EPA/600/R-18/305

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)

November 2018

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

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.

AEC	absolute eosinophil counts
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
CHO	Chinese hamster ovary
DAF	dosimetric adjustment factor
DNA	deoxyribonucleic acid
ECP	eosinophilic cationic protein
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
IUR	inhalation unit risk
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
OSF	oral slope factor
PBPK	physiologically based pharmacokinetic
PECO	populations, comparators, exposures, and outcomes
PFAS	per- and polyfluoroalkyl substances
PFBS	perfluorobutane sulfonic acid
PFHxA	perfluorohexanoic acid
PFOA	perfluorooctanoic acid
PFOS	perfluorooctane sulfonic acid

PND	postnatal day
POD	point of departure
POD_{HED}	point of departure human equivalent dose
PPARα	peroxisome proliferator-activated receptor alpha
PPRTV	provisional peer-reviewed toxicity value
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
TG	test guideline
TSCA	Toxic Substances Control Act
TSCATS1	Toxic Substances Control Act Test Submissions 1
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. Five independent external peer reviewers reviewed the draft assessment, and their comments are presented with the EPA's responses. Appendix A includes the full comments from 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 public comment. The EPA is publishing these toxicity values to facilitate further 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 [T4], total T4, and thyroid stimulating hormone), 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).

In consideration of peer reviewers' comments and further evaluation of the database, the PFBS toxicity assessment was revised and is being released for public review and comment. Overall, the kidney was consistently identified as a target of PFBS toxicity in adult rats (Lieder et al., 2009a; Lieder 2009b), and effects on the kidney were considered as a candidate critical effect for RfD derivation. Across all life stages evaluated, the thyroid was identified as the most sensitive target of PFBS toxicity. Additionally, developmental effects were often observed in animals in which thyroid effects occurred (Feng et al., 2017). However, the developmental effects appeared to be less sensitive; thus, effects on the thyroid were identified as a candidate critical effect for RfD derivation. Candidate subchronic and chronic RfDs were derived for both kidney and thyroid effects.

The candidate subchronic RfD for K^+PFBS associated with thyroid effects 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 9), yielding a candidate subchronic RfD of 4×10^{-2} 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 same value of 4×10^{-2} mg/kg-day for the candidate subchronic RfD (thyroid effects) for PFBS (free acid).

The candidate subchronic RfD for K^+PFBS associated with kidney effects was calculated by dividing the POD_{HED} for increased papillary epithelial tubular/ductal hyperplasia in P_0 female rats by a composite uncertainty factor (UF_C) of 100 to account for extrapolation from rats 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 candidate subchronic RfD of 1×10^{-1} 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 same value of 1×10^{-1} mg/kg-day for the candidate subchronic RfD (kidney effect) for PFBS (free acid).

The candidate chronic RfD for K^+PFBS associated with thyroid effects 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 14), yielding a chronic RfD of 1×10^{-2} mg/kg-day. Like the candidate subchronic RfD for thyroid, based on the data for K^+PFBS, a candidate chronic RfD for PFBS (free acid) of 1×10^{-2} mg/kg-day was derived.

The candidate chronic RfD for K^+PFBS associated with kidney effects was calculated by dividing the POD_{HED} for increased papillary epithelial tubular/ductal hyperplasia in P_0 female rats by a UF_C of 1,000 to account for extrapolation from rats to humans (UF_A of 3), for interindividual differences in human susceptibility (UF_H of 10), to account for less than chronic-duration exposure (UF_S of 10) and deficiencies in the toxicity database (UF_D of 3) (a value of 1 was applied for UF_L) (see Table 15), yielding a candidate chronic RfD of 1×10^{-2} mg/kg-day. Like the candidate subchronic RfD for kidney, based on the data for K^+PFBS, a candidate chronic RfD for PFBS (free acid) of 1×10^{-2} mg/kg-day was derived.

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

- choices of critical studies;
- choices of critical effects;
- benchmark dose modeling;
- determination of an HED dose using BW scaling;
- UF application; and
- cancer classification.

The peer reviewers made several clarifying comments. Specifically, the more substantive comments include considerations for:

- further attention to the overall database for thyroid effects including the inclusion of the 28-day adult rat thyroid data (NTP, 2018) for consideration as principal study/critical effect;
- the identification of free T4 rather than total T4 as the potential critical effect;
- use of 80 kilograms for human BW for allometric scaling; and
- greater support for a link between decreased thyroid hormone during developmental life stages and associated neurotoxicity and differential thyroid hormone reserve capacities between infants and adults.

Minor comments and editorial suggestions were reviewed and are addressed directly in this document. Specific responses to major comments are provided under each respective section/question.

SECTION I: TECHNICAL CHARGE TO EXTERNAL REVIEWERS

Technical Charge to External Peer Reviewers
Contract No. EP-C-17-017
Task Order 0008
June 2018

External Peer Review of EPA's Draft Human Health Toxicity Values for Perfluorobutane Sulfonic Acid and Related Compound Potassium Perfluorobutane Sulfonate

BACKGROUND

The U.S. Environmental Protection Agency (EPA) is issuing developmental, subchronic, and chronic toxicity values (i.e., reference doses, or RfDs) for perfluorobutane sulfonic acid (PFBS) CASRN 375-73-5) and its potassium salt (CASRN 29420-49-3).

The EPA developed a toxicity assessment of PFBS to provide the health effects information used as the basis for derivation of RfDs. The toxicity assessment identifies and characterizes the health hazards of PFBS and includes information to address the first two steps of the human health risk assessment paradigm: hazard identification and dose-response assessment. The EPA is publishing these toxicity values to facilitate decision-making by our Programmatic, Regional, and/or State partners, associated with contamination concerns in a variety of exposure scenarios.

This document was developed, in part, by building on a previously completed assessment. Specifically, in 2014, the EPA released a provisional peer-reviewed toxicity value (PPRTV) assessment for PFBS and its potassium salt. PPRTV assessments derive human health toxicity values specifically for use in the EPA's Superfund program. In this assessment effort, the EPA applied systematic review procedures to updating the literature, evaluating the latest science, and updating the toxicity values, as appropriate.

The PFBS literature identified by the search of publicly available sources and submitted studies are available through the EPA's Health & Environmental Research Online (HERO) website[1]: https://heronet.epa.gov/heronet/index.cfm/project/page/project_id/2610.

REVIEW MATERIALS PROVIDED BY THE EPA

- Draft *Human Health Toxicity Values for Perfluorobutane Sulfonic Acid (CASRN 375-73-5) and Related Compound Potassium Perfluorobutane Sulfonate (CASRN 29420-49-3)* (USEPA, 2018)
- References and supporting documentation (see below)
    - *Benchmark Dose Technical Guidance* (USEPA, 2012)

[1] Access to full-text references was provided to external peer-reviewers through the HERONet website, an internal database of bibliographic information and scientific studies. Due to copyright laws/regulations, access to the copyright protected materials that are stored in the HERONet database are prohibited from public dissemination. However, a list of all citations and access to publicly available references is provided via the public HERO website (https://hero.epa.gov/hero/index.cfm/project/page/project_id/2610).

- Background references are available on the EPA's HERONet[1]. The URL above will take you directly to the references needed.

CHARGE QUESTIONS

Perfluorobutane sulfonic acid (PFBS) and its potassium salt (K+PFBS)

1. The document describes and applies a systematic review process for identifying and selecting pertinent studies. Please comment on whether the literature search strategy, study selection considerations, and study evaluation are appropriate and clearly described. Please identify additional peer-reviewed studies that the assessment should consider.

2. For PFBS the key study chosen for determining the subchronic and chronic RfDs is the Lieder et al. (2009) 90-day rat study and the critical effect is increased incidence of kidney hyperplasia in female rats. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and defensible?

    a. If so, please explain your justification.
    b. If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the subchronic and chronic RfDs.
    c. In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described.

3. In addition, for PFBS, a RfD associated specifically with a developmental lifestage is derived. The gestational exposure mouse study by Feng et al. (2017) is chosen as the key study and the critical effect is decreased total thyroxine (T4) in offspring. Is the selection of the key study and critical effect for the derivation of this developmental RfD for PFBS scientifically justified and defensible?

    a. If so, please explain your justification.
    b. If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the developmental RfD.
    c. In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described and whether the critical effect is associated with clinically relevant health outcome(s).

4. EPA employed benchmark dose modeling (USEPA, 2012) in the identification of points-of-departure (PODs) for PFBS, one based on kidney hyperplasia and the other based on a decrease in total T4 levels in offspring.

[1] Access to full-text references was provided to external peer-reviewers through the HERONet website, an internal database of bibliographic information and scientific studies. Due to copyright laws/regulations, access to the copyright protected materials that are stored in the HERONet database are prohibited from public dissemination. However, a list of all citations and access to publicly available references is provided via the public HERO website (https://hero.epa.gov/hero/index.cfm/project/page/project_id/2610).

   a. Are the modeling approaches used, selection of benchmark response levels, and the selected models used to identify each POD for RfD derivation scientifically justified and defensible?
   b. Specifically, considering species and/or lifestage specific differences in thyroid economy (e.g., differential reserve capacities for thyroid hormone in infants compared to adults and mice compared to humans), comment on how EPA addresses these factors in the choice of a biologically based benchmark response level (i.e., level of change that characterizes the lower limit of biological significance compared with normal background responses).

5. Given what is known and not known about the interspecies differences in toxicokinetics of PFBS, EPA applied body weight to the ¾ allometric scaling to adjust the POD to estimate a human equivalent dose (HED) in the derivation of the respective RfDs (USEPA, 2011).

   a. Is applying the body weight to the ¾ for PFBS scientifically justified and defensible? If not, please provide your rationale and detail the alternative approach you would use.
   b. 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?

6. 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.

   a. Has uncertainty been adequately accounted for in the derivation of the RfDs? Please describe and provide suggestions, if needed.
   b. Does the provided scientific rationale support the application of the selected uncertainty factors? Please explain.

