therefore, it's difficult to make a conclusion in humans regarding what percent of hormone dysfunction is adverse, as those kinds of data are not generated. Additionally, there are no conclusive values from animal studies regarding to what degree of T4 reduction is adverse. This is due to several factors, including the existence of multiple thyroid-dependent processes in the brain, which likely have differing spatiotemporal sensitivities. But there are studies that show how graded reductions in T4 can lead to neuronal heterotopia Gilbert et al. (2014), synaptic transmission defects Gilbert and Sui (2008) and differential gene expression. When unable to estimate the BMR based on an evidentiary approach, the EPA takes a default approach as outlined in the Benchmark Dose Technical Guidance (U.S. EPA, 2012). Given the limited and inconsistent information available from human and animal studies, the assessment uses a standard deviation definition for the BMR (see Toxicological Review, Section 5.2.1 for justification for using BMRs). The Mechanistic Support and Supplemental Information section for thyroid effects

has been revised to more thoroughly discuss the nature of the available human and animal evidence.

E.4.3. Tier 2 Suggestions

Comment: Related to the Tier 1 recommendation above, the same reviewer re-iterated that the thyroid findings in rats are relevant to humans but proposing how to relate rat thyroid effects to humans in a quantitative manner (dose-response) is the challenge. This reviewer suggested a brief discussion on this topic would be useful.

EPA Response: Please see the responses provided to the Tier 1 Recommendation above as well as the public comment on this topic below.

Comment: One reviewer noted that, although Butenhoff et al. (2012) do not report statistically significant changes in serum TSH related to PFBA exposure, the coefficient of variation for TSH measurements in controls ranges from 40% to 72% in the two studies this group reports. The reviewer suggested that this means that detecting relatively small changes in serum TSH is difficult with this assay assembled from the reagents provided by the NIH Pituitary Program and that it is more likely, as pointed out in the Agency report, that the effect of PFBA exposure on thyroid histology reflects an increase in serum TSH that was not detected in this assay. This reviewer suggested the EPA consider evaluating the issue of the TSH assay using the reagents characterized in the Butenhoff studies and rather than emphasize the comparison with "hypothyroxinemia" (a clinical term that doesn't translate perfectly to animal studies), describe the situation as one in which serum T4 is low and TSH levels have increased.

EPA Response: The description of PFBA exposure on thyroid effects now includes discussion of the potential lack of detectable increases in serum TSH levels with corresponding decreased T4. Lack of detectable increase in serum TSH levels may be related to the specific assay utilized in the Butenhoff et al. (2012) study, and text has been added to the assessment noting this possibility.

Comment: One reviewer suggested that it is important to discriminate between "thyroid function" and "thyroid hormone action." This reviewer noted that evidence presented in Butenhoff et al. (2012) shows that PFBA decreases both serum total and free T_4 while also showing an increase in thyroid hormone action in the liver. This reviewer further notes that this scenario is highly reminiscent of the effects of PCB and PBDE exposure on TH signaling in which measures of TH action indicate an increase in TH action in the liver despite a reduction in serum total and free T_4 (Giera et al., 2011; Bansal et al., 2014). In the case of PCBs, this pattern of effects on serum hormones and gene expression in the liver was coincident with a complex pattern of effects on thyroid hormone action in brain (Zoeller et

al., 2000; Bansal and Zoeller, 2008). Finally, this reviewer noted that Butenhoff et al. 2012 reports that PFBA exposure also decreases serum cholesterol (Table 7), and thyroid hormone is known to reduce serum cholesterol in rodents and humans (Mullur et al., 2014). Ultimately this reviewer suggested that "[a]lthough paradoxical … that PFBA reduces serum total and free T4 while the liver appears to be responding to increased thyroid hormone action, it is recommended that the Agency incorporate all measured endpoints of thyroid hormone action in their analysis."

EPA Response: A brief description of reduced thyroid function (decreased serum total and free T_4) and increased thyroid hormone action in the liver (decreased serum cholesterol) has been added to Section 3.2.1 (Thyroid Effects).

Comment: One reviewer agreed that decreases in serum total and free T_4 are very clearly associated with developmental and physiological deficits: "In both humans and experimental animals, low TH is related to permanent neural and cognitive deficits (e.g., (Zoeller and Rovet, 2004; Rovet, 2014; Stagnaro-Green and Rovet, 2016a). This is likely to be true for several (if not all) organs including heart, bone, lung and intestine (Bizzarro and Gross, 2004; Bassett et al., 2007; Mochizuki et al., 2007; Wexler and Sharretts, 2007). Much of the experimental literature on this topic makes use of models of severe hypothyroidism (Crofton et al., 2005; Crofton and Zoeller, 2005). However, graded effects of thyroid hormone insufficiency have become the focus of increased attention in experimental systems (Gilbert et al., 2020). This experimental interest reflects the degree to which subclinical hypothyroidism should be viewed as a disease state (Cooper, 2001). Thus, a preponderance of information indicates that even a small degree of thyroid hormone insufficiency is associated with cognitive deficits in children (Haddow et al 1999; Rose et al., 2006; Nakamizo et al., 2007; Oerbeck et al., 2007; Korevaar et al., 2016)." This reviewer suggested that the EPA document more fully the sensitivity of the human brain to thyroid hormone insufficiency as it may strengthen the support for this choice.

EPA Response: Additional references which support an association between subclinical hypothyroidism and cognitive deficits have been added to Section 3.2.1 (Thyroid Effects).

Comment: One reviewer suggested that the EPA could consider a broader context of PFBA impacts on the thyroid system, including considerations on genetic deficits in specific proteins related to the thyroid hormone system.

EPA Response: Text was added to Section 3.2.1 (Thyroid Effects) which includes a brief discussion of PFBA effects and genetic differences in thyroid hormone system-related proteins.

E.4.4. Public Comments

Comment: One commenter noted that the significance of changes in T_4 levels in rodents to human risk assessment has been questioned by the National Academy of Sciences (NAS) and others because of the significant differences in binding proteins and affinities among species. These differences in binding proteins, binding affinities of the proteins for the hormones, turnover rates of the hormones, and thyroid stimulation lead to important quantitative differences between rats and humans. As a result, NAS concluded that "rats are much more sensitive to agents that disturb thyroid function than are humans, so the relevance of rat studies in quantitative terms to humans is limited." NAS further noted that "[t]he committee does not agree that transient changes in serum thyroid hormone or TSH concentrations are adverse health effects; they are simply biochemical changes that might precede adverse effects.

EPA Response: As noted above, six of the seven peer reviewers indicated the T_4 effects are significant and relevant to humans. One of the external peer reviewers further specified that decreases in serum total and free T_4 are very clearly associated with developmental and physiological deficits: "In both humans and experimental animals, low TH is related to permanent neural and cognitive deficits (e.g., Stagnaro-Green and Rovet (2016); Zoeller and Rovet (2004)). This is likely to be true for several (if not all) organs including heart, bone, lung, and intestine (Bassett et al. (2007); Mochizuki et al. (2007); Wexler and Sharretts (2007); Bizzarro and Gross (2004)). Much of the experimental literature on this topic makes use of models of severe hypothyroidism (Crofton and Zoeller, 2005; Crofton et al., 2005). However, graded effects of thyroid hormone insufficiency have become the focus of increased attention in experimental systems (Gilbert et al., 2020). This experimental interest reflects the degree to which subclinical hypothyroidism should be viewed as a disease state (Cooper, 2001). Thus, a preponderance of information indicates that even a small degree of thyroid hormone insufficiency is associated with cognitive deficits in children (Korevaar et al., 2016; Nakamizo et al., 2007; Oerbeck, 2007; Haddow et al., 1999). The majority of the panel concluded that the conclusions for the thyroid effects to be clearly described and scientifically, with another panel member stating that "the data leading to the conclusion that the thyroid effects are considered relevant to humans were appropriately synthesized and their strengths and limitations were adequately described." In response to this comment, however, the assessment has been revised to include a more detailed discussion on the relationship between thyroid hormone status and neurotoxicological effects to further support assessment conclusions.

Regarding the NAS (2005) report on the Health Implications of Perchlorate Ingestion, the conclusions of that report should be considered narrowly in the context of health effects due to exposure to perchlorate and not generally as broad statements on the human

relevance of thyroid effects observed in animals. The statement that the "committee does not agree that the transient changes in serum thyroid hormone or TSH concentrations are adverse health effects; they are simply biochemical changes that might precede adverse effects" is made regarding what endpoints to base the reference dose value on for perchlorate given the presumed chemical-specific mode of action. This mode of action is not pertinent to PFBA or other PFAS.

Comment: EPA failed to consider studies where quantitative histomorphometric analysis on thyroid function after PFBA exposure did not report statistically significant changes. Histomorphometric analyses of thyroid follicles provide a more quantitative indication of thyroid response than histopathological assessments. To conduct a thorough and defensible risk assessment for PFBA, EPA must consider this and similar studies.

EPA Response: The quantitative histomorphometric analyses from Butenhoff et al. (2012) are discussed in Section 3.2.1 (Thyroid Effects).

E.5. CHARGE QUESTION 3 – DEVELOPMENTAL EFFECTS

For each health effect considered in the assessment and outlined below, please comment on whether the available data have been clearly and appropriately synthesized to describe the strengths and limitations. For each, please also comment on whether the weight-of-evidence decisions for hazard identification have been clearly described and scientifically justified.

- *For developmental effects, the Toxicological Review concludes that the available evidence indicates PFBA exposure is likely to cause developmental effects in humans given relevant exposure circumstances, on the basis of a coherent pattern of delays in acquisition of three different developmental milestones in a single study in mice, with the findings presumed relevant to humans in the absence of evidence to suggest otherwise. The assessment discusses similar effects observed for structurally related PFAS*.

E.5.1. Overarching External Peer Reviewer Comments on Developmental Effects

"Most reviewers agreed that the available data were clearly and appropriately synthesized and concurred with EPA's conclusion. [One reviewer] commented that "in several but not all ways" the available data on developmental effects were clearly and appropriately synthesized to describe the strengths and limitations but noted that the document lacked a discussion of "relevant mechanistic information or information on the conserved biological processes causing the developmental effects observed in mice to be considered relevant to humans." [The same reviewer] noted that such discussions "would inevitably lead to a more complete and transparent description of the strengths and limitations of the available data relevant to developmental effects.""

