EPA/600/R-20/345R

Response to Peer Review Comments on the Draft Human Health Toxicity Values for Perfluorobutane Sulfonic Acid (CASRN 375-73-5) and Related Compound Potassium Perfluorobutane Sulfonate (CASRN 29420-49-3)

October 2020

Prepared by:
U.S. Environmental Protection Agency
Office of Research and Development
Center for Public Health and Environmental Assessment

DISCLAIMER

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

CONTENTS

ACRONYMS

Acronyms are not consistently defined throughout this document, as much of the text was extracted in its original format from charge questions and reviewer comments. This table provides acronym definitions.

BMD	benchmark dose
BMDL	benchmark dose lower limit
BMR	benchmark response
BW	body weight
BWa	body weight animal
BWh	body weight human
CASRN	Chemical Abstracts Service Registry Number
DAF	dosimetric adjustment factor
ELISA	enzyme-linked immunosorbent assay
EPA/USEPA	U.S. Environmental Protection Agency
FT4	free thyroxine
HAWC	Health Assessment Workspace Collaborative
HED	human equivalent dose
HERO	Health & Environmental Research Online
K+	potassium salt
kg	kilogram
LOAEL	lowest observed adverse effect level
µl	microliter
µM	micromole
mg/kg/day	milligrams per kilogram per day
NOAEL	no observed adverse effect level
NTP	National Toxicology Program
PBPK	physiologically based pharmacokinetic
PECO	populations, comparators, exposures, and outcomes
PFAS	per- and polyfluoroalkyl substances
PFBS	perfluorobutane sulfonic acid
PFOA	perfluorooctanoic acid
PFOS	perfluorooctane sulfonic acid
PND	postnatal day

POD	point of departure
POD_{HED}	point of departure human equivalent dose
RfD	reference dose
rT3	reverse total triiodothyronine
SEM	standard error of measurement
$T_{1/2}$	half-life
T3/TT3	total triiodothyronine
T4/TT4	total thyroxine
TSH	thyroid-stimulating hormone
UF	uncertainty factor
UF_A	interspecies uncertainty factor
UF_D	database uncertainty factor
UF_H	intraspecies uncertainty factor
UF_L	LOAEL to NOAEL extrapolation uncertainty factor
UF_S	extrapolation from subchronic to a chronic exposure duration uncertainty factor

INTRODUCTION

This document was prepared under the U.S. Environmental Protection Agency (EPA) Contract Number EP-C-17-017, Task Order 0008 with Eastern Research Group, Inc. Seven independent external peer reviewers reviewed the draft assessment (five of seven reviewers also reviewed the previous external review draft [circa July-August 2018]), and their comments are presented with the EPA's responses under each respective charge question. Appendix A includes the full comments from each of the reviewers.

The EPA is issuing draft subchronic and chronic oral toxicity values (i.e., reference doses, or RfDs) for perfluorobutane sulfonic acid (PFBS) (Chemical Abstracts Service Registry Number [CASRN] 375-73-5) and the related compound potassium perfluorobutane sulfonate (K^+PFBS) (CASRN 29420-49-3) for Interagency and agency comment. The EPA is publishing these toxicity values to facilitate decision-making by the Agency's programmatic, regional, and/or state partners associated with contamination concerns in a variety of exposure scenarios when they are finalized. The EPA developed this toxicity assessment to provide the health effects information used as the basis for derivation of these RfDs for PFBS.

The oral exposure database used to derive these RfDs for PFBS and its potassium salt includes multiple short-term and subchronic-duration toxicity studies in rats or mice, a two-generation reproductive toxicity study in rats, and multiple developmental toxicity studies in rats or mice. Information identifying health effects from inhalation exposure was not located, and dermal studies of PFBS exposure are limited. Further, no PFBS studies evaluating potential cancer effects were identified for any route of exposure. Thus, the PFBS assessment applies only to noncancer health outcomes via the oral route of exposure. Health outcomes evaluated across available oral PFBS studies include effects on the thyroid (decreased thyroid hormones such as triiodothyronine [T3], free thyroxine [FT4], total T4 [T4], and thyroid stimulating hormone [TSH]), reproductive organs, tissues, and health (decreased maternal feed consumption, body-weight (BW) gain, and gravid uterine weight), developing offspring (delayed eye opening, vaginal opening, final estrous, and decreased BW in pups), kidneys (increased kidney weight and histopathological foci [e.g., hyperplasia and focal papillary edema]), liver (increased liver weight), and lipids and lipoproteins (decreased hepatic lipase and triglycerides).

Across the body of evidence supporting hazards via the oral exposure route and across all life stages evaluated, the thyroid was identified as the most sensitive target of PFBS toxicity. Specifically, decreased thyroid hormone levels (i.e., total thyroxine [T4]) in newborn mice was identified as the critical effect from a single generation developmental study (Feng et al., 2017). Dose-response for this effect in newborn mice served as the basis for identification of a point-of-departure and derivation of a subchronic and chronic RfD.

The subchronic RfD for K^+PFBS was calculated by dividing the POD_{HED} for decreased serum total T4 observed in newborn (PND 1) mice by a composite uncertainty factor (UF_C) of 100 to account for extrapolation from mice to humans (an interspecies UF, or UF_A, of 3), for interindividual differences in human susceptibility (intraspecies UF, or UF_H, of 10), and for deficiencies in the toxicity database (database UF, or UF_D, of 3) (a value of 1 was applied for subchronic-to-chronic UF, or UF_S, and LOAEL-to-NOAEL UF, or UF_L) (see Table 10), yielding a subchronic RfD of 2×10^{-3} mg/kg-day. As K^+PFBS is fully dissociated in water at the environmental pH range of 4−9, data for K^+PFBS were used to derive a subchronic RfD for the free acid (PFBS) by adjusting for differences in molecular weight (MW) between K^+PFBS (338.19) and PFBS (300.10), yielding the value of 1×10^{-3} mg/kg-day for the subchronic RfD for PFBS (free acid).

The chronic RfD for K^+PFBS was calculated by dividing the POD_{HED} for decreased serum total T4 observed in newborn (PND 1) mice by a UF_C of 300 to account for extrapolation from mice to humans (UF_A of 3), for interindividual differences in human susceptibility (UF_H of 10), and deficiencies in the toxicity database (UF_D of 10) (a value of 1 was applied for UF_S and UF_L) (see Table 12), yielding a chronic RfD of 5×10^{-4} mg/kg-day. Like the derivation of the subchronic RfD, based on the data for K^+PFBS, a chronic RfD for PFBS (free acid) of 3×10^{-4} mg/kg-day was derived.

Overall, the peer reviewers agreed with the EPA's decisions regarding the:

- choice of the Feng et al. (2017) developmental mouse study as the principal study;
- choice of decreased total thyroxine (T4) in newborn (PND1) mice as the critical effect;
- benchmark dose modeling;
- use of a data-derived dosimetric adjustment approach to calculate human equivalent doses; and
- UF application

Some peer reviewers identified one topic specifically pertaining to uncertainty in interspecies extrapolation (i.e., suggested an increase in UF_A). This reviewer opinion was based on limited human toxicokinetic data in sensitive subpopulations (e.g., pregnancy, children, neonates) and lack of information to quantify relative cross-species sensitivity in thyroid hormone toxicodynamics. The reviewers also provided minor comments primarily regarding textual clarifications in Chapter 6, including:

- Need for additional clarifying language on the physiology and function of thyroid hormones during pregnancy and further description of the clinical condition "hypothyroxinemia";
- Explanation for the dose-response modeling using dose group sizes based on number of litters and fetuses in the principal study;
- Status of the *in press* manuscript providing mouse half-life data used to support calculation of the data-derived dosimetric adjustment factors;
- Need for further acknowledgment of the residual uncertainty in intra- and interspecies variability;
- Additional references proposed for consideration pertaining to:
    - human observational studies of PFAS and thyroid hormone alterations
    - associations between clinical hypothyroid conditions and neurobehavioral outcomes in progeny
    - cross-species comparison in HPT-axis physiology

The comment regarding interspecies uncertainty described above and other comments including minor comments and editorial suggestions were reviewed and are addressed directly in the draft assessment. Specific responses to major comments are provided under each respective charge question below. In consideration of the external peer reviewers' comments, the PFBS toxicity assessment was revised and is being released for interagency and agency review and comment.

SECTION I: TECHNICAL CHARGE TO EXTERNAL REVIEWERS

Technical Charge to External Peer Reviewers

Contract No. EP-C-17-017
Task Order 68HERH20F0097 (ERG Task Order 37)
February 2020

External Letter Peer Review of EPA's Draft Human Health Toxicity Values Assessment for Perfluorobutane Sulfonate (PFBS) (CASRN 375-73-5 [acid]) and Related Compound Potassium PFBS (CASRN 29420-49-3)

BACKGROUND

EPA has revised the draft PFBS assessment in response to public comments and relevant new data. EPA is requesting a second external peer review of the substantive changes made in the revised draft.

In the draft assessment released for public comment, the EPA stated that the weight-of-evidence for thyroid or kidney effects following oral PFBS exposure supports both of these effect domains as hazards. EPA presented both cases for RfD derivation in the public comment draft assessment and solicited feedback from public commenters on the suitability, or not, for each case. Based on the consistency and coherence in thyroid effects evidence in rats and mice across different lifestages, dose-response sensitivity of this effect domain compared to other candidate hazards, as well as consideration of public comment, thyroid effects were identified as the critical effect for derivation of subchronic and chronic RfDs in the revised draft assessment. The thyroid effect, more specifically decreased total thyroxine (T4) chosen as the critical effect for RfD derivation, is consistent with the human clinical condition known as 'hypothyroxinemia' where decreases in thyroid hormone (e.g., total T4) occur in the absence of reflex increases in TSH and thyroid tissue alterations (e.g., weight, histology), in contrast to traditional hypothyroidism.

CHARGE QUESTIONS

1. The key study chosen for determining the subchronic and chronic RfDs is the gestational exposure mouse study by Feng et al. (2017) and the critical effect is decreased total T4 in postnatal Day 1 (PND1) offspring.

    1a. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and clearly described?

        i. If so, please explain your reasoning.
        ii. If not, please provide your rationale and identify an alternative key study and/or critical effect to support the derivation of the subchronic and chronic RfDs and provide the scientific support for the alternative choice.

    1b. Is the selection of total T4 an appropriate biomarker/metric as it relates to clinically relevant hypothyroxinemia during pregnancy? Is such a measure applicable to both experimental test animals and humans?

        i. If so, please explain your rationale.
        ii. If not, are there other measures related to hypothyroxinemia that may be more useful for informing hazard potential during pregnancy? What are those measures?

    1c. Has EPA clearly articulated the challenges associated with extrapolating the PFBS-induced decrease in thyroid hormone (e.g., total T4) in rodents to humans?

        i. If so, please explain your reasoning.
        ii. If not, please provide your rationale.

1d. Has EPA clearly articulated what is and is not known about the clinical implications of changes in T4 in women during pregnancy on neonates and infants?
    i. If so, please explain your reasoning.
    ii. If not, please provide your rationale.

2. In the public review draft PFBS assessment, EPA employed benchmark dose modeling (U.S. EPA, 2012) in the identification of a point-of-departure (POD) for derivation of RfD values, based on a decrease in total T4 levels in PND1 offspring. The 20% Relative Deviation (RD) Benchmark Response Rate (BMR) used in the public review draft is no longer being considered for BMD modeling of thyroid hormone dose-response data. As a result of extensive public review comments on the BMD approach (and thyroid hormone endpoint) used in the previous draft, and because a clear or consistent biological threshold for T4 changes associated with untoward developmental health outcomes has not be identified in the available literature, EPA has identified a new BMR of 0.5 SD (standard deviation change over controls) as a default in the revised PFBS draft assessment for the thyroid hormone alterations in mouse neonates/offspring. A 1 SD BMR is also being presented as the standardized basis for comparison as recommended in the EPA BMD Technical Guidance (U.S. EPA, 2012).

    2a. Are the dose-response modeling approaches, selection of benchmark response rate (BMR), and the selected models used to identify the thyroid effect-related POD for RfD derivation scientifically justified and clearly described?
        i. If so, please explain your reasoning.
        ii. If not, please provide your rationale and identify alternative approaches, BMRs, and/or dose-response models that support the identification of alternative candidate POD(s) for the derivation of subchronic and chronic RfDs and provide the scientific support for the alternative choice(s).

3. Due to the availability of new toxicokinetic data in mice noting significant interspecies differences in toxicokinetics of PFBS, and, recommendations from public commenters, EPA has applied a data-informed approach to convert the oral dose-rate in animals to a human equivalent dose (HED) in the identification of candidate points-of-departure (PODs) considered for the derivation of the RfDs (U.S. EPA, 2014; see https://www.epa.gov/sites/production/files/2015-01/documents/ddef-final.pdf). In considering the new evidence for serum half-life in mice published in Lau (in press), EPA concluded that the toxicokinetic data for PFBS are adequate to support calculation of data-derived dosimetric adjustment factors (DAF), where the ratio of elimination half-life in animals to that in humans, $T0.5_A/T0.5_H$, is used to adjust candidate PODs. By using in vivo animal and human half-life data to calculate POD(HEDs) that account for differences in toxicokinetics between rodents and humans, the potential role of interspecies differences in processes such as renal resorption, hepatic transport, and enterohepatic recirculation is reflected. Further, by using a data-derived approach the uncertainty in interspecies toxicokinetic scaling (UF_A) has been reduced from a 3 to a 1; however, residual uncertainty (due to the lack of information) pertaining to toxicodynamics exists and is acknowledged in the assessment in the description for applying a UF_A of 3.

    3a. Is applying the data-informed dosimetric adjustment that utilizes the ratio of the PFBS elimination half-life in mice to that in the human scientifically justified and clearly described?
        i. If so, please explain your reasoning.
        ii. If not, please provide your rationale and identify an alternative approach to scale PFBS doses between rodents and humans and provide scientific support for the alternative choice.

4. 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 PFBS.

4a. Does the provided qualitative scientific rationale support the application of the selected uncertainty factors? If not, please explain.

4b. Has quantitative uncertainty been adequately accounted for in the derivation of the RfDs? Please describe and provide suggestions, if needed.

4c. Do the methods used to derive the RfDs for PFBS appropriately account for uncertainties in evaluating the toxicokinetic differences between the experimental animal data and humans? If not, please explain.

SECTION II: REVIEWER COMMENTS ORGANIZED BY CHARGE QUESTION

CHARGE QUESTION 1

1. **The key study chosen for determining the subchronic and chronic RfDs is the gestational exposure mouse study by Feng et al. (2017) and the critical effect is decreased total T4 in postnatal Day 1 (PND1) offspring.**
   - **1a. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and clearly described?**
     - **i. If so, please explain your reasoning.**
     - **ii. If not, please provide your rationale and identify an alternative key study and/or critical effect to support the derivation of the subchronic and chronic RfDs and provide the scientific support for the alternative choice.**

Chou

The key studies are appropriately selected, based on the existing data and critical effects are appropriately identified. The reasoning process of evaluating the quality of the studies, data uncertainties in the existing studies, as well as the validity of the rodent model for the thyroid function in human are thoroughly considered and presented in the document.

> **EPA Response: The reviewer agrees with the selection of the key study and critical effect. No revisions needed to address this comment.**

Hattis

The key study of Feng et al. (2017) is reasonable enough, as far as it goes. However, the paper reports only a single experimental run on a single group of mice. This is a little thin as a basis of a U.S. national regulatory action.

Recently published human epidemiological observations bolster the evidence. Recently epidemiological observations of Reardon et al. (2019)* in pregnant Canadian women have indicated an inverse association between serum perfluoroalkyl acids and lower FT4 levels, supporting the basis for concern for human exposures to perfluoroalkyl acids.

The full abstract for this paper is:

*Longitudinal Analysis Reveals Early-Pregnancy Associations Between Perfluoroalkyl Sulfonates and Thyroid Hormone Status in a Canadian Prospective Birth Cohort. Environ Int Vol. 129 pp. 389-399. Aug 2019

Anthony J F Reardon 1, Elham Khodayari Moez 2, Irina Dinu 2, Susan Goruk 3, Catherine J Field 3, David W Kinniburgh4, Amy M MacDonald4, Jonathan W Martin5, APrON Study

Affiliations expand

PMID: 31150980

PMCID: PMC6859374 (available on 2020-08-01)

DOI: 10.1016/j.envint.2019.04.023

Abstract

Serum perfluoroalkyl acids (PFAAs) have been linked to disruption of maternal thyroid hormone homeostasis, but results have varied between studies which we hypothesized was due to timing of the thyroid hormone measurements, variability in PFAA isomer patterns, or presence of other stressors. In a longitudinal study design, we investigated the time-dependency of associations between PFAA isomers and thyroid hormones during pregnancy and post-partum while considering thyroid peroxidase antibody (TPOAb) status and mercury (Hg) co-exposure. In participants of a prospective Canadian birth cohort (n = 494), free thyroxine (FT4), free triiodothyronine (FT3), thyroid stimulating hormone (TSH) and TPOAb were quantified in maternal plasma collected in each trimester and 3-months postpartum, and 25 PFAAs (15 linear and 10 branched) and Hg were quantified in samples collected during the second trimester. Perfluorohexane sulfonate (PFHxS) and total branched isomers of perfluorooctane sulfonate (PFOS) were positively associated with TSH in mixed-effect models, with strongest associations early in gestation. Throughout pregnancy and post-partum, PFHxS was inversely associated with FT4, consistent with elevated TSH, while Hg was inversely associated with FT3. In TPOAb-positive women, negative associations were found between PFUnA and FT4, and 1m-PFOS and TSH, supporting previous studies that thyroid disorder could increase susceptibility to PFAA-mediated hormone dysregulation. Hg did not confound associations but was a significant interaction term, revealing further positive associations between PFOS isomers (∑3m+4m-PFOS) and TSH. Higher perfluoroalkyl sulfonate exposures were associated with higher TSH and/or lower FT4, strongly suggestive that PFHxS and branched PFOS isomers are risk factors for subclinical maternal hypothyroidism. Isomer-specific analysis is important in future studies, as crude measures of 'total-PFOS' masked the associations of branched isomers. A concerning result was for PFHxS which had consistent negative associations with FT4 at all time points and a positive association with TSH in early pregnancy when fetal development is most sensitive to disruption.

Keywords: Longitudinal study design; Perfluoroalkyl acids; Perfluoroalkyl carboxylates; Perfluoroalkyl sulfonates; Pregnancy; Thyroid hormones.

I did a literature search and identified the following papers as potentially helpful for further study:

- Perfluoroalkyl and Polyfluoroalkyl Substances and Measures of Human Fertility: A Systematic Review Cathrine Carlsen Bach 1 2, Anne Vested 3 4, Kristian Tore Jørgensen 5, Jens Peter Ellekilde Bonde 5, Tine Brink Henriksen 1 6, Gunnar Toft 7 DOI: 10.1080/10408444.2016.1182117

Abstract

Perfluoroalkyl and polyfluoroalkyl substances (PFASs) are found widespread in the environment and humans. The relation of PFASs to fertility has now been examined in a relatively large number of epidemiologic studies and a synthesis is in order. The aim of this study was to assess the current human epidemiologic evidence on the association between exposure to PFASs and measures of human fertility, with particular emphasis on perfluorooctane sulfonate (PFOS) and perfluorooctanoate (PFOA). Systematic literature searches were initially conducted in MEDLINE and EMBASE and subsequently in references and citations of included papers. Studies were included if they assessed exposure to PFASs in biological samples in relation to reproductive hormones, semen characteristics, or time to pregnancy (TTP). Study characteristics and results were abstracted to predefined forms, and the studies were assessed for the risk of bias and confounding. Sixteen studies investigated the association between PFAS exposure in men and semen parameters, reproductive hormone levels, or TTP. There was a lack of consistent results among the numerous investigated exposure-outcome combinations. However, subtle associations between higher PFOS and lower testosterone or abnormal semen morphology cannot be excluded. Eleven studies assessed the association between PFAS exposure in women and TTP or reproductive hormones levels. Four of eight studies found prolonged TTP

with higher PFOS or PFOA, but only one study found an association when restricting to nulliparous women. In men, there is little evidence of an association between PFAS exposure and semen quality or levels of reproductive hormones. For PFOS and PFOA, the literature indicates an association with female fecundability in parous women, which is most likely not causal.

Keywords: Epidemiology; fecundability; fecundity; fertility; humans; perfluorinated compounds; perfluoroalkyl and polyfluoroalkyl substances; perfluorooctane sulfonate; perfluorooctanoate; semen quality; time to pregnancy.

Similar articles:

- Perfluoroalkyl and polyfluoroalkyl substances and measures of human fertility: a systematic review. Bach CC, et al. Crit Rev Toxicol 2016 - Review. PMID 27268162
- Perfluoroalkyl and polyfluoroalkyl substances and human fetal growth: a systematic review. Bach CC, et al. Crit Rev Toxicol 2015 - Review. PMID 25372700
- Perfluoroalkyl substances and time to pregnancy in couples from Greenland, Poland and Ukraine. Jørgensen KT, et al. Environ Health 2014. Among authors: Bach CC. PMID 25533644 Free PMC article.
- Maternal Exposure to Perfluorinated Chemicals and Reduced Fecundity: The MIREC Study. MP Vélez et al. Hum Reprod 30 (3), 701-9. Mar 2015. PMID 25567616.

  The cumulative probabilities of pregnancy at 1, 6 and 12 months were 0.42 (95% confidence interval (CI) 0.40-0.45), 0.81 (95% CI 0.79-0.83) and 0.90 (95% CI 0.89-0.92), …
- Association of Perfluoroalkyl and Polyfluoroalkyl Substances With Premature Ovarian Insufficiency in Chinese Women. S Zhang et al. J Clin Endocrinol Metab 103 (7), 2543-2551. 2018. PMID 29986037.

  High exposure to PFOA, PFOS, and PFHxS is associated with increased risk of POI in humans.
- Prenatal Exposure to Perfluoroalkyl Substances and Birth Outcomes in a Spanish Birth Cohort. CB Manzano-Salgado et al. Environ Int 108, 278-284. Nov 2017. PMID 28917208.

  In this study, PFAS showed little association with birth outcomes. Higher PFHxS, PFOA, and PFNA concentrations were non-significantly associated with reduced birth weight …
- Exposure to Perfluoroalkyl Substances and Thyroid Function in Pregnant Women and Children: A Systematic Review of Epidemiologic Studies. V Ballesteros et al. Environ Int 99, 15-28. Feb 2017. PMID 27884404.

  Although there is a small number of studies with comparable data, we found some consistency of a positive association between maternal or teenage male exposure to some PF …
- Profiles of Emerging and Legacy Per-/Polyfluoroalkyl Substances in Matched Serum and Semen Samples: New Implications for Human Semen Quality. Y Pan et al. Environ Health Perspect 127 (12), 127005. Dec 2019. PMID 31841032.

  Our results suggest the potential for deleterious effects following exposure to 6:2 Cl-PFESA and other PFASs. Compared with serum PFAS levels, the much clearer association…
- Toxicokinetics of 8:2 Fluorotelomer Alcohol (8:2-FTOH) in Male and Female Hsd:Sprague Dawley SD Rats After Intravenous and Gavage Administration. MC Huang et al. Toxicol Rep 6, 924-932. 2019. PMID 31516843.

  Fluorotelomer alcohols (FTOHs) are used in the production of persistent per- and polyfluorinated alkyl substances (PFAS). Rodents and humans metabolize FTOHs to …

- Early Pregnancy Serum Levels of Perfluoroalkyl Substances and Risk of Preeclampsia in Swedish Women. S Wikström et al. Sci Rep 9 (1), 9179. 2019. PMID 31235847.

  Preeclampsia is a major cause of maternal and fetal morbidity. Emerging research shows an association with environmental exposures. The present aim was to investigate …
- Exposure to Perfluoroalkyl Substances During Fetal Life and Pubertal Development in Boys and Girls From the Danish National Birth Cohort. A Ernst et al. Environ Health Perspect 127 (1), 17004. Jan 2019. PMID 30628845.

  Our population-based cohort study suggests sex-specific associations of altered pubertal development with prenatal exposure to PFASs. These findings are novel, and …
- Conditioning on Parity in Studies of Perfluoroalkyl Acids and Time to Pregnancy: An Example From the Danish National Birth Cohort. C Bach et al. Environ Health Perspect 126 (11), 117003. Nov 2018. PMID 30417653.