7. The draft assessment for PFBS identifies several potential human hazards. Please comment on whether the available data have been clearly and appropriately synthesized for each toxicological hazard. Please comment on whether the weight of evidence for hazard identification has been clearly described and scientifically justified.

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

9. **Editorial or Additional Comments:** Please provide any editorial or additional comments you would like to make here. These should be any comments that are not in direct response to the technical charge questions above.

REFERENCES

USEPA (U.S. Environmental Protection Agency). 2011. *Recommended Use of Body Weight $^{3/4}$ as the Default Method in Derivation of the Oral Reference Dose*. EPA/100/R11/0001. USEPA, Office of the Science Advisor, Risk Assessment Forum, Washington, DC. Accessed May 2018. https://www.epa.gov/sites/production/files/2013-09/documents/recommended-use-of-bw34.pdf.

USEPA (U.S. Environmental Protection Agency). 2012. *Benchmark Dose Technical Guidance*. EPA/100/R-12/001. USEPA, Risk Assessment Forum, Washington, DC. Accessed May 2018. https://www.epa.gov/risk/benchmark-dose-technical-guidance

SECTION II: REVIEWER COMMENTS ORGANIZED BY CHARGE QUESTION

CHARGE QUESTION 1

The document describes and applies a systematic review process for identifying and selecting pertinent studies. Please comment on whether the literature search strategy, study selection considerations, and study evaluation are appropriate and clearly described. Please identify additional peer-reviewed studies that the assessment should consider.

Chou

The literature search strategy and study evaluation considerations are clearly described and appear to be appropriate. The reviewer does not have any suggestions for additional studies.

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

Kamendulis

Yes, the literature search strategy, study selection and evaluation considerations were very well presented and sufficiently clear. The process used was described well and was a very thorough and transparent approach to systematically evaluate each of the available scientific studies that described the health effects of PFBS/K+PFBS.

I am unaware of other peer-reviewed studies that should be included in this assessment.

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

Leung

The draft report describes the systematic approach taken by the EPA toward the identification and selection of pertinent studies on this topic. The search strategy is overall easy to understand and transparent. The amount of detail provided in describing the screening process of potentially useful studies, the number of reviewers who completed each step, as well as the availability of each study's associated details in HAWC, are particularly appreciated. Overall, the strategy is appropriately comprehensive, and there do not appear to be any other peer-reviewed studies which need to be considered.

However, two epidemiologic studies (Bao 2017 and Kim 2016) were excluded due to their large number of samples (96% and 72%, respectively) below their limits of detection (Table 4 in the draft report). Although this appears reasonable for these particular studies, it would be beneficial to describe what a reasonable threshold for inclusion might be for potential future assessments.

> **EPA Response:** Because there are many assessment- and study-specific factors in addition to the detection frequency that influence the exposure measurement and study sensitivity (e.g., exposure levels and contrast, sample size, and statistical analysis approach), a "default" threshold for inclusion is not used. We required a sufficient number of the exposure data measurements to be above the limit of quantification for the assay, based on expert judgment and considering the context of the study. This is consistent with approaches used by others (NTP PFOS/PFOA assessment; NTP Handbook for Preparing Report on Carcinogens Monographs; and LaKind et al., 2014). In general, when evaluating a study with less than

20%–30% of samples above the limit of detection, there would be serious concerns about the study's ability to address the research question. This information would be used in conjunction with the aforementioned considerations when conducting study quality evaluations.

Slitt

The literature search strategy was appropriate and thorough, with the overall method being very thorough and objective. The method was described and included clear criteria for the inclusion and exclusion of studies. The databases utilized (i.e., PubMed, Web of Science, Toxline, and TSCATS via Toxline) are appropriate and the search terms were comprehensive in nature. The methods in appendices A-C used to evaluate study quality were systematic and thorough. The metrics and criteria applied for Animal and *in vitro* toxicity studies were exceedingly thorough and well defined. The weighting and relative important used for weighting the criteria was appropriate. Specifically, the approach for evaluating epidemiological and animal toxicology studies was well described and rationalized in several well-organized diagrams (Figures 3-5 and Tables 3 & 4). The overall framework for judging the health effect was systematic, objective, and unbiased.

EPA Response: No revisions needed to address this comment.

Warren

Pages 17-31 and Appendices A-D clearly describe an appropriate, but laborious process by which literature, peer-reviewed and otherwise, is identified and evaluated for the purpose of toxicity assessment. Use of the word "process" is clearly appropriate, as the toxicity assessment describes a series of progressive and interdependent steps as a means to an end. The process requires considerable skill if done well, and I refrain from using the phrase "done correctly" in recognition of the subjective judgments that remain inherent to it. The PECO criteria and additional exclusion criteria are straightforward, and were obviously effective at severely reducing all studies to a manageable number of relevant ones. The listed domains for epidemiological and animal studies should make for a comprehensive evaluation, and the domain-specific questions in Appendix C can be of assistance to even the most experienced reviewer. While Figures 5 and 6 are informative, the toxicity assessment fails to disclose the rules by which domain ratings are combined to reach an overall study classification. Such information would be a welcomed addition to the toxicity assessment document. And while HAWC undoubtedly benefits the toxicity assessment process, it is important that the written work product be sufficiently informative to the end user that web access is purely optional. Such is the case for PFBS and as such, I support the study identification and evaluation process used, while encouraging its refinement over time. As for additional peer-reviewed studies, three were located though no attempt was made to evaluate them against PECO criteria.

1. Xu et al., 2017. Effects of perfluoroalkyl substances on neurosteroid synthetic enzymes in the rat. *Chemico-Biological Interactions*, Vol. 272, pp. 182-187.
2. Chen et al., 2018. Multigenerational Disruption of the Thyroid Endocrine System in Marine Medaka after a Life-Cycle Exposure to Perfluorobutane sulfonate. *Environ. Sci. Technol*., 52(7), pp. 4432-4439.

3. Gyllenhammar et al., 2018. Perfluoroalkyl acids (PFAAs) in serum from 2-4 month-old infants: Influence of maternal serum concentration, gestational age, breast-feeding, and contaminated drinking water. *Environ. Sci. Technol.*52 (12), pp. 7101-7110.

   EPA Response: The EPA appreciates the observations on the use of HAWC as we transition to increased use of web-based information management resources to complement our assessment documents. In the PFBS assessment, as the reviewer notes, we have taken the approach of making sure all the key content from HAWC is included in the assessment as figures or appendices that do not require web access. The reviewer's general comments on encouraging refinement and monitoring user accessibility are appreciated and will be considered as we move forward with greater usage of HAWC. We have added information to the appendix on how domain ratings are combined to reach overall study confidence.

   In addition, the recommended studies were evaluated against PECO criteria and not determined to be relevant to the PECO (e.g., *in vitro* cell-based results only, ecological [fish] study, and exposure only/no reported health effects). These references were added to the HERO database as additional screened references obtained from "Other Sources." Figure 4 was updated in the assessment to illustrate the addition.

CHARGE QUESTION 2

For PFBS the key study chosen for determining the subchronic and chronic RfDs is the Lieder et al. (2009) 90-day rat study and the critical effect is increased incidence of kidney hyperplasia in female rats. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and defensible?

a. **If so, please explain your justification.**
b. **If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the subchronic and chronic RfDs.**
c. **In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described.**

Chou

a. **If so, please explain your justification.**

N/A

EPA Response: No revisions needed to address this comment.

b. **If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the subchronic and chronic RfDs.**

The selection of the key study by Lieder et al. (2009) and critical effect for the derivation of the subchronic and chronic RfDs for PFBS, as it is justified on p. 81 and p. 84, is highly questionable.

The draft provides two reasons for not using the Total T4 or Free T4 data from NTP (2018) study as critical effects. Here is the justification from the draft, from P. 80 Line 14 to P. 81, Line 2: (1)

"because exposure durations were shorter than subchronic-duration" and (2) "the 28-day exposure study in rats was not from a peer-reviewed publication." The reviewer disagrees to these two reasons.

When NTP (2018) was evaluated specifically for its quality and feasibility, and specifically for this assessment, it had determined that this is a High Confidence study. See the following two excerpts from the Draft that support the reviewer's argument that NTP (2018) should be the principal study for the subchronic oral RfD assessment.

> P. 17: "Although a peer-reviewed NTP Technical Report for the PFBS study is not yet available, this information was included in the assessment because these data had undergone normal NTP quality assurance/control processing and are publicly available. … During the process of deriving toxicity values, EPA conducted further quantitative analyses (e.g., BMD modeling) beyond what was reported by the NTP."
>
> P. 33: The NTP (2018) study is identified as a "high confidence study" on p. 33, under Evidence Synthesis.

Furthermore, effects on thyroid hormonal balance is the most consistent findings supported by several other studies.

In summary, the results of NTP (2018) has been peer reviewed and the data are available. The study results are so new that they are yet to be published. In addition, the most relevant target, thyroid imbalance, is not even examined by Lieder et al. (2009a). When the effects on thyroid imbalance is already evident after 28-day exposure, the results should not be criticized and disregard simply because the exposure period is too short.

> **EPA Response:** The comment is appreciated. The reviewer correctly pointed out that the EPA considers the NTP (2018) study/data to be of high confidence. The rationale for why the NTP thyroid data were not considered for the subchronic RfD (e.g., 28-day exposure duration and peer-review status) has been edited in the revised assessment. The NTP (2018) thyroid study/data are now integrated into the considerations for principal study and critical effect in the narrative associated with derivation of a subchronic and chronic RfD.

c. In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described.

The conclusions regarding adversity are scientifically supported and clearly described.