E.5.2. Tier 1 Recommendations

Comment: One reviewer stated that EPA should add information supporting the human relevance of the developmental effects to the assessment (e.g., the Evidence Integration Summary) in order to "… more fully and clearly support applicability of the weight-of-evidence decision." This reviewer noted that, currently, the text of the assessment lacks sufficient discussion of "…the conserved biological processes between mice and humans that the EPA considers relevant to the observed developmental effects (e.g., for delayed vaginal opening and preputial separation), whether the mouse has been shown to be a good laboratory animal model for assessing potential human developmental effects, or what human developmental endpoints (e.g., delayed onset of puberty) may be presumed to be correlates of some of the PFBA-induced developmental effects observed in the single mouse study (e.g., delays in vaginal opening and preputial separation in Das et al. 2008)." This reviewer further noted that some of this information could found later in the document (i.e., Section 5.2.1) and this information could be included and expanded on in Section 3. Public comments were also received regarding whether delays in vaginal opening and preputial separation whether "these endpoints accurately reflect pubertal development" and stating that "the biological basis for this assumption is lacking and reviews suggest that these measurements accurately reflect pubertal development in the rat but not the mouse." Public comments further stated that "mouse "puberty" has been used for POD calculations in some of the PFAS documents and I question that this is biologically correct, and the Agency may want to reconsider this. Clearly these are valid developmental landmarks in the mouse like eye-opening and etc. but they may not be valid indices of puberty."

EPA Response: An expanded discussion of the conserved biological processes between mice and humans that is considered relevant to the observed developmental effects (e.g., for delayed vaginal opening and preputial separation) has been included in Section 3 and updated in Section 5.2.1. to include supporting references. "The onset of puberty in humans is driven by surges in the levels of estrogen in females and testosterone in males, so the timing of puberty can be altered by exposure to endocrine disrupting chemicals that mimic or antagonize these hormones In female rodents, pubertal markers include vaginal opening (indicative of the first ovulation in rats, but not mice) and the subsequent first estrus and onset of regular estrous cyclicity (rats and mice) [Prevot, V., *Puberty in Mice and Rats*, in Knobil and Neill's *Physiology of Reproduction*, T.M. Plant and A.J. Zeleznik, Editors. 2015 p. 1395–1439]. Since Das et al. (2008), found delayed vaginal opening in mice (not rats), this is not a direct correlate to puberty in humans. However, the Reproductive Guidelines state that both accelerations and delays in the timing of reproductive milestones can be considered adverse.

Comment: One reviewer recommended that references to the developmental toxicity effects of other PFAS should be documented in a chart and/or link to a chart/table.

EPA Response: This information can be found in three recently published reports Carlson et al. (2022); Radke et al. (2022),and Pelch et al. (2022), and that outline the available references of developmental effects following hundreds of other PFAS. Additionally, Table 4-2 has been added to the assessment (see Section 4.1) to facilitate comparisons of developmental toxicity hazard conclusions across EPA PFAS assessments.

Comment: One reviewer recommended emphasizing statements on the observation of fetal effects in the absence of maternal toxicity and data gaps regarding information on the thyroid and nervous system following gestational exposures. **EPA Response:** This data gap is now discussed in the evidence integration summary of Section 3.2.3. (Developmental Effects); specifically, Section 3.2.3 notes that developmental delays consistent with delayed sexual maturation are observed in the absence of body weight or maternal effects, thus strengthening the certainty that the observed effects are adverse fetal effects.

E.5.3. Tier 2 Suggestions

Comment: One reviewer suggested that the EPA could opine on a thyroid-mediated mode-of-action in young animals based on evidence in adult rats (i.e., thyroid histopathology and serum T_4 decreases in adult animals).

EPA Response: There is insufficient available data supporting a direct link between a thyroid-mediated mode-of-action in young animals based on evidence in adult rats.

Comment: One reviewer commented that "[w]hile the rationale [for developmental effects] is well-explained, the tone of the report implies notably weaker support than for thyroid or hepatic effects. It seems odd to put them into the same ultimate bin when there is such a sharp contrast in the evidence and even in the way that the evidence is interpreted and expressed." This reviewer suggested further explanation in the report why developmental effects "ended up in the same level" despite a weaker evidence base.

EPA Response: Additional references have been incorporated in Section 3.2.3 to further support the rationale for the evidence integration judgment for developmental effects. Please note that the methods for drawing the evidence integration judgments and determining the levels of certainty regarding potential hazard are outlined in Appendix A. These methods do allow for a range of evidence scenarios with some variation in "strength" to ultimately lead to the same overall judgments.

E.5.4. Public Comments

Comment: The discussion of the effects of PFOA on timing of vaginal opening and preputial separation in mice reported by Lau et al. (2006) is incomplete and appears to be inappropriate. Lau et al. (2006) reported a non-monotonic dose-response for both day of vaginal opening and day of preputial separation in mouse offspring from dams treated with PFOA on gestational day (GD) 1–17. Vaginal opening was accelerated at the lowest dose (1 mg/kg/day) and delayed, at higher doses (3–20 mg/kg/day), with a greater delay with increasing dose. More notably, preputial separation was accelerated, rather than delayed, at all doses (1, 3, 5, and 10 mg/kg/day) except the highest dose (20 mg/kg/day). The acceleration of preputial separation was greatest at the lowest dose (1 mg/kg/day) and the magnitude of this effect decreased with increasing dose from 1–10 mg/kg/day. At 20 mg/kg/day, the only dose at which preputial separation was delayed, there was severe toxicity including full litter resorptions in 88% of dams and approximately10% survival (i.e., ~90% mortality) of offspring at postnatal day (PND) 23. As such, the statement that the observations in male offspring treated with PFOA from Lau et al. (2006) are part of a "consistent pattern of delayed pubertal milestones…following exposure to related PFAS" does not appear to be supportable.

EPA Response: The relevant text in the Toxicological Review has been updated to indicate, where appropriate, markers of sexual maturation were delayed or accelerated. Specifically, to the comment that the dose-response for vaginal opening observed in Lau et al. (2006) was non-monotonic: while the mean value of time to vaginal opening was slightly lower in the lowest exposure group (compared to controls), this difference was not statistically significant and is more appropriate characterized as “no effect” rather than “accelerated”.

Comment: The Reproductive Toxicity Guidelines (U.S. EPA,1996) state that alterations in the age at puberty should be considered as adverse effects. Citing this, the Toxicology Review of PFBA uses delays in vaginal opening and preputial separation in the mouse for RfD, assuming that these endpoints accurately reflect pubertal development. However, the biological basis for this assumption is lacking and reviews suggest that these measurements accurately reflect pubertal development in the rat but not the mouse. This species difference is described in Chapter 30 “Puberty in Rats and Mice” by Vincent Prevot (Knobil and Neil’s Physiology of Reproduction. Elsevier Science and Technology, 2015) and in a publication in Scientific Reports (2017) by Gaytan et al. (DOI:10.1038/srep46381).

Related Comment: In rats, but not mice, vaginal opening is normally associated with maturation of the HP axis, ovulation and the initiation of reproductive cycles. Mouse “puberty” has been used for POD calculations in some of the PFAS documents and this may not be biologically correct and the Agency may want to reconsider this. Clearly these are

valid developmental landmarks in the mouse like eye-opening but they may not be valid indices of puberty.

EPA Response: As noted by the commenter, alterations in the timing of reproductive development (whether delayed or accelerated) are considered adverse in accordance with the EPA Reproductive Guidelines. While vaginal opening in mice may not be a direct correlate to puberty in humans (as it is in rats), alterations in reaching reproductive or developmental milestones are considered adverse and relevant to human health. However, the linkage of these delays in acquisition of these milestones to "puberty" has been removed in this revised assessment.

E.6. CHARGE QUESTION 3 – REPRODUCTIVE AND OTHER EFFECTS

For each health effect considered in the assessment and outlined below, please comment on whether the available data have been clearly and appropriately synthesized to describe the strengths and limitations. For each, please also comment on whether the weight-of-evidence decisions for hazard identification have been clearly described and scientifically justified.

- *For reproductive effects and other noncancer effects (i.e., cardiometabolic effects, renal effects, ocular effects, body weight), the Toxicological Review concludes there is inadequate evidence to determine whether PFBA exposure has the potential to cause these effects in humans on the basis of the sparsity of available evidence.*

E.6.1. Overarching External Peer Reviewer Comments on Reproductive and Other Effects

"All reviewers concurred with EPA's conclusion that there is insufficient evidence to determine if PFBA can cause reproductive or other noncancer effects."

E.6.2. Tier 1 Recommendations

Reviewers had no Tier 1 Recommendations.

E.6.3. Tier 2 Suggestions

Comment: One reviewer suggested that EPA consider adding a recent study by Ou et al. (2021) that indicates some PFAS may increase the risk of heart defects.

EPA Response: While the Ou et al. (2021) study does suggest that some PFAS are associated with increased odds of septal defects, for PFBA specifically, odds ratios were not statistically significantly increased or decreased for septal defects, conotruncal defects, and all congenital heart defects. Hence, given that no defects were observed to be associated with PFBA exposure, discussion of this study was not added to the assessment given that consideration of its observations would not materially change the evidence integration conclusion for reproductive or developmental effects.

Comment: One reviewer suggested that EPA consider including supporting mechanistic evidence or supplemental information for reproductive and other noncancer effects for a fuller description of the strength and limitations of the available information for reproductive and noncancer effects.

EPA Response: There is ***inadequate evidence*** to determine whether PFBA exposure has the potential to cause human reproductive effects (aside from the delays in sexual and/or reproductive development discussed in Section 3.2.3: Developmental Effects) or the "other noncancer effects" discussed in the draft assessment. In general, and as in this case here where there is only limited (or none) supplemental information on these outcomes, there is not clear added value in delineating the potentially relevant mechanistic information on health effects for which toxicological evidence is generally lacking. Therefore, no additional discussions of mechanistic or supporting evidence for reproductive and other noncancer effects were added to this revised assessment.