  Associations between PFAAs and TTP in parous women may be biased by confounders related to previous pregnancies and exposure measurement error. To avoid these biases, …

EPA Response: The reviewer agreed with the selection of Feng et al. (2017) as the key study however no comment was offered pertaining to the selection of decreased total T4 in offspring as the critical effect. The reviewer did question why a single study (Feng et al., 2017) could serve as the basis for a U.S. National Regulation. Importantly, the PFBS human health assessment does not represent a regulatory action but rather may in part inform risk remediation activities. The Feng et al. (2017) publication is a robust high-confidence mouse study that presents health outcomes consistent with a broader body of evidence demonstrating an exposure-effect relationship between oral PFBS and thyroid hormone perturbations. The identification of decreased thyroid hormone (e.g., total T4) as a hazard for oral PFBS exposure is based on the entirety of the relevant study landscape (i.e., across rats and mice, different sexes, different exposure durations and lifestages); from amongst this body of evidence, the Feng et al. (2017) publication provided the highest confidence dose-response dataset on which to base identification of a POD for RfD derivation.

The reviewer identified 13 additional human epidemiology publications (one study was a duplicate; Bach et al., 2016) and one rodent TK study of potential relevance to the draft PFBS assessment. Each publication was screened for direct relevance to PFBS hazard identification or dose-response assessment and/or to inform, in general, health outcome domains discussed in the draft assessment. Only one of the proposed studies contained information potentially relevant to the draft PFBS assessment (e.g., Pan et al., 2019) however upon closer review the levels of PFBS detected in human serum or sperm in the cohort were so low (many were below the level of quantitation for the assay) that the authors did not further evaluate this PFAS in the study; therefore, this publication was not integrated into the draft assessment. All other publications were focused on health outcomes associated with PFAS ≥ C6 (e.g., PFHxS, PFHpA, PFOA/PFOS, PFNA, PFDA, etc.), and as such were not considered further.

Kamendulis

The Feng et al., 2017 study was selected as the key study for the derivation of subchronic and chronic RfDs for PFBS, based on findings of PFBS-mediated decreases in total T3, total T4, and free T4. PFBS-induced alterations of the thyroid (decreases in total T3, total T4, and free T4) was selected as the critical effects, and was consistently observed across two species, sexes, life stages, and exposure durations in two independent, studies (NTP, 2019; Feng et al., 2017) that following systematic evaluation, were determined to be of "high-confidence". The Feng et al., 2017 study was a gestational study and identified adverse effects in PND1 thyroid that is considered appropriate for selection as the key study. The information pertaining to the study selection and identification of decreases in T4 as the critical effects have been clearly described and is scientifically based.

> **EPA Response: The reviewer agrees with the selection of the key study and critical effect. No revisions needed to address this comment.**

Leung

From the initial EPA draft assessment, the two organ systems demonstrating adverse effects from PFBS exposure with the highest level of confidence were the kidney and the thyroid gland. Table 6 of the current report summarizes the available studies regarding noncancer effects following oral PFBS administration. Unfortunately, there are no human pregnancy data in this area. Regarding animal data, the Feng et al 2017 and NTP 2019 studies both demonstrate the development of biochemical hypothyroidism following PFBS exposure. Between them, only the Feng et al 2017 mouse study examined this in mothers and their offspring, which are the vulnerable population subgroups of interest. Thus, I agree that it is the appropriate key study.

Measured thyroid biomarkers from the Feng et al 2017 were serum TSH, TT3, TT4, and FT4 in both dams and pups. Figure 4 in this study showed significant decreases in serum TT3 and TT4 levels at PNDs 1, 30, and 60 at the maternal 200 and 500 mg/kg/day PFBS doses (but not the 50 mg/kg/day dose) among pups. The paper does not report the pups' serum FT4 response, but presumably these data are available as per their methods section. The caveats between TT4 and FT4, and between rat and human thyroid physiology, should be noted, as outlined in my responses to questions 1b and 1c below. Taken together though, although not ideal, serum total T4 concentrations at Postnatal Day 1 would a reasonable critical effect from these animal data.

> **EPA Response: The reviewer agrees with the selection of the key study. Regarding the critical effect (decreased total T4 in PND1 mice), the reviewer suggests textual clarification in the assessment to address caveats pertaining to total and free T4 differences between rodents and humans. As suggested by the reviewer, this issue will be addressed in responses to same/similar comments provided under charges 1b and 1c below.**

Slitt

i. The selection of the total thyroxine, free thyroxine, and total triiodothyronine are well justified critical effects. The document explains very well the effects of thyroid hormone disruption on health endpoints and makes a solid justification for thyroid disruption as a critical concern. Thyroid hormone serves many

functions during development and throughout the life span. With regard to development, thyroid hormone is thought impact the neuronal, reproductive, hepatic, and immune system. It is also known to influence brain development. Feng et al. (2017) is the key study that describes decreased T4 in PND1 offspring as the critical effect. The study was considered to be of high quality based on the study design metrics evaluated. Strengths of the study were that it was well powered (n=10 dams per treatment), 3 doses included, appropriate statistical analysis, and additional endpoints measured. The work reports changes in both maternal and offspring T4 and TSH. Because this work reported decreased serum T4 in the offspring with a rebound increase in TSH, it is felt to be a clearer observation to use for alteration in thyroid hormone.

ii. I do support the observed decrease in serum T4 levels as observed in the NTP, 2019 as an alternative key critical effect, even with the lack of observed increase in serum TSH levels because thyroid hormone has pleiotropic effects, and decreased T4 is associated with numerous health poor outcomes. The NTP, 2019 study is rigorously described and provides a higher quality study to utilize despite the lack of rebound TSH. All aspects of the study are well described and documented.

EPA Response: The reviewer agrees with the selection of thyroid hormone decrements as the critical effect. In selecting the principal study and POD for RfD derivation, decreased total T4 in adult rats from the 28-day NTP (2019) study was considered as a candidate critical effect, however uncertainties associated with the adversity of decreased hormone levels in the absence of overt signs of hypothyroidism (e.g., thyroid gland weight; histopathology) in adults precluded selection over decreased total T4 in PND1 mice from Feng et al. (2017). Specifically, while NTP (2019) did observe profound decreases in total and free T4, as well as T3, they reported no significant changes in thyroid gland weight, histopathology, or TSH levels following 28-days of oral exposure in adult rats. Further, the adult thyroid has compensatory abilities not present in early life stages (e.g., larger thyroid hormone reserve capacity), making fetal/neonatal populations particularly sensitive to perturbations in thyroid hormone economy. As such, it is unclear what health risk(s) decreased thyroid hormones poses in adults (i.e., what are the health implications for adults, sans overt signs of clinical hypothyroidism?).

Warren

Yes, the revised draft clearly and thoroughly provides the scientific justification for selection of Feng et al. (2017) as principal study and decreased T4 in PND1 offspring as critical effect. Importantly, it explains the rationale behind these preferences over other candidate studies and effects (e.g., citing comparative sensitivity to the renal hyperplasia observed in adult rats (Lieder et al., 2009a,b) and questions about the biological significance of decreased T4 in adult rats in the absence of overt thyroid toxicity (NTP, 2019)). Compared to those in the original draft assessment (July 2018), the revised subchronic and chronic RfDs are one to two orders of magnitude lower. This is also the case when the revised RfDs are compared to the developmental RfD calculated in the original draft, despite use of the same principal study and critical effect. Thus, the revised toxicity values are more health conservative than those in the original draft, and more importantly, reflect a stricter adherence to U.S. EPA methodologies for toxicity value derivation.

EPA Response: The reviewer agrees with the selection of the key study and critical effect. No revisions needed to address this comment.

Zoeller

The Agency has clearly described the choice of Feng et al. (2017) as the key study and the decrease in serum total T4 in postnatal day 1 offspring as the critical effect. There were two general reasons for this. First, three hazards of PFBS exposure were identified, including serum total T4, renal toxicity and developmental. The thyroid endpoints were chosen both because there is more confidence that it represents a hazard to human health compared to the others, and because effects were observed at a lower dose. These considerations were very well described in the report.

One concern about the Feng et al. study is that serum T4 in P1 control pups is reported to be at the level of sensitivity of the assay as they report in the Methods section. This was not obvious since they reported the sensitivity in terms of ng/mL and report serum total T4 levels in figure 4B in terms of µg/dL. I was not able to obtain the specification sheets from the manufacturer to ensure that there was no error in reporting. A LOQ for mouse serum total T4 of 1.4 µg/dL is similar to other kits. However, if this is correct, it means that the measurement of "reduced" serum total T4 in treated animals would be below the LOQ.

The scientific justification was also well reasoned by the Agency. First, it is clear that thyroid hormone is chemically identical among all vertebrates. Thyroid hormone is essential for normal brain development in all mammals including mice and humans. Moreover, the Agency made cogent arguments both for the use of total T4 as the index of adverse effect and for choosing the neonatal period as being most relevant.

EPA Response: The reviewer agrees with the selection of the key study and critical effect but questioned the reporting of the total T4 (TT4) hormone levels from PND1 mice. The reviewer identified a potential discrepancy between the LOQ/sensitivity of the ELISA kit (ng/mL) used by Feng et al. (2017) and the units used to report the TT4 in Fig. 4B of that study (µg/dL); the reviewer further posited that if the units for TT4 in Fig. 4B are correct that reduced TT4 in PND1 mice would be below the LOQ for the assay. Upon re-examination of the Feng et al. (2017) study, and conversion of the presented data to a consistent unit of measure, there does not appear to be a discrepancy in the TT4 reported for PND1 mice in Fig.4B and the reported LOQ of the ELISA kit. For example, the high dose PFBS PND1 TT4 level in Fig. 4B is approximately 0.75 µg/dL which converts to 7.5 ng/mL which is above the LOQ for the ELISA assay. No revisions needed based on this comment.

1b. Is the selection of total T4 an appropriate biomarker/metric as it relates to clinically relevant hypothyroxinemia during pregnancy? Is such a measure applicable to both experimental test animals and humans?

i. **If so, please explain your rationale.**
ii. **If not, are there other measures related to hypothyroxinemia that may be more useful for informing hazard potential during pregnancy? What are those measures?**

Chou

The reviewer is not certain about the main point of this charge question. If the question is about total T4 vs. free T4, i.e. "Is total T4 an appropriate stand-alone biomarker?" Total T4 is selected as the critical effect in this assessment based on the collective evidence of TSH, total T4, and total T3; therefore, it is appropriately selected because it is supported by additional evidence of hypothyroidism. In addition, plasma-protein bound thyroid hormones is likely to be equally important as, if not more important than, the free thyroid hormones in the blood for the following reason: Due to the high lipophilicity, free T4 preferentially partition into lipid environment of membrane it first come in contact with, thus minimize its availability at other target cells (Rabah et al. 2019). Plasma protein binding serves as a distributing vehicle to ensure the availability of T4 at target cells and to prevent excessive free T4 in the blood circulation.

If the question is about whether total T4 is clinically important in diagnosing hypothyroxinemia in pregnancy, yes, it is an important and relevant test in pregnancy associated hypothyroidism. The level of total T4 is consider together with the levels of TSH, thyroperoxidase antibodies and other parameters for differential diagnosis of the etiology of clinical hypothyroidism in pregnancy.

> **EPA Response: The reviewer agrees with the identification of total T4 as an appropriate metric for clinically relevant hypothyroxinemia during pregnancy. The reviewer also agrees with total T4 as a measure applicable to both experimental test animals and humans, within the context of the overall hormone economy landscape associated with hypothyroidism.**

Hattis

Yes, and

Yes. T4 is a reasonable indicator that is often affected by chemicals thought to be important in influencing thyroid function.

> **EPA Response: The reviewer agrees with the selection of total T4 as a cross-species- and clinically- relevant thyroid hormone measure. No revisions needed to address this comment.**

Kamendulis

In general, I agree that T4 is an appropriate biomarker to be used to derive RfDs since decreases in this parameter (coupled with normal TSH levels) are clinically relevant to hypothyroxinemia in pregnancy. During development, many organ systems are affected by altered thyroid homeostasis as the maintenance of adequate thyroid hormone levels are needed for their normal growth and development. As described in the document, rodents are considered to be a good model for evaluating the potential effects of chemicals on thyroid function in humans (Zoeller et al., 2007), and the pattern of decreased thyroid hormones in the absence of TSH changes and thyroid tissue weight and/or histology, observed in PFBS studies (e.g., (Feng et al., 2017), are consistent with the human clinical condition referred to as "hypothyroxinemia". The document could be more specific however, in stating that this is a clinical condition observed in *human pregnancy* in section 5. The evidence and data presented in section 6 clearly provides support for the clinical relevance of hypothyroxinemia during pregnancy and its relevance to developmental outcomes in both animals and humans.

EPA Response: The reviewer agrees with the selection of total T4 as a cross-species- and clinically- relevant thyroid hormone measure. As per the reviewer's suggestion, a qualifying statement was added to text at the end of section 5.1 to clarify that hypothyroxinemia is commonly associated with pregnancy in humans.

Leung

It is noted that this question refers to serum TT4 levels among *offspring* of exposed mothers. There are two points to address in regard to this question:

1. Offspring T4 versus T3: It should be clarified that although it may be minimal, there is likely some T3 transport across the placenta. It is not completely absent, as stated on page 21 of the EPA responses to the previous draft report: "Keep in mind that TSH and T3 are not transported across the placenta", and in several areas of Section 6.1.1. in the current draft report. Please see some references:

   Visser T. Thyroid hormone transport across the placenta. *Ann Endocrinol (Paris)* 2016;77:680-3. (https://www.ncbi.nlm.nih.gov/pubmed/27659266)

   Porterfield SP et al. The role of thyroid hormones in prenatal and neonatal neurological development--current perspectives. *Endocr Rev* 1993 Feb;14(1):94-106. (https://www.ncbi.nlm.nih.gov/pubmed/8491157)

   Calvo R et al. Congenital hypothyroidism, as studied in rats. Crucial role of maternal thyroxine but not of 3,5,3'-triiodothyronine in the protection of the fetal brain. *J Clin Invest* 1990 Sep;86(3):889-99. (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC296808/)

   James et al. Placental transport of thyroid hormone. *Best Pract Res Clin Endocrinol Metab* 2007 Jun;21(2):253-64. (https://www.ncbi.nlm.nih.gov/pubmed/17574007)

   Huang SA. Physiology and pathophysiology of type 3 deiodinase in humans. *Thyroid* 2005 Aug;15(8):875-81. (https://www.ncbi.nlm.nih.gov/pubmed/16131330)

   However, I agree that the contribution of T3 toward overall thyroid status in the developing fetus is minimal, and it is well-accepted that T4 (whether total or free) is a much better marker than T3 of low thyroid status, including during pregnancy.

2. Offspring TT4 versus FT4: Both serum TT4 and FT4 concentrations are associated with inherent challenges in their interpretation, particularly during pregnancy. There are even less data of what the appropriate extrapolated measure of this would be for the offspring of pregnant mothers. The draft report (Section 6.1.1) addresses some of these issues.

   - For the *last* part of human pregnancy, the American Thyroid Association (ATA) guidelines for the management of thyroid disease in pregnancy recommend that serum TT4 is a more accurate measurement of thyroid status during this period, since the effect of thyroid binding proteins is less of an issue in later gestation (Recommendation 3; Alexander et al. 2017 Guidelines of the American

Thyroid Association for the diagnosis and management of thyroid disease during pregnancy and the postpartum. *Thyroid* 2017:27:315-389 https://www.ncbi.nlm.nih.gov/pubmed/28056690)

- However, the most vulnerable window of thyroid-dependent neurodevelopment is *very early* pregnancy (i.e. beginning as early as gestational weeks 3-4 in humans and most critically from gestational day 18 to postnatal days 21-25 in rats; see Bernal J. Thyroid hormone receptors in brain development and function. *Nat Rev Endocrinol*,2007;3:249-259 (https://www.ncbi.nlm.nih.gov/pubmed/17315033). In the *earlier* stages of pregnancy, the ATA also states that serum TT4 can be used and assessed in reference to an increasing upward bound that is dependent on gestational age, based on a study of 20 women (Weeke J et al. A longitudinal study of serum TSH, and total and free iodothyronines during normal pregnancy. *Acta Endocrinologica* 1982;101:531. https://www.ncbi.nlm.nih.gov/pubmed/7158229). However, early pregnancy is accompanied by a rapid rise of thyroid binding proteins that must be interpreted alongside serum TT4 levels (Glinoer D et al. Regulation of maternal thyroid during pregnancy.*J Clin Endocrinol Metabol* 1990;71:276-287. https://www.ncbi.nlm.nih.gov/pubmed/2116437), thus there are concerns that TT4 may not be the best thyroid biomarker during this critical window. Additionally, more recent evidence shows that the serum TT4 variability is greater than that of serum FT4 during the first half of pregnancy, and importantly, that only FT4 (not TT4) was associated with several adverse birth outcomes in a cohort of 5,647 mother-child pairs (Korevaar TI et al. Maternal total T4 during the first half of pregnancy: physiologic aspects and the risk of adverse outcomes in comparison with free T4. *Clin Endocrinol (Oxf)* 2016 Nov;85(5):757-763. https://www.ncbi.nlm.nih.gov/pubmed/27187054). Although these issues may be lessened since TBG levels are relatively lower in rodents than humans, the majority of thyroid hormone is still in the bound form (to transthyretin and albumin) rodents; overall, only less than 1% of the thyroid hormones are in the unbound form among both species (Choksi NY et al. Role of thyroid hormones in human and laboratory animal reproductive health. *Birth Defects Res B Dev Reprod Toxicol*; 2003 Dec;68(6):479-91. https://www.ncbi.nlm.nih.gov/pubmed/14745982)

- Specifically, regarding placental transport of thyroid hormones during pregnancy, Section 6.1.1 of the draft report states: “Due to placental barrier functionality, free T4 levels in a pregnant dam might not be entirely representative of actual T4 status in a developing fetus.” I agree that this might be true, but maternal FT4 is still probably the best available representation of offspring FT4 status. This is supported by the understanding that fetal T4 status is determined by several factors: Placental type 3 deiodinase inactivates maternal T4; of the small amount of remaining T4, the unbound portion is then transported across the placenta by both passive diffusion and active mechanisms, the latter via various transport proteins. See James et al. Placental transport of thyroid hormone. Best Pract Res Clin Endocrinol ; 2007 Jun;21(2):253-64 https://www.ncbi.nlm.nih.gov/pubmed/17574007. The relative contribution of the passive and active mechanisms to fetal T4 status is not well understood. As such, at present we can only rely of the best available measure of maternal status, particularly during early pregnancy, which is maternal FT4 for the reasons below.

- The complexities between animal and human thyroid physiology have also been recently summarized in the following ATA guideline: Bianco AC et al. American Thyroid Association Guide to

investigating thyroid hormone economy and action in rodent and cell models. *Thyroid* 2014 Jan;24(1):88-168 https://www.ncbi.nlm.nih.gov/pubmed/24001133.

- Taken together, there is no perfect assessment of thyroid function during *early* pregnancy, when the effects of maternal PFBS exposure would be the most critical, and thus the accuracy of extrapolating these measurements to their offspring is similarly incompletely understood. However, given that FT4 appears to be less affected by binding proteins and maternal FT4 has been associated with adverse clinical outcomes in offspring, it appears that it would be the better representation of thyroid status among offspring of exposed mothers. It may be worthwhile to assess whether data for pups' serum FT4 levels are available from the Feng et al 2017 study, as FT4 would be a better measurement of hypothyroxinemia during pregnancy. If not, serum TT4 would be an alternate reasonable, albeit imperfect, marker of thyroid status during early pregnancy. It is noted that there is unfortunately a paucity of available data on this topic.

EPA Response: EPA appreciates the constructive and comprehensive comments concerning thyroid hormones during pregnancy. The first point made by the reviewer pertains to the transfer of T3 across the placenta during pregnancy. The draft PFBS assessment states that T3 does not cross the placental barrier (e.g., section 6.1.1 of the draft assessment). The reviewer questioned this statement and provided 5 references to inform this issue. Review of these publications and others identified during draft assessment development did not reveal a definitive answer on if or how much free T3 might cross the placenta in experimental animals or humans. The landscape of publications on this specific topic seems to have shifted over the years where earlier publications indicated minimal transplacental transfer of free T3, but since the early 2000's, dozens of studies and reports have been published suggesting that free T3 does not cross the placenta. Considering the mixed evidence on this issue, text was modified on pg. 64 (paragraph 1) of the draft PFBS assessment to the following "little if any maternal T3 is transferred across the placenta primarily due to high levels of deiodinase 3 activity that catabolizes T3 to a biologically inactive form" to address the reviewer's comment.

The second portion of the reviewer's comment provides an overview of considerations that make the case for free T4 (FT4) as the hormone measure of interest in pregnant dams and their offspring. Issues that the reviewer raises include specificity/sensitivity of in utero lifestages (e.g., early/1st trimester vs. later/3rd trimester) to FT4 as opposed to TT4 or T3, and dynamic shifts in thyroid carrier proteins during pregnancy that may impact interpretations on TT4. The proposal to identify FT4 as the critical effect is understood and appreciated. Further, the significant issues and considerations provided by the reviewer, in part 2 of the comment above, are all presented in Chapter 6 of the draft PFBS assessment; EPA acknowledges the complexity of thyroid hormone economy and dynamics during pregnancy. As presented in Chapter 6 of the draft assessment, EPA selected TT4 as the critical effect instead of FT4 for the following reasons:

(1) FT4 presented to the placenta is subject to catabolism via deiodinases 2 (D2) and 3 (D3) resulting in the formation of active free T3 or inactive reverse T3 (rT3), respectively. Importantly, the density/distribution/activity of D3 in the placenta is disproportionately increased over D2 in the placenta; as such there is preferential deactivation of FT4 to rT3 during pregnancy particularly during early gestation. This is a critical gatekeeping function of the placenta as extremely low

concentrations of active free T3 are required in the embryonic/fetal compartment to drive basal development.

(2) As the reviewer points out, FT4 represents just a fraction (typically ≤ 1%) of the total T4 pool in an adult. However, by using TT4 as the critical effect, it is presumed that the assessment indirectly accounts for FT4 as TT4 represents the aggregate of the circulating T4 pool in dams or offspring. Lastly, the American Thyroid Association recommends that women on T4 supplementation/prophylaxis increase their daily dose by 50% once they become pregnant. This suggests that in human clinical settings, the objective for thyroid hormone economy during pregnancy is to increase the circulating "total" T4 pool. As such, identification of TT4 as the critical effect for PFBS is well supported. No further revisions have been made to the assessment based on part 2 of the reviewer's comment.

Slitt

Yes to both questions. clinically relevant hypothyroxinemia during pregnancy is a relevant biomarker. There are multiple human clinical studies that cite hypothyroxinemia as a potential issue for worse outcomes for the pregnancy, as well as for the offspring. Here are some examples of recent studies. Hypothyroxinemia during pregnancy has been associated with altered reaction in 5-6 year olds (Finken et al., *J Clin Endocrinol Metab.* 2013); Lower non-verbal IQ in children 5-8 years old (Levie et al., *J Clin Endocrinol Metab.* 2018), adverse neuropsychological function of the child at 5 years of age. Additionally, marked hypothyroidism was has been associated with motor function and executive and behavior problems (Andersen et al., J Clin Endocrinol Metab. 2018). Studies also point to rodent models that induce hypothyroidism during pregnancy can have adverse effects of the development of the nervous system (Berbel et al., Cereb Cortex. 2010; Wei et al., Environ Toxicol. 2015), autism (Sadamatsu et al., Congenit Anom, 2006).

EPA Response: The reviewer agrees with the selection of total T4 as a cross-species and clinically- relevant thyroid hormone measure. The reviewer offered seven additional studies for consideration pertaining to associations between clinical hypothyroid conditions and neurobehavioral outcomes in progeny. The Fenken et al. (2016) study is already incorporated into Chapter 6 of the draft PFBS assessment. The Levie et al. (2018) citation was added to text pertaining to observed associations between decreased T4 and neurobehavioral conditions in offspring on pg. 64 of the draft PFBS assessment. Information from the Andersen et al. (2018) publication was also added to pg. 64 of the draft PFBS assessment regarding hypothyroxinemia in mothers and neurobehavior in offspring. The remaining suggested publications are redundant to references already assembled in Chapter 6 pertaining specifically to this topic area.