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

Kamendulis

I agree with the selection of the critical study selected for deriving RfD's for K+PFBS (Lieder et al., 2009). The Lieder et al. (2009) study was a 90-day oral gavage study that reported mild to moderate hyperplasia in the kidneys of both male and female Sprague-Dawley rats administered K^+PFBS. The subchronic and chronic RfDs for K^+PFBS were derived from papillary epithelial tubular/ductal

hyperplasia in female rats. The selection of this study and endpoint was scientifically justified and clearly described.

EPA Response: No revisions needed to address this comment.

Leung

The section summarizing the included studies and Table 7 are well-organized and thoughtfully prepared. Separation of the human vs animal studies in each organ system, as well as presentation of their findings via interactive graphics, greatly enhances the readability of the document. From the available data, the two organ systems demonstrating adverse effects from PFBS exposure with the highest level of confidence are the kidney and the thyroid gland. Reproductive, liver, lipid, and immune effects were all considered to be equivocal, which appears appropriate, given the paucity of data in these areas.

Thyroid: There are no human studies regarding thyroidal effects from PFBS exposure, but the available animal studies support a consistent trend of associated hypothyroidism, which notably includes an up to 97% reduction in serum thyroid hormone concentrations among groups exposed to very high PFBS doses.

Kidney: Animal data also show that PFBS exposure is associated with increased renal weight and abnormal histopathologic findings (mostly renal hyperplasia). The one available human renal study suggesting that uric acid levels may be elevated in exposed boys is not particularly robust.

Overall, the selection of Lieder et al 2009 as the key study for determining subchronic and chronic RfDs (corresponding to the critical effect of abnormal renal histopathology in primarily female rats) is reasonably supported by these data.

However, the adverse thyroidal effects from exposure are substantial, and hypothyroidism can arguably be considered also as another critical effect. Specifically, the NTP 2018 study was notable for the following points:

1. In the high-dose PFBS group (500mg/kg/day) after 28 days, striking reductions of approximately 92% in serum total thyroxine (TT4), 85% in free thyroxine (FT4), and 56% in total triiodothyronine (TT3) concentrations among male rats, and 69% in serum TT4, 65% in FT4, and 43% in TT3 among female rats were observed. The one female rat which was exposed to 1000 mcg/kg/day was found to have even more profound hypothyroidism, in accordance with a monotonic dose-response effect. The associated standard error of measurement (SEM) values in all groups were reasonably small to support these overall trends.

   [I was not able to access the NTP 2011 study to examine it greater detail, and there is no associated link in the references, but it is also described in the draft report to have produced similarly dramatic reductions in serum thyroid hormone levels.]

2. In this timeframe of exposure (28 days), the lack of significant abnormalities related to thyroid weight and histopathology is not unexpected. However, one would have expected a greater

elevation of serum TSH than what was observed for this degree of hypothyroidism within this timeframe.

3. In mild hypothyroidism, individuals with an intact hypothalamic-pituitary-thyroid axis can compensate by the action of increased TSH production to stimulate thyroid hormone production at the thyroid gland. However, for the greatest severity of hypothyroidism seen in the NTP 2011 and 2018 studies, normal physiologic adaptive processes would be insufficient to restore normal thyroid function. Thus, given the importance of normal thyroid hormone levels in multiple systems, the significant reduction of thyroid hormone availability can justifiably be considered a critical effect of PFBS exposure.

    EPA Response: The reviewer summarizes a significant case for including thyroid effects in the deliberations for identifying a critical effect for derivation purposes. Indeed, in the revised draft assessment, both the NTP (2018) and the Feng et al. (2017) studies are integrated into the narrative leading to identification of principal study and critical effect(s) because of their strong support for thyroid as a sensitive health outcome of concern associated with oral PFBS exposure. In regard to the reviewer not being able to access the NTP (2011) citation, the EPA includes a reference to the 2011 document because it contains the study protocol associated with the NTP (2018) study results. They are not independent studies or experiments.

Slitt

a. If so, please explain your justification.

The Leider et al., 2009a study was considered to meet the criteria to be included. This was a 90-day gavage study for adult male and female rats. The study methods are well described, and the study had a reasonable "n" of 10 animals per treatment group to statistically detect effects. The Leider et al., 2009a study for kidney effects is also supported by NTP, 2018, 2011; Leider et al., 2009; 3M, 2001; 200d. These studies describe alterations in renal weight and some evidence of histopathological changes (i.e., inflammation) that were dose dependent. Thus, the renal effects being considered to be the critical effect is supported.

EPA Response: No revisions needed to address this comment.

b. If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the subchronic and chronic RfDs.

I do not have a suggestion for alternative studies or effects to be selected to support the RfD. The Leider et al., 2009a study was thoroughly designed and met the evaluation criteria to be included. Along with renal, that publication considered multiple other endpoints, such as other organs, clinical chemistry, hematology, and histopathology for four other tissues. Although liver effects were also noted among several studies (incuding Leider et al., 2009a), none noted evidence or cytotoxicity or single-cell necrosis.

EPA Response: No revisions needed to address this comment.

c. In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described.

Yes, the conclusions regarding adversity about supported and described. Table 7 on pages 62 and 63 presents the findings from six high confidence studies that there are various kidney effects observed in rats, with histopathology changes in kidney for female rats being the primary effect.

EPA Response: No revisions needed to address this comment.

Warren

Considering the database limitations for PFBS, I support selection of Lieder et al. (2009a) as the critical study and kidney hyperplasia in females as the critical effect. However, while these selections can be scientifically justified and defended, an argument can also be made for alternatives (i.e., reduced T4 in mice and rats from Feng et al. (2017) and NTP (2018), respectively) that when subjected to BMD modeling, result in lower BMDLs. While Lieder et al.'s evaluation resulted in its characterization as "good" (Figure 6) or "high confidence" (Table 6), this alone does not distinguish it from several other studies that underwent consideration. The decision ultimately came down to what was considered of most importance - basing RfD derivation on the most sensitive endpoint in adults (thyroid hormone perturbation) or reliance on a peer-reviewed study of subchronic duration. Clearly, the Agency felt as though peer-review and adequate exposure duration trumped sensitivity, despite NTP's reputation for generating high quality data and reliance on Feng et al. (2017) for developmental RfD derivation. As I am unable to make a reasoned choice between the aforementioned alternatives, I support the Agency's decision. It seems clear that given database limitations, the choice of target tissue was limited to the thyroid and kidney. The toxicity assessment does an adequate job of discussing the scientific support for both as hazards.

EPA Response: Dr. Warren's comment is consistent with concerns raised by other reviewers (e.g., Drs. Chou and Leung) regarding the exclusion of thyroid effects data from NTP (2018) and/or Feng et al. (2017) for consideration in the derivation of subchronic and chronic RfDs. As noted by the reviewer, the EPA's confidence in the NTP (2018), Feng et al. (2017), and Leider et al. (2009a) studies are approximately equivalent (e.g., high). As such, the reviewer's supposition that study confidence is not an appropriate delineation in ultimate selection of principal study is well received. The reviewer also pointed out that thyroid effects are more sensitive than kidney effects, which is also correct. As such, thyroid studies/data (i.e., NTP [2018] and Feng et al. [2017]) were reconsidered and integrated into the identification of principal study and critical effect in the subchronic/chronic RfD derivation section of the revised draft assessment.

CHARGE QUESTION 3

In addition, for PFBS, a RfD associated specifically with a developmental lifestage is derived. The gestational exposure mouse study by Feng et al. (2017) is chosen as the key study and the critical effect is decreased total thyroxine (T4) in offspring. Is the selection of the key study and critical effect for the derivation of this developmental RfD for PFBS scientifically justified and defensible?

a. **If so, please explain your justification.**
b. **If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the developmental RfD.**
c. **In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described and whether the critical effect is associated with clinically relevant health outcome(s).**

Chou

a. **If so, please explain your justification.**

The selection of the key study by Feng et al. (2017) and the critical effect for the derivation of the subchronic and chronic RfDs for PFBS is well justified and clearly stated on p. 69-71 and p. 74-75. The approaches used for the selection of PODs and the justification of rodent models are well stated. The reviewer has no additional justifications.

EPA Response: No revisions needed to address this comment.

b. **If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the developmental RfD.**

N/A

EPA Response: No revisions needed to address this comment.

c. **In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described and whether the critical effect is associated with clinically relevant health outcome(s).**

The conclusions regarding adversity of PFBS on thyroid hormone T3 and T4 are scientifically supported and clearly described. These effects are supported by clinically relevant health outcome(s) observed in epidemiological studies.

EPA Response: No revisions needed to address this comment.

Kamendulis

The concern of relevance to developmental toxicity was identified as a prolonged decrease in serum thyroxine (PND 1-60) in mice that were exposed to PFBS in utero. This gestational exposure study in mice (Feng et al., 2017) was selected as the principal study for derivation of the developmental reference dose using decreased serum total thyroxine (T4) in newborn (PND1) mice as the critical effect. As adequate levels of thyroid hormones are required for the development of organ systems for normal growth and development in early lifestages, I agree with the selection of the Feng et al., study and decreased T4 as the critical effect. The conclusions were scientifically justified and clearly presented in this document.

EPA Response: No revisions needed to address this comment.

Leung

Although the statement "The selection of total T4 as the critical effect is based on the consideration that this represents the aggregate of potential thyroid endocrine signaling (i.e., free T4 + protein bound T4) at any given time" (page 75, lines 18-20) is accurate, potential thyroid hormone availability is less relevant for the present issue. Free T4 is the form of thyroid hormone which is biologically active, and thus is the relevant measure for assessing the effects of thyroid toxicant exposures. In addition to the interspecies differences in free vs total T4 physiology as described in the draft report, total T4 is a less optimal surrogate for free FT4 and would be valid only if all serum thyroid binding protein concentrations (which were not measured in the Feng 2017 study) are within their normal ranges.