E.7. CHARGE QUESTION 5 – CANCER HAZARD

The draft assessment concludes there is inadequate evidence to assess carcinogenic potential for PFBA and that this descriptor applies to oral and inhalation routes of human exposure. Please comment on whether the available animal and mechanistic studies, and the analysis presented in the Toxicological Review, support this conclusion.

E.7.1. Overarching External Peer Reviewer Comments on Cancer Hazard

"All reviewers concurred with EPA's conclusion that there is inadequate evidence to assess carcinogenic potential for PFBA for either oral or oral inhalation exposure. [One reviewer] noted that the "evidence that PFBA is carcinogenic is sparse and continues to be sparse according to a recent review of PFAS and cancer (Steenland and Winquist 2021)."

E.7.2. Tier 1 Recommendations

Comment: One reviewer commented unlike the other sections with limited data, no discussion was included in the carcinogenicity section about how other studies of PFBA-related compounds could or could not inform the data gaps for carcinogenicity or genotoxicity. This reviewer recommended that EPA should include a short section to address this missing component in a manner similar to the other sections of the report.

EPA Response: A short explanation that the evidence for carcinogenicity due to PFBA exposure is non-existent and is limited for other related PFAS has been added to Section 3.3. In addition, the carcinogenicity conclusions from other EPA PFAS analyses that have been finalized have been included in Table 4.2, although they are not influential to the carcinogenicity judgment for PFBA. Conclusions from other ongoing IRIS assessments have

not been added because they have not yet been finalized. Although not currently available, future efforts on PFAS (see PFAS Strategic Roadmap {U.S. EPA, 2021, 10002133@@author-year), potentially including additional studies on PFBA, may help to inform this data gap.

E.7.3. Tier 2 Suggestions

Reviewers had no Tier 2 suggestions.

E.8. CHARGE QUESTION 6 – NONCANCER TOXICITY VALUE DATA SELECTION

For PFBA, no RfC was derived. The Butenhoff et al. (2012) 90-day rat study was the study chosen for use in deriving the RfD on the basis of an increased incidence of hepatocellular hyperplasia and decreased total T_4 in male rats. Is the selection of this study and these effects for use in deriving the RfD for PFBA scientifically justified?

a. *If so, please provide an explanation.*

b. *If not, please provide an alternative study(ies) or effect(s) that should be used to support the derivation of the RfD and detail the rationale for use of such an alternative.*

c. *As part of the recommendations in "a" or "b" above, please comment on whether the effects selected are appropriate for use in deriving the RfD, including considerations regarding adversity (or appropriateness in representing an adverse change) and the scientific support for their selection. More specifically, Appendix A identifies interpreting the adversity of certain outcomes observed in rodents, including some hepatic effects, as a key science issue. Please consider in your recommendation the narrative in the Toxicological Review related to the decision that the observed hepatocellular hypertrophy, when considered within the broader constellation of effects, is representative of an adverse change in the organ.*

d. *Given the lack of studies on inhalation exposure to PFBA, no reference concentration (RfC) is derived. Please comment on this decision*.

E.8.1. Overarching External Peer Reviewer Comments on Noncancer Toxicity Value Data Selection

"All reviewers concurred: (1) with the selection of the Butenhoff et al. (2012) study as scientifically justified for derivation of an RfD for PFBA; (2) that the critical effects selected were appropriate for use in deriving the RfD; and (3) that the decision to not derive a reference concentration (RfC) was justified, given the lack of studies on inhalation exposures to PFBA."

E.8.2. Tier 1 Recommendations

Comment: One reviewer recommended adding a discussion of other carcinogenicity studies across the structurally related perfluorinated compounds to Section 5.2.

EPA Response: A short explanation that the evidence for carcinogenicity due to PFBA exposure is non-existent and is limited for other related PFAS has been added to Section 3.3. In addition, the carcinogenicity conclusions from other EPA PFAS analyses that have been finalized have been included in Table 4.2.

E.8.3. Tier 2 Suggestions

Comment: One reviewer commented that identification of hepatic effects observed in rodents as a key scientific issue is a very conservative approach and that the conclusion is in contradiction with Hall et al. (2012), stating "The results of this workshop concluded that hepatomegaly as a consequence of hepatocellular hypertrophy without histologic or clinical pathology alterations indicative of liver toxicity was considered an adaptive and a non-adverse reaction." This reviewer stated that this conclusion should be reached by an integrative weight of evidence approach.

EPA Response: An integrative weight of evidence approach is the approach taken in Section 3.2.2. The conclusions in Section 3.2.2 are entirely in-line with the recommendations of {Hall, 2012, 2718645@@author-year} that coincident histological evidence of liver injury/damage can be used to support the conclusion that liver weight and hypertrophic changes are adverse. Specifically, for PFBA, liver weight changes and hypertrophic lesions are accompanied by other lesions such as necrosis and vacuolation, supporting their adversity along a progression to more severe effects. Thus, no revision was made. See related responses to comments on charge question 3.

E.8.4. Public Comments

Comment: The Das study was not considered as the critical study because it did not measure serum thyroid hormones. This decision does not withstand scrutiny from the scientific community. Serum thyroid hormones are subject to a great degree of variability due to the assay issues, diurnal variations, and even husbandry conditions, and thus are not suitable grounds upon which to exclude a relevant study. The Das study evaluated important key functional aspects of pregnancy outcomes as well as neonatal development into young adults and should have been considered as the critical study in the Draft IRIS Review for PFBA. Further, given the current Draft IRIS Review's emphasis on developmental toxicity with exposure to PFBA and the availability of relevant studies, EPA should rely on its own developmental study for PFBA as the critical study to develop the reference value for PFBA.

EPA Response: It is not accurate to state that the Das study was not considered as the critical study because it did not measure serum thyroid hormones. Endpoints from both Das et al. (2008) and Butenhoff et al. (2012) were considered for POD derivation. Ultimately, as

liver and thyroid effects from (Butenhoff et al., 2012) were interpreted with the highest confidence and resulted in lower PODs, the candidate values for these endpoints were selected for the RfD. Additionally, the candidate value for developmental delays from Das et al. (2008) was selected for the subchronic RfD (see below).

Comment: It is suggested that the rationale for not considering hepatic effects in male mice exposed to PFBA for 28 days in Foreman et al. (2009) as the basis for Reference Dose (RfD) development be reconsidered. It (i.e., the assessment) is stated that the subchronic (90 day) study in rats (Butenhoff et al., 2012b) and the developmental study in mice with 17 days of exposure (GD 1-17; Das et al., 2008) were considered for RfD development because these study designs can "estimate potential effects of lifetime exposure, as compared to short-term [i.e., 28 day] or acute studies." However, the preference for developmental studies over short term (i.e., 28 day) studies does not appear to be supportable for RfDs based on systemic effects such as increased relative liver weight or histopathological changes in the liver. Specifically, Table 3-5 shows that both wild-type and humanized PPAR-alpha male mice exposed to PFBA for 28 days (Foreman et al., 2009) are more sensitive to increased relative liver weight (i.e., a greater increase at the same or similar dose) than pregnant and non-pregnant female mice exposed for 17 days (Das et al., 2008) and are also more sensitive than male rats exposed for 28 and 90 days (Butenhoff et al., 2012b).

EPA Response: For the purpose of deriving chronic non-cancer reference values, subchronic exposure studies are preferred over short-term studies when chronic studies are lacking in the toxicity database. Also, developmental toxicity studies are useful for evaluating the potential for increased susceptibility in pregnant animals or their offspring during this sensitive lifestage, when short periods of exposure during critical windows of development can be considered more relevant to identifying sensitive health effects from a lifetime of exposure than subchronic or chronic exposure durations. This latter consideration does not necessarily apply to all health endpoints, likely including the hepatic effects of PFBA. A greater amount of uncertainty exists in extrapolating from short-term to chronic durations and thus, given the availability of preferred studies in the PFBA database, the results of short-term studies are only considered as supporting information for toxicity value derivation.

E.9. CHARGE QUESTION 7 – SUBCHRONIC REFERENCE DOSE

In addition, for PFBA, an RfD for less-than-lifetime ("subchronic") exposures is derived. No "subchronic" RfC was derived. The study chosen for use in deriving the subchronic RfD is the gestational exposure mouse study by Das et al. (2008) with the RfD based on delayed acquisition of developmental milestones, as indicated by delayed time to vaginal opening, eye opening, and preputial

separation in exposed male and female offspring. Is the selection of this study and these effects for the derivation of the subchronic RfD for PFBA scientifically justified?

a. *If so, please provide an explanation.*

b. *If not, please provide an alternative study(ies) or effect(s) that should be used to support the derivation of the subchronic RfD and detail the rationale for use of such an alternative.*

c. *As part of the recommendations in "a" or "b" above, please comment on whether the effects selected are appropriate for use in deriving the RfD, including considerations regarding adversity (or appropriateness in representing an adverse change) and the scientific support for their selection.*

d. *Given the lack of studies on inhalation exposure to PFBA, no "subchronic" RfC is derived. Please comment on this decision.*

E.9.1. Overarching External Peer Reviewer Comments on the Subchronic Reference Dose

"All reviewers concurred that selection of the Das et al. (2008) study is scientifically justified for derivation of the subchronic RfD. [One reviewer] noted that a subchronic RfD based on developmental effects resulting from a shorter exposure duration "will provide a useful risk assessment complement to the chronic RfD, furthering risk assessment and risk communication."

"Most reviewers concurred that the selection of effects were justified for derivation of the RfD, although [one reviewer] commented that EPA should consider improving the scientific justification to the extent possible for certain effects."

"All reviewers concurred that the decision to not derive a subchronic RfC was justified."

E.9.2. Tier 1 Recommendations

Comment: One reviewer recommended that a discussion be included regarding how inclusion of delayed eye opening would or would not change the RfD.

EPA Response: The candidate value for delayed eye opening is included in the endpoints considered for final RfD selection in Table 5-7 of the assessment. As can be seen from the values presented in this table, selection of this endpoint would result in an RfD 33% higher than currently selected (8×10^{-3} mg/kg-day vs. 6×10^{-3} mg/kg-day) and was not interpreted with greater confidence, and thus would be inadequately protective of human health. A brief discussion to this effect has been added.