Warren

Yes on both accounts. Biomarker selection in the present context is a challenge given the complexity of thyroid physiology, its species variability, multiple mechanistic possibilities by which PFBS might perturb thyroid hormone homeostasis (e.g., increased hepatic T4 glucuronidation; increased thyroidal conversion of T4 to T3), and the diverse array of adverse developmental endpoints under the control of one or more thyroid hormones. As such, the selection of total T4 appears to be the most appropriate biomarker-of-effect since, as

stated in the revised draft, it *represents the aggregate of potential endocrine thyroid signaling (i.e., free T4 + protein bound T4) at any given time*. Furthermore, since similar patterns of decreases in total T3, total T4 and free T4 were observed in the principal study and that of NTP (2019), selecting total T4 when one of the two alternatives would have been more appropriate is of lesser consequence than if the three candidate biomarkers had been differentially affected by PFBS. Selection of total T4 as a biomarker is also supported by evidence that T3 is unable to cross the blood-brain barrier during fetal development. As a result, all T3 in the fetal brain is locally derived from T4 by deiodination. Interestingly, deiodinase-deficient mice do not generally exhibit altered brain development or functional deficits and the predominant isoform of thyroid hormone receptor in brain responds to both T3 and T4. This suggests that T4 may play a more active role in brain physiology than has been previously accepted. As to whether total T4 is applicable to both experimental animals and humans, the highly conserved structure and function of the thyroid among mammalian species suggest so. So too does the considerable concordance in the adverse effects observed secondary to hypothyroxinemia in humans and animals (e.g., see Crofton (2004) on the relationship between decreased total T4 and hearing loss). This is not to say species differences (e.g., metabolic turnover rates; windows of susceptibility; dose-response relationships between hormonal disruption and toxicity) can't impact the interpretation of rodent thyroid toxicity data in terms of predicting effects in humans. Rather, such differences must be appreciated and accounted for by acknowledging their contribution to uncertainty in the derivation of toxicity values.

EPA Response: The reviewer agrees with the selection of total T4 as a cross-species and clinically- relevant thyroid hormone measure. No revisions needed to address this comment.

Zoeller

Serum total T4 in the mouse pup is an appropriate biomarker/metric as it relates to clinically relevant hypothyroxinemia both during pregnancy in humans and in the human neonate. Thyroid hormone is clearly essential for brain development and growth in both rodents (mouse) and humans. It is also clear that thyroid hormone in both rodents and humans exert different actions on the brain as development proceeds. Although mice are born at a time that is equivalent roughly to the human third trimester, thyroid hormone insufficiency is relevant throughout human pregnancy and the first period of postnatal human development. In addition, serum total T4 is a good reflection of thyroid homeostasis in both human pregnancy and in the mouse. (Note that serum total T4 increases by about 50% during human pregnancy, but serum free T4 does not change. Based on this, the American Thyroid Association recommends that women on T4 supplementation before pregnancy increase their dose by 50% once they become pregnant. In other words, serum total T4 is the basis for clinical recommendations.) This measurement is applicable, therefore, to both the experimental animal paradigm as well as humans.

EPA Response: The reviewer agrees with the selection of total T4 as a cross-species- and clinically- relevant thyroid hormone measure. No revisions needed to address this comment.

1c. Has EPA clearly articulated the challenges associated with extrapolating the PFBS-induced decrease in thyroid hormone (e.g., total T4) in rodents to humans?

i. If so, please explain your reasoning.
ii. If not, please provide your rationale.

Chou

The draft assessment document (Draft) clearly stated that (1) thyroid hormonal and development effects of PFBS are more sensitive than kidney effects (2) the lack of information on PFBS effect on thyroid, developing offspring or renal system in human studies.

The reviewer believes that the study by Zhang (2018) is not a valid study to be included in this assessment for hazard identification or additional discussion on human studies in this document, because the concentrations of PFBS in patents with POI is the same as the concentration in the control subjects. Please see Table 3 on p. 2547 of the original publication by Zhang et al. (2018). No result or conclusion on the effect of PFBS should be reported from the database used in this study.

> **EPA Response: Zhang et al. (2018) is a study primarily designed to assess whether PFAS levels, including PFBS, differ in women with premature ovarian insufficiency (POI) versus control women. The fact that PFBS levels were similar in these populations does not invalidate its findings, but rather suggest that on initial analysis, there is not an association between PFBS and POI in this population. Other analyses (reproductive hormones) in Zhang et al. (2018) are stratified by case status, and thus relative exposure levels or the presence/absence of an association with POI would not influence the results. No revisions needed in response to this comment.**

Hattis

There has been a reasonable first effort at this.

Part of the challenge that could be discussed is whether or not there are differences in baseline T4 between rodents and humans, and how this affects the interspecies projection of effects on T4.

> **EPA Response: The challenge with trying to elaborate on what the reviewer proposes is that thyroid hormones are a dynamic pool, constantly in flux based on diurnal variations, sex, lifestage, pregnancy, diet (i.e., iodine intake), hormone reserve capacity and turnover rates, and thyroid carrier protein profiles. These factors are complex and convolute any meaningful cross-species interpretation of basal/baseline levels of T4. In part, the UF_A of 3 quantitatively accounts for uncertainty in cross-species thyroid dynamics. Lastly, an additional reference suggested by another reviewer (D.A. Warren) provides a comprehensive overview of cross species similarities and differences in HPT-axis physiology; this report has been added as a reference to both sections 5.1 and 6.1.1 of the draft PFBS assessment to provide readers with an opportunity to gather further details as needed. No revisions needed in response to this comment.**

Kamendulis

The document provided some information and one literature citation concerning challenges with extrapolation from rodent to human. The document stated that "*Although there are some differences in hypothalamic-pituitary-thyroid (HPT) regulation across species (e.g., serum hormone-binding proteins, hormone turnover rates, and timing of in utero thyroid development), rodents are generally considered to be a good model for*

evaluating the potential for thyroid effects of chemicals in humans (Zoeller et al., 2007)." While these statements were included in the text of section 5. specifics on what the differences in serum hormone-binding proteins, hormone turnover rates, and timing of *in utero* thyroid development were not specified in that section. Section 6 contained significant details on these endpoints and their potential significance and differences between rodents and humans. It would be useful to the reader to indicate that additional details are provided in section 6.

EPA Response: A statement directing the reader to a new comprehensive report suggested by another reviewer (D.A. Warren) for further details pertaining to HPT dynamics and the similarities/differences between rodents and humans was added to section 5.1.

Leung

The draft report has been carefully organized, and the challenges of interpreting thyroid physiology during pregnancy across species are particularly well-described in Section 6.1.1. It may be also helpful to note that the newborn rat is developmentally equivalent to the human 4-5-month-old fetus, thus there would be important differences regarding the relative contribution of maternal thyroid hormones to the developing fetus at similar PND1. See: Bernal J. Thyroid hormone receptors in brain development and function. *Nat Rev Endocrinol* 2007;3:249-259 https://www.ncbi.nlm.nih.gov/pubmed/17315033.

EPA Response: Text relating to the reviewer's suggestion can be found in section 6.1.1.

Slitt

In my opinion, the "challenges" could be more clearly articulated. The information provided on pages 71 and 72 are very detailed with regard to laying out a foundational knowledge of T4, T3, and thyroid hormone metabolism regulation during pregnancy. However, the document could further expand on outcomes and mechanisms that are similar between rodents and humans, versus any proposed differences that could be an issue interpreting data between species. It would be good to have a paragraph that specifically addresses this concern with very pointed writing.

EPA Response: The objective of the narrative on pp. 71-73 of the draft PFBS assessment is to provide the reader with perspective on the relative similarities and differences between humans and rodents during early lifestages, including in utero development. The HPT-axis in general is functionally similar between rodents and humans, however there are key considerations across species specifically pertaining to thyroid tissue development and hormone synthesis that merit discussion. This is the focus of the text particularly on pp. 72-73 of the draft assessment. To better delineate the text addressing the similarities between rodent and human thyroid hormone economy, a paragraph break was inserted near the bottom of pg. 72 of the draft assessment. Further, in the 2nd paragraph on pg. 72, reference to a new citation that provides a comprehensive overview of the cross-species similarities and differences in thyroid hormone economy was inserted; the new reference is as follows: 'A Literature Review of the Current State of the Science Regarding Species Differences in the Control of, and Response to, Thyroid Hormone Perturbations. Part 1: A Human Health Perspective' (Regulatory Science Associates, 2018)).

Warren

Yes. The revised draft includes considerable discussion of interspecies (human vs. rodent) differences, as well as commonalities. For example, species differences in the time course of HPT axis development and regulation are noted, as are differences in the fraction of gestation during which fetal development is entirely dependent on maternal thyroid hormone. On the other hand, there is also a brief discussion supporting the lack of a significant species difference in the hormonal reserve capacity between human and rodent neonates. In addition to the rodent studies of Feng et al. (2017) and NTP (2019), several human epidemiological studies of pregnant women with decreased thyroid hormone levels are discussed. Though the neurodevelopmental status of their offspring was examined as well, neither the rodent nor human studies were sufficient to identify a BMR with any degree of certainty. However, the magnitude of T4 decrease associated with developmental sequelae in both species, albeit based on limited data, appears to be roughly comparable. While the challenges of interspecies extrapolation are clearly articulated in the revised draft, consideration might be given to referencing the following report, recently published and comprehensive:

> *A literature review of the state of the science regarding species differences in the control of, and response to, thyroid hormone perturbations, Part 1: Human health perspective*. Report prepared for Sponsor: European Crop Protection Association, Prepared by: Regulatory Science Associates, Regulatory Science Ltd1, Kip Marina, Inverkip, Renfrewshire, PA16 0AS, APRIL 2018.

Also, featuring more prominently the studies of Yang et al. (2016) and Wang et al. (2014) of PFAS other than PFBS, might strengthen support for the extrapolation of rodent data to pregnant women and their offspring (and potentially, selection of total T4 as a biomarker). Lastly, among the statements in the revised draft that speak to the issue of interspecies extrapolation are the following, all of which this reviewer considers supported:

> "*rodents are generally considered to be a good model for evaluating the potential for thyroid effects of chemicals in humans (Zoeller et al., 2007),"*
>
> "*these interrelated developmental effects in mice (i.e., delays and hormonal changes) are coherent with effects on the thyroid and presumed to be directly relevant to similar processes in humans; however, studies evaluating these outcomes in humans are not available,*"

and

> "*the selection of total T4 as the critical effect is based on a number of key considerations that account for cross-species correlations in thyroid physiology and hormone dynamics particularly within the context of a developmental life stage*."
>
> **EPA Response: The reviewer agreed that the EPA clearly articulated the challenges associated with extrapolating the PFBS-induced decrease in thyroid hormone (e.g., total T4) in rodents to humans. A new comprehensive report was suggested that provides an overview of HPT-axis physiology and function and compares and contrasts experimental animals and humans. Text has been added to sections 5.1 and 6.1.1 of the draft PFBS assessment with reference to this informative report as suggested by this reviewer and to address the comments of additional external peer reviewers. The reviewer also suggested that the Yang et al. (2016) and Wang et al.**

(2014) publications be leveraged more heavily to augment support for comparisons between pregnant rodents and humans as it pertains to thyroid hormone economy. The text in section 6.1.1 uses these two studies in a supportive evidence role for cross species similarity.

Zoeller

There are two components to this extrapolation. First is the relative efficacy of PFBS in humans and animals with respect to T4 suppression. The second is the efficacy of T4 suppression to adverse outcome. The Agency has made clear that this extrapolation was essentially described by the following equation:

$$HED = \frac{POD}{CL_A/CL_H} = POD \times \frac{CL_H}{CL_A}$$

Where CL is the clearance rate in animals (A) and humans (H). The problem is that there are no estimates of clearance in humans. However, because clearance rates are similar between rodents and monkey, and half life is inversely related to clearance, the Agency made rational estimates to extrapolate PFBS-induced decrease in thyroid hormone in rodents to humans. As the agency presents in Table 9, similar patterns of decreases in serum total T3, total T4, and free T4 were observed in PFBS-exposed pregnant mice, nonpregnant adult female rats, adult male rats, and gestationally exposed female mouse offspring. The magnitude of decrease was deemed concerning (~20% in dams and ~50% in offspring), and more importantly, they were shown to persist at least 60 days after gestational exposure in offspring, and they exhibited a clear dose dependence.

EPA Response: This reviewer interpreted the charge question differently from the other reviewers however offered perspective on and agreement with the kinetic extrapolation between rodents and humans. No revisions needed to address this comment.

1d. Has EPA clearly articulated what is and is not known about the clinical implications of changes in T4 in women during pregnancy on neonates and infants?

- **i. If so, please explain your reasoning.**
- **ii. If not, please provide your rationale.**

Chou

The reviewer believes that the draft document has provided sufficient information for the purpose of identifying potential hazard and critical effects that are relevant to humans. Following is the reviewer's additional rationale for the purpose of communication.

In general, pregnancy require additional thyroid hormones in most hypothyroid patients. The increases in thyroid binding globulin and decreases in albumin concentrations during pregnancy further complicate the clinical implications of changes in T4 levels during different trimesters of the pregnancy. The importance of sufficient maternal thyroid hormone levels during pregnancy on neonates and infants is well demonstrated. Nonetheless, the cutoff level of T4 used to define sufficient and insufficient level is not well established in the clinical practice. The diagnostic approach and management of pregnant women with subclinical hypothyroidism remain to be an area of current pursue by researchers.

The reviewer understands that the dose-response marker used in this assessment, total T4, is not the same as the frequently used clinical diagnostic tests (TSH and FT4) in humans. When stand alone, each of these measurements has its limitations, and none qualifies as a marker without being accompanied by additional information. In the case of this assessment, the effect on hypothyroidism is concluded by the collective observations of many animals, from multiple studies, challenged by the same chemical, and based on several parameters, i.e. TSH, FT4, total T4 and Total T3. On the other hand, in clinical settings, the conclusion is drawn from the values measured in a single individual, with unknow causes. In addition, the criteria used in selecting a critical effect in dose-response assessment are different from the criteria used in selecting a diagnostic marker. In addition to being etiologically relevant, the critical effect is also selected based on the consideration of dose-response sensitivity.

> **EPA Response: The reviewer agreed that the EPA clearly articulated what is and is not known about the clinical implications of changes in T4 in women during pregnancy on neonates and infants. No revisions are needed to address this comment.**

Hattis

The relevant section of the document for this would seem to be 4.1. However, I do not see discussion there of pregnancy-related implications of changes in T4. Some evidence is reported in the literature that some chemical exposures can cause disruptions in thyroid hormone levels [see Patrick, LND "Thyroid Disruption: Mechanisms and Clinical Implications in Human Health" Alternative Medicine Review 14(4):326-346], however the journal source of this paper ("alternative medicine review") leaves room for doubt on this apparent conclusion. "Alternative medicine" suggests that the journal source identifies itself as not in the main line of medical thought.

> **EPA Response: As described in the introduction to Chapter 4, this chapter presents the evidence base identified in the systematic literature review for each potential health effect from human and animal studies that are potentially relevant to the derivation of RfD values. Section 4.1 describes the database of human and animal studies relevant to thyroid effects, but does not provide the evidence integration analyses and overall judgments on the hazard, clinical implication of health effects, or discussion of the dose-response analyses (including selection of critical effect, principal study, POD), which is found in Chapters 5 and 6. Rather, portions of Chapter 6 discuss in great detail, and include associated citations, information pertaining to the clinical implications of thyroid hormone decrements as mentioned by other external peer reviewers.**

Kamendulis

In general, the clinical implications of changes in T4 during pregnancy on neonates and infants was described. The document provided statements and citations describing that adequate levels of thyroid hormones are needed for normal growth and development in early life stages (Forhead and Fowden, 2014; Gilbert and Zoeller, 2010; Hulbert, 2000). Additional implications for thyroid hormone disruption and adverse developmental consequences were well described in section 6. Further, the document discussed that the presence of sufficient thyroid hormones during the gestational and neonatal period is essential for brain

development and maturation. Importantly, the document identified that while altered thyroid hormone levels may be expected to impact neurodevelopment, no studies have evaluated the effect of PFBS on neurodevelopment, therefore there is uncertainty as to the potential developmental consequences of PFBS exposure.

> **EPA Response: The reviewer agreed that the EPA clearly articulated what is and is not known about the clinical implications of changes in T4 in women during pregnancy on neonates and infants. No revisions are needed to address this comment.**

Leung

It may be helpful to emphasize that the conventional definition of hypothyroxinemia is really utilized only in pregnancy and based on subnormal serum FT4 levels in the setting of normal serum TSH concentrations; outside of pregnancy, the clinical relevance of this entity is unknown. For example, outside of pregnancy, hypothyroxinemia has been mostly described in some premature infants and has not been rigorously studied in other age groups or populations (See Rapaport R et al. Hypothyroxinemia in the preterm infant: the benefits and risks of thyroxine treatment. *J Pediatr* 2001 Aug;139(2):182-8. https://www.ncbi.nlm.nih.gov/pubmed/11487741)

Thus, using hypothyroxinemia as a critical effect among offspring is based on our understanding of this condition in their mothers. As such, hypothyroxinemia (in pregnancy) is usually not defined by TT4 levels, since the predominant (and even perhaps all available) studies assessing the adverse clinical consequences of hypothyroxinemia in pregnancy have only been based on FT4. Please see Negro R et al. Hypothyroxinemia and pregnancy. *Endocr Pract* 2011 May-Jun;17(3):422-9 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3637943/) as an excellent review on this topic. Additionally, in the draft report Section 6.1.1., it might be better to thus not describe hypothyroxinemia as a form of hypothyroidism, which is traditionally understood as either subclinical or overt biochemical hypothyroidism.

> **EPA Response: The reviewer suggests that 'hypothyroxinemia' is only utilized as a condition description within the context of pregnancy. While it is true that pregnant women represent a large proportion of the population in which hypothyroxinemia is diagnosed, there are several other conditions in male and female adults where this specific thyroid hormone economy profile occurs. Hypothyroxinemia has been observed/diagnosed following exposure to certain medications such as salicylates, NSAIDs, L-asparaginase, danazol, niacin, furosemide (given to people in some state of kidney disease), and mefenamic acid; hereditary conditions that cause a drop in thyroid binding globulin (TBG) such as Cushing's syndrome and other X-linked disorders; or people suffering from kidney disease that involves an excessive loss of serum proteins (such as TBG) via urinary excretion.**
>
> **The reviewer also suggested that hypothyroxinemia is contingent upon or defined by a decrease in free T4. The clinical definition of hypothyroxinemia is "the presence of an abnormally low concentration of thyroxine in the blood with no change in thyrotropin levels"; more specifically, several medical resources define "euthyroid hypothyroxinemia" as a decrease in *total thyroxine***

(T4) **and triiodothyronine (T3) with normal levels of TSH. The circulating levels of free T4 (FT4) are clearly of tremendous import for diagnosing perturbations in HPT function no matter the human condition, pregnant or not. In clinical settings FT4 (and/or T3) is often the target hormone for diagnostic purposes as this represents the active form available for bioactivity (via deiodination to T3) in tissues; however, measures of total T4 are indicative of the entire T4 pool including FT4. The American Thyroid Association recommends that women on T4 supplementation/prophylaxis increase their daily dose by 50% once they become pregnant. This suggests that in human clinical settings, the objective for thyroid hormone economy during pregnancy is to increase the circulating "total" T4 pool. As such, identification of TT4 as the critical effect for PFBS is well supported. No further revisions have been made to the assessment based on this aspect of the reviewer's comment.**

The reviewer suggested that hypothyroxinemia not be referred to as a form of hypothyroidism in the draft PFBS assessment. It is not clear why this is problematic as myriad resources and publications exist that consider hypothyroxinemia a part or subcategory of hypothyroidism. No revisions have been made to the assessment based on this aspect of the reviewer's comment.

Slitt

Yes. The information provided on pages 71 and 72 describe in detail the role of T4 hormone during pregnancy and the relationship between maternal T4 levels and outcomes in the offspring. The writing is detailed and well cited.

EPA Response: The reviewer agreed that the EPA clearly articulated what is and is not known about the clinical implications of changes in T4 in women during pregnancy on neonates and infants. No revisions are needed to address this comment.

Warren

Yes, pages 63-64 present a succinct description of human epidemiological studies of pregnant women with decreased thyroid hormone levels and the neurodevelopmental status of their offspring. The revised draft clearly makes the following points: 1) associations between thyroid hormone levels in pregnant mothers and neurodevelopment in their offspring are inconsistent; 2) the inconsistency may be associated with variable timing of hypothyroxinemia during pregnancy; 3) the inconsistency may also be associated with variable types of maternal hypothyroidism, only one of which involves a subnormal T4 concentration; 4) the magnitude of T4 decrease associated with developmental sequelae, albeit based on limited data, appears roughly comparable in humans and rodents; and 5) ultimately, the database does not allow, to a reasonable degree of certainty, identification of the minimum extent of T4 decrease necessary for adverse developmental outcomes. Perhaps one addition to the discussion of clinical studies might be in order – that is, noting that most rely on maternal free T4 as a measure of thyroid hormone status rather than total T4 in PND1 offspring on which the POD is based.

EPA Response: The reviewer agreed that the EPA clearly articulated what is and is not known about the clinical implications of changes in T4 in women during pregnancy on neonates and

infants. The reviewer suggested that text pertaining to human clinical studies of thyroid hormone highlight that free T4 is the most common measure. The text in Chapter 6 delineates free T4 when the authors specified as such; when results and conclusions did not specify free vs. total T4, the text reflects "T4". For further rationale on the general issue of relying on total vs. free T4 as the basis for determination of clinically relevant concern over thyroid hormone perturbations is provided in the response to Leung above (under this same charge question).

Zoeller

The Agency has developed a strong argument for the importance of thyroid hormone in child health. They identified the critical effect from the Feng et al. (2017) study as decreased serum total thyroxine (T4) in newborn (PND 1) mice. Further, they state that T4 and T3 are essential for normal growth of developing offspring across animal species, and that previous studies show that exposure to other PFAS during pregnancy results in lower T4 and T3 levels in pregnant women and fetuses or neonates. The selection of total T4 as the critical effect is based on a number of key considerations that account for cross-species correlations in thyroid physiology and hormone dynamics particularly within the context of development.

The Agency argues that a key issue for the focus on total T4 is that it "*represents the aggregate of potential thyroid endocrine signaling (i.e., free T4 + protein bound T4) at any given time*." It is true that although T3 is the "hormonally active" form at the receptor, it is T4 that gains access to tissues (e.g., brain and fetal compartment) and that *de novo* conversion to T3 is part of the signaling pathway. It is somewhat confusing that the Agency focused on the type 3 deiodinase in placenta. The Agency states that, "*The placenta has transporters and deiodinases that collectively act as a gatekeeper to maintain an optimal T4 microenvironment in the fetal compartment*." which is true enough. However, their example deiodinase 3 (D3), which is *"highly expressed in human uterus, placenta, and amniotic membrane, where it serves a critical role of regulating thyroid hormone transfer to the fetus through the deiodination of T4 to transcriptionally inactive reverse triiodothyronine (rT3) or T3 to inactive 3,5-diiodo-L-thyronine (T2)*". This is also true, but it is unclear what relevance this has to the issue at hand. Moreover, the Agency states that, "*Further, the Dio3 gene that encodes D3 has been shown to be imprinted in the mouse (Hernandez et al., 2002), suggesting a pivotal role for this specific deiodinase in the mouse as well*." However, Hernandez et al. showed that the paternal *Dio3* gene is preferentially expressed in the offspring. It is not clear how this indicates a pivotal role for D3, nor that T4 degradation should be the focus. But it is true that the human and rodent placenta have been shown to be similarly permeable to T4 and T3 (Fisher, 1997; Calvo et al., 1992). Finally, the Agency concludes that "*Due to placental barrier functionality, free T4 levels in a pregnant dam might not be entirely representative of actual T4 status in a developing fetus. Thus, decreased total T4 in offspring is expected to be more representative of PFBS-mediated thyroid effects and potentially associative developmental effects*." Although I agree with this conclusion, I don't really follow the argument, which seems discursive. Rather, I would focus on the fact that serum total T4 increases by about 50% in pregnant women without a concomitant increase in serum free T4, and that hypothyroid women should increase their daily dose of T4 to reflect this if they become pregnant. Thus, total T4 is an important index at this life stage.

This argument was a prelude to the development of the idea that the clinical manifestation of low T4 in pregnancy (but also in the neonate) results in neurocognitive deficits in humans and animals. This is a complex field because while it is clear that the human brain is sensitive to thyroid hormone insufficiency, the disconnect

between the timing of T4 measurement in the pregnant woman, and the cognitive domains tested in the offspring do not always match. This weakness is revealed by the work and writing of Professor Joanne Rovet (Rovet 2014) who clearly described the temporal relationship between thyroid hormone insufficiency and the cognitive domain affected. However, the Agency did a thorough job of articulating the clinical relevance of total T4 insufficiency and its relation to adverse cognitive outcome in humans.