The summary of early gestational thyroid hormone levels as a determinant of later health outcomes is accurately described in the draft report. It is well-accepted that low serum free T4 levels (hypothyroxinemia) during development is associated with impaired somatic growth and neurocognitive deficits. The use of gestational serum free T4 concentrations (particularly during early gestation) as a critical effect would be a well-supported measure of clinically relevant health outcomes.

> **EPA Response:** It is recognized that free T4 (FT4) is a common measure of thyroid hormone status in human clinical medicine as well as in many experimental animal studies, as is total T4 (TT4). It should be noted, however, that dynamics of hormone distribution in mammals under certain conditions or life stages (e.g., developing offspring) complicate interpretation of FT4. Specifically, the ultimate population of concern in the principal study (Feng et al., 2017) is the birthed fetus/neonate. Thyroid hormone levels in this particular life stage (i.e., offspring) are comprised of a mix of maternal (transplacental) hormone and some, albeit low, amount of offspring-dependent hormone synthesis. Importantly, during gestation the placenta acts as a gatekeeper in an attempt to maintain optimal levels of T4 (and through deiodination at target tissues [e.g., brain], T3). In this gatekeeper role, the placenta employs a milieu of deiodinases and membrane transport proteins to regulate T4 transfer to the fetus (keep in mind that TSH and T3 are not transported across the placenta). Placental deiodinases in particular (deiodinase 1 and 3, D1 and D3) are partly responsible for the metabolism of FT4 to reverse T3 (rT3), which is a biologically inactive form of hormone. Protein-bound T4 (i.e., non-free T4) would presumably be impervious to this deiodination at the placenta. As such, there is a concern about the interpretation of FT4 levels at the site of concern (i.e., developing/birthed fetus or neonate) due to the dynamics associated with placental transfer/metabolism of FT4. Therefore, TT4 as measured in offspring was and is considered a more accurate measure of hormone status in the exposed unit/population of concern. Furthermore, FT4 was unfortunately not measured in the mouse offspring in the principal study (Feng et al., 2017). Text has been added to the narrative in the RfD derivation section of the revised assessment for clarity.

Slitt

a. If so, please explain your justification.

After evaluation of the three key developmental studies, it does appear that the gestational exposure mouse study by Feng et al. (2017) is the key study to select for critical developmental effects. All studies seem to have similar limitations in that the PFBS used is not ultrapure and there was no blinding in the studies. However, even with the limitations, the authors did use a sufficient "n" and the study meets the criteria as "high criteria". There are other studies cited (NTP, 2018 and NTP, 2011) in the document that support thyroid hormone changes as an adverse effect – Figure 7 and Table 7 outline other studies that support the effect.

EPA Response: No revisions needed to address this comment.

b. If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the developmental RfD.

Both Leider et al., 2009a and York, 2003 meet the criteria described and are considered "high". Both describe other effects at 200-300 mg/kg/day, which is in the general range of Feng et al., 2017. Other effects to consider would be increased liver weight described by York, 2003. However, like thyroid measures it also doesn't appear to be dose responsive. Liver weight without evidence of cytotoxicity is not considered an adverse effect (Hall et al., 2012).

EPA Response: No revisions needed to address this comment.

c. In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described and whether the critical effect is associated with clinically relevant health outcome(s).

The document makes a sound argument for thyroid disruption as an adverse endpoint. Feng et al., 2017 is supported by other studies in the document that cite decreased serum T3, T4 and increased TSH (NTP, 2018; NTP 2011). 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.

However, Table 7, page 58, states "rodents are considered good models for human thyroid effects". There needs to be discussion about this along with some support from the literature that this is correct. A quick search on this revealed that it is not obvious. This should be addressed to support thyroid as the critical measure.

EPA Response: The textual rationale provided in the RfD derivation section has been expanded to further clarify the similarities and differences between humans and rodents as it pertains to thyroid hormone economy. Although no longer explicitly stated, the rationale provides a weight of evidence suggesting that rodents are a good model of thyroid toxicity, specifically in early life stages (e.g., newborn/neonate).

Warren

I fully support selection of the "high confidence" study of Feng et al. (2017) as the critical study and decreased total T4 as the critical effect. These selections are scientifically justified and defensible.

The text on pp. 69-75 provides a good discussion of the maternal-fetal unit's role in thyroid hormone homeostasis, the importance of thyroid hormone to developmental integrity, and why total T4 is the metric of choice. It also conveys the considerable uncertainties surrounding the degree of thyroid hormone reduction required for developmental insult. However, it does not address the unique sensitivity of the rat thyroid to perturbation by a host of toxicants, a consideration given the use of mice in the Feng et al. (2017) study. Overall, the toxicity assessment does an adequate job of discussing the scientific support for the thyroid as a hazard. However, it is perhaps worthy of mention that numerous epidemiological and clinical studies report that even subclinical maternal hypothyroidism during pregnancy can have neurotoxicological consequences measurable at birth and for years thereafter.

> **EPA Response:** Text regarding associations between decreased thyroid hormone during pregnancy and untoward neurodevelopmental outcomes in offspring was provided in the external review draft, and the reviewer asked the Agency to expand this section to include "subclinical" or mild hypothyroidism. Text with new citations has been added to the narrative in the subchronic RfD derivation section of the revised draft assessment to better characterize the relationship between mild/modest perturbations in maternal thyroid hormone (e.g., hypothyroxinemia versus overt hypothyroidism) and developmental outcomes.

CHARGE QUESTION 4

EPA employed benchmark dose modeling (USEPA, 2012) in the identification of points-of-departure (PODs) for PFBS, one based on kidney hyperplasia and the other based on a decrease in total T4 levels in offspring.

a. **Are the modeling approaches used, selection of benchmark response levels, and the selected models used to identify each POD for RfD derivation scientifically justified and defensible?**
b. **Specifically, considering species and/or lifestage specific differences in thyroid economy (e.g., differential reserve capacities for thyroid hormone in infants compared to adults and mice compared to humans), comment on how EPA addresses these factors in the choice of a biologically based benchmark response level (i.e., level of change that characterizes the lower limit of biological significance compared with normal background responses).**

Chou

a. **Are the modeling approaches used, selection of benchmark response levels, and the selected models used to identify each POD for RfD derivation scientifically justified and defensible?**

The selection of benchmark response levels and the selection models used to identify each PODs are justified and defensible.

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

b. Specifically, considering species and/or lifestage specific differences in thyroid economy (e.g., differential reserve capacities for thyroid hormone in infants compared to adults and mice compared to humans), comment on how EPA addresses these factors in the choice of a biologically based benchmark response level (i.e., level of change that characterizes the lower limit of biological significance compared with normal background responses).

The explanation provided in the draft for the choice of using PND1's total T4 as the critical effect was acceptable. The reviewer accepts the choice of use $BMDL_{20}$ and $BMDL_{1SD}$, as described on p. 125. The reviewer also accepts the resultant values of $MBDL_{20}$ and $BMDL_{1SD}$ shown in Table 9. The selection of a protective value of 4.4 mg/kg-d is a reasonable approach for a development stage that is expected to be very sensitive to the impact of thyroid hormone imbalance.

EPA Response: No revisions needed to address this comment.

Kamendulis

a. Are the modeling approaches used, selection of benchmark response levels, and the selected models used to identify each POD for RfD derivation scientifically justified and defensible?

Yes, I agree with the approach used. As there are no biologically based dose-response models available for K+PFBS, benchmark dose modeling was used, and was consistent with EPA's guidance document (USEPA 2012). The approach was adequately described and scientifically justified in the document.

EPA Response: No revisions needed to address this comment.

b. Specifically, considering species and/or lifestage specific differences in thyroid economy (e.g., differential reserve capacities for thyroid hormone in infants compared to adults and mice compared to humans), comment on how EPA addresses these factors in the choice of a biologically based benchmark response level (i.e., level of change that characterizes the lower limit of biological significance compared with normal background responses).

As noted in the review document, there are differences in the development and functional maturation of the hypothalamic-pituitary-thyroid axis during early lifestages between humans and rodents. Those differences are clearly described in the document. However, as pointed out, the impact on dynamic reserve capacity of T4 between species may not be as significant. Human neonates have a serum half-life of T4 of approximately 3 days (compared to 0.5-1 day in rodents) and tissue stores of T4 are approximately <1 day. Comparatively, rodents do not begin producing to produce T4 until late in gestation, therefore, newborn rodent T4 levels are primarily a reflection of transplacentally translocated maternal hormone. Thus, using data from the Feng et al. study to derive an RfD for PFBS would be expected to be protective of human toxicity.

EPA Response: No revisions needed to address this comment.

Leung

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

Regarding the differences in thyroid physiology between species and lifestages, the draft report (pages 52 and 74) accurately describes the critical role of thyroid hormone in early gestation for brain and overall development, the known differences in rodent and human thyroidal physiology, and that rodents serve as a reliable model for human thyroid physiology. The references supplied in these sections are scientifically robust and appropriate. Overall, the available animal data are sufficient to support adverse thyroidal effects from PFBS exposures. I defer to the other reviewers on the appropriateness of the uncertainty factors chosen.

EPA Response: No revisions needed to address this comment.

Slitt

a. Are the modeling approaches used, selection of benchmark response levels, and the selected models used to identify each POD for RfD derivation scientifically justified and defensible?

I do not have the expertise to comment on modeling approaches or selection of benchmark response levels.

EPA Response: No revisions needed to address this comment.

b. Specifically, considering species and/or lifestage specific differences in thyroid economy (e.g., differential reserve capacities for thyroid hormone in infants compared to adults and mice compared to humans), comment on how EPA addresses these factors in the choice of a biologically based benchmark response level (i.e., level of change that characterizes the lower limit of biological significance compared with normal background responses).

This is again out of my area of expertise, but it the documents reads as appropriate and provides a solid rationale for the selection of life stage with regard to thyroid function as being the biological benchmark response level. Some aspects of the document read well regarding the influence of thyroid hormone impact on aspects of development. The document should add additional citations to address the "differential reserve capacities" in infants compared to adults. But, with these assumptions in place, the selection of this characterizing the lower limit of biological significance is justifiable.