Comment: One reviewer commented that they were surprised that the cumulative endpoint of "total affected implants" was not reported or evaluated. This reviewer

recommended that the Toxicological Review add a paragraph discussing further several of the decisions to not look at the endpoint of all affected, why the Rai and Van Ryzin model was not used and discuss the importance of looking at this cumulative endpoint. This paragraph should specifically discuss whether these suggestions would have changed the calculation of the subchronic RfD.

EPA Response: A short discussion of this topic, including an explanation for why the cumulative endpoint was not modeled, has been added to the assessment in Appendix D.1.2 and as a footnote to Table 5-1. The Rai and van Ryzin model are a nested dichotomous model that can account for intralitter similarity (via estimation of intralitter correlation and use of litter specific covariate parameters). Functionally, this model has the same capabilities of the nested Logistic model, which is currently implemented in BMDS 3.2 and was used to model the embryo/fetal mortality endpoint. How modeling "total affected implants" would ultimately impact the final RfD derived is difficult to characterize but given that fetuses that lived to parturition and experienced some delay in developmental milestones would be counted as "responding" alongside fetuses that died in utero, the total incidence of "affected fetuses" would be greater per dose group than for the individual endpoints, thus resulting in a lower BMD and BMDL. This would result in a lower POD.

Comment: One reviewer commented on the need to add a paragraph that expands the discussion of what the lack of a functional reproductive or neurodevelopmental assessment means for the application of the current uncertainty factors in this incomplete toxicological assessment package. The signals from the Das et al 2008 study should raise concern of this lack and require additional assessment of what this widespread environmentally relevant compound and metabolic common breakdown product means to the overall IRIS assessment report across the structurally related perfluorinated compounds.

EPA Response: The lack of a functional reproductive or neurodevelopmental study in relation to selecting the uncertainty factors is discussed in Section 5.2.1. Briefly, the lack of a functional reproductive study was not considered a key quantitative data gap in the PFBA database, in part given the general lack of evidence for sensitive reproductive effects for some other similar PFAS (see Table 4-2). Regarding the lack of a functional neurodevelopmental study, while this is identified as a key data gap in the PFBA evidence base, concerns over developmental neurotoxicological effects presumed to result from thyroid hormone insufficiency are mitigated given evidence from PFBS (i.e., PODs for effects in dams and offspring are almost identical for thyroid hormones). However, concern over developmental neurotoxicity independent of a thyroid hormone-mediated mechanism remains an important uncertainty accounted for via application of a $UF_D = 3$.

Comment: One reviewer reiterated their previous Tier 1 Recommendation (see Section E.5) that EPA should consider improving scientific justification (e.g., expanded to other developmental effects) to the extent possible, particularly for the critical effect. Specifically, this reviewer requested that EPA expand discussion on the conserved biological processes between mice and humans that the EPA considers relevant to the observed developmental effects (e.g., for delayed vaginal opening and preputial separation), whether the mouse has been shown to be a good laboratory animal model for assessing potential human developmental effects, or what human developmental endpoints (e.g., delayed onset of puberty) may be presumed to be correlates of some of the PFBA-induced developmental effects observed in the single mouse study (e.g., delays in vaginal opening and preputial separation in Das et al. (2008)).

EPA Response: A discussion on the known similarities and differences in the development of these outcomes between rodents and humans has been added to the assessment (see Section 3.2.3). It is important to emphasize, as described in EPA guidelines, that it is not expected that the effects that manifest in animal studies will manifest similarly in humans, although outcomes that are strongly correlated across species can provide strong evidence. Likewise, in the absence of evidence to the contrary, effects observed in animal models are considered relevant to humans.

Comment: One reviewer noted that it appeared to them that the "…document supports use of delayed eye opening and embryo/fetal mortality more strongly than other developmental effects. While these effects provide candidate subchronic RfD values similar to that selected for the proposed final subchronic RfD, delays in vaginal opening observed in Das et al. (2008) apparently provides the most specific basis for that value (Table 5-10, p. 5-24), the adversity of which does not appear to be fully addressed in the document." This reviewer recommended that relevant information to support the adversity of the critical effect ultimately selected for the determination of the subchronic RfD be added to the assessment.

EPA Response: A discussion regarding the adversity of the observed developmental delays has been added to the assessment in Section 3.2.3.

E.9.3. Tier 2 Suggestions

Reviewers had no Tier 2 suggestions.

E.10. CHARGE QUESTION 8 – NONCANCER TOXICITY VALUE DOSE-RESPONSE MODELING

EPA used benchmark dose modeling (U.S. EPA, 2012) to identify points-of-departure (PODs) for oral exposure to PFBA. Are the modeling approaches used, selection and justification of benchmark

response levels, and the selected models used to identify each POD for toxicity value derivation scientifically justified?

E.10.1. Overarching External Peer Reviewer Comments on Noncancer Toxicity Value Dose-Response Modeling

"The reviewers who responded to this charge question supported the modeling approach used and the justification provided, with some reviewers noting enthusiastic support for the approach."

"Several reviewers noted that benchmark dose modeling is not their area of expertise and declined to comment."

E.10.2. Tier 1 Recommendations

Reviewers had no Tier 1 recommendations.

E.10.3. Tier 2 Suggestions

Comment: One reviewer suggested that EPA attempt to strengthen the benchmark response (BMR) justification for vaginal opening delays beyond historical precedence (to the extent possible) given that it is ultimately the critical effect used in determining the subchronic RfD.

EPA Response: The selection of a BMR = 5% relative deviation for delayed vaginal opening is not based on historical precedence, but rather is based on a determination of what level of change in this effect has been judged to be a relevant response level (i.e., a minimally biologically significant response) in the EPA Endocrine Disruptor Screening Program. In addition, this approach in defining a BMR for continuous endpoints based on biological information is consistent with the recommendations of the BMD Technical Guidance (U.S. EPA, 2012). A full discussion of the selection of the BMR for developmental delays is included in Table 5-2.

Comment: One reviewer commented that the model selected for liver hypertrophy should possibly be the log-logistic model vs the currently selected Weibull model based on equal Akaike Information Criterion values and a BMDL value that appears to better agree with the actual study data.

EPA Response: After consideration of the shape of the dose-response curve for this dataset, it was determined that the minimum-maximum characteristic of the data (i.e., response going from 0/10 to 9/10 between the mid- and high-dose groups precludes modeling this dataset (consistent with the BMD Technical Guidance (2012) document (U.S. EPA, 2012)). Thus, the NOAEL/LOAEL approach was used instead for this endpoint. This

change in POD derivation methodology does not have a large impact on the assessment as the NOAEL (6 mg/kg-day) is very close to the BMDL (5.4 mg/kg-day).

E.11. CHARGE QUESTION 9 - TOXICOKINETICS

Appendix A identifies the potential for toxicokinetic differences across species and sexes as a key science issue and lays out a hierarchy for using relevant toxicokinetic data in extrapolating doses between laboratory animals and humans. Given what is known and not known about the potential interspecies differences in toxicokinetics of PFBA, EPA used the ratio of human-to-animal serum clearance values to adjust the POD to estimate a human equivalent dose in the derivation of the respective RfDs.

a. *Is applying the ratio of human-to-animal serum clearance values for PFBA scientifically justified? If not, please provide an explanation and detail on a more appropriate approach.*

b. *Do the methods used to derive toxicity values for PFBA appropriately account for uncertainties in evaluating the toxicokinetic differences between the experimental animal data and humans?*

E.11.1. Overarching External Peer Reviewer Comments on Toxicokinetics

"The reviewers concurred that the application of the ratio of human-to-animal serum clearance values for PFBA was scientifically justified. [One reviewer] commented that applying the ratio is a more appropriate choice than scaling doses allometrically using body weight $(BW)^{3/4}$ methods."

E.11.2. Tier 1 Recommendations

Comment: One reviewer suggested that another approach to help justify or strengthen the animal to human extrapolation is to use kidney filtration (GFR) and that comparison of excretion ratios (i.e., clearance/GFR) could be compared to other PFASs where data exist. This reviewer stated that these analyses may be insightful for characterizing the degree of renal/hepatic reabsorption and ultimately increase the confidence in the calculated human CL.

EPA Response: Two paragraphs were added to Section 3.1.5 comparing the clearance values to species-specific average GFR values for mice, rats, and humans, with or without including the impact of serum binding (i.e., GFR*free). A comparison across PFAS was not made in the revised PFBA assessment (a preliminary analysis of results across draft and final assessments indicates that there is not a consistent, predictable pattern, so any such discussion would not be straightforward or brief). However, this is taken as a useful research note for future applications.

Comment: One reviewer recommended that EPA clarify sections on PFBA toxicokinetics with respect to the consistency of the linearity assumptions with the presence of saturation processes. This reviewer specifically noted that EPA stated that dosimetric adjustments are made "assuming the exposure being evaluated is low enough to be in the linear (or first-order) range of clearance" in one part of Section 5, but in another part stated that "results for both male and female mice [from Chang et al., 2008] show a dose-dependent increase in clearance across all dose levels, consistent with the hypothesis of saturable renal resorption."

EPA Response: Given the limited PK data available, an extensive analysis of this issue is not possible. However, a paragraph has been added to Section 3.1.5 explaining that the mouse data appear to be reasonably consistent with constant CL at ≤ 30 mg/kg-day, and the plotted rat data for the same dose in Chang et al. (2008) appear likewise, so linearity will be assumed valid for dose levels below that, not above. A statement was also added in Section 5, where the approach for extrapolation is described.

E.11.3. Tier 2 Suggestions

Comment: One reviewer suggested that EPA could use dosimetrically adjusted doses rather than a ratio of clearance values, and that this could be considered to see if such calculations would make significant differences.

EPA Response: If the dosimetric adjustment still involves use of a single clearance value, then changing the order of dosimetric adjustment and dose-response analysis will not change the outcome. Such an approach could make a significant difference if the adjustment is dose-dependent. However, such an approach would significantly complicate the analysis. Clearance will not be a function of dose directly, but rather a function of blood concentration or internal dose, which in turn both depends on dose and time. Thus, a nonlinear PK model would need to be developed, validated, and applied to properly account for the time- and dose-dependence in a bioassay. The comment suggests consideration of a simpler approach, with clearance just assumed to be a nonlinear function of dose, but this would require picking a function form and then using it to interpolate between and extrapolate above doses used in the mouse PK studies, which has its own uncertainties. Further, since the PK study for rats only used a single dose level, the approach can't be used directly for that species. Given these considerations, development and application of a nonlinear PK model was not attempted.