EPA Response: The reviewer agreed in general that the Agency did a thorough job of articulating the clinical relevance of total T4 insufficiency and its relation to adverse cognitive outcomes in humans. A clarification is warranted here to address a question raised by the reviewer. Specifically, the reviewer did not understand the focus on deiodinase 3 in the Chapter 6 narrative. Free T4 presented to the placenta is subject to catabolism via deiodinases 2 (D2) and 3 (D3) resulting in the formation of active free T3 or inactive reverse T3 (rT3), respectively. The key point of the D3 narrative in Chapter 6 is that the density/distribution/activity of D3 in the placenta is disproportionately increased over D2 in the placenta; as such there is preferential deactivation of free T4 to rT3 during pregnancy particularly during early gestation. This is a critical gatekeeping function of the placenta as extremely low concentrations of active free T3 are required in the embryonic/fetal compartment to drive basal development. Why is this key? Because unbound FT4 that arrives at the placenta is subject to catabolism via deiodination (again, primarily D3), whereas some fraction of protein bound T4 (represented as "total T4") traverses the placenta unmodified and subsequently becomes available to deiodinases in the fetal compartment.

CHARGE QUESTION 2

2. In the public review draft PFBS assessment, EPA employed benchmark dose modeling (U.S. EPA, 2012) in the identification of a point-of-departure (POD) for derivation of RfD values, based on a decrease in total T4 levels in PND1 offspring. The 20% Relative Deviation (RD) Benchmark Response Rate (BMR) used in the public review draft is no longer being considered for BMD modeling of thyroid hormone dose-response data. As a result of extensive public review comments on the BMD approach (and thyroid hormone endpoint) used in the previous draft, and because a clear or consistent biological threshold for T4 changes associated with untoward developmental health outcomes has not be identified in the available literature, EPA has identified a new BMR of 0.5 SD (standard deviation change over controls) as a default in the revised PFBS draft assessment for the thyroid hormone alterations in mouse neonates/offspring. A 1 SD BMR is also being presented as the standardized basis for comparison as recommended in the EPA BMD Technical Guidance (U.S. EPA, 2012).

2a. Are the dose-response modeling approaches, selection of benchmark response rate (BMR), and the selected models used to identify the thyroid effect-related POD for RfD derivation scientifically justified and clearly described?

- **i. If so, please explain your reasoning.**
- **ii. If not, please provide your rationale and identify alternative approaches, BMRs, and/or dose-response models that support the identification of alternative candidate POD(s) for the derivation of subchronic and chronic RfDs and provide the scientific support for the alternative choice(s).**

Chou

The reviewer believes that additional information or modification of the writing in Paragraph 1 of page 69 may clarify the meaning of these sentences. For example, what does this sentence say, "A primary delineating feature between adult animals and developing offspring is that adults have a considerable reserve thyroid hormone capacity" (p. 69, first paragraph)? Perhaps this is also trying to say that fetal thyroid hormones depend on the supply from maternal T4?

Nonetheless, the dose-response modeling approaches, selection of benchmark response rate, and the selected models used to identify the thyroid effect-related POD for RfD derivation are scientifically justified and clearly described. The review agrees to the selection of the BMRs.

> **EPA Response: The reviewer agrees with the selection of BMRs. As suggested by the reviewer, the sentences describing the selection of BMR for adult rodents were clarified.**

Hattis

I don't see a clear justification by EPA for any specific choice of benchmark response rate. I don't have further comments on these aspects of the analysis.

> **EPA Response: Justification for the specific choice of benchmark response rates is included in Section 6.1.1 of the draft PFBS assessment. All other reviewers agreed with the selected dose-response modeling approaches and selected BMRs in the external review draft.**

Kamendulis

In general, the approaches used to derive an RfD for PFBS were scientifically justified and well described. The document provides a clear line of evidence describing that in the existing data, there is no clear or consistent biological threshold for T4 changes associated with untoward developmental health outcomes. While BMD guidance would indicate a BMR of 1SD from control, (EPA 2012), a BMR of 0.5 SD was used as the default when performing BMD modeling on thyroid hormone and the potential developmental outcomes in offspring. A BMR of 0.5 SD from control is justified as effects in developing offspring, including thyroid hormone changes, should be used for effects occurring in a sensitive life stage.

> **EPA Response: The reviewer agrees with the selection of the dose-response modeling approaches, selection of benchmark response rates, and the selected models. No revisions needed to address this comment.**

Leung

Benchmark dose modeling is not my area of expertise; thus I defer to the other reviewers.

> **EPA Response: No revisions needed to address this comment.**

Slitt

This is outside of my area of expertise.

EPA Response: No revisions needed to address this comment.

Warren

Yes, the revised draft provides a clear description and adequate scientific justification for dose-response modeling approaches (including the NOAEL/LOAEL approach for data not amenable to benchmark dose modeling), use of several BMRs for continuous and dichotomous data (BMR of 1 SD from control mean, BMR of 0.5 SD from control mean, BMR of 10% extra risk), and the ultimate selection of the exponential 4 model based on it returning the lowest BMDL (Table F-2). Not addressed in the text, however, is the revised draft's consideration of both litter and individual fetuses as the experimental unit, with the former being amenable to BMD modeling, but not the latter. The revised draft is written with transparency clearly in mind, as renal hyperplasia (from the original draft's principal study, no less) and developmental delay data were modeled and PODs presented for comparative purposes. The ultimate selection of total T4 in PND1 offspring (and 0.5 SD from control mean) as the BMR, as noted, is consistent with U.S. EPA policy given the uncertainty surrounding the response level to consider adverse and the use of data from a particularly susceptible lifestage. The discussion of reserve thyroid hormone capacity was particularly effective as partial justification for the selection of PND1 mice, especially as some model-derived PODs based on total T4 in adults were at or below that selected for RfD derivation.

EPA Response: The reviewer agreed with the selection of BMRs across endpoints. In the draft PFBS assessment, EPA modeled total T4 and developmental outcomes in PND1 mice using dose group sizes based on both the total number of fetuses (i.e., fetal n) or dams (i.e., litter n). In the study by Feng et al. (2017), it is unclear if the study-reported standard errors pertain to litters or fetuses. By alternatively modeling fetal endpoints using litter n or fetal n, these two modeling results bracket the "true" variance among all fetuses in a dose group. Individual animal data were requested from the study authors, but the EPA was unable to obtain. Results of both modeling approaches were provided in the draft and HAWC. To clarify, a footnote was added to Table 9 and Appendix F of the draft PFBS assessment to explain this approach.

Zoeller

First, I am not expert on the issue of benchmark dose modeling. However, the Agency's argument for use of the 0.5 SD over controls was reasonable to me. In particular, it is true that there is no identified "threshold" of total T4 insufficiency that is clearly causative in the production of cognitive – and other developmental – deficits. Thus, the Agency needed to formalize an approach that would be science-based and reflect a rational approach to identifying the POD. One argument that could have further strengthened this approach is to address the issue of "compensation". This concept is that as serum total T4 declines, endogenous "adaptive" responses are triggered – both in tissues and in the blood – to ameliorate the adverse consequences of low T4. One study examined this issue specifically (Sharlin et al. 2010) finding that if these adaptive responses are compensatory, they occur at a level of T4 insufficiency that is not measurable. Thus, using a 0.5SD cut-off appears reasonable and not overly protective.

EPA Response: The reviewer supported the selection of BMRs across endpoints, specifically the BMR of 0.5 SD used for developmental and thyroid hormone changes in neonatal mice. In

regard to the proposed involvement of compensatory mechanisms in response to thyroid hormone perturbations, it is unclear if such mechanisms are active/relevant following PFBS exposure. Specifically, the Sharlin et al. (2010) study (suggested by the reviewer) used propylthiouracil (PTU) to disrupt the HPT-axis. PTU inhibits thyroid hormone production directly in the gland by inhibiting the enzyme thyroid peroxidase, which converts iodide to iodine; this is a critical step in hormone production as the iodine molecules are requisite for incorporation with the amino acid tyrosine. As reported by Sharlin et al. (2010), PTU caused a dose-dependent decrease in T4 and corresponding increase in TSH, consistent with prototypical clinical hypothyroidism. In contrast, PFBS caused a decrease in T4 without a corresponding reflex increase in TSH in mice (Feng et al., 2017) and rats (NTP, 2019), consistent with clinical hypothyroxinemia. While both conditions involve a decrease in circulating T4, the associated reflex responses or compensatory mechanisms are diverse and may not be activated or engaged in a conserved manner (e.g., increased T4 production via elevated TSH signaling [with PTU exposure] vs. lack of apparent upregulation of T4 due to static TSH levels [with PFBS exposure]) across conditions. Further, the Sharlin et al. (2010) study observed biological perturbations that do not support the concept of compensation, with a specific focus on the developing brain, associated with decreased circulating T4. For example, the mRNA encoding RC3/neurogranin, a direct target of T3 action, exhibited a strong negative linear correlation with serum total T4 despite the activation of adaptive responses. In addition, RC3 mRNA levels in cortical neurons of developing rats demonstrated that the co-expression of a T3-specific transporter (MCT8) did not alter the relationship between RC3 mRNA and serum T4, suggesting the lack of a relationship between compensatory mechanisms in the developing brain and circulating levels of T4 (Sharlin et al. 2010).

CHARGE QUESTION 3

3. **Due to the availability of new toxicokinetic data in mice noting significant interspecies differences in toxicokinetics of PFBS, and, recommendations from public commenters, EPA has applied a data-informed approach to convert the oral dose-rate in animals to a human equivalent dose (HED) in the identification of candidate points-of-departure (PODs) considered for the derivation of the RfDs (U.S. EPA, 2014; see https://www.epa.gov/sites/production/files/2015-01/documents/ddef-final.pdf). In considering the new evidence for serum half-life in mice published in Lau (in press), EPA concluded that the toxicokinetic data for PFBS are adequate to support calculation of data-derived dosimetric adjustment factors (DAF), where the ratio of elimination half-life in animals to that in humans, T0.5A/T0.5H, is used to adjust candidate PODs. By using in vivo animal and human half-life data to calculate POD(HEDs) that account for differences in toxicokinetics between rodents and humans, the potential role of interspecies differences in processes such as renal resorption, hepatic transport, and enterohepatic recirculation is reflected. Further, by using a data-derived approach the uncertainty in interspecies toxicokinetic scaling (UF_A) has been reduced from a 3 to a 1; however, residual uncertainty (due to the lack of information) pertaining to toxicodynamics exists and is acknowledged in the assessment in the description for applying a UF_A of 3.**

3a. Is applying the data-informed dosimetric adjustment that utilizes the ratio of the PFBS elimination half-life in mice to that in the human scientifically justified and clearly described?

- **i. If so, please explain your reasoning.**

ii. If not, please provide your rationale and identify an alternative approach to scale PFBS doses between rodents and humans and provide scientific support for the alternative choice.

Chou

When substance-specific empirical data are available, Data-Derived Extrapolation Factor (DDEF) described in the EPA Guidance (2014) and applied in this assessment is an appropriate and improved methods for interspecies toxicokinetic extrapolation. The method is clearly described in the assessment document (p. 65-67) and the EPA Guidance (2014).

EPA Response: The reviewer agrees with the application of data-derived dosimetric adjustment. No revisions needed to address this comment.

Hattis

The choice suggested by EPA in this case seems arbitrary. Better thought, and explorations with examples of other thyroid-acting chemicals may be helpful for deriving a widely applicable projection rule.

I have now been provided with the unpublished Lau et al. paper that was the basis for the derivation of the proposed RfD value. This is a mouse study that involved two relatively widely spaced doses (30 and 300 mg/kg-day).

I find that, far from providing justification and documentation of the sufficiency of the proposed RfD, the paper simply does not contain a detailed justification for a proposed RfD, let alone the extraordinary proposed reduction of the UFA from 3 to 1. The proposed RfD therefore is not justified by the current document and should be revised downward.

I just did not see that the ratio used (PFBS elimination half lives in mice to humans) sufficiently removed uncertainty in interspecies projection to justify the reduction of the safety factor from 3 to 1.

EPA Response: EPA appreciates the request for further clarification of the data-derived dosimetric adjustment approach and interspecies uncertainty factor (UF_A) application. EPA relied on the Lau et al. publication for serum terminal half-life measures for PFBS in mice, which, along with human half-life data, were used to calculate data-derived dosimetric adjustment factors (DAF) to adjust candidate PODs to a human equivalent dose (HED). EPA applies a UF_A of 3 ($10^{0.5}$) to account for residual uncertainty in characterizing the toxicokinetic and toxicodynamic differences between mice and humans following oral K^+PFBS/PFBS exposure as recommended by EPA guidance when calculating a HED. Application of the data-derived adjustment factors reduces some uncertainty in the differences in toxicokinetics between rodents and humans and the potential role of interspecies differences in processes such as renal resorption, hepatic transport, and enterohepatic recirculation; however, residual uncertainty associated with cross-species toxicokinetics and toxicodynamics remains. Therefore, EPA maintains the UF_A of 3, consistent with EPA guidance and practice.

Kamendulis

I agree that how the DAF was derived was clearly described in the document. However, the incorporation of new data for mouse serum $t_{1/2}$ was included (Lau, in press) and heavily relied upon to support this approach. While it is acknowledged that the reviewers were provided a preprint of this manuscript, at the time of submission of this review, several points remain unclear: 1) whether the manuscript been accepted for publication; 2) the stature and rigor to which journal was the manuscript submitted; and 3) the appropriateness of using data from a manuscript that is not yet publicly available. In addition, the derivation of a DAF using the approach presented herein is not my area of expertise, therefore I cannot fully comment of the scientific validity of this approach. However, in addition to reliance on the Lau manuscript for animal data, another concern with this approach is the limited human elimination data that is available, in particular for females. Although methods were applied to account for small sample sizes, the overall appropriateness of this approach is questioned.

> **EPA Response: EPA acknowledges the challenges presented by the reviewer in the data used to support the derivation of data-derived dosimetric adjustment factors (DAFs). First, the new data for mouse serum half-life that is used in the draft PFBS assessment are included in a manuscript that has not reached final publication stage. This manuscript has been accepted with minor revisions by the journal Toxicology, an international, peer-reviewed journal with a current impact factor of 3.547. The EPA anticipates that the Lau et al. manuscript will be published and publicly available at the time of the finalization of the PFBS toxicity assessment. Second, the human half-life for PFBS used to support the derivation of the data-derived DAFs is the combined male and female geometric mean serum elimination half-life calculated from a group of six workers with occupational exposure, including five men and one woman, from Olsen et al. (2009). This study by Olsen et al. (2009) is the only evaluation of the elimination of human serum K^+PFBS and therefore provides the only available human serum half-life values. EPA acknowledges this database deficiency as an uncertainty, but concludes that the use of data-derived dosimetric adjustment factors better accounts for differences in toxicokinetics between rodents and humans and the potential role of interspecies differences in processes such as renal resorption, hepatic transport, and enterohepatic recirculation, compared to use of a default BW3/4 adjustment. To further acknowledge this, text was added to Tables 11 and 13 (i.e., RfD confidence descriptors) to point out uncertainties in the toxicokinetic database.**

Leung

This is not an area that I am able to comment on, thus I defer to the other reviewers.

> **EPA Response: No revisions needed to address this comment.**

Slitt

Yes, this approach is reasonable and does provide an adjustment for the marked differences in species toxicokinetics. As mentioned, the processes that dictate renal resorption, hepatic transport, and enterohepatic recirculation doe have species differences with regard to transporter affinity, function, and even localization. This is the most reasonable method to scale from mouse to human.

EPA Response: The reviewer agrees with the application of data-derived dosimetric adjustment. No revisions needed to address this comment.

Warren

Yes. In what is another radical departure from the original draft, the default extrapolation procedure of body weight$^{3/4}$ has been superseded by application of a data-informed adjustment factor based on the ratio of animal to human serum PFBS elimination half-lives. This is scientifically justified, consistent with U.S. EPA's hierarchy of approaches to dosimetric adjustment in the derivation of RfDs and is appropriate given the lack of confidence in the one existing PBPK model for extrapolation purposes. Pages 65-67 clearly describe the availability of clearance and half-life data, including the recent addition of sex-specific half-life data in mice. On the subject of sex-specific half-lives, it is noteworthy that the PFBS half-life measured in the one female subject in Olsen et al. (2009) was nearly twice that of the mean half-life of the five male subjects. Thus, use of the geometric mean value of the six human subjects to calculate the DAF creates the possibility that animal doses were not adjusted sufficiently downward. Lastly, it is worth emphasizing that some uncertainty in the derived RfDs stems from two common assumptions severely lacking in empirical validation - 1) that total T4 concentration in humans and mice will respond with equal sensitivity to the same internal or target tissue dose of PFBS, and 2) that the average serum concentration of PFBS over time is the dose metric mechanistically linked to thyroid hormone economy.

EPA Response: The reviewer supported the use of the data-derived dosimetric adjustment, and EPA acknowledges the residual uncertainty in characterizing the toxicokinetic and toxicodynamic differences between mice and humans following oral K^+PFBS/PFBS exposure. To account for these uncertainties EPA has applied the UF_A of 3. These residual uncertainties include those pointed out by the reviewer, such as the lack of data for quantifying the "relative cross-species sensitivity in toxicodynamics (e.g., thyroid signaling)" and limited toxicokinetic data in the susceptible human population. Further, data are not available to discern the exact dose metric that is mechanistically linked to thyroid hormone effects. Despite these residual uncertainties, by using a PFBS-specific data-derived dosimetric adjustment to develop HEDs, EPA decreases the uncertainty in some aspects of the cross-species extrapolation of toxicokinetic and toxicodynamic processes, including the potential role of interspecies differences in processes such as renal resorption, hepatic transport, and enterohepatic recirculation. Thus, EPA retains the UF_A of 3 for use in deriving a subchronic and a chronic RfD as recommended by EPA guidance when calculating a HED. To further acknowledge this, text was added to Tables 11 and 13 (i.e., RfD confidence descriptors) to point out limitations in the toxicokinetic database.

Zoeller

The rationale provided by the Agency for dosimetric adjustment was scientifically rational and clearly described. It is also somewhat reasonable to make the assumption that these measures are an overall reflection of processes that eliminate PFBS from the system. Other elements of this response are described in question 1.

EPA Response: The reviewer agrees with the application of data-derived dosimetric adjustment. No revisions needed to address this comment.

CHARGE QUESTION 4

4. 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 PFBS.

4a. Does the provided qualitative scientific rationale support the application of the selected uncertainty factors? If not, please explain.

Chou

Yes, the document provides qualitative scientific rationale support the application of the selected uncertainty factors.

EPA Response: The reviewer agrees with the selected uncertainty factors. No revisions needed to address this comment.

Hattis

I do not find the proffered explanation sufficiently clear to be convincing. As it stands the rationale is not credible and cannot stand.

An uncertainty factor of 1 (rather than the usual 3) suggests that there is no remaining uncertainty in the interspecies projection. I just disagree.

EPA Response: EPA appreciates the request for further clarification of the qualitative rationale for the application of the interspecies uncertainty factor (UF_A). For clarification, EPA applies a UF_A of 3 ($10^{0.5}$) to account for residual uncertainty in characterizing the toxicokinetic and toxicodynamic differences between mice and humans following oral K^+PFBS/PFBS exposure as recommended by EPA guidance when calculating a human equivalent dose (HED). Application of the data-derived adjustment factors reduces some uncertainty in the differences in toxicokinetics between rodents and humans and the potential role of interspecies differences in processes such as renal resorption, hepatic transport, and enterohepatic recirculation; however, residual uncertainty remains. Therefore, EPA maintains the UF_A of 3 and has added language in Tables 11 and 13 to acknowledge the residual uncertainty in characterizing the toxicokinetic and toxicodynamic differences between species.

Kamendulis

UF_A - Due to the concerns raised in question 3 concerning how the DAF was derived, there might be an overall concern using an UF of 3 for the derivation of both subchronic and chronic RfDs for PFBS. However, should the application of another approach be used to derive an HED, it would likely result in the use of an UF of 3 - so this might not be problematic.

UF_S, UF_D – chronic RfD – an UF-S of 1 was applied despite the lack of chronic studies. However, as it is stated (EPA, 1991) developmental period is recognized as a susceptible life stage in which exposure is more relevant to the induction of developmental effects than lifetime exposure, therefore, an UF of 1 is justified. Further, an UF of 10 was applied for database limitations, therefore accounting for uncertainty for less than lifetime exposures.

EPA Response: The reviewer agrees with the selected uncertainty factors. No revisions needed to address this comment.

Leung

This is not an area that I am able to comment on, thus I defer to the other reviewers.

EPA Response: No revisions needed to address this comment.

Slitt

Yes. Overall, the factors that have been accounted for and described in Table 10 support the application of uncertainty factors. There are gaps in our knowledge regarding toxicokinetics for newborns, interindividual variability in the toxicokinetic and toxicodynamic response, and lack of literature.

EPA Response: The reviewer agrees with the selected uncertainty factors. No revisions needed to address this comment.

Warren

Yes, the explanatory text in Tables 10 and 12 provides scientific rationale appropriate to each of the individual UFs. I have no issues with the UFs for the subchronic or chronic RfDs.

EPA Response: The reviewer agrees with the selected uncertainty factors. No revisions needed to address this comment.

Zoeller

In general, I think that the qualitative rationale in support of uncertainty factors was well-reasoned. However, I don't believe that a UF_A of 3 is fully justified. There are several reasons for this. First, the study of Feng et al. (2017) showed that serum total T4 was diminished by PFBS at PND1 but also at PND 30 and PND 60 – 30 and 60 days after cessation of exposure. If the toxicokinetic data of Lau et al. (unpublished) is correct, PFBS was fully eliminated from the animals by the 30 and 60-day timepoints. These data indicate that human neonates may experience T4 suppression for much longer than the fetal/neonatal period. Moreover, the human neonate is quite sensitive to thyroid hormone insufficiency for a minimum of 2 years. Finally, it is likely that PFBS will contaminate breast milk since other PFAS are found in breast milk (e.g., (Beser et al. 2019). Therefore, it is scientifically justified to expect that PFBS will suppress serum T4 both early in development as well as perhaps many months after birth. The uncertainty of these likely impacts would justify a UF_A of 10.

EPA Response: The reviewer suggests increasing the UF_A to 10 to account for additional uncertainty in PFBS exposure and effects in early childhood. The data from Feng et al. (2017) indicates significantly decreased total T3 and T4 at birth and at PNDs 30 and 60 in female pups, despite the exposure to PFBS occuring during gestation only. As the reviewer points out, this indicates the continuation of thyroid hormone decrements in infants beyond the original window of exposure. To examine the potential for temporal sensitivity of effects, thyroid hormone decrements were dose-response modeled to obtain potential points of departure at all reported timepoints (PND 1, 30, and 60). The *in utero* stage measured at PND1 was the most sensitive to PFBS-induced alterations in total T4, compared to PNDs 30 or 60. Therefore, as a function of dose-response, changes in thyroid hormone economy became less sensitive with age, and later timepoints following cessation of (*in utero*) PFBS exposure would be protected by the current RfD derived from PND1 pup hormone changes. The reviewer also proposes that PFBS may transfer to the infant via breast milk, resulting in another exposure pathway during childhood. The potential source contribution of PFBS in breast milk to the infant is beyond the scope of the current toxicity assessment and would be included in an exposure assessment for this chemical. Considering the extended single generation developmental toxicity study protocol from Feng et al. (2017), the exposure of F1 mice would account for potential lactational transfer of PFBS through PND 21 when the pups were weaned.

4b. Has quantitative uncertainty been adequately accounted for in the derivation of the RfDs? Please describe and provide suggestions, if needed.

Chou

Yes, the reviewer believes that the quantitative uncertainty has been adequately account for in the derivation of the RfDs.

EPA Response: The reviewer agrees with the selected uncertainty factors. No revisions needed to address this comment.

Hattis

No. It seems to me that there is still plenty of uncertainty in the true quantitative difference between (likely inbred) mouse susceptibility and the susceptibilities of the diverse arrays of humans who will be exposed.

EPA Response: EPA acknowledges the residual uncertainty in characterizing the toxicokinetic and toxicodynamic differences between mice and the diverse human population following oral K^+PFBS/PFBS exposure. To account for these uncertainties in interspecies sensitivity and intraspecies variability EPA has applied the UF_A of 3 and UF_H of 10. Residual uncertainties in extrapolation from rodents to humans include the lack of data for quantifying the "relative cross-species sensitivity in toxicodynamics (e.g., thyroid signaling)" and limited toxicokinetic data in the susceptible human population. Despite these residual uncertainties, by using a PFBS-specific data-derived dosimetric adjustment to develop HEDs, EPA decreases the uncertainty in some aspects of the cross-species extrapolation of toxicokinetic and toxicodynamic processes, including the potential role of interspecies differences in processes

such as renal resorption, hepatic transport, and enterohepatic recirculation. To further account for the differential susceptibility of the diverse human population, EPA applies a UF_H of 10, which is intended to account for the variability in the responses within the human populations because of both intrinsic (toxicokinetic, toxicodynamic, genetic, life stage, and health status) and extrinsic (life style) factors that can influence the response to dose; clarifying text has been added to the UF_H descriptors in Tables 10 and 12. Text has been added to Tables 11 and 13 (i.e., RfD confidence descriptors) to further acknowledge limitations in the toxicokinetic database.