EPA Response: Additional text and new citations have been included in the section on subchronic RfD derivation to address the topic of differential reserve capacities between life stages.

Warren

One aspect of toxicity assessments that is particularly admirable is their presentation of BMD modeling results of multiple datasets or the same dataset using different BMRs. Such is the case with

the present assessment that presents numerous BMDLs for consideration in deriving a developmental RfD, and somewhat fewer BMDLs for subchronic and chronic RfD derivation. The presentation allows the end user (and reviewer) to conduct comparative analyses, thereby increasing confidence in the toxicity assessment's end results. For PFBS, the modeling approaches appear consistent with USEPA's *Benchmark Dose Technical Guidance* document based on the text and modeling output in Appendix F. However, while it is recognized that modeled data and output are available via embedded links within the toxicity assessment, summary tables that normally allow for the quick application of model choice decision logic are absent from the document itself. For those unfamiliar with the HAWC interface, this makes an assessment of the scientific justification and defensibility of the modeling effort more difficult (though the modeling results in F.3.1 and F.3.2 are somewhat useful in this regard). Nonetheless, I see no reason to doubt the validity of the modeling effort. As for the BMR level selected as the POD, pp. 69-70 are convincing that a BMR of 20% relative deviation from the control mean is most appropriate (compared to 1 SD) given studies that have examined the relationship between T4 deficiencies and adverse developmental outcomes.

EPA Response: The EPA appreciates the observations on the use of HAWC as we transition to increased use of web-based information management resources to complement our assessment documents. In the PFBS assessment, as the reviewer notes, we have taken the approach of making sure all the key content from HAWC is included in the assessment as figures or appendices so availability to this information does not require web access.

CHARGE QUESTION 5

Given what is known and not known about the interspecies differences in toxicokinetics of PFBS, EPA applied body weight to the ¾ allometric scaling to adjust the POD to estimate a human equivalent dose (HED) in the derivation of the respective RfDs (USEPA, 2011).

a. **Is applying the body weight to the ¾ for PFBS scientifically justified and defensible? If not, please provide your rationale and detail the alternative approach you would use.**
b. **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?**

Chou

a. **Is applying the body weight to the ¾ for PFBS scientifically justified and defensible? If not, please provide your rationale and detail the alternative approach you would use.**

The reviewer does not find any reason for deviating from the ¾ allometric scaling approach.

EPA Response: No revisions needed to address this comment.

b. **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?**

To the best of my knowledge, the methods used to derive RfDs and the applied UFs are appropriate.

EPA Response: No revisions needed to address this comment.

Kamendulis

a. Is applying the body weight to the ¾ for PFBS scientifically justified and defensible? If not, please provide your rationale and detail the alternative approach you would use.

Yes, the application of ¾ allometric scaling was justified and adequately presented. Any uncertainty in using this approach is accounted for by application of uncertainty factors (see answers to question 6 below).

EPA Response: No revisions needed to address this comment.

b. 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?

Yes, the methods that were used to derive the RfD's for PFAS took into consideration toxicokinetic differences between animals and humans. Further, any uncertainties are accounted for by the application of uncertainty factors (see question 6 below).

EPA Response: No revisions needed to address this comment.

Leung

This is not my area of expertise, thus I defer to the other reviewers.

EPA Response: No revisions needed to address this comment.

Slitt

a. Is applying the body weight to the ¾ for PFBS scientifically justified and defensible? If not, please provide your rationale and detail the alternative approach you would use.

This is not within my expertise to respond. However, the guidance supplied (USEPA, 2011b) outlines a strong scientific rationale for the appropriateness of body weight to the ¾. There is no clear reason to not apply this criteria for PFBS. I have no suggestions for alternatives.

EPA Response: No revisions needed to address this comment.

b. 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?

Yes, the methods used do account for the appropriate uncertainties.

EPA Response: No revisions needed to address this comment.

Warren

Yes, use of the default dosimetric adjustment factor (DAF) of bw to the ¾ is scientifically justified and defensible. Such is the case given the lack of a PBPK model for interspecies dose extrapolation and the dearth of mechanistic information to inform the issue of how internal dose relates to the nature, magnitude, and time-course of biological effects. Despite the existence of some data (pp. 9-14) that allow for interspecies comparisons of elimination half-lives, the toxicokinetic database is inadequate to derive equivalent human oral exposures to those posing a hazard to the thyroid of PND1 female mice or the kidneys of adult female rats. Furthermore, half-life comparisons between species are based on elimination rates after single dose administration in humans and may not apply to repeated dose conditions that more accurately reflect real-world exposure scenarios. The toxicity assessment provides additional justification for bw to the ¾ allometric scaling by noting PFBS's lack of metabolism and likely parent compound-mediated toxicities. This is consistent with the discussion in Lieder et al. (2009a) that points to PFBS as a strong surfactant with irritant properties, making its contact with renal tubule cells a viable explanation for the resulting histopathological effects. Such a mode of action is supported by the presence of a large percentage of the administered dose of PFBS in urinary output, making elimination half-life a potential determinant of species sensitivity. For example, one might predict that a species with a longer elimination half-life (and lower dose rate to the kidney) would be less susceptible to renal injury, though such a prediction is admittedly, most uncertain. Defaulting to bw to the ¾ allometric scaling is consistent with the hierarchy of approaches for interspecies extrapolation clearly expressed by USEPA in multiple documents, including *Recommended Use of Body Weight3/4 as the Default Method in Derivation of the Oral Reference Dose* and *Harmonization in Interspecies Extrapolation: Use of BW3/4 as Default Method in Derivation of the Oral RfD*. The toxicity assessment's repeated use of a UF of 3 for interspecies differences is also consistent with Agency guidance, as the default DAF appropriately addresses some, but not all, of the considerable cross-species uncertainties in both PFBS toxicokinetics and toxicodynamics. Lastly, some toxicity assessments, including that for GenX, are using an updated bw of 80 kg for adult humans in the DAF equation. While resulting in only a minor change to the DAF and human equivalent doses compared to the use of a 70 kg default, the Agency should consider standardization of this parameter across toxicity assessments.

> **EPA Response:** The EPA appreciates the support for the application of the body weight$^{3/4}$ default approach for calculating human equivalent doses. As suggested, the Agency has updated the default human body weight used in calculating the dosimetric adjustment factors to 80 kg, consistent with the adult human body weight applied in the assessment of GenX chemicals and the updated *Exposure Factors Handbook* (USEPA, 2011). Updated human equivalent doses were input into HAWC and the assessment, and new benchmark dose modeling was conducted with the updated values.

CHARGE QUESTION 6

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.

a. **Has uncertainty been adequately accounted for in the derivation of the RfDs? Please describe and provide suggestions, if needed.**
b. **Does the provided scientific rationale support the application of the selected uncertainty factors? Please explain.**

Chou

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

Uncertainty factors are appropriately applied. I have no additional suggestions.

EPA Response: No revisions needed to address this comment.

b. **Does the provided scientific rationale support the application of the selected uncertainty factors? Please explain.**

The draft provided sufficient rationale for the application of the selected uncertainty factors.

EPA Response: No revisions needed to address this comment.

Kamendulis

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

Yes, uncertainty has been adequately accounted for in the derivation of RfDs for PFBS. For all 3 RfD derivations, an interspecies uncertainty factor of 3 was applied to account for uncertainty in extrapolating from laboratory animals to humans. A factor of 3 was used in lieu of a 10 since a POD_{HED} was derived from the BMDL as specified in EPA's guidance document (USEPA 2011b). As the allometric scaling accounts for some aspects of species extrapolation, some uncertainty remains. Thus, the application of an UF of 3 appears appropriate.

For all 3 RfD derivations, an intraspecies uncertainty factor (UF_H) of 10 was assigned to account for variability in the responses within the human populations. This is also appropriate.

For the developmental RfD, a UF_D of 3 was applied. Although developmental toxicity studies in mice and rats are available, decreased thyroid hormone is critical during development, including neurodevelopment. Since the effects of PFBS on developmental neurotoxicity have not been performed, this represents a database weakness. A UF of 3 was applied to account for this uncertainty and appears justified. For the subchronic and chronic RfDs, a UF_D of 3 was applied since no peer-reviewed subchronic or chronic studies have evaluated the effects of PFBS on thyroid function. Since the thyroid was the most sensitive effect in developing offspring, the lack of toxicity assessment of this endpoint represents a data gap which justifies the used of a UF of 3.

For the chronic RfD derivation a UF_S of 10 was applied since data from a sub-chronic study was used – this is appropriate.

EPA Response: No revisions needed to address this comment.

b. Does the provided scientific rationale support the application of the selected uncertainty factors? Please explain.

This was addressed in the responses to 6a above - the application of the UF's used to derive RfD's for K+PFBS was scientifically based and well described.

EPA Response: No revisions needed to address this comment.

Leung

This is not my area of expertise, thus I defer to the other reviewers.

EPA Response: No revisions needed to address this comment.

Slitt

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

Yes. The use of uncertainty factors was used according to EPA guidance (US EPA, 2011b and US EPA, 1991). Based on guidance from "Recommended Use of Body Weight3/4 as the Default Method in Derivation of the Oral Reference Dose, US EPA, 2011b" the application of UF3 is appropriate based on guidance on pages ix and x in that guidance. The application of a UF of 3 for interspecies differences is acceptable although the interspecies differences listed in the document were fairly similar regarding dose. With NOEALs ranging from 50 mg/kd-day (Thyroid hormone changes; Feng et al., 2017) to 100 mg/kg-day in rats (kidney effects in P0 and F1 generation), UF of 3 is in line with the variation observed within the species for NOAEL. A UFH of 10 is appropriate given the lack of data for K^+PFBS or PFBS in humans. UFL of 1 is appropriate because a NOAEL was used and and a UFs of 1 because the POD was from Feng et al., 2017 which was a developmental study in mice and represents a susceptible life stage.