Comment: One reviewer noted that, while the Chang et al. (2008) occupational exposure data provide important insights into PFBA pharmacokinetic behavior, these data are not sufficiently robust for calculating pharmacokinetic parameters, and that EPA uses a half-life

that is sufficiently long to encompass most adults from the study. This reviewer suggested showing how the half-lives were determined with the data and the logic for selecting a specific value.

EPA Response: Details of the analysis were previously provided in Appendix C.2. The rationale for excluding subjects for whom the second measurement was below the LLOQ were provided in Section 5.2.1 but are now restated in Appendix C.2. A table listing the subjects used in EPA's analysis has been added to Appendix C.2. Of those eight subjects, four had half-lives greater than the estimated half-life, 67.9 hours, and four were below. Of the four excluded subjects, one had an estimated half-life greater than 67.9 hours. Hence, while it may be true that 7 of the 12 total subjects (i.e., "most") had half-lives lower than EPA's estimate, this is only 58% of the study population. Therefore, the assessment concludes that the estimated mean in this estimate is reasonable.

While it is recognized that there is uncertainty inherent in the use of these data, but they are nonetheless human elimination data. The alternative to use of these human elimination data is to estimate the human half-life based on $BW^{¾}$ scaling, which leads to predicted half-lives of 38 hours in men and 7 hours in women, which are within a factor of 2 of the value from EPA's analysis (67.9 hours) for men and within a factor of 10 of that estimates for women. Given that PFAS are generally known to be subject to renal resorption, use of 67.9 hours is therefore considered both biologically plausible and modestly health-protective compared to use of $BW^{¾}$ scaling. The presumption is that it provides an average elimination, hence clearance rate for humans and that uncertainty in this value is addressed by the portion of UF_H assigned (i.e., factor of 3).

Comment: One reviewer commented that, under steady-state exposure to humans (e.g., drinking contaminated water) the bioaccumulation would depend on the half-life but no matter how quickly it's cleared, it is steadily replaced. This reviewer further stated that humans don't receive a single dose and stop being exposed, so slow or rapid clearance would be less influential than in the typical laboratory situation. This reviewer suggested that this could be explained in the report, indicating how the steady-state exposure to humans bears on the evaluation.

EPA Response: If humans are exposed to a regular (daily) dose, D, then use of the estimated human clearance (CL_H) leads to a prediction of an ongoing blood concentration equal to D/CL_H; i.e., that is the steady-state or average blood concentration given the daily dose, D. Hence, this evaluation assumes that the steady-state level increases or decreases in direct proportion to D, with $1/CL_H$ being the proportionality constant. This is now stated in the section on dose extrapolation.

E.11.4. Public Comments

Comment: The assessment uses a worst-case estimate for the serum excretion half-life (t1/2) of PFBA in humans which is confounded by potential co-exposure to other PFAS molecules. The draft IRIS assessment assumes a t1/2 of 67.9 hours based on a study of workers exposed to other PFAS believed to be metabolized to PFBA by Chang et al. (2008). The authors [i.e., Chang et al., 2008] note, however, that the data need to be interpreted cautiously, because the workers were not exposed directly to PFBA and were exposed to materials that, given their chemical structures, likely are metabolized to PFBA via oxidation or hydrolysis. Moreover, the draft assessment rejects the default body weight (BW0.75) approach which would generate a half-life of 37.8 hours – within a factor of two of that derived from Chang et al., 2008.

EPA Response: Given the lack of controlled PK studies of PFBA in humans, the only empirical data one can use are those potentially confounded with possible ongoing exposures or metabolic production from precursors. Environmental epidemiological analyses, in which observed blood concentrations are correlated with estimated exposure levels, have an inherent uncertainty in that the exposure is not exactly known. Given that any estimate of human elimination has uncertainty, the EPA considers the use of a health-protective estimate to be reasonable, while noting that 67.9 hours is in fact not an upper bound of the Chang et al. (2008) data: five of the 12 subjects had estimated half-lives greater than that, as noted above. It is also noted that for subjects 1 and 2 of the Cottage Grove group in Chang, for whom multiple time-points are shown in Figure 6, the decline is very close to log-linear. If there was significant ongoing exposure or metabolic production, the decline would asymptotically approach a plateau, the steady-state level given that exposure. Further, the t1/2 values for these two subjects are above the average obtained. In summary, the results do not represent a worst-case scenario, which would involve use of an upper-bound half-life estimate from these data. While it is possible that the true average is lower, there are no specific data to support this possibility, for example that PFBA elimination is reduced in this population due to co-exposures to other PFAS. Further, EPA guidelines U.S. EPA (2011) support the use of such empirical data over default $BW^{¾}$ scaling when such data exist.

Comment: The approach taken in the draft assessment for calculating a dose adjustment factor (DAF) for the animal data overestimates serum concentrations which leads to an over prediction of the toxicity of PFBA. Although using the ratio of clearance rates to generate the DAF is preferable to using the ratio based on half-life, the uncertainty around the t1/2 estimate raises significant questions about the decision to use it to derive the human equivalent dose (HED). Given this uncertainty, using the default body-weight scaling

method to develop the DAF is more appropriate. Failing that, the assessment should include additional discussion of the uncertainty in the clearance model that is chosen.

EPA Response: As noted above, EPA guidelines U.S. EPA (2011) state that extrapolation based on chemical-specific data is preferred over default scaling. Discussion of the uncertainty in the human half-life estimate has been augmented in the revised assessment (see Sections 3.1.4 and 5.2.1), including discussion of the potential for ongoing exposure or metabolic production of PFBA, although noting that there is no clear indication of this occurring. The 2-fold difference between the estimated clearance and that predicted from BW^¾ is well within the range of uncertainty that would be expected for either value.

Comment: The body weight values for the animals in the study that is being evaluated, if available, should be used instead of default body weight values to derive dosimetric adjustment factors (DAFs) based on body weight$^{3/4}$.

EPA Response: In general, study-specific body weight (BW) should be used when extrapolating dosimetry for a given toxicological observation. However, use of a DAF based on clearance builds in a BW adjustment, in that clearance is a rate per kg BW. While some intra-species variation in CL may occur with BW, the data available are not sufficient to demonstrate such variation, such variation will not have a significant effect on the outcome, and as addressed above, BW$^{3/4}$ scaling was not determined to be the best approach for PFBA.

Comment: A recent paper, Abraham et al. (2021), that investigated the distribution of PFBA in human tissues and came to conclusions that differ from those of Pérez et al. (2013). should be cited. See: Abraham et al. (2021). Perfluorobutanoic acid (PFBA): No high-level accumulation in human lung and kidney tissue. International Journal of Hygiene and Environmental Health, 237, 113830.

EPA Response: First, the term "accumulation" is interpreted as an increase in tissue levels over time, given ongoing exposure. The determination of whether or not accumulation in that sense occurs requires longitudinal samples of the same individual over time while carefully monitoring exposure. Both Pérez et al. (2013) and Abraham et al. (2021) only reported single time-point samples; hence, neither paper contains data that can be used to demonstrate or refute accumulation (e.g., a high tissue concentration can occur due to high exposure without accumulation). Second, the subjects of Pérez et al. (2013) were from Tarragona County (Catalonia, Spain), while those analyzed by Abraham et al. (2021) were from France, collected 3–6 years later than Pérez, so there were likely differences in their exposure levels. That being said, the enormous differences in reported tissue levels certainly raises questions about the respective analytic methods, and so the results of

Abraham are now also provided in the revised assessment (see Section 3.1.2), noting the discrepancy given these caveats about unknown differences in exposure. A recently published paper by EPA authors Bangma et al. (2021) did identify an endogenous compound in placenta that is a likely analytic interferent, and this has been added to the document. The interfering compound identified by Bangma et al. (2021) would have to be present in human tissues but not in the pig tissues used for QA by Pérez et al. (2013) in order to explain the discrepancy.

Importantly, as noted in the revised assessment, Abraham et al. (2021) paper does not report some necessary methodological information including whether matrix-matched calibration curves were used or what QA/QC measures were taken (i.e., duplicates, method blanks, continuous calibration verification, etc.). Pérez et al. (2013) state that matrix matched calibration was used for each tissue type using pig brain, liver, bone, and kidney; that reagent blanks, sample blanks, a repeated measures were conducted. The Pérez et al. (2013) study also states that their method was validated, which is not reported by (Abraham et al., 2021). Hence, based on details provided in the published papers, EPA's evaluation yields much higher confidence in (Pérez et al., 2013) than (Abraham et al., 2021). Since the results of Pérez et al. (2013) are not used in the quantitative assessment, this notation will only have a qualitative impact on the Toxicological Review.

Comment: The toxicokinetic differences between wild-type, PPAR-alpha null, and humanized PPAR-alpha mice reported by Foreman et al. (2009) are unlikely to result from differences in PPAR-alpha status and are potentially relevant to interpretation of differences in susceptibility to toxicity among these strains. Suggest adding serum and liver concentration data from Foreman et al. (2009) to Table 3-1.

EPA Response: While some of the differences between strains reported by Foreman et al. (2009) are not easily attributable to PPRAR-alpha status, this is not true for all of the differences. For example, the lower liver concentrations in null mice compared to wild-type is explainable as the lack of binding to PPAR-alpha in the null mouse liver. Other differences may be secondary to differences in hepatotoxicity (i.e., that this toxicity may be more severe in PPAR-alpha-carrying mice, but not in null mice). That both serum and liver concentrations in the humanized mice at 175 mg/kg were lower than at 35 mg/kg is more difficult to explain, as it suggests significantly higher clearance in this strain at that dose compared to lower doses. The data indicate, though, that PPAR-alphas status has a minimal effect on liver:serum distribution. This now stated and the data have been added to Table 3-1.