Kamendulis

Yes, this was addressed

EPA Response: The reviewer agrees with the selected uncertainty factors. No revisions needed to address this comment.

Leung

This is not an area that I am able to comment on, thus I defer to the other reviewers.

EPA Response: No revisions needed to address this comment.

Slitt

UF_H – UF of 10 is appropriate. Examples are evidence of polymorphisms in xenobiotic transporters, such as OATPs and OATs, in which PFAS are likely a substrate (Yee et al. Clin Pharmacol Ther. 2018 Nov; 104(5): 803–817). Renal function in the elderly or disease also can be impacted. UF_D – UF of 10 is appropriate. The literature is quite limited and lacking for some health effects known to occur with other PFAS. UF_L - UF of 1 is appropriate as rationalized in the document. UF_S – 1 is properly justified based on (U.S. EPA, 1991b).

EPA Response: The reviewer agrees with the selected uncertainty factors. No revisions needed to address this comment.

Warren

Yes, the composite UFs of 100 and 300 applied for derivation of the subchronic and chronic RfDs, respectively, are appropriate. Again, I have no issues with the UFs for either toxicity value.

EPA Response: The reviewer agrees with the selected uncertainty factors. No revisions needed to address this comment.

Zoeller

As described in 4a, I do not believe the quantitative uncertainty has been fully accounted for in the derivation of the RfDs. Specifically, the UF_A should be 10, not 3.

EPA Response: As discussed in the response to 4a, EPA acknowledges the residual uncertainty in extrapolation from animals to humans when considering PFBS exposure and effects in early childhood. The reviewer proposed increasing the UF_A to 10 to capture gaps in understanding potential lactational transfer of PFBS to infants and potential continuation of thyroid hormone decrements in infants beyond the original window of exposure. These residual uncertainties are not intended to be addressed by the UF_A, which covers animal to human extrapolation. With respect to the gap in understanding lactation as a source of exposure to PFBS in infancy, the PFBS toxicity assessment does not evaluate the potential exposure source contributions of PFBS, which would instead be covered in an exposure assessment. Once final, the PFBS toxicity assessment can be used, along with specific exposure and other relevant information, to determine potential risk associated with human exposures to PFBS. With respect to uncertainty in the continuation of effects after the cessation of exposure, thyroid hormone decrements were modeled as potential points of departure at all reported timepoints (PND 1, 30, and 60), including after the window of exposure. The *in utero* stage measured at PND1 was the most sensitive to PFBS-induced alterations in total T4. Therefore, changes in thyroid hormone economy became less sensitive with age and later timepoints after cessation of exposure would be protected by the current RfD derived from PND1 hormone changes.

4c. Do the methods used to derive the RfDs for PFBS appropriately account for uncertainties in evaluating the toxicokinetic differences between the experimental animal data and humans? If not, please explain.

Chou

Yes, the methods used to derive the RfDs for PFBS appropriately account for uncertainties in evaluating the toxicokinetic differences between the experimental animal data and humans.

EPA Response: The reviewer agrees with the derivation of the RfDs. No revisions needed to address this comment.

Hattis

No. EPA guidance on the application of interspecies uncertainty factors is:

> "The default value for UFH is 10-fold; the default value for interspecies uncertainty factor (UFA) is apportioned into a TD component valued at one-half order of magnitude and a TK component addressed via default inhalation dosimetry methods (U.S. EPA, 1994) or body-weight scaling for orally encountered compounds (U.S. EPA, 2011). DDEFs fall within this hierarchical range of approaches."
>
> (quote from "Guidance for Applying Quantitative Data to Develop Data-Derived Extrapolation Factors for Interspecies and Intraspecies Extrapolation" Office of the Science Advisor, Risk Assessment Forum, U.S. Environmental Protection Agency, EPA/100/R-14/002F, September 2014).

A single factor of 3 appears to have been used in this case. The factor should be increased, or additional reasoning provided to justify it.

It seems to me that it is for EPA to justify its substantial departure from previous practices on interspecies projection.

> **EPA Response: Consistent with previous practice and as recommended by EPA guidance when calculating a human equivalent dose (HED), EPA applies a UF_A of 3 ($10^{0.5}$) to account for residual uncertainty in characterizing the toxicokinetic and toxicodynamic differences between mice and humans following oral K^+PFBS/PFBS exposure. The UF_A is generally presumed to include both toxicokinetic (i.e., absorption, distribution, metabolism, and elimination) and toxicodynamic (i.e., MOA) aspects. EPA guidance documents indicate a hierarchy in approaches for interspecies toxicokinetic (TK) extrapolation where PBPK models are the preferred approach. As no PFBS-specific PBPK model is available, EPA recommends the use of chemical specific TK data, when available to calculate a data-derived adjustment factor. As such, default allometric body weight scaling dosimetric adjustment approaches are superseded when more detailed information on tissue dosimetry can be developed. Application of the data-derived adjustment factors reduces some uncertainty in the differences in toxicokinetics between rodents and humans and the potential role of interspecies differences in processes such as renal resorption, hepatic transport, and enterohepatic recirculation; however, residual uncertainty remains. Therefore, EPA maintains the UF_A of 3.**

Kamendulis

Yes, this was addressed

> **EPA Response: The reviewer agrees with the derivation of the RfDs. No revisions needed to address this comment.**

Leung

This is not an area that I am able to comment on, thus I defer to the other reviewers.

> **EPA Response: No revisions needed to address this comment.**

Slitt

No it does not. The UF of 3 might be too generous given the limited knowledge about PFBS kinetics in neonates or children. Given that PFBS is cleared through renal clearance to urine (Bogdanska et al., 2014), there should be extra considering given to the differences between infants and adults with regard to renal clearance. Given that the main critical effect was observed at PND1, this is a relevant consideration. As reviewed by Rodieux et al., Clin Pharmacokinet. 2015; 54: 1183–1204, the effect of kidney function on children with regard to pharmacokinetics and dosing is largely unknown, with ~80% of drugs given to children not having been studied. The impact on renal clearance with regard to other PFAS infants is also largely unknown. Given the likely clearance of PFBS by human kidney and the uncertainty of kidney function with regard to PFBS clearance in children, I believe this UF could be higher.

> **EPA Response: EPA acknowledges the uncertainty in the potential kinetic differences within the human population, including differences between infants and adults. As presented by the**

reviewer, there are no data to use to characterize the differential clearance of PFBS between infants and adults. To account for the uncertainty in the variability in response among individual members of the general population and potential greater susceptibility to PFBS toxicity in sensitive subpopulations, such as infants and children, EPA applies an intraspecies uncertainty factor (UF_H) of 10. Additional clarifying language has been added to the "Justification" for the UF_H of 10 in Tables 10 and 12 to describe the considerations and uncertainties accounted for by this UF, including variability in toxicokinetic factors in the human population.

Warren

Yes, an UF_A of 3 is applied for both subchronic and chronic RfD derivation. This value seems appropriate given the use of the mouse:human ratio of PFBS half-life (0.0073) as the DAF on one hand, and the absence of toxicokinetic data during sensitive life stages on the other.

EPA Response: The reviewer agrees with the derivation of the RfDs. No revisions needed to address this comment.

Zoeller

In general I agree that they do. However, the literature does not fully inform the toxicokinetic (and dynamic) differences between experimental animals and humans. Importantly, the mechanism by which PFBS causes serum total T4 reduction is not clear. As a result, it is not clear how potent PFBS would be on those molecular targets responsible for this decline. Thus, while the toxicokinetic difference is estimated in a rational way, the toxicodynamic difference cannot actually be estimated. Thus, a UF_D of 10 would appear warranted.

EPA Response: The reviewer supported the approach used to estimate toxicokinetic differences between mice and humans following oral K^+PFBS/PFBS exposure. By using a PFBS-specific data-derived dosimetric adjustment to develop HEDs, EPA decreases the uncertainty in some aspects of the cross-species extrapolation of toxicokinetic and toxicodynamic processes. EPA acknowledges the residual uncertainty in characterizing the toxicokinetic and toxicodynamic differences between species in Tables 10 and 12, including those pointed out by the reviewer, such as the lack of data to inform relative cross-species sensitivity in toxicodynamics (e.g., thyroid signaling) and molecular targets of effect. The reviewer suggests accounting for these uncertainties in cross-species extrapolation in the database uncertainty factor (UF_D) as well. However, the UF_D accounts for deficiencies in the database to characterize toxicity, both data lacking and data available. EPA applies a UF_D of 3 when deriving the Subchronic RfD and a UF_D of 10 when deriving the Chronic RfD, to account for deficiencies in the toxicity database as described in Tables 10 and 12. EPA maintains that it would not be appropriate to account for the residual uncertainty in interspecies extrapolation in both the UF_A and UF_D.

SECTION III: REVIEWER ADDITIONAL AND EDITORIAL COMMENTS

Chou

Reviewer's Additional Suggestions for Considerations:

p. 3, Line 39: Is the word "chemical" necessary? If the resistance could also be caused biological reasons, you may consider deleting the word "chemical".

References:

Rabah, S. A., I. L. Gowan, M. Pagnin, N. Osman & S. J. Richardson (2019) Thyroid Hormone Distributor Proteins During Development in Vertebrates. *Frontiers in Endocrinology*, 10.

EPA Response: The word "chemical" is not necessary and has been removed from pg. 3 (opening paragraph of section 1.3.1) in the revised PFBS assessment document. The additional reference (Rabah et al., 2019) proposed by the reviewer has been added to the draft assessment on pg. 72, line 42.

Zoeller

References

Beser MI, Pardo O, Beltran J, Yusa V. 2019. Determination of 21 perfluoroalkyl substances and organophosphorus compounds in breast milk by liquid chromatography coupled to orbitrap high-resolution mass spectrometry. Analytica chimica acta 1049: 123-132.

Rovet JF. 2014. The role of thyroid hormones for brain development and cognitive function. Endocrine development 26: 26-43.

Sharlin DS, Gilbert ME, Taylor MA, Ferguson DC, Zoeller RT. 2010. The nature of the compensatory response to low thyroid hormone in the developing brain. J Neuroendocrinol 22(3): 153-165.

EPA Response: The additional references were screened for information pertinent to PFBS and/or further details or insights pertaining to health hazards of concern (e.g., thyroid hormones and development). The Beser et al. (2019) study provides sparse data specifically related to PFBS and only presents occurrence in biological samples without relation to effects of concern. As such this specific reference was not considered further. The JF Rovet (2014) review provides a comprehensive overview of the relationship between thyroid hormones and early life stage neurodevelopment. Although text in the draft assessment on this topic is already well supported with multiple citations, this reference was added to pg. 63, lines 37-38 for completeness. Lastly, the Sharlin et al. (2010) study was already considered and addressed in the response to Dr. Zoeller's comment under charge question 2a. This specific reference was not considered further.

APPENDIX A: INDIVIDUAL REVIEWER COMMENTS

COMMENTS SUBMITTED BY

Karen Chou, Ph.D.
Associate Professor, Department of Animal Science
Michigan State University
East Lansing, Michigan

External Letter Peer Review of EPA's Draft Human Health Toxicity Values Assessment for Perfluorobutane Sulfonate (PFBS) (CASRN 375-73-5 [acid]) and Related Compound Potassium Perfluorobutane Sulfonate (CASRN 29420-49-3)

1. The key study chosen for determining the subchronic and chronic RfDs is the gestational exposure mouse study by Feng et al. (2017) and the critical effect is decreased total T4 in postnatal Day 1 (PND1) offspring.

1a. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and clearly described?

- **i. If so, please explain your reasoning.**
- **ii. If not, please provide your rationale and identify an alternative key study and/or critical effect to support the derivation of the subchronic and chronic RfDs and provide the scientific support for the alternative choice.**

The key studies are appropriately selected, based on the existing data and critical effects are appropriately identified. The reasoning process of evaluating the quality of the studies, data uncertainties in the existing studies, as well as the validity of the rodent model for the thyroid function in human are thoroughly considered and presented in the document.

1b. Is the selection of total T4 an appropriate biomarker/metric as it relates to clinically relevant hypothyroxinemia during pregnancy? Is such a measure applicable to both experimental test animals and humans?

- **i. If so, please explain your rationale.**
- **ii. If not, are there other measures related to hypothyroxinemia that may be more useful for informing hazard potential during pregnancy? What are those measures?**

The reviewer is not certain about the main point of this charge question. If the question is about total T4 vs. free T4, i.e. "Is total T4 an appropriate stand-alone biomarker?" Total T4 is selected as the critical effect in this assessment based on the collective evidence of TSH, total T4, and total T3; therefore, it is appropriately selected because it is supported by additional evidence of hypothyroidism. In addition, plasma-protein bound thyroid hormones is likely to be equally important as, if not more important than, the free thyroid hormones in the blood for the following reason: Due to the high lipophilicity, free T4 preferentially partition into lipid environment of membrane it first come in contact with, thus minimize its availability at other target cells (Rabah et al. 2019). Plasma protein binding serves as a distributing vehicle to ensure the availability of T4 at target cells and to prevent excessive free T4 in the blood circulation.

If the question is about whether total T4 is clinically important in diagnosing hypothyroxinemia in pregnancy, yes, it is an important and relevant test in pregnancy associated hypothyroidism. The level of total T4 is consider together with the levels of TSH, thyroperoxidase antibodies and other parameters for differential diagnosis of the etiology of clinical hypothyroidism in pregnancy.

1c. Has EPA clearly articulated the challenges associated with extrapolating the PFBS-induced decrease in thyroid hormone (e.g., total T4) in rodents to humans?

- **i. If so, please explain your reasoning.**
- **ii. If not, please provide your rationale.**

The draft assessment document (Draft) clearly stated that (1) thyroid hormonal and development effects of PFBS are more sensitive than kidney effects (2) the lack of information on PFBS effect on thyroid, developing offspring or renal system in human studies.

The reviewer believes that the study by Zhang (2018) is not a valid study to be included in this assessment for hazard identification or additional discussion on human studies in this document, because the concentrations of PFBS in patents with POI is the same as the concentration in the control subjects. Please see Table 3 on p. 2547 of the original publication by Zhang et al. (2018). No result or conclusion on the effect of PFBS should be reported from the database used in this study.

1d. Has EPA clearly articulated what is and is not known about the clinical implications of changes in T4 in women during pregnancy on neonates and infants?

i. If so, please explain your reasoning.

ii. If not, please provide your rationale.

The reviewer believes that the draft document has provided sufficient information for the purpose of identifying potential hazard and critical effects that are relevant to humans. Following is the reviewer's additional rationale for the purpose of communication.

In general, pregnancy require additional thyroid hormones in most hypothyroid patients. The increases in thyroid binding globulin and decreases in albumin concentrations during pregnancy further complicate the clinical implications of changes in T4 levels during different trimesters of the pregnancy. The importance of sufficient maternal thyroid hormone levels during pregnancy on neonates and infants is well demonstrated. Nonetheless, the cutoff level of T4 used to define sufficient and insufficient level is not well established in the clinical practice. The diagnostic approach and management of pregnant women with subclinical hypothyroidism remain to be an area of current pursue by researchers.

The reviewer understands that the dose-response marker used in this assessment, total T4, is not the same as the frequently used clinical diagnostic tests (TSH and FT4) in humans. When stand alone, each of these measurements has its limitations, and none qualifies as a marker without being accompanied by additional information. In the case of this assessment, the effect on hypothyroidism is concluded by the collective observations of many animals, from multiple studies, challenged by the same chemical, and based on several parameters, i.e. TSH, FT4, total T4 and Total T3. On the other hand, in clinical settings, the conclusion is drawn from the values measured in a single individual, with unknow causes. In addition, the criteria used in selecting a critical effect in dose-response assessment are different from the criteria used in selecting a diagnostic marker. In addition to being etiologically relevant, the critical effect is also selected based on the consideration of dose-response sensitivity.

2. In the public review draft PFBS assessment, EPA employed benchmark dose modeling (U.S. EPA, 2012) in the identification of a point-of-departure (POD) for derivation of RfD values, based on a decrease in total T4 levels in PND1 offspring. The 20% Relative Deviation (RD) Benchmark Response Rate (BMR) used in the public review draft is no longer being considered for BMD modeling of thyroid hormone dose-response data. As a result of extensive public review comments on the BMD approach (and thyroid hormone endpoint) used in the previous draft, and because a clear or consistent biological threshold for T4 changes associated with untoward developmental health outcomes has not be identified in the available literature, EPA has identified a new BMR of

0.5 SD (standard deviation change over controls) as a default in the revised PFBS draft assessment for the thyroid hormone alterations in mouse neonates/offspring. A 1 SD BMR is also being presented as the standardized basis for comparison as recommended in the EPA BMD Technical Guidance (U.S. EPA, 2012).

2a. Are the dose-response modeling approaches, selection of benchmark response rate (BMR), and the selected models used to identify the thyroid effect-related POD for RfD derivation scientifically justified and clearly described?

i. If so, please explain your reasoning.

ii. If not, please provide your rationale and identify alternative approaches, BMRs, and/or dose-response models that support the identification of alternative candidate POD(s) for the derivation of subchronic and chronic RfDs and provide the scientific support for the alternative choice(s).

The reviewer believes that additional information or modification of the writing in Paragraph 1 of page 69 may clarify the meaning of these sentences. For example, what does this sentence say, "A primary delineating feature between adult animals and developing offspring is that adults have a considerable reserve thyroid hormone capacity" (p. 69, first paragraph)? Perhaps this is also trying to say that fetal thyroid hormones depend on the supply from maternal T4?

Nonetheless, the dose-response modeling approaches, selection of benchmark response rate, and the selected models used to identify the thyroid effect-related POD for RfD derivation are scientifically justified and clearly described. The review agrees to the selection of the BMRs.

3. Due to the availability of new toxicokinetic data in mice noting significant interspecies differences in toxicokinetics of PFBS, and, recommendations from public commenters, EPA has applied a data-informed approach to convert the oral dose-rate in animals to a human equivalent dose (HED) in the identification of candidate points-of-departure (PODs) considered for the derivation of the RfDs (U.S. EPA, 2014; see https://www.epa.gov/sites/production/files/2015-01/documents/ddef-final.pdf). In considering the new evidence for serum half-life in mice published in Lau (in press), EPA concluded that the toxicokinetic data for PFBS are adequate to support calculation of data-derived dosimetric adjustment factors (DAF), where the ratio of elimination half-life in animals to that in humans, $T0.5_A/T0.5_H$, is used to adjust candidate PODs. By using in vivo animal and human half-life data to calculate POD(HEDs) that account for differences in toxicokinetics between rodents and humans, the potential role of interspecies differences in processes such as renal resorption, hepatic transport, and enterohepatic recirculation is reflected. Further, by using a data-derived approach the uncertainty in interspecies toxicokinetic scaling (UF_A) has been reduced from a 3 to a 1; however, residual uncertainty (due to the lack of information) pertaining to toxicodynamics exists and is acknowledged in the assessment in the description for applying a UF_A of 3.

3a. Is applying the data-informed dosimetric adjustment that utilizes the ratio of the PFBS elimination half-life in mice to that in the human scientifically justified and clearly described?

i. If so, please explain your reasoning.

ii. If not, please provide your rationale and identify an alternative approach to scale PFBS doses between rodents and humans and provide scientific support for the alternative choice.

When substance-specific empirical data are available, Data-Derived Extrapolation Factor (DDEF) described in the EPA Guidance (2014) and applied in this assessment is an appropriate and improved methods for

interspecies toxicokinetic extrapolation. The method is clearly described in the assessment document (p. 65-67) and the EPA Guidance (2014).

4. 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 PFBS.

4a. Does the provided qualitative scientific rationale support the application of the selected uncertainty factors? If not, please explain.

Yes, the document provides qualitative scientific rationale support the application of the selected uncertainty factors.

4b. Has quantitative uncertainty been adequately accounted for in the derivation of the RfDs? Please describe and provide suggestions, if needed.

Yes, the reviewer believes that the quantitative uncertainty has been adequately account for in the derivation of the RffDs.

4c. Do the methods used to derive the RfDs for PFBS appropriately account for uncertainties in evaluating the toxicokinetic differences between the experimental animal data and humans? If not, please explain.

Yes, the methods used to derive the RfDs for PFBS appropriately account for uncertainties in evaluating the toxicokinetic differences between the experimental animal data and humans.

Reviewer's Additional Suggestions for Considerations:

p. 3, Line 39: Is the word "chemical" necessary? If the resistance could also be caused biological reasons, you may consider deleting the word "chemical".

References:

Rabah, S. A., I. L. Gowan, M. Pagnin, N. Osman & S. J. Richardson (2019) Thyroid Hormone Distributor Proteins During Development in Vertebrates. *Frontiers in Endocrinology*, 10.

COMMENTS SUBMITTED BY

Dale Hattis, Ph.D.
Research Professor
George Perkins Marsh Institute
Clark University
Worcester, Massachusetts

External Letter Peer Review of EPA's Draft Human Health Toxicity Values Assessment for Perfluorobutane Sulfonate (PFBS) (CASRN 375-73-5 [acid]) and Related Compound Potassium Perfluorobutane Sulfonate (CASRN 29420-49-3)

1. The key study chosen for determining the subchronic and chronic RfDs is the gestational exposure mouse study by Feng et al. (2017) and the critical effect is decreased total T4 in postnatal Day 1 (PND1) offspring.

1a. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and clearly described?

- **i. If so, please explain your reasoning.**
- **ii. If not, please provide your rationale and identify an alternative key study and/or critical effect to support the derivation of the subchronic and chronic RfDs and provide the scientific support for the alternative choice.**

The key study of Feng et al. (2017) is reasonable enough, as far as it goes. However, the paper reports only a single experimental run on a single group of mice. This is a little thin as a basis of a U.S. national regulatory action.

Recently published human epidemiological observations bolster the evidence. Recently epidemiological observations of Reardon et al. (2019)* in pregnant Canadian women have indicated an inverse association between serum perfluoroalkyl acids and lower FT4 levels, supporting the basis for concern for human exposures to perfluoroalkyl acids.

The full abstract for this paper is:

*Longitudinal Analysis Reveals Early-Pregnancy Associations Between Perfluoroalkyl Sulfonates and Thyroid Hormone Status in a Canadian Prospective Birth Cohort. Environ Int Vol. 129 pp. 389-399. Aug 2019

Anthony J F Reardon 1, Elham Khodayari Moez 2, Irina Dinu 2, Susan Goruk 3, Catherine J Field 3, David W Kinniburgh4, Amy M MacDonald4, Jonathan W Martin5, APrON Study

Affiliations expand

PMID: 31150980

PMCID: PMC6859374 (available on 2020-08-01)

DOI: 10.1016/j.envint.2019.04.023

Abstract

Serum perfluoroalkyl acids (PFAAs) have been linked to disruption of maternal thyroid hormone homeostasis, but results have varied between studies which we hypothesized was due to timing of the thyroid hormone measurements, variability in PFAA isomer patterns, or presence of other stressors. In a longitudinal study design, we investigated the time-dependency of associations between PFAA isomers and thyroid hormones during pregnancy and post-partum while considering thyroid peroxidase antibody (TPOAb) status and mercury (Hg) co-exposure. In participants of a prospective Canadian birth cohort (n = 494), free thyroxine (FT4), free triiodothyronine (FT3), thyroid stimulating hormone (TSH) and TPOAb were quantified in maternal plasma collected in each trimester and 3-months postpartum, and 25 PFAAs (15 linear and 10 branched) and Hg were quantified in samples collected during the second trimester. Perfluorohexane

sulfonate (PFHxS) and total branched isomers of perfluorooctane sulfonate (PFOS) were positively associated with TSH in mixed-effect models, with strongest associations early in gestation. Throughout pregnancy and post-partum, PFHxS was inversely associated with FT4, consistent with elevated TSH, while Hg was inversely associated with FT3. In TPOAb-positive women, negative associations were found between PFUnA and FT4, and 1m-PFOS and TSH, supporting previous studies that thyroid disorder could increase susceptibility to PFAA-mediated hormone dysregulation. Hg did not confound associations but was a significant interaction term, revealing further positive associations between PFOS isomers (∑3m+4m-PFOS) and TSH. Higher perfluoroalkyl sulfonate exposures were associated with higher TSH and/or lower FT4, strongly suggestive that PFHxS and branched PFOS isomers are risk factors for subclinical maternal hypothyroidism. Isomer-specific analysis is important in future studies, as crude measures of 'total-PFOS' masked the associations of branched isomers. A concerning result was for PFHxS which had consistent negative associations with FT4 at all time points and a positive association with TSH in early pregnancy when fetal development is most sensitive to disruption.