EPA Response: No revisions needed to address this comment.

b. Does the provided scientific rationale support the application of the selected uncertainty factors? Please explain.

Yes, Table 10 outlines and provide sufficient rationales for the application of the selected uncertainty factors. See response in part a.

EPA Response: No revisions needed to address this comment.

Warren

Yes, the considerable uncertainty surrounding the risk of PFBS exposure has been adequately accounted for in the toxicity assessment. Tables 10, 13 and 15 do a good job of explaining the rationale behind each individual UF value and I agree with the composite UFs of 100, 100 and 1,000 used for derivation of the developmental, subchronic and chronic RfDs, respectively. Admittedly, a composite UF of 1,000 is unsettling, but the conduct of a single, well-designed chronic study could result in its substantial reduction. Though peripheral to the question at hand, the thyroid hormone

perturbations reported in NTP (2018) are a reminder of the greater sensitivity frequently seen in rats to such effects compared to mice or humans. Such differential sensitivity must be a consideration should NTP (2018) ever find use in a risk assessment context, and I am reminded that some have argued for an interspecies UF less than one to account for it.

EPA Response: No revisions needed to address this comment.

CHARGE QUESTION 7

The draft assessment for PFBS identifies several potential human hazards. Please comment on whether the available data have been clearly and appropriately synthesized for each toxicological hazard. Please comment on whether the weight of evidence for hazard identification has been clearly described and scientifically justified.

Chou

The organization and description of the potential human hazards are very clearly stated in this draft.

EPA Response: No revisions needed to address this comment.

Kamendulis

The available scientific literature includes animal studies (repeat-dose) oral exposure studies using the potassium salt of PFBS (K^+PFBS). No chronic studies are available, as such, only non-cancer endpoints have been described. The rat and mouse studies have consistently identified developmental, renal, and thyroid effects as potential health outcomes following repeated exposures in utero and/or during adulthood. Other potential toxicities (hepatic, reproductive parameters, lipid/lipoprotein homeostasis, spleen, and hematology) were evaluated by were not associated with PFBS exposure. The available data for each endpoint was thoroughly and clearly presented in the document. For each endpoint, the data was contextually described, and the weight of evidence was adequately discussed and scientifically justified.

EPA Response: No revisions needed to address this comment.

Leung

Available data of PFBS exposures in human studies is overall scarce, as described in the draft report. Table 7 provides a clear summary of available data as separated by human vs animal studies. There are no human studies regarding thyroid, developmental, and hepatic effects, and the available human studies describing reproductive, renal, lipid, asthma, and atopic dermatitis effects are appropriately categorized as generally low confidence with insufficient ability to assess the current exposures of interest.

EPA Response: No revisions needed to address this comment.

Slitt

The document is based on literature recovered from very thorough searches and then applied rigorous criteria from which the recovered literature was allowed to be used to identify potential hazardous effects. From this process, several key studies emerged that pointed to developmental and chronic effects for thyroid hormone and kidney. In addition to those studies being utilized each study (Feng et al., 2017 and Leider et al., 2009a) were supported by additional studies supporting related or similar adverse effect. Overall the weight of the evidence provided does support thyroid hormone disruption and renal effects in well conducted rodent studies.

EPA Response: No revisions needed to address this comment.

Warren

The toxicity assessment does an admirable job of synthesizing the available data on PFBS' effects on the thyroid, reproduction, development, kidneys, liver, lipids/lipoproteins, immunity and genotoxicity/mutagenicity. It is skillfully written and its liberal use of tables (e.g., Tables 6) makes a substantial contribution to its organizational clarity. Initially, I thought the Evidence Synthesis and Evidence Integration sections were too repetitious and should be combined for the sake of brevity. I am no longer of that opinion, and realize the existence of both forces the toxicity assessment to logically progress through a series of interdependent steps as intended. The criteria for making a judgment as to whether the evidence supports a hazard are clearly described (Table 3), while Table 7 summarizes the application of these criteria to each study and allows for a transparent process. In my opinion, the most challenging part of the toxicity assessment process is Evidence Integration and Hazard Characterization (pp. 49-66). In the case of PFBS, the challenge was more than met.

EPA Response: No revisions needed to address this comment.

CHARGE QUESTION 8

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

Chou

The reviewer agrees that there is no sufficient information to assess carcinogenic potential of PFBS.

EPA Response: No revisions needed to address this comment.

Kamendulis

No studies exist that have evaluated the carcinogenicity of PFBS or K^+PFBS in humans or animals by any route of exposure. Therefore, an oral (oral slope factor [OSF]) or inhalation (inhalation unit risk [IUR]) cannot be derived. I agree with the conclusion that there is inadequate evidence to assess the carcinogenic potential of PFBS.

EPA Response: No revisions needed to address this comment.

Leung

This is an appropriate statement; the currently available data have not assessed carcinogenicity of PFBS exposure and are thus insufficient to make a determination regarding this potential effect.

EPA Response: No revisions needed to address this comment.

Slitt

Yes, this assessment is correct given the literature search did not reveal any studies that assessed the carcinogenic potential of PFBS. Several studies (i.e., mutagenicity test, Ames, DNA damage, CHO chromosomal aberration, micronucleus assay) failed to detect mutagenicity or genotoxicity. There were no mechanistic studies that met the criteria or recovered from the literature search performed that were included, so no mechanisms were proposed. There are a couple of studies by EPA scientists that do warrant being included that would address mechanistic aspects of PFBS signaling. The following study "Activation of mouse and human peroxisome proliferator-activated receptor alpha by perfluoroalkyl acids of different functional groups and chain lengths, Wolf CJ, Takacs ML, Schmid JE, Lau C, Abbott BD" was not included in the report but does seem to meet the criteria. This study does describe PFBS activation of both human and mouse PPAR-α at 30 μM and 150 μM, respectively and compares PFBS activity to other PFASs known to induce PPAR-α activity. A second paper by EPA scientists, entitled "Evaluation of perfluoroalkyl acid activity using primary mouse and human hepatocytes" by Rosen et al., 2013 evaluated transcriptional pathways induced in human and mouse hepatocytes by 12 different PFASs, including PFBS. This publication should be reviewed to determine if it meets the criteria and the findings could be presented in the report.

EPA Response: The EPA appreciates the suggestion of additional references for consideration in the PFBS assessment. The recommended studies were evaluated against PECO criteria and not determined to be relevant to the PECO (e.g., specifically focused on mechanistic details; in vitro only). These references were added to the HERO database as additional screened references obtained from "Other Sources". Figure 4 was updated in the assessment to illustrate the addition.

Warren

I agree that the descriptor, "Inadequate evidence to assess carcinogenic potential" is appropriate for PFBS and its potassium salt for both the oral and inhalation exposure routes. As there is a complete absence of data on the carcinogenicity of these compounds, the brief weight-of-evidence narrative afforded the issue is likewise, appropriate. It is noteworthy, however, that the critical effects on which RfD derivations are based are at least, suggestive of carcinogenic potential. Pathways to this end might include high-dose cytotoxicity followed by compensatory renal cell division and an increased incidence of renal cell tumors. Likewise, thyroid overstimulation by TSH in response to a diminution in circulating thyroid hormones may lead to hyperplasia and eventually the development of cancer.

EPA Response: No revisions needed to address this comment.

SECTION III: REVIEWER ADDITIONAL AND EDITORIAL COMMENTS

****All comments/suggestions/edits offered by reviewers in this section have been carefully considered and addressed appropriately in the revised draft assessment****

Chou

Editorial Comments: Clearly describing study design and results in a few sentences is a complex mental exercise. Author(s) of this manuscript did an excellent job. Following are a few minor points for improving study descriptions:

1) P. 33, Line 7-11: When were these observations performed in dams, PND 21?
2) P. 36, Line 8-11: Please state when (what age) sperm measurements were performed.
3) P. 53, Line 16, please indicates the study was conducted in mice.
4) P. 36, Line 12: these two studies appear to provide different results, but one would not call this "discrepancies" in endocrine toxicity. As stated in this paragraph, the animals in these two studies were exposure during different stages of life. When an underline biology involves hormonal feedback control mechanisms, exposure at different life stages or observations conducted at different times after exposure could give results appear to be conflicting, but they may not be. For example, chemically induced thyroid toxicity may go from hyperthyroidism to hypothyroidism.
5) P. 80, Table 12: Please mark this part of the Table with "(cont.)"

Kamendulis

In general, the document was very thorough and well written.

P32, line 16 – a word appears to be missing from this sentence.

P69, line 26 – a word appears to be missing from this sentence.

Leung

Page 74, lines 25-26: In addition to TBG and transthyretin, serum albumin is also considered to be an important thyroid binding protein and is usually included in this list.

Page 75, lines 21-23: Since free (or unbound) T4 is the form of thyroid hormone which is biologically active, it would be the driver of toxicant-mediated effects on the thyroid, rather than total T4.

Slitt

The document distills the information well and is comprehensive.

In the writing (for example page 75, lines 23-34), the authors need to state the dose used in the Feng et al., 2017 that is being used for to derive the $BMDL_{20}$(HED). A reader experienced reading these types of documents would be able to go through the tables of the document and find this, but it should be stated. For example, in Table 9, perhaps there could be another column that lists the dose used in the publication that is then used to derive the POD(HED) for each endpoint. Showing the calculation from start to the final RfD would be be helpful to the reader.

There are some details missing from Feng et al., 2017 that should be considered. The authors do not detail the method of blood collection for the adult mice. Multiple ELISA assays were conducted

along with using 100 µl for PFBS quantification. The vendor for the ELISA assays is not a globally recognized company in these assays and fact sheets/directions cannot be easily accessed. For this reason, it is hard to determine whether the serum might have been diluted or how much serum was used per assay. The concern is whether there was enough serum collected to run so many assays, especially from young mice.