E.12. CHARGE QUESTION 10 – UNCERTAINTY FACTOR APPLICATION

EPA has evaluated and applied where appropriate uncertainty factors to account for intraspecies variability (UF_H), interspecies differences (UF_A), database limitations (UF_D), duration (UF_S), and LOAEL-to-NOAEL extrapolation (UF_L) for PFBA.

a. *Has uncertainty been adequately accounted for in the derivation of the toxicity values? Please describe and provide suggestions, if needed.*

b. *For uncertainty in interspecies differences (UFA), a value of 3 is applied to extrapolate between effects in laboratory animals and in humans. Although PPARα dependence might support a value of UFA = 1 if that were the sole mode of action, evidence for non-PPARα MOAs is available in the PFBA (and larger PFAS) database. Thus, uncertainty remains regarding the potential differences in sensitivity across species due to the involvement of both PPARα-dependent and PPARα-independent mechanisms. Further, data are lacking to determine with confidence the relative contribution of these competing MOAs. As such, the Toxicological Review concludes the available data are not adequate to determine if humans are likely to be equally or less sensitive than laboratory animals with respect to the observed hepatic effects and that a value of UFA=3 is warranted to account for the residual uncertainty in toxicodynamic differences across species. Please comment on whether the available animal and mechanistic studies support this conclusion and whether the analysis presented in the Toxicological Review is clearly documented.*

c. *For uncertainty in extrapolating from subchronic to chronic exposure scenarios (UFS), a default value of 10 is applied. The assessment concludes there is conflicting evidence on whether effects manifest at lower exposure levels or are more severe at equivalent exposure levels when comparing findings across short-term and subchronic exposure durations. Thus, to account for the potential for some effects to worsen with longer durations of exposure (subchronic vs. short-term) and the lack of data on whether effects from subchronic exposures might worsen in a chronic exposure scenario, a UFS=10 is applied in the Toxicological Review. Does the provided scientific rationale support this decision? Please explain.*

d. *To inform uncertainty in intraspecies variability (UFH), the assessment evaluates and considers the available evidence on potential susceptibility to PFBA within different populations or lifestages, including any potential human health impacts from early life exposure. Are the available information and data appropriately considered and the resultant UFH values scientifically justified and clearly described?*

e. *Does the provided scientific rationale support the application of the remaining uncertainty factors (UFL, UFD)? Please explain.*

E.12.1. Tier 1 Recommendations

Reviewers had no Tier 1 recommendations.

E.12.2. Tier 2 Suggestions

Comment: One reviewer suggested that the selection of a value of 3 for UF_A should be evaluated and considered as an alternative to a value of 10 (rather than a value of 1) as our

current understanding of interspecies differences in PFAS toxicokinetics and toxicodynamics has very significant gaps.

EPA Response: A value of $UF_A = 10$ is not supported given that interspecies differences in toxicokinetics are accounted for in the application of the dosimetric adjustment factor (DAF). In the case of the PFBA assessment, this DAF explicitly accounts for differences in the toxicokinetics (i.e., serum clearance values) between rodents and humans. However, the data gaps in understanding of PFBA toxicokinetics remain highlighted in the revised assessment as an area deserving of additional research.

Comment: One reviewer noted that, while hepatocellular hypertrophy exhibited exposure duration dependence when comparing 28- and 90-day results, there were no apparent increased sensitivity with longer exposure durations for liver weight or thyroid hormone measures. This reviewer commented that a UF_S of 10 for liver weight and thyroid hormones may then not be justified and their corresponding candidate RfDs may be unjustifiably low. This reviewer suggested that the EPA consider additional justification for the selection of the UF_S for these endpoints.

EPA Response: Although no increase of effect was noted in liver weight endpoints or thyroid hormone levels when comparing short-term (28-day) and subchronic (90-day exposures), this is not sufficient evidence to conclude that effects would not worsen at the same exposure level or become evident at lower exposure levels with chronic exposure. It should be noted that the increase in exposure duration is approximately 8-fold when comparing chronic exposures to subchronic exposures and only 3-fold when comparing subchronic exposures to short-term exposures. A short discussion of this matter has been added to Section 5.2.1 of the assessment and a value of $UF_S = 10$ is retained for all non-developmental toxicity endpoints.

Comment: One reviewer commented that the UF_A of 3 did not account for interspecies differences in thyroid toxicity or developmental effects related to thyroid hormone insufficiency. This reviewer suggested that the EPA consider increasing the UF_A from 3 to 10.

EPA Response: As noted in Table 5-5, after application of the chemical-specific DAF (based explicitly on differences in rodent and human serum clearance values), residual uncertainty regarding toxicodynamics remains and is accounted for in the application of a $UF_A = 3$. EPA is unaware of any chemical-specific data that would support increasing the UF_A to 10 to account for differences in toxicodynamics regarding PFBA-induced thyroid or developmental effects.

E.12.3. Public Comments

Comment: The draft assessment applies a total uncertainty factor of 1,000 to generate the chronic reference dose (RfD). This includes a 10-fold adjustment for a subchronic-to-chronic exposure adjustment (UFS), and a 3-fold adjustment for database uncertainty (UFD). The data from the study by Butenhoff et al. suggest, however, that PFBA levels have reached steady state conditions in rat livers after 28 days. As a result, the draft assessment notes that "[i]ncreased duration of exposure might not elicit increased effects in the target tissue." In the absence of a chronic study, however, the draft concludes that liver effects may increase with prolonged exposure and that includes a UFS of 10. EPA's conclusion is based on evidence that hepatocellular hypertrophy was observed at lower doses after 90 days when compared to the 28-day results, despite the fact that the lowest observed adverse effect level (LOAEL) for increased liver weight (likely resulting from hypertrophy) was the same in both the 28 and 90 day study and the clinical chemistry was inconsistent across endpoints and durations of exposure. EPA's conclusion is also not supported by the studies with perfluorooctanoic acid (PFOA), a structurally similar compound. While hypertrophy was observed in chronic studies of male rats exposed to PFOA, the concentrations required were 10-fold higher than those eliciting a response in the subchronic studies. Consequently, the addition of a UFS of 10 for liver effects is not supported by the available evidence. While a study of chronic exposure to PFBA is not available, the chronic data from studies with PFOA indicate that a UFS of 1 is more appropriate.

EPA Response: For PFBA, there is chemical-specific data that demonstrates, for some endpoints (i.e., hepatocellular hypertrophy), an increase in duration results in observation of effects at lower doses. While this data is missing for other effects (liver weight, thyroid hormone levels), it is consistent with EPA guidelines U.S. EPA (2002) to apply a UF_S = 10 to account for the possibility of increased effects at lower doses when considering chronic exposures. Application of a UF_S = 10 for liver effects in adult animals is consistent with other recent EPA assessments of PFAS, including GenX U.S. EPA (2021), with the justification for this decision documented in Section 5.2.1 of the revised assessment.

Comment: For the candidate RfDs for both chronic and subchronic effects, the draft assessment includes a UFD of 3 based on concern for neurodevelopmental effects independent of a thyroid hormone-related mechanism. Agency guidance explains, however, that a database uncertainty factor is applied when reproductive and developmental toxicity studies are missing since they have been found to provide useful information for establishing the lowest no adverse effect level. The guidance notes that, for a reference dose (RfD) based on animal data, a factor of 3 is often applied if either a prenatal toxicity study or a two-generation reproduction study is missing, or a factor of 10 may be applied if both are missing. In deciding whether to apply an UFD, EPA advises that the assessor should

consider both the data lacking and the data available for a particular organ system as well as life stages. As noted in the draft assessment, a high confidence developmental study is available for PFBA; the draft also notes that the lack of a multigenerational reproductive study "is not considered a major concern." Therefore, EPA's proposal to add an uncertainty factor to address concern about neurodevelopmental effects is not supported by its own analysis.

EPA Response: EPA's *A Review of the Reference Dose and Reference Concentration Processes* (U.S. EPA, 2002) states that the "database UF is intended to account for the potential for deriving an underprotective RfD/RfC as a result of an incomplete characterization of the chemical's toxicity." This comment itself acknowledges that the RfD/RfC document recommends "…the assessor should consider both the data lacking and the data available for particular organ systems as well as life stages" when determining the value of the UF_D. Therefore, it is wrong to conclude that the recommendations for application of the UF_D state that this uncertainty factor is intended to *only* account for the lack of developmental or reproductive studies. Given residual uncertainties regarding the potential for reproductive, developmental neurotoxicity, immunotoxicity, or mammary gland effects, a $UF_D = 3$ is applied and this is consistent with EPA's RfD/RfC recommendations.

Comment: Uncertainties are not appropriately accounted for. In the recent Human Health Toxicity Values derivation for PFBS, EPA included a database uncertainty factor of 10, citing a lack of chronic studies and neurodevelopmental and immunotoxicity studies as well as a lack of mammary gland studies. The same deficits were noted by EPA for PFBA. It is therefore unclear why EPA drew a different conclusion in the draft toxicological review of PFBA, deciding to only apply a partial database uncertainty factor of 3.

EPA Response: The rationale for the selected $UF_D = 3$ in the PFBA assessment is extensively discussed in Section 5.2.1 and the considerations that inform the final selection of $UF_D = 3$ (lack of developmental neurotoxicity, immunotoxicity, and mammary gland toxicity studies) are consistent with the rationale for a higher UF_D provided in the PFBS assessment. The PFBS assessment additionally considers the lack of a chronic study in the final selection of a $UF_D = 10$; this is considered and, as necessary, addressed by the UF_S in the PFBA assessment. For PFBA, the selection of $UF_D = 3$ was supported by the external peer reviewers and is retained in the revised assessment.

Comment: Biomonitoring studies demonstrate that Americans have chronic exposure to multiple PFAS chemicals throughout their lifetimes. Therefore, it is impossible to be exposed to PFBA and no other PFAS chemicals. CDC's NHANES studies reveal that nearly every American has detectable concentrations of four PFAS chemicals in their bloodstream (PFOS, PFOA, PFHxS and PFNA). Multiple other PFAS have been detected in NHANES and

state biomonitoring programs. Toxicity assessment (i.e, the PFBA assessment) should account for simultaneous exposure to other PFAS chemicals that impact the same target organs. EPA must promote similar assessments for other PFAS related health outcomes with potential for additive toxicity, including kidney and liver toxicity, lipid metabolism, birth outcomes, immunotoxicity and developmental effects. At the very least, EPA should add an additional uncertainty factor to account for the high likelihood of additive effects with other PFAS.