Keywords: Longitudinal study design; Perfluoroalkyl acids; Perfluoroalkyl carboxylates; Perfluoroalkyl sulfonates; Pregnancy; Thyroid hormones.

I did a literature search and identified the following papers as potentially helpful for further study:

- Perfluoroalkyl and Polyfluoroalkyl Substances and Measures of Human Fertility: A Systematic Review Cathrine Carlsen Bach 1 2, Anne Vested 3 4, Kristian Tore Jørgensen 5, Jens Peter Ellekilde Bonde 5, Tine Brink Henriksen 1 6, Gunnar Toft 7 DOI: 10.1080/10408444.2016.1182117

Abstract

Perfluoroalkyl and polyfluoroalkyl substances (PFASs) are found widespread in the environment and humans. The relation of PFASs to fertility has now been examined in a relatively large number of epidemiologic studies and a synthesis is in order. The aim of this study was to assess the current human epidemiologic evidence on the association between exposure to PFASs and measures of human fertility, with particular emphasis on perfluorooctane sulfonate (PFOS) and perfluorooctanoate (PFOA). Systematic literature searches were initially conducted in MEDLINE and EMBASE and subsequently in references and citations of included papers. Studies were included if they assessed exposure to PFASs in biological samples in relation to reproductive hormones, semen characteristics, or time to pregnancy (TTP). Study characteristics and results were abstracted to predefined forms, and the studies were assessed for the risk of bias and confounding. Sixteen studies investigated the association between PFAS exposure in men and semen parameters, reproductive hormone levels, or TTP. There was a lack of consistent results among the numerous investigated exposure-outcome combinations. However, subtle associations between higher PFOS and lower testosterone or abnormal semen morphology cannot be excluded. Eleven studies assessed the association between PFAS exposure in women and TTP or reproductive hormones levels. Four of eight studies found prolonged TTP with higher PFOS or PFOA, but only one study found an association when restricting to nulliparous women. In men, there is little evidence of an association between PFAS exposure and semen quality or levels of reproductive hormones. For PFOS and PFOA, the literature indicates an association with female fecundability in parous women, which is most likely not causal.

Keywords: Epidemiology; fecundability; fecundity; fertility; humans; perfluorinated compounds; perfluoroalkyl and polyfluoroalkyl substances; perfluorooctane sulfonate; perfluorooctanoate; semen quality; time to pregnancy.

Similar articles:

- Perfluoroalkyl and polyfluoroalkyl substances and measures of human fertility: a systematic review. Bach CC, et al. Crit Rev Toxicol 2016 - Review. PMID 27268162
- Perfluoroalkyl and polyfluoroalkyl substances and human fetal growth: a systematic review. Bach CC, et al. Crit Rev Toxicol 2015 - Review. PMID 25372700
- Perfluoroalkyl substances and time to pregnancy in couples from Greenland, Poland and Ukraine. Jørgensen KT, et al. Environ Health 2014. Among authors: Bach CC. PMID 25533644 Free PMC article.
- Maternal Exposure to Perfluorinated Chemicals and Reduced Fecundity: The MIREC Study. MP Vélez et al. Hum Reprod 30 (3), 701-9. Mar 2015. PMID 25567616.

  The cumulative probabilities of pregnancy at 1, 6 and 12 months were 0.42 (95% confidence interval (CI) 0.40-0.45), 0.81 (95% CI 0.79-0.83) and 0.90 (95% CI 0.89-0.92), …

- Association of Perfluoroalkyl and Polyfluoroalkyl Substances With Premature Ovarian Insufficiency in Chinese Women. S Zhang et al. J Clin Endocrinol Metab 103 (7), 2543-2551. 2018. PMID 29986037.

  High exposure to PFOA, PFOS, and PFHxS is associated with increased risk of POI in humans.

- Prenatal Exposure to Perfluoroalkyl Substances and Birth Outcomes in a Spanish Birth Cohort. CB Manzano-Salgado et al. Environ Int 108, 278-284. Nov 2017. PMID 28917208.

  In this study, PFAS showed little association with birth outcomes. Higher PFHxS, PFOA, and PFNA concentrations were non-significantly associated with reduced birth weight …

- Exposure to Perfluoroalkyl Substances and Thyroid Function in Pregnant Women and Children: A Systematic Review of Epidemiologic Studies. V Ballesteros et al. Environ Int 99, 15-28. Feb 2017. PMID 27884404.

  Although there is a small number of studies with comparable data, we found some consistency of a positive association between maternal or teenage male exposure to some PF …

- Profiles of Emerging and Legacy Per-/Polyfluoroalkyl Substances in Matched Serum and Semen Samples: New Implications for Human Semen Quality. Y Pan et al. Environ Health Perspect 127 (12), 127005. Dec 2019. PMID 31841032.

  Our results suggest the potential for deleterious effects following exposure to 6:2 Cl-PFESA and other PFASs. Compared with serum PFAS levels, the much clearer association…

- Toxicokinetics of 8:2 Fluorotelomer Alcohol (8:2-FTOH) in Male and Female Hsd:Sprague Dawley SD Rats After Intravenous and Gavage Administration. MC Huang et al. Toxicol Rep 6, 924-932. 2019. PMID 31516843.

  Fluorotelomer alcohols (FTOHs) are used in the production of persistent per- and polyfluorinated alkyl substances (PFAS). Rodents and humans metabolize FTOHs to …

- Early Pregnancy Serum Levels of Perfluoroalkyl Substances and Risk of Preeclampsia in Swedish Women. S Wikström et al. Sci Rep 9 (1), 9179. 2019. PMID 31235847.

  Preeclampsia is a major cause of maternal and fetal morbidity. Emerging research shows an association with environmental exposures. The present aim was to investigate …

- Exposure to Perfluoroalkyl Substances During Fetal Life and Pubertal Development in Boys and Girls From the Danish National Birth Cohort. A Ernst et al. Environ Health Perspect 127 (1), 17004. Jan 2019. PMID 30628845.

  Our population-based cohort study suggests sex-specific associations of altered pubertal development with prenatal exposure to PFASs. These findings are novel, and …
- Conditioning on Parity in Studies of Perfluoroalkyl Acids and Time to Pregnancy: An Example From the Danish National Birth Cohort. C Bach et al. Environ Health Perspect 126 (11), 117003. Nov 2018. PMID 30417653.

  Associations between PFAAs and TTP in parous women may be biased by confounders related to previous pregnancies and exposure measurement error. To avoid these biases, …

1b. Is the selection of total T4 an appropriate biomarker/metric as it relates to clinically relevant hypothyroxinemia during pregnancy? Is such a measure applicable to both experimental test animals and humans?

i. If so, please explain your rationale.

ii. If not, are there other measures related to hypothyroxinemia that may be more useful for informing hazard potential during pregnancy? What are those measures?

Yes, and

Yes. T4 is a reasonable indicator that is often affected by chemicals thought to be important in influencing thyroid function.

1c. Has EPA clearly articulated the challenges associated with extrapolating the PFBS-induced decrease in thyroid hormone (e.g., total T4) in rodents to humans?

i. If so, please explain your reasoning.

ii. If not, please provide your rationale.

There has been a reasonable first effort at this.

Part of the challenge that could be discussed is whether or not there are differences in baseline T4 between rodents and humans, and how this affects the interspecies projection of effects on T4.

1d. Has EPA clearly articulated what is and is not known about the clinical implications of changes in T4 in women during pregnancy on neonates and infants?

i. If so, please explain your reasoning.

ii. If not, please provide your rationale.

The relevant section of the document for this would seem to be 4.1. However, I do not see discussion there of pregnancy-related implications of changes in T4. Some evidence is reported in the literature that some chemical exposures can cause disruptions in thyroid hormone levels [see Patrick, LND "Thyroid Disruption: Mechanisms and Clinical Implications in Human Health" Alternative Medicine Review 14(4):326-346], however the journal source of this paper ("alternative medicine review") leaves room for doubt on this apparent conclusion.

Some evidence is reported in the literature that some chemical exposures can cause disruptions in thyroid hormone levels [see Patrick, LND "Thyroid Disruption: Mechanisms and Clinical Implications in Human Health" Alternative Medicine Review 14(4):326-346], however the journal source of this paper ("alternative

medicine review") leaves room for doubt on this apparent conclusion. "Alternative medicine" suggests that the journal source identifies itself as not in the main line of medical thought.

2a. Are the dose-response modeling approaches, selection of benchmark response rate (BMR), and the selected models used to identify the thyroid effect-related POD for RfD derivation scientifically justified and clearly described?

- **i. If so, please explain your reasoning.**
- **ii. If not, please provide your rationale and identify alternative approaches, BMRs, and/or dose- response models that support the identification of alternative candidate POD(s) for the derivation of subchronic and chronic RfDs and provide the scientific support for the alternative choice(s).**

I don't see a clear justification by EPA for any specific choice of benchmark response rate.

I don't have further comments on these aspects of the analysis.

3a. Is applying the data-informed dosimetric adjustment that utilizes the ratio of the PFBS elimination half-life in mice to that in the human scientifically justified and clearly described?

- **i. If so, please explain your reasoning.**
- **ii. If not, please provide your rationale and identify an alternative approach to scale PFBS doses scientific support for the alternative choice.**

The choice suggested by EPA in this case seems arbitrary. Better thought, and explorations with examples of other thyroid-acting chemicals may be helpful for deriving a widely applicable projection rule.

I just did not see that the ratio used (PFBS elimination half lives in mice to humans) sufficiently removed uncertainty in interspecies projection to justify the reduction of the safety factor from 3 to 1.

I have now been provided with the unpublished Lau et al. paper that was the basis for the derivation of the proposed RfD value. This is a mouse study that involved two relatively widely spaced doses (30 and 300 mg/kg-day).

I find that, far from providing justification and documentation of the sufficiency of the proposed RfD, the paper simply does not contain a detailed justification for a proposed RfD, let alone the extraordinary proposed reduction of the UFA from 3 to 1. The proposed RfD therefore is not justified by the current document and should be revised downward.

4. EPA has evaluated and applied, where appropriate, uncertainty factors to account for intraspecies variability (UFH), interspecies differences (UFA), database limitations (UFD), duration (UFS), and LOAEL-to-NOAEL extrapolation (UFL) for PFBS.

4a. Does the provided qualitative scientific rationale support the application of the selected uncertainty factors? If not, please explain.

I do not find the proffered explanation sufficiently clear to be convincing. As it stands the rationale is not credible and cannot stand.

An uncertainty factor of 1 (rather than the usual 3) suggests that there is no remaining uncertainty in the interspecies projection. I just disagree.

4b. Has quantitative uncertainty been adequately accounted for in the derivation of the RfDs? Please describe and provide suggestions, if needed.

No.

It seems to me that there is still plenty of uncertainty in the true quantitative difference between (likely inbred) mouse susceptibility and the susceptibilities of the diverse arrays of humans who will be exposed.

4c. Do the methods used to derive the RfDs for PFBS appropriately account for uncertainties in evaluating the toxicokinetic differences between the experimental animal data and humans? If not, please explain.

No. EPA guidance on the application of interspecies uncertainty factors is:

> "The default value for UFH is 10-fold; the default value for interspecies uncertainty factor (UFA) is apportioned into a TD component valued at one-half order of magnitude and a TK component addressed via default inhalation dosimetry methods (U.S. EPA, 1994) or body-weight scaling for orally encountered compounds (U.S. EPA, 2011). DDEFs fall within this hierarchical range of approaches."
>
> (quote from "Guidance for Applying Quantitative Data to Develop Data-Derived Extrapolation Factors for Interspecies and Intraspecies Extrapolation" Office of the Science Advisor, Risk Assessment Forum, U.S. Environmental Protection Agency, EPA/100/R-14/002F, September 2014).

A single factor of 3 appears to have been used in this case. The factor should be increased, or additional reasoning provided to justify it.

It seems to me that it is for EPA to justify its substantial departure from previous practices on interspecies projection.

COMMENTS SUBMITTED BY

Lisa M. Kamendulis, Ph.D.
Associate Professor
and
Core Director, Oxidative Stress and Environmental Analysis Core
Department of Environmental Health
School of Public Health
Indiana University
Bloomington, Indiana

External Letter Peer Review of EPA's Draft Human Health Toxicity Values Assessment for Perfluorobutane Sulfonate (PFBS) (CASRN 375-73-5 [acid]) and Related Compound Potassium Perfluorobutane Sulfonate (CASRN 29420-49-3)

1. The key study chosen for determining the subchronic and chronic RfDs is the gestational exposure mouse study by Feng et al. (2017) and the critical effect is decreased total T4 in postnatal Day 1 (PND1) offspring.

1a. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and clearly described?

- **i. If so, please explain your reasoning.**
- **ii. If not, please provide your rationale and identify an alternative key study and/or critical effect to support the derivation of the subchronic and chronic RfDs and provide the scientific support for the alternative choice.**

The Feng et al., 2017 study was selected as the key study for the derivation of subchronic and chronic RfDs for PFBS, based on findings of PFBS-mediated decreases in total T3, total T4, and free T4. PFBS-induced alterations of the thyroid (decreases in total T3, total T4, and free T4) was selected as the critical effects, and was consistently observed across two species, sexes, life stages, and exposure durations in two independent, studies (NTP, 2019; Feng et al., 2017) that following systematic evaluation, were determined to be of "high-confidence". The Feng et al., 2017 study was a gestational study and identified adverse effects in PND1 thyroid that is considered appropriate for selection as the key study. The information pertaining to the study selection and identification of decreases in T4 as the critical effects have been clearly described and is scientifically based.

1b. Is the selection of total T4 an appropriate biomarker/metric as it relates to clinically relevant hypothyroxinemia during pregnancy? Is such a measure applicable to both experimental test animals and humans?

- **I. If so, please explain your rationale.**
- **II. If not, are there other measures related to hypothyroxinemia that may be more useful for informing hazard potential during pregnancy? What are those measures?**

In general, I agree that T4 is an appropriate biomarker to be used to derive RfDs since decreases in this parameter (coupled with normal TSH levels) are clinically relevant to hypothyroxinemia in pregnancy. During development, many organ systems are affected by altered thyroid homeostasis as the maintenance of adequate thyroid hormone levels are needed for their normal growth and development. As described in the document, rodents are considered to be a good model for evaluating the potential effects of chemicals on thyroid function in humans (Zoeller et al., 2007), and the pattern of decreased thyroid hormones in the absence of TSH changes and thyroid tissue weight and/or histology, observed in PFBS studies (e.g., (Feng et al., 2017), are consistent with the human clinical condition referred to as "hypothyroxinemia". The document could be more specific however, in stating that this is a clinical condition observed in *human pregnancy* in section 5. The evidence and data presented in section 6 clearly provides support for the clinical relevance of hypothyroxinemia during pregnancy and its relevance to developmental outcomes in both animals and humans.

1c. Has EPA clearly articulated the challenges associated with extrapolating the PFBS-induced decrease in thyroid hormone (e.g., total T4) in rodents to humans?

i. **If so, please explain your reasoning.**
ii. **If not, please provide your rationale.**

The document provided some information and one literature citation concerning challenges with extrapolation from rodent to human. The document stated that "*Although there are some differences in hypothalamic-pituitary-thyroid (HPT) regulation across species (e.g., serum hormone-binding proteins, hormone turnover rates, and timing of in utero thyroid development), rodents are generally considered to be a good model for evaluating the potential for thyroid effects of chemicals in humans (Zoeller et al., 2007).*" While these statements were included in the text of section 5. specifics on what the differences in serum hormone-binding proteins, hormone turnover rates, and timing of *in utero* thyroid development were not specified in that section. Section 6 contained significant details on these endpoints and their potential significance and differences between rodents and humans. It would be useful to the reader to indicate that additional details are provided in section 6.

1d. Has EPA clearly articulated what is and is not known about the clinical implications of changes in T4 in women during pregnancy on neonates and infants?

i. **If so, please explain your reasoning.**
ii. **If not, please provide your rationale.**

In general, the clinical implications of changes in T4 during pregnancy on neonates and infants was described. The document provided statements and citations describing that adequate levels of thyroid hormones are needed for normal growth and development in early life stages (Forhead and Fowden, 2014; Gilbert and Zoeller, 2010; Hulbert, 2000). Additional implications for thyroid hormone disruption and adverse developmental consequences were well described in section 6. Further, the document discussed that the presence of sufficient thyroid hormones during the gestational and neonatal period is essential for brain development and maturation. Importantly, the document identified that while altered thyroid hormone levels may be expected to impact neurodevelopment, no studies have evaluated the effect of PFBS on neurodevelopment, therefore there is uncertainty as to the potential developmental consequences of PFBS exposure.

2. In the public review draft PFBS assessment, EPA employed benchmark dose modeling (U.S. EPA, 2012) in the identification of a point-of-departure (POD) for derivation of RfD values, based on a decrease in total T4 levels in PND1 offspring. The 20% Relative Deviation (RD) Benchmark Response Rate (BMR) used in the public review draft is no longer being considered for BMD modeling of thyroid hormone dose-response data. As a result of extensive public review comments on the BMD approach (and thyroid hormone endpoint) used in the previous draft, and because a clear or consistent biological threshold for T4 changes associated with untoward developmental health outcomes has not be identified in the available literature, EPA has identified a new BMR of 0.5 SD (standard deviation change over controls) as a default in the revised PFBS draft assessment for the thyroid hormone alterations in mouse neonates/offspring. A 1 SD BMR is also being presented as the standardized basis for comparison as recommended in the EPA BMD Technical Guidance (U.S. EPA, 2012).

2a. Are the dose-response modeling approaches, selection of benchmark response rate (BMR), and the selected models used to identify the thyroid effect-related POD for RfD derivation scientifically justified and clearly described?

i. **If so, please explain your reasoning.**

ii. **If not, please provide your rationale and identify alternative approaches, BMRs, and/or dose-response models that support the identification of alternative candidate POD(s) for the derivation of subchronic and chronic RfDs and provide the scientific support for the alternative choice(s).**

In general, the approaches used to derive an RfD for PFBS were scientifically justified and well described. The document provides a clear line of evidence describing that in the existing data, there is no clear or consistent biological threshold for T4 changes associated with untoward developmental health outcomes. While BMD guidance would indicate a BMR of 1SD from control, (EPA 2012), a BMR of 0.5 SD was used as the default when performing BMD modeling on thyroid hormone and the potential developmental outcomes in offspring. A BMR of 0.5 SD from control is justified as effects in developing offspring, including thyroid hormone changes, should be used for effects occurring in a sensitive life stage.

3. **Due to the availability of new toxicokinetic data in mice noting significant interspecies differences in toxicokinetics of PFBS, and, recommendations from public commenters, EPA has applied a data-informed approach to convert the oral dose-rate in animals to a human equivalent dose (HED) in the identification of candidate points-of-departure (PODs) considered for the derivation of the RfDs (U.S. EPA, 2014; see https://www.epa.gov/sites/production/files/2015-01/documents/ddef-final.pdf). In considering the new evidence for serum half-life in mice published in Lau (in press), EPA concluded that the toxicokinetic data for PFBS are adequate to support calculation of data-derived dosimetric adjustment factors (DAF), where the ratio of elimination half-life in animals to that in humans, $T0.5_A/T0.5_H$, is used to adjust candidate PODs. By using in vivo animal and human half-life data to calculate POD(HEDs) that account for differences in toxicokinetics between rodents and humans, the potential role of interspecies differences in processes such as renal resorption, hepatic transport, and enterohepatic recirculation is reflected. Further, by using a data-derived approach the uncertainty in interspecies toxicokinetic scaling (UF_A) has been reduced from a 3 to a 1; however, residual uncertainty (due to the lack of information) pertaining to toxicodynamics exists and is acknowledged in the assessment in the description for applying a UF_A of 3.**

3a. **Is applying the data-informed dosimetric adjustment that utilizes the ratio of the PFBS elimination half-life in mice to that in the human scientifically justified and clearly described?**
   i. **If so, please explain your reasoning.**
   ii. **If not, please provide your rationale and identify an alternative approach to scale PFBS doses between rodents and humans and provide scientific support for the alternative choice.**

I agree that how the DAF was derived was clearly described in the document. However, the incorporation of new data for mouse serum $t_{1/2}$ was included (Lau, in press) and heavily relied upon to support this approach. While it is acknowledged that the reviewers were provided a preprint of this manuscript, at the time of submission of this review, several points remain unclear: 1) whether the manuscript been accepted for publication; 2) the stature and rigor to which journal was the manuscript submitted; and 3) the appropriateness of using data from a manuscript that is not yet publicly available. In addition, the derivation of a DAF using the approach presented herein is not my area of expertise, therefore I cannot fully comment of the scientific validity of this approach. However, in addition to reliance on the Lau manuscript for animal data, another concern with this approach is the limited human elimination data that is available, in particular for females. Although methods were applied to account for small sample sizes, the overall appropriateness of this approach is questioned.

4. 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 PFBS.

4a. Does the provided qualitative scientific rationale support the application of the selected uncertainty factors? If not, please explain.

UF_A - Due to the concerns raised in question 3 concerning how the DAF was derived, there might be an overall concern using an UF of 3 for the derivation of both subchronic and chronic RfDs for PFBS. However, should the application of another approach be used to derive an HED, it would likely result in the use of an UF of 3 - so this might not be problematic.

UF_S, UF_D – chronic RfD – an UF-S of 1 was applied despite the lack of chronic studies. However, as it is stated (EPA, 1991) developmental period is recognized as a susceptible life stage in which exposure is more relevant to the induction of developmental effects than lifetime exposure, therefore, an UF of 1 is justified. Further, an UF of 10 was applied for database limitations, therefore accounting for uncertainty for less than lifetime exposures.

4b. Has quantitative uncertainty been adequately accounted for in the derivation of the RfDs? Please describe and provide suggestions, if needed.

Yes, this was addressed.

4c. Do the methods used to derive the RfDs for PFBS appropriately account for uncertainties in evaluating the toxicokinetic differences between the experimental animal data and humans? If not, please explain.

Yes, this was addressed.

COMMENTS SUBMITTED BY

Angela M. Leung, MD
Health Sciences Clinical Assistant Professor of Medicine
David Geffen School of Medicine
University of California Los Angeles
and
Division of Endocrinology, Diabetes, and Metabolism
Department of Medicine
VA Greater Los Angeles Healthcare System
Los Angeles, California

External Letter Peer Review of EPA's Draft Human Health Toxicity Values Assessment for Perfluorobutane Sulfonate (PFBS) (CASRN 375-73-5 [acid]) and Related Compound Potassium Perfluorobutane Sulfonate (CASRN 29420-49-3)

1. The key study chosen for determining the subchronic and chronic RfDs is the gestational exposure mouse study by Feng et al. (2017) and the critical effect is decreased total T4 in postnatal Day 1 (PND1) offspring.

1a. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and clearly described?

- **i. If so, please explain your reasoning.**
- **ii. If not, please provide your rationale and identify an alternative key study and/or critical effect to support the derivation of the subchronic and chronic RfDs and provide the scientific support for the alternative choice.**

From the initial EPA draft assessment, the two organ systems demonstrating adverse effects from PFBS exposure with the highest level of confidence were the kidney and the thyroid gland. Table 6 of the current report summarizes the available studies regarding noncancer effects following oral PFBS administration. Unfortunately, there are no human pregnancy data in this area. Regarding animal data, the Feng et al 2017 and NTP 2019 studies both demonstrate the development of biochemical hypothyroidism following PFBS exposure. Between them, only the Feng et al 2017 mouse study examined this in mothers and their offspring, which are the vulnerable population subgroups of interest. Thus, I agree that it is the appropriate key study.

Measured thyroid biomarkers from the Feng et al 2017 were serum TSH, TT3, TT4, and FT4 in both dams and pups. Figure 4 in this study showed significant decreases in serum TT3 and TT4 levels at PNDs 1, 30, and 60 at the maternal 200 and 500 mg/kg/day PFBS doses (but not the 50 mg/kg/day dose) among pups. The paper does not report the pups' serum FT4 response, but presumably these data are available as per their methods section. The caveats between TT4 and FT4, and between rat and human thyroid physiology, should be noted, as outlined in my responses to questions 1b and 1c below. Taken together though, although not ideal, serum total T4 concentrations at Postnatal Day 1 would a reasonable critical effect from these animal data.