Warren

Though most USEPA work products that I have reviewed are of high quality, the present effort is truly superior in terms of editorial quality. It is clearly written by author(s) having considerable skill in synthesizing results and integrating appropriate evidence to support judgments of hazard, including selection of critical studies and effects. Nonetheless, as would be expected of any document of such length and complexity, there are minor errors of grammar and technical accuracy, none of which detract significantly from the effort or make its decisions suspect (see recommendations for minor editorial changes below). One additional issue I wish to raise is the treatment of NTP (2018), data from which the Agency included in its toxicity assessment "*because these data had undergone normal NTP quality assurance/control processing and are publicly available"* (p. 17). It is thus puzzling why the Agency would essentially acknowledge NTP's efforts to ensure research validity, then disqualify these data from serving as the basis of a RfD based, in part, on the lack of their appearance in a peer-reviewed publication (pp. 80-81).

1. line 3, p. 3: change "was" to "were"
2. line 8, p. 8: delete "a" before "groundwater"
3. next to last line, p. 12: pluralize "value"
4. line 5, p. 13: delete the comma following "K^+PFBS"
5. line 15, p. 13: change the comma to a semi-colon
6. lines 22 and 29, p. 13: consider defining PFHxA (perfluorohexanoic acid) in the text or list of abbreviations and acronyms
7. line 17, p. 14: delete "elimination" following "K+PFBS"
8. line 11, p. 20: delete "the" prior to "Appendix C"
9. line 23, p. 25: change "was" to "were"
10. lines 15, 23 and 32, p. 32: "evidence" is grammatically singular, not plural as used
11. line 13, p. 34: delete "able to be"
12. line 10, p. 36: delete the comma
13. line 33, p. 36: add commas following "production/levels" and "progesterone"
14. line 35, p. 40: change "were" to "was"
15. line 35, p. 40: Figure E-4 presents data on developmental effects (eye opening); E-4 should be changed to E-7 and/or E-8.

16. line 8, p. 44: consider defining AEC (absolute eosinophil counts) and ECP (eosinophilic cationic protein) in the text or list of abbreviations and acronyms
17. Table 5, p. 46: In the description of Material and methods for Bomhard and Loser (1996), change "species" to "strain"
18. Table 5, p. 47: In the description of Materials and methods for 3M (2002a), delete "was" before "0.5 g K+PFBS"
19. Table 6, p. 50: Superscript the "b" under Doses tested for NTP (2018); NTP (2011)
20. Table 6, p. 50: For the study of Feng et al. (2017), decreased TSH is listed as an effect at the LOAEL; however, pages 412 and 413 of the Feng et al. publication indicate, as one might predict given reductions in thyroid hormone levels, increased TSH in offspring and dams
21. lines 24 and 31, p. 55: the two-generation study of Lieder et al. should be designated Lieder et al., 2009b rather than Leider et al., 2009a; this mistake likely occurs at other locations within the text as well
22. line 33, p. 55: change "increase" to "increased"
23. lines 31-36, p. 56: though a rare occurrence, the sentence beginning with "While it is recognized…" might be characterized as an error in syntax; as such, consideration should be given to restructuring it
24. Table 7, p. 59: In Summary of findings, delete "either" in the sentence beginning with "Endpoints relating to…"
25. Table 7, pp. 61, 62, 63: there are multiple locations where Lieber et al., 2009a should be changed to Lieber et al., 2009b
26. Table 7, p. 63: The Summary of findings makes reference to the "two longest duration studies", but only one is cited (i.e., Lieder et al., 2009a)
27. Line 26, p. 69: Change "in myriad physiological" to "in a myriad of physiological"
28. Line 20, p. 70 and line 18, p. 78: Change Appendix G to Appendix F
29. Line 39, p. 70 and line 37, p. 78: Figure 2 is the conceptual model that summarizes data availability; might Figure 3 be the decision process to which the assessment refers?
30. Table 9, Footnote a, p. 73 and Table 12, Footnote a, p. 80 and Table F-1, footnote a, p. 122: a reference human body weight of 70 kg was used to calculate the DAF; why not use a BWh of 80 kg, described by the Agency in its GenX toxicity assessment as an update to the 70 kg default?
31. Table 12, p. 79: Similar to Table 9, consider placing information related to the selected POD in bold-faced type.
32. Table 12, Footnote b, p. 80: Convert animal body weight (BWa) to kg for DAF calculation (i.e., use 0.211 kg rather than 211 g).
33. Line 20, p. 86: change "are" to "is"

34. Line 11, p. 87: singularize “generations”
35. Line 12, p. 93 and line 9, p. 109: change “is” to “are”
36. Table F-1, p. 122: Similar to that done for the developmental POD, place information related to the subchronic and chronic PODs in bold-faced type.

APPENDIX A: INDIVIDUAL REVIEWER COMMENTS

COMMENTS SUBMITTED BY

Karen Chou, PhD
Associate Professor, Department of Animal Science
Michigan State University
East Lansing, Michigan

External Peer Review of EPA's Draft
Human Health Toxicity Values for Perfluorobutane Sulfonic Acid and
Related Compound Potassium Perfluorobutane Sulfonate

1. **The document describes and applies a systematic review process for identifying and selecting pertinent studies. Please comment on whether the literature search strategy, study selection considerations, and study evaluation are appropriate and clearly described. Please identify additional peer-reviewed studies that the assessment should consider.**

   The literature search strategy and study evaluation considerations are clearly described and appear to be appropriate. The reviewer does not have any suggestions for additional studies.

2. **For PFBS the key study chosen for determining the subchronic and chronic RfDs is the Lieder et al. (2009) 90-day rat study and the critical effect is increased incidence of kidney hyperplasia in female rats. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and defensible?**

   a. **If so, please explain your justification.**

   N/A

   b. **If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the subchronic and chronic RfDs.**

   The selection of the key study by Lieder et al, (2009) and critical effect for the derivation of the subchronic and chronic RfDs for PFBS, as it is justified on p. 81 and p. 84, is highly questionable.

   The draft provides two reasons for not using the Total T4 or Free T4 data from NTP (2018) study as critical effects. Here is the justification from the draft, from P. 80 Line 14 to P. 81, Line 2: (1) "because exposure durations were shorter than subchronic-duration" and (2) "the 28-day exposure study in rats was not from a peer-reviewed publication." The reviewer disagrees to these two reasons.

   When NTP (2018) was evaluated specifically for its quality and feasibility, and specifically for this assessment, it had determined that this is a High Confidence study. See the following two excerpts from the Draft that support the reviewer's argument that NTP (2018) should be the principal study for the subchronic oral RfD assessment.

   > P. 17: "Although a peer-reviewed NTP Technical Report for the PFBS study is not yet available, this information was included in the assessment because these data had undergone normal NTP quality assurance/control processing and are publicly available. … During the process of deriving toxicity values, EPA conducted further quantitative analyses (e.g., BMD modeling) beyond what was reported by the NTP."

P. 33: The NTP (2018) study is identified as a "high confidence study" on p. 33, under Evidence Synthesis.

Furthermore, effects on thyroid hormonal balance is the most consistent findings supported by several other studies.

In summary, the results of NTP (2018) has been peer reviewed and the data are available. The study results are so new that they are yet to be published. In addition, the most relevant target, thyroid imbalance, is not even examined by Lieder et al. (2009a). When the effects on thyroid imbalance is already evident after 28-day exposure, the results should not be criticized and disregard simply because the exposure period is too short.

c. In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described.

The conclusions regarding adversity are scientifically supported and clearly described.

3. In addition, for PFBS, a RfD associated specifically with a developmental lifestage is derived. The gestational exposure mouse study by Feng et al. (2017) is chosen as the key study and the critical effect is decreased total thyroxine (T4) in offspring. Is the selection of the key study and critical effect for the derivation of this developmental RfD for PFBS scientifically justified and defensible?

a. If so, please explain your justification.

The selection of the key study by Feng et al. (2017) and the critical effect for the derivation of the subchronic and chronic RfDs for PFBS is well justified and clearly stated on p. 69-71 and p. 74-75. The approaches used for the selection of PODs and the justification of rodent models are well stated. The reviewer has no additional justifications.

b. If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the developmental RfD.

N/A

c. In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described and whether the critical effect is associated with clinically relevant health outcome(s).

The conclusions regarding adversity of PFBS on thyroid hormone T3 and T4 are scientifically supported and clearly described. These effects are supported by clinically relevant health outcome(s) observed in epidemiological studies.

4. EPA employed benchmark dose modeling (USEPA, 2012) in the identification of points-of-departure (PODs) for PFBS, one based on kidney hyperplasia and the other based on a decrease in total T4 levels in offspring.

a. Are the modeling approaches used, selection of benchmark response levels, and the selected models used to identify each POD for RfD derivation scientifically justified and defensible?

The selection of benchmark response levels and the selection models used to identify each PODs are justified and defensible.

b. Specifically, considering species and/or lifestage specific differences in thyroid economy (e.g., differential reserve capacities for thyroid hormone in infants compared to adults and mice compared to humans), comment on how EPA addresses these factors in the choice of a biologically based benchmark response level (i.e., level of change that characterizes the lower limit of biological significance compared with normal background responses).

The explanation provided in the draft for the choice of using PND1's total T4 as the critical effect was acceptable. The reviewer accepts the choice of use $BMDL_{20}$ and $BMDL_{1SD}$, as described on p. 125. The reviewer also accepts the resultant values of $MBDL_{20}$ and $BMDL_{1SD}$ shown in Table 9. The selection of a protective value of 4.4 mg/kg-d is a reasonable approach for a development stage that is expected to be very sensitive to the impact of thyroid hormone imbalance.