EPA Response: The PFBA assessment derives organ or system-specific reference doses and states "… these toxicity values might be useful in some contexts (e.g., when assessing the potential cumulative effects of multiple chemical exposures occurring simultaneously)." Therefore, when assessing cumulative risk, the values presented in the PFBA assessment can be used by risk assessors in conjunction with values in other PFAS human health risk assessments (and separately conducted exposure assessments) to account for exposures to multiple PFAS simultaneously.

E.13. CHARGE QUESTION 11 – CANCER TOXICITY VALUES

Given the conclusion there was inadequate evidence to assess carcinogenic potential for PFBA (Charge Question 5), the Toxicological Review does not derive quantitative estimates for cancer effects for oral or inhalation exposures. Is this decision scientifically justified?

E.13.1. Overarching External Peer Reviewer Comments on Cancer Toxicity Values

"All reviewers concurred with the decision to not derive quantitative estimates for cancer effects due to inadequate evidence to assess carcinogenic potential for PFBA. Additionally, [one reviewer] commented that "no robust scientific foundation has been laid, critically reviewed and broadly accepted by the scientific community for the use of any surrogate PFAS with carcinogenicity data (e.g., PFOA) for this purpose.""

E.13.2. Tier 1 Recommendations

Reviewers had no Tier 1 recommendations.

E.13.3. Tier 2 Suggestions

Reviewers had no Tier 2 suggestions.

APPENDIX F. QUALITY ASSURANCE FOR THE IRIS TOXICOLOGICAL REVIEW OF PERFLUOROBUTANOIC ACID AND RELATED COMPOUND AMMONIUM PERFLUOROBUTANOATE

This assessment was prepared under the auspices of the U.S. Environmental Protection Agency's (EPA's) Integrated Risk Information System (IRIS) Program. The IRIS Program is housed within the Office of Research and Development (ORD) in the Center for Public Health and Environmental Assessment (CPHEA). EPA has an agency-wide quality assurance policy, and that policy is outlined in the *EPA Quality Manual for Environmental Programs* (see CIO 2105-P-01.1) and follows the specifications outlined in EPA Order CIO 2105.1.

As required by CIO 2105.1, ORD maintains a Quality Management Program, which is documented in an internal Quality Management Plan (QMP). The latest version was developed in 2013 using *Guidance for Developing Quality Systems for Environmental Programs (QA/G-1)*. An NCEA/CPHEA-specific QMP also was developed in 2013 as an appendix to the ORD QMP. Quality assurance for products developed within CPHEA is managed under the ORD QMP and applicable appendices.

The *IRIS Toxicological Review of Perfluorobutanoic Acid and Related Salts* has been designated as Influential Scientific Information (ISI) and is classified as QA Category A. Category A designations require reporting of all critical QA activities, including audits. IRIS assessments are developed through a seven-step process. Documentation of this process is available on the IRIS website: https://www.epa.gov/iris/basic-information-about-integrated-risk-information-system#process.

Specific management of quality assurance within the IRIS Program is documented in a Programmatic Quality Assurance Project Plan (PQAPP). A PQAPP was developed using the EPA *Guidance for Quality Assurance Project Plans (QA/G-5),* and the latest approved version is dated June 2022. All IRIS assessments follow the IRIS PQAPP, and all assessment leads and team members are required to receive QA training on the IRIS PQAPP. During assessment development, additional QAPPs may be applied for quality assurance management. They include:

Title	Document Number	Date
Program Quality Assurance Project Plan (PQAPP) for the Integrated Risk Information System (IRIS) Program	L-CPAD-0030729-QP-1-5	June 2022
An Umbrella Quality Assurance Project Plan (QAPP) for Dosimetry and Mechanism-Based Models (PBPK)	L-CPAD-0032188-QP-1-2	December 2020
Quality Assurance Project Plan (QAPP) for Enhancements to Benchmark Dose Software (BMDS)	L-HEEAD-0032189-QP-1-2	September 2020
Umbrella Quality Assurance Project Plan for CPHEA PFAS Toxicity Assessments	L-CPAD-0031652-QP-1-4	October 2021

During assessment development, this project underwent four quality audits during assessment development including:

Date	Type of audit	Major findings	Actions taken
August 2022	Technical System Audit	No findings	None
July 2021	Technical System Audit	No findings	None
August 2020	Technical System Audit	No findings	None
August 2019	Technical System Audit	No findings	None

During Step 3 of the IRIS Process, the IRIS Toxicological Review was subjected to external reviews by other federal agency partners including the Executive Offices of the White House. Comments during these IRIS Process steps are available in the Docket EPA-HQ-ORD-2020-0675 on http:/www.regulations.gov.

During Step 4 assessment development, the *IRIS Toxicological Review of Perfluorobutanoic Acid and Related* underwent public comment from August 23, 2021, to November 8, 2021. Following this comment period, the toxicological review underwent external peer review by a contractor-led panel performed by ERG from October 2021 to June 2022. The peer-review report is available on the peer review website. All public and peer-review comments are available in the docket EPA-HQ-ORD-2020-0675.

Prior to release (Step 7 of the IRIS process), the final toxicological review is submitted to management and QA clearance. During this step the CPHEA QA Director and QA Managers review the project QA documentation and ensure that EPA QA requirements are met.

REFERENCES

Abraham, K; El-Khatib, AH; Schwerdtle, T; Monien, BH. (2021). Perfluorobutanoic acid (PFBA): No high-level accumulation in human lung and kidney tissue. Int J Hyg Environ Health 237: 113830. http://dx.doi.org/10.1016/j.ijheh.2021.113830.

Bangma, JT; Reiner, J; Fry, RC; Manuck, T; Mccord, J; Strynar, MJ. (2021). Identification of an Analytical Method Interference for Perfluorobutanoic Acid in Biological Samples. Environ Sci Technol Lett 8: 1085-1090. http://dx.doi.org/10.1021/acs.estlett.1c00901.

Bassett, JHD; Clement-Lacroix, P; Samarut, J; Chassande, O; Williams, GR; O'Shea, PJ; Sriskantharajah, S; Rabier, B; Boyde, A; Howell, PGT; Weiss, R, oyE; Roux, JP; Malaval, L, uc. (2007). Thyroid hormone excess rather than thyrotropin deficiency induces osteoporosis in hyperthyroidism. Mol Endocrinol1095-1107. http://dx.doi.org/10.1210/me.2007-0033.

Bizzarro, MJ; Gross, I. (2004). Effects of hormones on fetal lung development [Review]. Obstet Gynecol Clin North Am 31: 949-961, xii. http://dx.doi.org/10.1016/j.ogc.2004.08.001.

Butenhoff, JL; Bjork, JA; Chang, SC; Ehresman, DJ; Parker, GA; Das, K; Lau, C; Lieder, PH; van Otterdijk, FM; Wallace, KB. (2012). Toxicological evaluation of ammonium perfluorobutyrate in rats: twenty-eight-day and ninety-day oral gavage studies. Reprod Toxicol 33: 513-530. http://dx.doi.org/10.1016/j.reprotox.2011.08.004.

Carlson, LA; Angrish, M; Shirke, AV; Radke, EG; Schulz, B; Kraft, A; Judson, R; Patlewicz, G; Blain, R; Lin, C; Vetter, N; Lemeris, C; Hartman, P; Hubbard, H; Arzuaga, X; Davis, A; Dishaw, LV; Druwe, IL; Hollinger, H; Jones, R; Kaiser, JP; Lizarraga, L; Noyes, PD; Taylor, M; Shapiro, AJ; Williams, AJ; Thayer, KA. (2022). Systematic evidence map for over one hundred and fifty per- and polyfluoroalkyl substances (PFAS) [Review]. Environ Health Perspect 130: 56001. http://dx.doi.org/10.1289/EHP10343.

Chang, S; Das, K; Ehresman, DJ; Ellefson, ME; Gorman, GS; Hart, JA; Noker, PE; Tan, Y; Lieder, PH; Lau, C; Olsen, GW; Butenhoff, JL. (2008). Comparative pharmacokinetics of perfluorobutyrate in rats, mice, monkeys, and humans and relevance to human exposure via drinking water. Toxicol Sci 104: 40-53. http://dx.doi.org/10.1093/toxsci/kfn057.

Cooper, DS. (2001). Clinical practice. Subclinical hypothyroidism [Review]. N Engl J Med 345: 260-265. http://dx.doi.org/10.1056/NEJM200107263450406.

Crofton, KM; Craft, ES; Hedge, JM; Gennings, C; Simmons, JE; Carchman, RA; Carter, WH, Jr; Devito, MJ. (2005). Thyroid-hormone-disrupting chemicals: Evidence for dose-dependent additivity or synergism. Environ Health Perspect 113: 1549-1554. http://dx.doi.org/10.1289/ehp.8195.

Crofton, KM; Zoeller, RT. (2005). Mode of action: neurotoxicity induced by thyroid hormone disruption during development--hearing loss resulting from exposure to PHAHs [Review]. Crit Rev Toxicol 35: 757-769. http://dx.doi.org/10.1080/10408440591007304.

Crump, KS. (1995). Calculation of benchmark doses from continuous data. Risk Anal 15: 79-89. http://dx.doi.org/10.1111/j.1539-6924.1995.tb00095.x.

Das, KP; Grey, BE; Zehr, RD; Wood, CR; Butenhoff, JL; Chang, SC; Ehresman, DJ; Tan, YM; Lau, C. (2008). Effects of perfluorobutyrate exposure during pregnancy in the mouse. Toxicol Sci 105: 173-181. http://dx.doi.org/10.1093/toxsci/kfn099.

Foreman, JE; Chang, SC; Ehresman, DJ; Butenhoff, JL; Anderson, CR; Palkar, PS; Kang, BH; Gonzalez, FJ; Peters, JM. (2009). Differential Hepatic Effects of Perfluorobutyrate Mediated by Mouse and Human PPAR-alpha. Toxicol Sci 110: 204-211. http://dx.doi.org/10.1093/toxsci/kfp077.