1b. Is the selection of [offspring] total T4 an appropriate biomarker/metric as it relates to clinically relevant hypothyroxinemia during pregnancy? Is such a measure applicable to both experimental test animals and humans?

- **i. If so, please explain your rationale.**
- **ii. If not, are there other measures related to hypothyroxinemia that may be more useful for informing hazard potential during pregnancy? What are those measures?**

It is noted that this question refers to serum TT4 levels among *offspring* of exposed mothers. There are two points to address in regard to this question:

1. Offspring T4 versus T3: It should be clarified that although it may be minimal, there is likely some T3 transport across the placenta. It is not completely absent, as stated on page 21 of the EPA responses to the previous draft report: "Keep in mind that TSH and T3 are not transported across the placenta", and in several areas of Section 6.1.1. in the current draft report. Please see some references:

Visser T. Thyroid hormone transport across the placenta. *Ann Endocrinol (Paris)* 2016;77:680-3. (https://www.ncbi.nlm.nih.gov/pubmed/27659266)

Porterfield SP et al. The role of thyroid hormones in prenatal and neonatal neurological development--current perspectives. *Endocr Rev* 1993 Feb;14(1):94-106. (https://www.ncbi.nlm.nih.gov/pubmed/8491157)

Calvo R et al. Congenital hypothyroidism, as studied in rats. Crucial role of maternal thyroxine but not of 3,5,3'-triiodothyronine in the protection of the fetal brain. *J Clin Invest* 1990 Sep;86(3):889-99. (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC296808/)

James et al. Placental transport of thyroid hormone. *Best Pract Res Clin Endocrinol Metab* 2007 Jun;21(2):253-64. (https://www.ncbi.nlm.nih.gov/pubmed/17574007)

Huang SA. Physiology and pathophysiology of type 3 deiodinase in humans. *Thyroid* 2005 Aug;15(8):875-81. (https://www.ncbi.nlm.nih.gov/pubmed/16131330)

However, I agree that the contribution of T3 toward overall thyroid status in the developing fetus is minimal, and it is well-accepted that T4 (whether total or free) is a much better marker than T3 of low thyroid status, including during pregnancy.

2. Offspring TT4 versus FT4: Both serum TT4 and FT4 concentrations are associated with inherent challenges in their interpretation, particularly during pregnancy. There are even less data of what the appropriate extrapolated measure of this would be for the offspring of pregnant mothers. The draft report (Section 6.1.1) addresses some of these issues.

   - For the *last* part of human pregnancy, the American Thyroid Association (ATA) guidelines for the management of thyroid disease in pregnancy recommend that serum TT4 is a more accurate measurement of thyroid status during this period, since the effect of thyroid binding proteins is less of an issue in later gestation (Recommendation 3; Alexander et al. 2017 Guidelines of the American Thyroid Association for the diagnosis and management of thyroid disease during pregnancy and the postpartum. *Thyroid* 2017:27:315-389 https://www.ncbi.nlm.nih.gov/pubmed/28056690)

   - However, the most vulnerable window of thyroid-dependent neurodevelopment is *very early* pregnancy (i.e. beginning as early as gestational weeks 3-4 in humans and most critically from gestational day 18 to postnatal days 21-25 in rats; see Bernal J. Thyroid hormone receptors in brain development and function. *Nat Rev Endocrinol*,2007;3:249-259 (https://www.ncbi.nlm.nih.gov/pubmed/17315033). In the *earlier* stages of pregnancy, the ATA also states that serum TT4 can be used and assessed in reference to an increasing upward bound that is dependent on gestational age, based on a study of 20 women (Weeke J et al. A longitudinal study of serum TSH, and total and free iodothyronines during normal pregnancy. *Acta Endocrinologica* 1982;101:531. https://www.ncbi.nlm.nih.gov/pubmed/7158229). However, early pregnancy is accompanied by a rapid rise of thyroid binding proteins that must be interpreted alongside serum TT4 levels (Glinoer D et al. Regulation of maternal thyroid during pregnancy.*J Clin Endocrinol Metabol* 1990;71:276-287. https://www.ncbi.nlm.nih.gov/pubmed/2116437), thus there are concerns that TT4 may not be the best thyroid biomarker during this critical window. Additionally, more recent evidence shows that

the serum TT4 variability is greater than that of serum FT4 during the first half of pregnancy, and importantly, that only FT4 (not TT4) was associated with several adverse birth outcomes in a cohort of 5,647 mother-child pairs (Korevaar TI et al. Maternal total T4 during the first half of pregnancy: physiologic aspects and the risk of adverse outcomes in comparison with free T4. *Clin Endocrinol (Oxf)* 2016 Nov;85(5):757-763. https://www.ncbi.nlm.nih.gov/pubmed/27187054). Although these issues may be lessened since TBG levels are relatively lower in rodents than humans, the majority of thyroid hormone is still in the bound form (to transthyretin and albumin) rodents; overall, only less than 1% of the thyroid hormones are in the unbound form among both species (Choksi NY et al. Role of thyroid hormones in human and laboratory animal reproductive health. *Birth Defects Res B Dev Reprod Toxicol*; 2003 Dec;68(6):479-91. https://www.ncbi.nlm.nih.gov/pubmed/14745982)

- Specifically, regarding placental transport of thyroid hormones during pregnancy, Section 6.1.1 of the draft report states: "Due to placental barrier functionality, free T4 levels in a pregnant dam might not be entirely representative of actual T4 status in a developing fetus." I agree that this might be true, but maternal FT4 is still probably the best available representation of offspring FT4 status. This is supported by the understanding that fetal T4 status is determined by several factors: Placental type 3 deiodinase inactivates maternal T4; of the small amount of remaining T4, the unbound portion is then transported across the placenta by both passive diffusion and active mechanisms, the latter via various transport proteins. See James et al. Placental transport of thyroid hormone. Best Pract Res Clin Endocrinol ; 2007 Jun;21(2):253-64 https://www.ncbi.nlm.nih.gov/pubmed/17574007. The relative contribution of the passive and active mechanisms to fetal T4 status is not well understood. As such, at present we can only rely of the best available measure of maternal status, particularly during early pregnancy, which is maternal FT4 for the reasons below.

- The complexities between animal and human thyroid physiology have also been recently summarized in the following ATA guideline: Bianco AC et al. American Thyroid Association Guide to investigating thyroid hormone economy and action in rodent and cell models. *Thyroid* 2014 Jan;24(1):88-168 https://www.ncbi.nlm.nih.gov/pubmed/24001133.

- Taken together, there is no perfect assessment of thyroid function during *early* pregnancy, when the effects of maternal PFBS exposure would be the most critical, and thus the accuracy of extrapolating these measurements to their offspring is similarly incompletely understood. However, given that FT4 appears to be less affected by binding proteins and maternal FT4 has been associated with adverse clinical outcomes in offspring, it appears that it would be the better representation of thyroid status among offspring of exposed mothers. It may be worthwhile to assess whether data for pups' serum FT4 levels are available from the Feng et al 2017 study, as FT4 would be a better measurement of hypothyroxinemia during pregnancy. If not, serum TT4 would be an alternate reasonable, albeit imperfect, marker of thyroid status during early pregnancy. It is noted that there is unfortunately a paucity of available data on this topic.

1c. Has EPA clearly articulated the challenges associated with extrapolating the PFBS-induced decrease in thyroid hormone (e.g., total T4) in rodents to humans?

i. If so, please explain your reasoning.

ii. If not, please provide your rationale.

The draft report has been carefully organized, and the challenges of interpreting thyroid physiology during pregnancy across species are particularly well-described in Section 6.1.1. It may be also helpful to note that the newborn rat is developmentally equivalent to the human 4-5-month-old fetus, thus there would be important differences regarding the relative contribution of maternal thyroid hormones to the developing fetus at similar PND1. See: Bernal J. Thyroid hormone receptors in brain development and function. *Nat Rev Endocrinol* 2007;3:249-259 https://www.ncbi.nlm.nih.gov/pubmed/17315033.

1d. Has EPA clearly articulated what is and is not known about the clinical implications of changes in T4 in women during pregnancy on neonates and infants?

i. If so, please explain your reasoning.

ii. If not, please provide your rationale.

It may be helpful to emphasize that the conventional definition of hypothyroxinemia is really utilized only in pregnancy and based on subnormal serum FT4 levels in the setting of normal serum TSH concentrations; outside of pregnancy, the clinical relevance of this entity is unknown. For example, outside of pregnancy, hypothyroxinemia has been mostly described in some premature infants and has not been rigorously studied in other age groups or populations (See Rapaport R et al. Hypothyroxinemia in the preterm infant: the benefits and risks of thyroxine treatment. *J Pediatr* 2001 Aug;139(2):182-8. https://www.ncbi.nlm.nih.gov/pubmed/11487741)

Thus, using hypothyroxinemia as a critical effect among offspring is based on our understanding of this condition in their mothers. As such, hypothyroxinemia (in pregnancy) is usually not defined by TT4 levels, since the predominant (and even perhaps all available) studies assessing the adverse clinical consequences of hypothyroxinemia in pregnancy have only been based on FT4. Please see Negro R et al. Hypothyroxinemia and pregnancy. *Endocr Pract* 2011 May-Jun;17(3):422-9 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3637943/) as an excellent review on this topic. Additionally, in the draft report Section 6.1.1., it might be better to thus not describe hypothyroxinemia as a form of hypothyroidism, which is traditionally understood as either subclinical or overt biochemical hypothyroidism.

2. In the public review draft PFBS assessment, EPA employed benchmark dose modeling (U.S. EPA, 2012) in the identification of a point-of-departure (POD) for derivation of RfD values, based on a decrease in total T4 levels in PND1 offspring. The 20% Relative Deviation (RD) Benchmark Response Rate (BMR) used in the public review draft is no longer being considered for BMD modeling of thyroid hormone dose-response data. As a result of extensive public review comments on the BMD approach (and thyroid hormone endpoint) used in the previous draft, and because a clear or consistent biological threshold for T4 changes associated with untoward developmental health outcomes has not be identified in the available literature, EPA has identified a new BMR of 0.5 SD (standard deviation change over controls) as a default in the revised PFBS draft assessment for the thyroid hormone alterations in mouse neonates/offspring. A 1 SD BMR is also being presented as the standardized basis for comparison as recommended in the EPA BMD Technical Guidance (U.S. EPA, 2012).

2a. Are the dose-response modeling approaches, selection of benchmark response rate (BMR), and the selected models used to identify the thyroid effect-related POD for RfD derivation scientifically justified and clearly described?

i. **If so, please explain your reasoning.**
ii. **If not, please provide your rationale and identify alternative approaches, BMRs, and/or dose-response models that support the identification of alternative candidate POD(s) for the derivation of subchronic and chronic RfDs and provide the scientific support for the alternative choice(s).**

Benchmark dose modeling is not my area of expertise; thus I defer to the other reviewers.

3. Due to the availability of new toxicokinetic data in mice noting significant interspecies differences in toxicokinetics of PFBS, and, recommendations from public commenters, EPA has applied a data-informed approach to convert the oral dose-rate in animals to a human equivalent dose (HED) in the identification of candidate points-of-departure (PODs) considered for the derivation of the RfDs (U.S. EPA, 2014; see https://www.epa.gov/sites/production/files/2015-01/documents/ddef-final.pdf). In considering the new evidence for serum half-life in mice published in Lau (in press), EPA concluded that the toxicokinetic data for PFBS are adequate to support calculation of data-derived dosimetric adjustment factors (DAF), where the ratio of elimination half-life in animals to that in humans, T0.5A/T0.5H, is used to adjust candidate PODs. By using in vivo animal and human half-life data to calculate POD(HEDs) that account for differences in toxicokinetics between rodents and humans, the potential role of interspecies differences in processes such as renal resorption, hepatic transport, and enterohepatic recirculation is reflected. Further, by using a data-derived approach the uncertainty in interspecies toxicokinetic scaling (UFA) has been reduced from a 3 to a 1; however, residual uncertainty (due to the lack of information) pertaining to toxicodynamics exists and is acknowledged in the assessment in the description for applying a UFA of 3.

3a. Is applying the data-informed dosimetric adjustment that utilizes the ratio of the PFBS elimination half-life in mice to that in the human scientifically justified and clearly described?
i. **If so, please explain your reasoning.**
ii. **If not, please provide your rationale and identify an alternative approach to scale PFBS doses between rodents and humans and provide scientific support for the alternative choice.**

This is not an area that I am able to comment on, thus I defer to the other reviewers.

4. 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 PFBS.

4a. Does the provided qualitative scientific rationale support the application of the selected uncertainty factors? If not, please explain.

This is not an area that I am able to comment on, thus I defer to the other reviewers.

4b. Has quantitative uncertainty been adequately accounted for in the derivation of the RfDs? Please describe and provide suggestions, if needed.

This is not an area that I am able to comment on, thus I defer to the other reviewers.

4c. Do the methods used to derive the RfDs for PFBS appropriately account for uncertainties in evaluating the toxicokinetic differences between the experimental animal data and humans? If not, please explain.

This is not an area that I am able to comment on, thus I defer to the other reviewers.

COMMENTS SUBMITTED BY

Angela L. Slitt, Ph.D.
Associate Professor
Department of Biomedical and Pharmaceutical Sciences
University of Rhode Island
Kingston, Rhode Island

External Letter Peer Review of EPA's Draft Human Health Toxicity Values Assessment for Perfluorobutane Sulfonate (PFBS) (CASRN 375-73-5 [acid]) and Related Compound Potassium Perfluorobutane Sulfonate (CASRN 29420-49-3)

1. The key study chosen for determining the subchronic and chronic RfDs is the gestational exposure mouse study by Feng et al. (2017) and the critical effect is decreased total T4 in postnatal Day 1 (PND1) offspring.

1a. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and clearly described?

i. If so, please explain your reasoning.

Response: The selection of the total thyroxine, free thyroxine, and total triiodothyronine are well justified critical effects. The document explains very well the effects of thyroid hormone disruption on health endpoints and makes a solid justification for thyroid disruption as a critical concern. Thyroid hormone serves many functions during development and throughout the life span. With regard to development, thyroid hormone is thought impact the neuronal, reproductive, hepatic, and immune system. It is also known to influence brain development. Feng et al. (2017) is the key study that describes decreased T4 in PND1 offspring as the critical effect. The study was considered to be of high quality based on the study design metrics evaluated. Strengths of the study were that it was well powered (n=10 dams per treatment), 3 doses included, appropriate statistical analysis, and additional endpoints measured. The work reports changes in both maternal and offspring T4 and TSH. Because this work reported decreased serum T4 in the offspring with a rebound increase in TSH, it is felt to be a clearer observation to use for alteration in thyroid hormone.

ii. If not, please provide your rationale and identify an alternative key study and/or critical effect to support the derivation of the subchronic and chronic RfDs and provide the scientific support for the alternative choice.

Response: I do support the observed decrease in serum T4 levels as observed in the NTP, 2019 as an alternative key critical effect, even with the lack of observed increase in serum TSH levels because thyroid hormone has pleiotropic effects, and decreased T4 is associated with numerous health poor outcomes. The NTP, 2019 study is rigorously described and provides a higher quality study to utilize despite the lack of rebound TSH. All aspects of the study are well described and documented.

1b. Is the selection of total T4 an appropriate biomarker/metric as it relates to clinically relevant hypothyroxinemia during pregnancy? Is such a measure applicable to both experimental test animals and humans?

Response: Yes to both questions. clinically relevant hypothyroxinemia during pregnancy is a relevant biomarker. There are multiple human clinical studies that cite hypothyroxinemia as a potential issue for worse outcomes for the pregnancy, as well as for the offspring. Here are some examples of recent studies. Hypothyroxinemia during pregnancy has been associated with altered reaction in 5-6 year olds (Finken et al., J Clin Endocrinol Metab. 2013); Lower non-verbal IQ in children 5-8 years old (Levie et al., J Clin Endocrinol Metab. 2018), adverse neuropsychological function of the child at 5 years of age. Additionally, marked hypothyroidism was has been associated with motor function and executive and behavior problems (Andersen et al., J Clin Endocrinol Metab. 2018). Studies also point to rodent models that induce

hypothyroidism during pregnancy can have adverse effects of the development of the nervous system (Berbel et al., Cereb Cortex. 2010; Wei et al., Environ Toxicol. 2015), autism (Sadamatsu et al., Congenit Anom, 2006).

1c. Has EPA clearly articulated the challenges associated with extrapolating the PFBS-induced decrease in thyroid hormone (e.g., total T4) in rodents to humans?

Response: In my opinion, the "challenges" could be more clearly articulated. The information provided on pages 71 and 72 are very detailed with regard to laying out a foundational knowledge of T4, T3, and thyroid hormone metabolism regulation during pregnancy. However, the document could further expand on outcomes and mechanisms that are similar between rodents and humans, versus any proposed differences that could be an issue interpreting data between species. It would be good to have a paragraph that specifically addresses this concern with very pointed writing.

1d. Has EPA clearly articulated what is and is not known about the clinical implications of changes in T4 in women during pregnancy on neonates and infants?

Response: Yes. The information provided on pages 71 and 72 describe in detail the role of T4 hormone during pregnancy and the relationship between maternal T4 levels and outcomes in the offspring. The writing is detailed and well cited.

2. In the public review draft PFBS assessment, EPA employed benchmark dose modeling (U.S. EPA, 2012) in the identification of a point-of-departure (POD) for derivation of RfD values, based on a decrease in total T4 levels in PND1 offspring. The 20% Relative Deviation (RD) Benchmark Response Rate (BMR) used in the public review draft is no longer being considered for BMD modeling of thyroid hormone dose-response data. As a result of extensive public review comments on the BMD approach (and thyroid hormone endpoint) used in the previous draft, and because a clear or consistent biological threshold for T4 changes associated with untoward developmental health outcomes has not be identified in the available literature, EPA has identified a new BMR of 0.5 SD (standard deviation change over controls) as a default in the revised PFBS draft assessment for the thyroid hormone alterations in mouse neonates/offspring. A 1 SD BMR is also being presented as the standardized basis for comparison as recommended in the EPA BMD Technical Guidance (U.S. EPA, 2012).

2A. Are the dose-response modeling approaches, selection of benchmark response rate (BMR), and the selected models used to identify the thyroid effect-related POD for RfD derivation scientifically justified and clearly described?

Response: This is outside of my area of expertise.

3. Due to the availability of new toxicokinetic data in mice noting significant interspecies differences in toxicokinetics of PFBS, and, recommendations from public commenters, EPA has applied a data-informed approach to convert the oral dose-rate in animals to a human equivalent dose (HED) in the identification of candidate points-of-departure (PODs) considered for the derivation of the RfDs (U.S. EPA, 2014; see https://www.epa.gov/sites/production/files/2015-01/documents/ddef-final.pdf). In considering the new evidence for serum half-life in mice published in Lau (in press), EPA concluded that the toxicokinetic data for PFBS are adequate to support calculation of data-derived dosimetric adjustment factors (DAF), where the ratio of elimination half-life in animals to that in humans, T0.5A/T0.5H, is used to adjust candidate PODs. By using in vivo animal and human half-life data to calculate POD(HEDs) that account for differences in toxicokinetics between rodents and humans, the potential role of interspecies differences in processes such as renal resorption, hepatic

transport, and enterohepatic recirculation is reflected. Further, by using a data-derived approach the uncertainty in interspecies toxicokinetic scaling (UFA) has been reduced from a 3 to a 1; however, residual uncertainty (due to the lack of information) pertaining to toxicodynamics exists and is acknowledged in the assessment in the description for applying a UFA of 3.

3A. Is applying the data-informed dosimetric adjustment that utilizes the ratio of the PFBS elimination half-life in mice to that in the human scientifically justified and clearly described?

Response: Yes, this approach is reasonable and does provide an adjustment for the marked differences in species toxicokinetics. As mentioned the processes that dictate renal resorption, hepatic transport, and enterohepatic recirculation doe have species differences with regard to transporter affinity, function, and even localization. This is the most reasonable method to scale from mouse to human.

4. EPA has evaluated and applied, where appropriate, uncertainty factors to account for intraspecies variability (UFH), interspecies differences (UFA), database limitations (UFD), duration (UFS), and LOAEL-to-NOAEL extrapolation (UFL) for PFBS.

4a. Does the provided qualitative scientific rationale support the application of the selected uncertainty factors? If not, please explain.

Response: Yes. Overall, the factors that have been accounted for and described in Table 10 support the application of uncertainty factors. There are gaps in our knowledge regarding toxicokinetics for newborns, interindividual variability in the toxicokinetic and toxicodynamic response, and lack of literature.

4b. Has quantitative uncertainty been adequately accounted for in the derivation of the RfDs? Please describe and provide suggestions, if needed.

Response: UF_H – UF of 10 is appropriate. Examples are evidence of polymorphisms in xenobiotic transporters, such as OATPs and OATs, in which PFAS are likely a substrate (Yee et al. Clin Pharmacol Ther. 2018 Nov; 104(5): 803–817). Renal function in the elderly or disease also can be impacted. UF_D – UF of 10 is appropriate. The literature is quite limited and lacking for some health effects known to occur with other PFAS. UF_L - UF of 1 is appropriate as rationalized in the document. UF_S – 1 is properly justified based on (U.S. EPA, 1991b).

4c. Do the methods used to derive the RfDs for PFBS appropriately account for uncertainties in evaluating the toxicokinetic differences between the experimental animal data and humans? If not, please explain.

Response: No, it does not. The UF of 3 might be too generous given the limited knowledge about PFBS kinetics in neonates or children. Given that PFBS is cleared through renal clearance to urine (Bogdanska et al., 2014), there should be extra considering given to the differences between infants and adults with regard to renal clearance. Given that the main critical effect was observed at PND1, this is a relevant consideration. As reviewed by Rodieux et al., Clin Pharmacokinet. 2015; 54: 1183–1204, the effect of kidney function on children with regard to pharmacokinetics and dosing is largely unknown, with ~80% of drugs given to children not having been studied. The impact on renal clearance with regard to other PFAS infants is also largely unknown. Given the likely clearance of PFBS by human kidney and the uncertainty of kidney function with regard to PFBS clearance in children, I believe this UF could be higher.

COMMENTS SUBMITTED BY

David Alan Warren, MPH, Ph.D.
Program Director, Environmental Health Science
University of South Carolina Beaufort
Beaufort, South Carolina

External Letter Peer Review of EPA's Draft Human Health Toxicity Values Assessment for Perfluorobutane Sulfonate (PFBS) (CASRN 375-73-5 [acid]) and Related Compound Potassium Perfluorobutane Sulfonate (CASRN 29420-49-3)

1. The key study chosen for determining the subchronic and chronic RfDs is the gestational exposure mouse study by Feng et al. (2017) and the critical effect is decreased total T4 in postnatal Day 1 (PND1) offspring.

1a. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and clearly described?

- **i. If so, please explain your reasoning.**
- **ii. If not, please provide your rationale and identify an alternative key study and/or critical effect to support the derivation of the subchronic and chronic RfDs and provide the scientific support for the alternative choice.**

Yes, the revised draft clearly and thoroughly provides the scientific justification for selection of Feng et al. (2017) as principal study and decreased T4 in PND1 offspring as critical effect. Importantly, it explains the rationale behind these preferences over other candidate studies and effects (e.g., citing comparative sensitivity to the renal hyperplasia observed in adult rats (Lieder et al., 2009a,b) and questions about the biological significance of decreased T4 in adult rats in the absence of overt thyroid toxicity (NTP, 2019)). Compared to those in the original draft assessment (July 2018), the revised subchronic and chronic RfDs are one to two orders of magnitude lower. This is also the case when the revised RfDs are compared to the developmental RfD calculated in the original draft, despite use of the same principal study and critical effect. Thus, the revised toxicity values are more health conservative than those in the original draft, and more importantly, reflect a stricter adherence to U.S. EPA methodologies for toxicity value derivation.

1b. Is the selection of total T4 an appropriate biomarker/metric as it relates to clinically relevant hypothyroxinemia during pregnancy? Is such a measure applicable to both experimental test animals and humans?

- **i. If so, please explain your rationale.**
- **ii. If not, are there other measures related to hypothyroxinemia that may be more useful for informing hazard potential during pregnancy? What are those measures?**

Yes on both accounts. Biomarker selection in the present context is a challenge given the complexity of thyroid physiology, its species variability, multiple mechanistic possibilities by which PFBS might perturb thyroid hormone homeostasis (e.g., increased hepatic T4 glucuronidation; increased thyroidal conversion of T4 to T3), and the diverse array of adverse developmental endpoints under the control of one or more thyroid hormones. As such, the selection of total T4 appears to be the most appropriate biomarker-of-effect since, as stated in the revised draft, it *represents the aggregate of potential endocrine thyroid signaling (i.e., free T4 + protein bound T4) at any given time*. Furthermore, since similar patterns of decreases in total T3, total T4 and free T4 were observed in the principal study and that of NTP (2019), selecting total T4 when one of the two alternatives would have been more appropriate is of lesser consequence than if the three candidate biomarkers had been differentially affected by PFBS. Selection of total T4 as a biomarker is also supported by evidence that T3 is unable to cross the blood-brain barrier during fetal development. As a result, all T3 in the fetal brain is locally derived from T4 by deiodination. Interestingly, deiodinase-deficient mice do not generally exhibit altered brain development or functional

deficits and the predominant isoform of thyroid hormone receptor in brain responds to both T3 and T4. This suggests that T4 may play a more active role in brain physiology than has been previously accepted. As to whether total T4 is applicable to both experimental animals and humans, the highly conserved structure and function of the thyroid among mammalian species suggest so. So too does the considerable concordance in the adverse effects observed secondary to hypothyroxinemia in humans and animals (e.g., see Crofton (2004) on the relationship between decreased total T4 and hearing loss). This is not to say species differences (e.g., metabolic turnover rates; windows of susceptibility; dose-response relationships between hormonal disruption and toxicity) can't impact the interpretation of rodent thyroid toxicity data in terms of predicting effects in humans. Rather, such differences must be appreciated and accounted for by acknowledging their contribution to uncertainty in the derivation of toxicity values.

1c. Has EPA clearly articulated the challenges associated with extrapolating the PFBS-induced decrease in thyroid hormone (e.g., total T4) in rodents to humans?

i. If so, please explain your reasoning.

ii. If not, please provide your rationale.

Yes. The revised draft includes considerable discussion of interspecies (human vs. rodent) differences, as well as commonalities. For example, species differences in the time course of HPT axis development and regulation are noted, as are differences in the fraction of gestation during which fetal development is entirely dependent on maternal thyroid hormone. On the other hand, there is also a brief discussion supporting the lack of a significant species difference in the hormonal reserve capacity between human and rodent neonates. In addition to the rodent studies of Feng et al. (2017) and NTP (2019), several human epidemiological studies of pregnant women with decreased thyroid hormone levels are discussed. Though the neurodevelopmental status of their offspring was examined as well, neither the rodent nor human studies were sufficient to identify a BMR with any degree of certainty. However, the magnitude of T4 decrease associated with developmental sequelae in both species, albeit based on limited data, appears to be roughly comparable. While the challenges of interspecies extrapolation are clearly articulated in the revised draft, consideration might be given to referencing the following report, recently published and comprehensive:

> *A literature review of the state of the science regarding species differences in the control of, and response to, thyroid hormone perturbations, Part 1: Human health perspective*. Report prepared for Sponsor: European Crop Protection Association, Prepared by: Regulatory Science Associates, Regulatory Science Ltd1, Kip Marina, Inverkip, Renfrewshire, PA16 0AS, APRIL 2018.

Also, featuring more prominently the studies of Yang et al. (2016) and Wang et al. (2014) of PFAS other than PFBS, might strengthen support for the extrapolation of rodent data to pregnant women and their offspring (and potentially, selection of total T4 as a biomarker). Lastly, among the statements in the revised draft that speak to the issue of interspecies extrapolation are the following, all of which this reviewer considers supported:

> "rodents are generally considered to be a good model for evaluating the potential for thyroid effects of chemicals in humans (Zoeller et al., 2007),"

"these interrelated developmental effects in mice (i.e., delays and hormonal changes) are coherent with effects on the thyroid and presumed to be directly relevant to similar processes in humans; however, studies evaluating these outcomes in humans are not available,"

and

"the selection of total T4 as the critical effect is based on a number of key considerations that account for cross-species correlations in thyroid physiology and hormone dynamics particularly within the context of a developmental life stage."

1d. Has EPA clearly articulated what is and is not known about the clinical implications of changes in T4 in women during pregnancy on neonates and infants?

i. If so, please explain your reasoning.

ii. If not, please provide your rationale.

Yes, pages 63-64 present a succinct description of human epidemiological studies of pregnant women with decreased thyroid hormone levels and the neurodevelopmental status of their offspring. The revised draft clearly makes the following points: 1) associations between thyroid hormone levels in pregnant mothers and neurodevelopment in their offspring are inconsistent; 2) the inconsistency may be associated with variable timing of hypothyroxinemia during pregnancy; 3) the inconsistency may also be associated with variable types of maternal hypothyroidism, only one of which involves a subnormal T4 concentration; 4) the magnitude of T4 decrease associated with developmental sequelae, albeit based on limited data, appears roughly comparable in humans and rodents; and 5) ultimately, the database does not allow, to a reasonable degree of certainty, identification of the minimum extent of T4 decrease necessary for adverse developmental outcomes. Perhaps one addition to the discussion of clinical studies might be in order – that is, noting that most rely on maternal free T4 as a measure of thyroid hormone status rather than total T4 in PND1 offspring on which the POD is based.

2. In the public review draft PFBS assessment, EPA employed benchmark dose modeling (U.S. EPA, 2012) in the identification of a point-of-departure (POD) for derivation of RfD values, based on a decrease in total T4 levels in PND1 offspring. The 20% Relative Deviation (RD) Benchmark Response Rate (BMR) used in the public review draft is no longer being considered for BMD modeling of thyroid hormone dose-response data. As a result of extensive public review comments on the BMD approach (and thyroid hormone endpoint) used in the previous draft, and because a clear or consistent biological threshold for T4 changes associated with untoward developmental health outcomes has not be identified in the available literature, EPA has identified a new BMR of 0.5 SD (standard deviation change over controls) as a default in the revised PFBS draft assessment for the thyroid hormone alterations in mouse neonates/offspring. A 1 SD BMR is also being presented as the standardized basis for comparison as recommended in the EPA BMD Technical Guidance (U.S. EPA, 2012).

2a. Are the dose-response modeling approaches, selection of benchmark response rate (BMR), and the selected models used to identify the thyroid effect-related POD for RfD derivation scientifically justified and clearly described?

i. If so, please explain your reasoning.

ii. If not, please provide your rationale and identify alternative approaches, BMRs, and/or dose-response models that support the identification of alternative candidate POD(s) for the derivation of subchronic and chronic RfDs and provide the scientific support for the alternative choice(s).

Yes, the revised draft provides a clear description and adequate scientific justification for dose-response modeling approaches (including the NOAEL/LOAEL approach for data not amenable to benchmark dose modeling), use of several BMRs for continuous and dichotomous data (BMR of 1 SD from control mean, BMR of 0.5 SD from control mean, BMR of 10% extra risk), and the ultimate selection of the exponential 4 model based on it returning the lowest BMDL (Table F-2). Not addressed in the text, however, is the revised draft's consideration of both litter and individual fetuses as the experimental unit, with the former being amenable to BMD modeling, but not the latter. The revised draft is written with transparency clearly in mind, as renal hyperplasia (from the original draft's principal study, no less) and developmental delay data were modeled and PODs presented for comparative purposes. The ultimate selection of total T4 in PND1 offspring (and 0.5 SD from control mean) as the BMR, as noted, is consistent with U.S. EPA policy given the uncertainty surrounding the response level to consider adverse and the use of data from a particularly susceptible lifestage. The discussion of reserve thyroid hormone capacity was particularly effective as partial justification for the selection of PND1 mice, especially as some model-derived PODs based on total T4 in adults were at or below that selected for RfD derivation.

3. Due to the availability of new toxicokinetic data in mice noting significant interspecies differences in toxicokinetics of PFBS, and, recommendations from public commenters, EPA has applied a data-informed approach to convert the oral dose-rate in animals to a human equivalent dose (HED) in the identification of candidate points-of-departure (PODs) considered for the derivation of the RfDs (U.S. EPA, 2014; see https://www.epa.gov/sites/production/files/2015-01/documents/ddef-final.pdf). In considering the new evidence for serum half-life in mice published in Lau (in press), EPA concluded that the toxicokinetic data for PFBS are adequate to support calculation of data-derived dosimetric adjustment factors (DAF), where the ratio of elimination half-life in animals to that in humans, $T0.5_A/T0.5_H$, is used to adjust candidate PODs. By using in vivo animal and human half-life data to calculate POD(HEDs) that account for differences in toxicokinetics between rodents and humans, the potential role of interspecies differences in processes such as renal resorption, hepatic transport, and enterohepatic recirculation is reflected. Further, by using a data-derived approach the uncertainty in interspecies toxicokinetic scaling (UF_A) has been reduced from a 3 to a 1; however, residual uncertainty (due to the lack of information) pertaining to toxicodynamics exists and is acknowledged in the assessment in the description for applying a UF_A of 3.

3a. Is applying the data-informed dosimetric adjustment that utilizes the ratio of the PFBS elimination half-life in mice to that in the human scientifically justified and clearly described?

- **i. If so, please explain your reasoning.**
- **ii. If not, please provide your rationale and identify an alternative approach to scale PFBS doses between rodents and humans and provide scientific support for the alternative choice.**

Yes. In what is another radical departure from the original draft, the default extrapolation procedure of body weight$^{3/4}$ has been superseded by application of a data-informed adjustment factor based on the ratio of animal to human serum PFBS elimination half-lives. This is scientifically justified, consistent with U.S. EPA's hierarchy of approaches to dosimetric adjustment in the derivation of RfDs and is appropriate given the lack of confidence in the one existing PBPK model for extrapolation purposes. Pages 65-67 clearly describe the availability of clearance and half-life data, including the recent addition of sex-specific half-life data in mice. On the subject of sex-specific half-lives, it is noteworthy that the PFBS half-life measured in the one female subject in Olsen et al. (2009) was nearly twice that of the mean half-

life of the five male subjects. Thus, use of the geometric mean value of the six human subjects to calculate the DAF creates the possibility that animal doses were not adjusted sufficiently downward. Lastly, it is worth emphasizing that some uncertainty in the derived RfDs stems from two common assumptions severely lacking in empirical validation - 1) that total T4 concentration in humans and mice will respond with equal sensitivity to the same internal or target tissue dose of PFBS, and 2) that the average serum concentration of PFBS over time is the dose metric mechanistically linked to thyroid hormone economy.

4. 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 PFBS.

4a. Does the provided qualitative scientific rationale support the application of the selected uncertainty factors? If not, please explain.

Yes, the explanatory text in Tables 10 and 12 provides scientific rationale appropriate to each of the individual UFs. I have no issues with the UFs for the subchronic or chronic RfDs.

4b. Has quantitative uncertainty been adequately accounted for in the derivation of the RfDs? Please describe and provide suggestions, if needed.

Yes, the composite UFs of 100 and 300 applied for derivation of the subchronic and chronic RfDs, respectively, are appropriate. Again, I have no issues with the UFs for either toxicity value.

4c. Do the methods used to derive the RfDs for PFBS appropriately account for uncertainties in evaluating the toxicokinetic differences between the experimental animal data and humans? If not, please explain.

Yes, an UF_A of 3 is applied for both subchronic and chronic RfD derivation. This value seems appropriate given the use of the mouse:human ratio of PFBS half-life (0.0073) as the DAF on one hand, and the absence of toxicokinetic data during sensitive life stages on the other.

COMMENTS SUBMITTED BY

R. Thomas Zoeller, Ph.D.
Professor Emeritus
Department of Biology
University of Massachusetts
Amherst, Massachusetts

External Letter Peer Review of EPA's Draft Human Health Toxicity Values Assessment for Perfluorobutane Sulfonate (PFBS) (CASRN 375-73-5 [acid]) and Related Compound Potassium Perfluorobutane Sulfonate (CASRN 29420-49-3)

1. The key study chosen for determining the subchronic and chronic RfDs is the gestational exposure mouse study by Feng et al. (2017) and the critical effect is decreased total T4 in postnatal Day 1 (PND1) offspring.

1a. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and clearly described?

The Agency has clearly described the choice of Feng et al. (2017) as the key study and the decrease in serum total T4 in postnatal day 1 offspring as the critical effect. There were two general reasons for this. First, three hazards of PFBS exposure were identified, including serum total T4, renal toxicity and developmental. The thyroid endpoints were chosen both because there is more confidence that it represents a hazard to human health compared to the others, and because effects were observed at a lower dose. These considerations were very well described in the report.

One concern about the Feng et al. study is that serum T4 in P1 control pups is reported to be at the level of sensitivity of the assay as they report in the Methods section. This was not obvious since they reported the sensitivity in terms of ng/mL and report serum total T4 levels in figure 4B in terms of µg/dL. I was not able to obtain the specification sheets from the manufacturer to ensure that there was no error in reporting. A LOQ for mouse serum total T4 of 1.4 µg/dL is similar to other kits. However, if this is correct, it means that the measurement of "reduced" serum total T4 in treated animals would be below the LOQ.

The scientific justification was also well reasoned by the Agency. First, it is clear that thyroid hormone is chemically identical among all vertebrates. Thyroid hormone is essential for normal brain development in all mammals including mice and humans. Moreover, the Agency made cogent arguments both for the use of total T4 as the index of adverse effect and for choosing the neonatal period as being most relevant.

1b. Is the selection of total T4 an appropriate biomarker/metric as it relates to clinically relevant hypothyroxinemia during pregnancy? Is such a measure applicable to both experimental test animals and humans?

Serum total T4 in the mouse pup is an appropriate biomarker/metric as it relates to clinically relevant hypothyroxinemia both during pregnancy in humans and in the human neonate. Thyroid hormone is clearly essential for brain development and growth in both rodents (mouse) and humans. It is also clear that thyroid hormone in both rodents and humans exert different actions on the brain as development proceeds. Although mice are born at a time that is equivalent roughly to the human third trimester, thyroid hormone insufficiency is relevant throughout human pregnancy and the first period of postnatal human development. In addition, serum total T4 is a good reflection of thyroid homeostasis in both human pregnancy and in the mouse. (Note that serum total T4 increases by about 50% during human pregnancy, but serum free T4 does not change. Based on this, the American Thyroid Association recommends that women on T4 supplementation before pregnancy increase their dose by 50% once they become pregnant.

In other words, serum total T4 is the basis for clinical recommendations.) This measurement is applicable, therefore, to both the experimental animal paradigm as well as humans.

1c. Has EPA clearly articulated the challenges associated with extrapolating the PFBS-induced decrease in thyroid hormone (e.g., total T4) in rodents to humans?

There are two components to this extrapolation. First is the relative efficacy of PFBS in humans and animals with respect to T4 suppression. The second is the efficacy of T4 suppression to adverse outcome. The Agency has made clear that this extrapolation was essentially described by the following equation:

$$HED = \frac{POD}{CL_A/CL_H} = POD \times \frac{CL_H}{CL_A}$$

Where CL is the clearance rate in animals (A) and humans (H). The problem is that there are no estimates of clearance in humans. However, because clearance rates are similar between rodents and monkey, and half life is inversely related to clearance, the Agency made rational estimates to extrapolate PFBS-induced decrease in thyroid hormone in rodents to humans. As the agency presents in Table 9, similar patterns of decreases in serum total T3, total T4, and free T4 were observed in PFBS-exposed pregnant mice, nonpregnant adult female rats, adult male rats, and gestationally exposed female mouse offspring. The magnitude of decrease was deemed concerning (~20% in dams and ~50% in offspring), and more importantly, they were shown to persist at least 60 days after gestational exposure in offspring, and they exhibited a clear dose dependence.

1d. Has EPA clearly articulated what is and is not known about the clinical implications of changes in T4 in women during pregnancy on neonates and infants?

The Agency has developed a strong argument for the importance of thyroid hormone in child health. They identified the critical effect from the Feng et al. (2017) study as decreased serum total thyroxine (T4) in newborn (PND 1) mice. Further, they state that T4 and T3 are essential for normal growth of developing offspring across animal species, and that previous studies show that exposure to other PFAS during pregnancy results in lower T4 and T3 levels in pregnant women and fetuses or neonates. The selection of total T4 as the critical effect is based on a number of key considerations that account for cross-species correlations in thyroid physiology and hormone dynamics particularly within the context of development.

The Agency argues that a key issue for the focus on total T4 is that it "*represents the aggregate of potential thyroid endocrine signaling (i.e., free T4 + protein bound T4) at any given time*." It is true that although T3 is the "hormonally active" form at the receptor, it is T4 that gains access to tissues (e.g., brain and fetal compartment) and that *de novo* conversion to T3 is part of the signaling pathway. It is somewhat confusing that the Agency focused on the type 3 deiodinase in placenta. The Agency states that, "*The placenta has transporters and deiodinases that collectively act as a gatekeeper to maintain an optimal T4 microenvironment in the fetal compartment*." which is true enough. However, their example deiodinase 3 (D3), which is "*highly expressed in human uterus, placenta, and amniotic membrane, where it serves a critical role of regulating thyroid hormone transfer to the fetus through the deiodination of T4 to transcriptionally inactive reverse triiodothyronine (rT3) or T3 to inactive 3,5-diiodo-L-thyronine (T2)*". This is also true, but it is unclear what relevance this has to the issue at hand. Moreover, the Agency states

that, "*Further, the Dio3 gene that encodes D3 has been shown to be imprinted in the mouse (Hernandez et al., 2002), suggesting a pivotal role for this specific deiodinase in the mouse as well.*" However, Hernandez et al. showed that the paternal *Dio3* gene is preferentially expressed in the offspring. It is not clear how this indicates a pivotal role for D3, nor that T4 degradation should be the focus. But it is true that the human and rodent placenta have been shown to be similarly permeable to T4 and T3 (Fisher, 1997; Calvo et al., 1992). Finally, the Agency concludes that "*Due to placental barrier functionality, free T4 levels in a pregnant dam might not be entirely representative of actual T4 status in a developing fetus. Thus, decreased total T4 in offspring is expected to be more representative of PFBS-mediated thyroid effects and potentially associative developmental effects*." Although I agree with this conclusion, I don't really follow the argument, which seems discursive. Rather, I would focus on the fact that serum total T4 increases by about 50% in pregnant women without a concomitant increase in serum free T4, and that hypothyroid women should increase their daily dose of T4 to reflect this if they become pregnant. Thus, total T4 is an important index at this life stage.

This argument was a prelude to the development of the idea that the clinical manifestation of low T4 in pregnancy (but also in the neonate) results in neurocognitive deficits in humans and animals. This is a complex field because while it is clear that the human brain is sensitive to thyroid hormone insufficiency, the disconnect between the timing of T4 measurement in the pregnant woman, and the cognitive domains tested in the offspring do not always match. This weakness is revealed by the work and writing of Professor Joanne Rovet (Rovet 2014) who clearly described the temporal relationship between thyroid hormone insufficiency and the cognitive domain affected. However, the Agency did a thorough job of articulating the clinical relevance of total T4 insufficiency and its relation to adverse cognitive outcome in humans.

2. In the public review draft PFBS assessment, EPA employed benchmark dose modeling (U.S. EPA, 2012) in the identification of a point-of-departure (POD) for derivation of RfD values, based on a decrease in total T4 levels in PND1 offspring. The 20% Relative Deviation (RD) Benchmark Response Rate (BMR) used in the public review draft is no longer being considered for BMD modeling of thyroid hormone dose-response data. As a result of extensive public review comments on the BMD approach (and thyroid hormone endpoint) used in the previous draft, and because a clear or consistent biological threshold for T4 changes associated with untoward developmental health outcomes has not be identified in the available literature, EPA has identified a new BMR of 0.5 SD (standard deviation change over controls) as a default in the revised PFBS draft assessment for the thyroid hormone alterations in mouse neonates/offspring. A 1 SD BMR is also being presented as the standardized basis for comparison as recommended in the EPA BMD Technical Guidance (U.S. EPA, 2012).

2a. Are the dose-response modeling approaches, selection of benchmark response rate (BMR), and the selected models used to identify the thyroid effect-related POD for RfD derivation scientifically justified and clearly described?

First, I am not expert on the issue of benchmark dose modeling. However, the Agency's argument for use of the 0.5 SD over controls was reasonable to me. In particular, it is true that there is no identified "threshold" of total T4 insufficiency that is clearly causative in the production of cognitive – and other developmental – deficits. Thus, the Agency needed to formalize an approach that would be science-based and reflect a rational approach to identifying the POD. One argument that could have further strengthened this approach is to address the issue of "compensation". This concept is that as serum total T4 declines,

endogenous "adaptive" responses are triggered – both in tissues and in the blood – to ameliorate the adverse consequences of low T4. One study examined this issue specifically (Sharlin et al. 2010) finding that if these adaptive responses are compensatory, they occur at a level of T4 insufficiency that is not measurable. Thus, using a 0.5SD cut-off appears reasonable and not overly protective.

3. Due to the availability of new toxicokinetic data in mice noting significant interspecies differences in toxicokinetics of PFBS, and, recommendations from public commenters, EPA has applied a data-informed approach to convert the oral dose-rate in animals to a human equivalent dose (HED) in the identification of candidate points-of-departure (PODs) considered for the derivation of the RfDs (U.S. EPA, 2014; see https://www.epa.gov/sites/production/files/2015-01/documents/ddef-final.pdf**). In considering the new evidence for serum half-life in mice published in Lau (in press), EPA concluded that the toxicokinetic data for PFBS are adequate to support calculation of data-derived dosimetric adjustment factors (DAF), where the ratio of elimination half-life in animals to that in humans, $T0.5_A/T0.5_H$, is used to adjust candidate PODs. By using in vivo animal and human half-life data to calculate POD(HEDs) that account for differences in toxicokinetics between rodents and humans, the potential role of interspecies differences in processes such as renal resorption, hepatic transport, and enterohepatic recirculation is reflected. Further, by using a data-derived approach the uncertainty in interspecies toxicokinetic scaling (UF_A) has been reduced from a 3 to a 1; however, residual uncertainty (due to the lack of information) pertaining to toxicodynamics exists and is acknowledged in the assessment in the description for applying a UF_A of 3.**

3a. Is applying the data-informed dosimetric adjustment that utilizes the ratio of the PFBS elimination half-life in mice to that in the human scientifically justified and clearly described?

The rationale provided by the Agency for dosimetric adjustment was scientifically rational and clearly described. It is also somewhat reasonable to make the assumption that these measures are an overall reflection of processes that eliminate PFBS from the system. Other elements of this response are described in question 1.

4. 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 PFBS.

4a. Does the provided qualitative scientific rationale support the application of the selected uncertainty factors? If not, please explain.

In general, I think that the qualitative rationale in support of uncertainty factors was well-reasoned. However, I don't believe that a UF_A of 3 is fully justified. There are several reasons for this. First, the study of Feng et al. (2017) showed that serum total T4 was diminished by PFBS at PND1 but also at PND 30 and PND 60 – 30 and 60 days after cessation of exposure. If the toxicokinetic data of Lau et al. (unpublished) is correct, PFBS was fully eliminated from the animals by the 30 and 60-day timepoints. These data indicate that human neonates may experience T4 suppression for much longer than the fetal/neonatal period. Moreover, the human neonate is quite sensitive to thyroid hormone insufficiency for a minimum of 2 years. Finally, it is likely that PFBS will contaminate breast milk since other PFAS are found in breast milk (e.g., (Beser et al. 2019). Therefore, it is scientifically justified to expect that PFBS will suppress serum T4 both early in development as well as perhaps many months after birth. The uncertainty of these likely impacts would justify a UF_A of 10.

4b. Has quantitative uncertainty been adequately accounted for in the derivation of the RfDs?

As described in 4a, I do not believe the quantitative uncertainty has been fully accounted for in the derivation of the RfDs. Specifically, the UF_A should be 10, not 3.

4c. Do the methods used to derive the RfDs for PFBS appropriately account for uncertainties in evaluating the toxicokinetic differences between the experimental animal data and humans?

In general I agree that they do. However, the literature does not fully inform the toxicokinetic (and dynamic) differences between experimental animals and humans. Importantly, the mechanism by which PFBS causes serum total T4 reduction is not clear. As a result, it is not clear how potent PFBS would be on those molecular targets responsible for this decline. Thus, while the toxicokinetic difference is estimated in a rational way, the toxicodynamic difference cannot actually be estimated. Thus, a UF_D of 10 would appear warranted.

References

Beser MI, Pardo O, Beltran J, Yusa V. 2019. Determination of 21 perfluoroalkyl substances and organophosphorus compounds in breast milk by liquid chromatography coupled to orbitrap high-resolution mass spectrometry. Analytica chimica acta 1049: 123-132.

Rovet JF. 2014. The role of thyroid hormones for brain development and cognitive function. Endocrine development 26: 26-43.

Sharlin DS, Gilbert ME, Taylor MA, Ferguson DC, Zoeller RT. 2010. The nature of the compensatory response to low thyroid hormone in the developing brain. J Neuroendocrinol 22(3): 153-165.