Gilbert, ME; O'Shaughnessy, KL; Axelstad, M. (2020). Regulation of Thyroid-disrupting Chemicals to Protect the Developing Brain [Review]. Endocrinology 161. http://dx.doi.org/10.1210/endocr/bqaa106.

Gilbert, ME; Ramos, RL; Mccloskey, DP; Goodman, JH. (2014). Subcortical band heterotopia in rat offspring following maternal hypothyroxinaemia: Structural and functional characteristics. J Neuroendocrinol 26: 528-541. http://dx.doi.org/10.1111/jne.12169.

Gilbert, ME; Sui, L. (2008). Developmental exposure to perchlorate alters synaptic transmission in hippocampus of the adult rat. Environ Health Perspect 116: 752-760. http://dx.doi.org/10.1289/ehp.11089.

Haddow, JE; Palomaki, GE; Allan, WC; Williams, JR; Knight, GJ; Gagnon, J; O'Heir, CE; Mitchell, ML; Hermos, RJ; Waisbren, SE; Faix, JD; Klein, RZ. (1999). Maternal thyroid deficiency during pregnancy and subsequent neuropsychological development of the child. N Engl J Med 341: 549-555. http://dx.doi.org/10.1056/NEJM199908193410801.

Hall, AP; Elcombe, CR; Foster, JR; Harada, T; Kaufmann, W; Knippel, A; Küttler, K; Malarkey, DE; Maronpot, RR; Nishikawa, A; Nolte, T; Schulte, A; Strauss, V; York, MJ. (2012). Liver hypertrophy: a review of adaptive (adverse and non-adverse) changes--conclusions from the 3rd International ESTP Expert Workshop [Review]. Toxicol Pathol 40: 971-994. http://dx.doi.org/10.1177/0192623312448935.

Jansen, TA; Korevaar, TIM; Mulder, TA; White, T; Muetzel, RL; Peeters, RP; Tiemeier, H. (2019). Maternal thyroid function during pregnancy and child brain morphology: a time window-specific analysis of a prospective cohort. Lancet Diabetes Endocrinol 7: 629-637. http://dx.doi.org/10.1016/S2213-8587(19)30153-6.

Korevaar, TIM; Muetzel, R; Medici, M; Chaker, L; Jaddoe, VWV; de Rijke, YB; Steegers, EAP; Visser, TJ; White, T; Tiemeier, H; Peeters, RP. (2016). Association of maternal thyroid function during early pregnancy with offspring IQ and brain morphology in childhood: A population-based prospective cohort study. Lancet Diabetes Endocrinol 4: 35-43. http://dx.doi.org/10.1016/S2213-8587(15)00327-7.

Lau, C; Thibodeaux, JR; Hanson, RG; Narotsky, MG; Rogers, JM; Lindstrom, AB; Strynar, MJ. (2006). Effects of perfluorooctanoic acid exposure during pregnancy in the mouse. Toxicol Sci 90: 510-518. http://dx.doi.org/10.1093/toxsci/kfj105.

Levie, D; Korevaar, TIM; Bath, SC; Dalmau-Bueno, A; Murcia, M; Espada, M; Dineva, M; Ibarluzea, JM; Sunyer, J; Tiemeier, H; Rebagliato, M; Rayman, MP; Peeters, RP; Guxens, M. (2018). Thyroid Function in Early Pregnancy, Child IQ, and Autistic Traits: A Meta-Analysis of Individual Participant Data. J Clin Endocrinol Metab 103: 2967-2979. http://dx.doi.org/10.1210/jc.2018-00224.

Li, Y; Fletcher, T; Mucs, D; Scott, K; Lindh, CH; Tallving, P; Jakobsson, K. (2018). Half-lives of PFOS, PFHxS and PFOA after end of exposure to contaminated drinking water. Occup Environ Med 75: 46-51. http://dx.doi.org/10.1136/oemed-2017-104651.

Mochizuki, K; Yagi, E; Sakaguchi, N; Mochizuki, H; Takabe, S; Kuranuki, S; Suzuki, T; Shimada, M; Goda, T. (2007). The critical period for thyroid hormone responsiveness through thyroid hormone receptor isoform alpha in the postnatal small intestine. Biochim Biophys Acta 1770: 609-616. http://dx.doi.org/10.1016/j.bbagen.2006.12.011.

Nakamizo, M; Toyabe, S; Asami, T; Akazawa, K. (2007). Mental development of infants with congenital hypothyroidism: a longitudinal study. Clin Pediatr (Phila) 46: 53-58. http://dx.doi.org/10.1177/0009922806289326.

Oerbeck, B. (2007). Young adults with severe congenital hypothroidism: Cognitive event related potentials (ERPs) and the significance of an early start of thyroxine treatment. Scand J Psychol 48: 61-67. http://dx.doi.org/10.1111/j.1467-9450.2006.00545.x.

Page, M; McKenzie, J; Bossuyt, P; Boutron, I; Hoffmann, T; Mulrow, Cea. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ Acad Ed 372: 1-9. http://dx.doi.org/10.1136/bmj.n71.

Pelch, KE; Reade, A; Kwiatkowski, CF; Merced-Nieves, FM; Cavalier, H; Schultz, K; Wolffe, T; Varshavsky, J. (2022). The PFAS-Tox Database: A systematic evidence map of health studies on 29 per- and polyfluoroalkyl substances. Environ Int 167: 107408. http://dx.doi.org/10.1016/j.envint.2022.107408.

Pérez, F; Nadal, M; Navarro-Ortega, A; Fàbrega, F; Domingo, JL; Barceló, D; Farré, M. (2013). Accumulation of perfluoroalkyl substances in human tissues. Environ Int 59: 354-362. http://dx.doi.org/10.1016/j.envint.2013.06.004.

Radke, E; Wright, MJ; Christensen, K; Lin, CJ; Goldstone, AE; Glenn, B; Thayer, K. (2022). Epidemiology evidence for health effects of 150 per- and polyfluoroalkyl substances: A systematic evidence map [Review]. Environ Health Perspect 130: 96003. http://dx.doi.org/10.1289/EHP11185.

Russell, MH; Himmelstein, MW; Buck, RC. (2015). Inhalation and oral toxicokinetics of 6:2 FTOH and its metabolites in mammals. Chemosphere 120: 328-335. http://dx.doi.org/10.1016/j.chemosphere.2014.07.092.

Stagnaro-Green, A; Rovet, J. (2016). Pregnancy: Maternal thyroid function in pregnancy - A tale of two tails [Comment]. Nat Rev Endocrinol 12: 10-11. http://dx.doi.org/10.1038/nrendo.2015.212.

U.S. EPA (U.S. Environmental Protection Agency). (1991). Guidelines for developmental toxicity risk assessment. Fed Reg 56: 63798-63826.

U.S. EPA (U.S. Environmental Protection Agency). (1998). Guidelines for neurotoxicity risk assessment [EPA Report] (pp. 1-89). (ISSN 0097-6326
EISSN 2167-2520
EPA/630/R-95/001F). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. http://www.epa.gov/risk/guidelines-neurotoxicity-risk-assessment.

U.S. EPA (U.S. Environmental Protection Agency). (2002). A review of the reference dose and reference concentration processes. (EPA630P02002F). Washington, DC. https://www.epa.gov/sites/production/files/2014-12/documents/rfd-final.pdf.

U.S. EPA (U.S. Environmental Protection Agency). (2005). Guidelines for carcinogen risk assessment [EPA Report]. (EPA630P03001F). Washington, DC. https://www.epa.gov/sites/production/files/2013-09/documents/cancer_guidelines_final_3-25-05.pdf.

U.S. EPA (U.S. Environmental Protection Agency). (2011). Recommended use of body weight 3/4 as the default method in derivation of the oral reference dose. (EPA100R110001). Washington, DC. https://www.epa.gov/sites/production/files/2013-09/documents/recommended-use-of-bw34.pdf.

U.S. EPA (U.S. Environmental Protection Agency). (2012). Benchmark dose technical guidance [EPA Report]. (EPA100R12001). Washington, DC: U.S. Environmental Protection Agency, Risk Assessment Forum. https://www.epa.gov/risk/benchmark-dose-technical-guidance.

U.S. EPA (U.S. Environmental Protection Agency). (2015). Peer review handbook [EPA Report] (4th ed.). (EPA/100/B-15/001). Washington, DC: U.S. Environmental Protection Agency, Science Policy Council. https://www.epa.gov/osa/peer-review-handbook-4th-edition-2015.

U.S. EPA (U.S. Environmental Protection Agency). (2021). Human health toxicity values for hexafluoropropylene oxide (HFPO) dimer acid and its ammonium salt (CASRN 13252-13-6 and CASRN 62037-80-3). Also known as "GenX chemicals." Final report [EPA Report]. (EPA-822R-21-010). Washington, DC: U.S. Environmental Protection Agency, Office of Water.

van Otterdijk, FM. (2007a). Repeated dose 28-day oral toxicity study with MTDID-8391 by daily gavage in the rat, followed by a 21-day recovery period. (Study Number 06-226). Maplewood, MN: 3M.

van Otterdijk, FM. (2007b). Repeated dose 90-day oral toxicity study with MTDID 8391 by daily gavage in the rat followed by a 3-week recovery period. (Study Number 06-398). Maplewood, MN: 3M.

Weatherly, LM; Shane, HL; Lukomska, E; Baur, R; Anderson, SE. (2021). Systemic toxicity induced by topical application of heptafluorobutyric acid (PFBA) in a murine model. Food Chem Toxicol 156: 112528. http://dx.doi.org/10.1016/j.fct.2021.112528.

Wexler, JA; Sharretts, J. (2007). Thyroid and bone [Review]. Endocrinol Metab Clin North Am 36: 673-705, vi. http://dx.doi.org/10.1016/j.ecl.2007.04.005.

Zoeller, RT; Rovet, J. (2004). Timing of thyroid hormone action in the developing brain: Clinical observations and experimental findings [Review]. J Neuroendocrinol 16: 809-818. http://dx.doi.org/10.1111/j.1365-2826.2004.01243.