5. Given what is known and not known about the interspecies differences in toxicokinetics of PFBS, EPA applied body weight to the ¾ allometric scaling to adjust the POD to estimate a human equivalent dose (HED) in the derivation of the respective RfDs (USEPA, 2011).

a. Is applying the body weight to the ¾ for PFBS scientifically justified and defensible? If not, please provide your rationale and detail the alternative approach you would use.

The reviewer does not find any reason for deviating from the ¾ allometric scaling approach.

b. 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?

To the best of my knowledge, the methods used to derive RfDs and the applied UFs are appropriate.

6. 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.

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

Uncertainty factors are appropriately applied. I have no additional suggestions.

b. Does the provided scientific rationale support the application of the selected uncertainty factors? Please explain.

The draft provided sufficient rationale for the application of the selected uncertainty factors.

7. The draft assessment for PFBS identifies several potential human hazards. Please comment on whether the available data have been clearly and appropriately synthesized for each toxicological hazard. Please comment on whether the weight of evidence for hazard identification has been clearly described and scientifically justified.

The organization and description of the potential human hazards are very clearly stated in this draft.

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

The reviewer agrees that there is no sufficient information to assess carcinogenic potential of PFBS.

9. Editorial or Additional Comments: Please provide any editorial or additional comments you would like to make here. These should be any comments that are not in direct response to the technical charge questions above.

Editorial Comments: Clearly describing study design and results in a few sentences is a complex mental exercise. Author(s) of this manuscript did an excellent job. Following are a few minor points for improving study descriptions:

1) P. 33, Line 7-11: When were these observations performed in dams, PND 21?
2) P. 36, Line 8-11: Please state when (what age) sperm measurements were performed.
3) P. 53, Line 16, please indicates the study was conducted in mice.
4) P. 36, Line 12: these two studies appear to provide different results, but one would not call this "discrepancies" in endocrine toxicity. As stated in this paragraph, the animals in these two studies were exposure during different stages of life. When an underline biology involves hormonal feedback control mechanisms, exposure at different life stages or observations conducted at different times after exposure could give results appear to be conflicting, but they may not be. For example, chemically induced thyroid toxicity may go from hyperthyroidism to hypothyroidism.
5) P. 80, Table 12: Please mark this part of the Table with "(cont.)"

COMMENTS SUBMITTED BY

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

External Peer Review of EPA's Draft
Human Health Toxicity Values for Perfluorobutane Sulfonic Acid and
Related Compound Potassium Perfluorobutane Sulfonate

1. **The document describes and applies a systematic review process for identifying and selecting pertinent studies. Please comment on whether the literature search strategy, study selection considerations, and study evaluation are appropriate and clearly described. Please identify additional peer-reviewed studies that the assessment should consider.**

   Yes, the literature search strategy, study selection and evaluation considerations were very well presented and sufficiently clear. The process used was described well and was a very thorough and transparent approach to systematically evaluate each of the available scientific studies that described the health effects of PFBS/K+PFBS.

   I am unaware of other peer-reviewed studies that should be included in this assessment.

2. **For PFBS the key study chosen for determining the subchronic and chronic RfDs is the Lieder et al. (2009) 90-day rat study and the critical effect is increased incidence of kidney hyperplasia in female rats. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and defensible?**

   a. **If so, please explain your justification.**
   b. **If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the subchronic and chronic RfDs.**
   c. **In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described.**

   I agree with the selection of the critical study selected for deriving RfD's for K+PFBS (Lieder et al., 2009). The Lieder et al. (2009) study was a 90-day oral gavage study that reported mild to moderate hyperplasia in the kidneys of both male and female Sprague-Dawley rats administered K^+PFBS. The subchronic and chronic RfDs for K^+PFBS were derived from papillary epithelial tubular/ductal hyperplasia in female rats. The selection of this study and endpoint was scientifically justified and clearly described.

3. **In addition, for PFBS, a RfD associated specifically with a developmental lifestage is derived. The gestational exposure mouse study by Feng et al. (2017) is chosen as the key study and the critical effect is decreased total thyroxine (T4) in offspring. Is the selection of the key study and critical effect for the derivation of this developmental RfD for PFBS scientifically justified and defensible?**

   a. **If so, please explain your justification.**
   b. **If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the developmental RfD.**

c. In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described and whether the critical effect is associated with clinically relevant health outcome(s).

The concern of relevance to developmental toxicity was identified as a prolonged decrease in serum thyroxine (PND 1-60) in mice that were exposed to PFBS in utero. This gestational exposure study in mice (Feng et al., 2017) was selected as the principal study for derivation of the developmental reference dose using decreased serum total thyroxine (T4) in newborn (PND1) mice as the critical effect. As adequate levels of thyroid hormones are required for the development of organ systems for normal growth and development in early lifestages, I agree with the selection of the Feng et al. study and decreased T4 as the critical effect. The conclusions were scientifically justified and clearly presented in this document.

4. EPA employed benchmark dose modeling (USEPA, 2012) in the identification of points-of-departure (PODs) for PFBS, one based on kidney hyperplasia and the other based on a decrease in total T4 levels in offspring.

a. Are the modeling approaches used, selection of benchmark response levels, and the selected models used to identify each POD for RfD derivation scientifically justified and defensible?

Yes, I agree with the approach used. As there are no biologically based dose-response models available for K+PFBS, benchmark dose modeling was used, and was consistent with EPA's guidance document (USEPA 2012). The approach was adequately described and scientifically justified in the document.

b. Specifically, considering species and/or lifestage specific differences in thyroid economy (e.g., differential reserve capacities for thyroid hormone in infants compared to adults and mice compared to humans), comment on how EPA addresses these factors in the choice of a biologically based benchmark response level (i.e., level of change that characterizes the lower limit of biological significance compared with normal background responses).

As noted in the review document, there are differences in the development and functional maturation of the hypothalamic-pituitary-thyroid axis during early lifestages between humans and rodents. Those differences are clearly described in the document. However, as pointed out, the impact on dynamic reserve capacity of T4 between species may not be as significant. Human neonates have a serum half-life of T4 of approximately 3 days (compared to 0.5-1 day in rodents)and tissue stores of T4 are approximately <1 day. Comparatively, rodents do not begin producing to produce T4 until late in gestation, therefore, newborn rodent T4 levels are primarily a reflection of transplacentally translocated maternal hormone. Thus, using data from the Feng et al. study to derive an RfD for PFBS would be expected to be protective of human toxicity.

5. Given what is known and not known about the interspecies differences in toxicokinetics of PFBS, EPA applied body weight to the ¾ allometric scaling to adjust the POD to estimate a human equivalent dose (HED) in the derivation of the respective RfDs (USEPA, 2011).

a. Is applying the body weight to the ¾ for PFBS scientifically justified and defensible? If not, please provide your rationale and detail the alternative approach you would use.

Yes, the application of ¾ allometric scaling was justified and adequately presented. Any uncertainty in using this approach is accounted for by application of uncertainty factors (see answers to question 6 below).

b. 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?

Yes, the methods that were used to derive the RfD's for PFAS took into consideration toxicokinetic differences between animals and humans. Further, any uncertainties are accounted for by the application of uncertainty factors (see question 6 below).

6. 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.

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

Yes, uncertainty has been adequately accounted for in the derivation of RfDs for PFBS. For all 3 RfD derivations, an interspecies uncertainty factor of 3 was applied to account for uncertainty in extrapolating from laboratory animals to humans. A factor of 3 was used in lieu of a 10 since a POD_{HED} was derived from the BMDL as specified in EPA's guidance document (USEPA 2011b). As the allometric scaling accounts for some aspects of species extrapolation, some uncertainty remains. Thus, the application of an UF of 3 appears appropriate.

For all 3 RfD derivations, an intraspecies uncertainty factor (UF_H) of 10 was assigned to account for variability in the responses within the human populations. This is also appropriate.

For the developmental RfD, a UF_D of 3 was applied. Although developmental toxicity studies in mice and rats are available, decreased thyroid hormone is critical during development, including neurodevelopment. Since the effects of PFBS on developmental neurotoxicity have not been performed, this represents a database weakness. A UF of 3 was applied to account for this uncertainty and appears justified. For the subchronic and chronic RfDs, a UF_D of 3 was applied since no peer-reviewed subchronic or chronic studies have evaluated the effects of PFBS on thyroid function. Since the thyroid was the most sensitive effect in developing offspring, the lack of toxicity assessment of this endpoint represents a data gap which justifies the used of a UF of 3.

For the chronic RfD derivation a UF_S of 10 was applied since data from a sub-chronic study was used – this is appropriate.

b. Does the provided scientific rationale support the application of the selected uncertainty factors? Please explain.

This was addressed in the responses to 6a above - the application of the UF's used to derive RfD's for K+PFBS was scientifically based and well described.

7. **The draft assessment for PFBS identifies several potential human hazards. Please comment on whether the available data have been clearly and appropriately synthesized for each toxicological hazard. Please comment on whether the weight of evidence for hazard identification has been clearly described and scientifically justified.**

   The available scientific literature includes animal studies (repeat-dose) oral exposure studies using the potassium salt of PFBS (K^+PFBS). No chronic studies are available, as such, only non-cancer endpoints have been described. The rat and mouse studies have consistently identified developmental, renal, and thyroid effects as potential health outcomes following repeated exposures in utero and/or during adulthood. Other potential toxicities (hepatic, reproductive parameters, lipid/lipoprotein homeostasis, spleen, and hematology) were evaluated by were not associated with PFBS exposure. The available data for each endpoint was thoroughly and clearly presented in the document. For each endpoint, the data was contextually described, and the weight of evidence was adequately discussed and scientifically justified.

8. **The draft assessment concludes that there is inadequate evidence to assess carcinogenic potential for PFBS and that this descriptor applies to oral and inhalation routes of human exposure. Please comment on whether the available animal and mechanistic studies support this conclusion.**

   No studies exist that have evaluated the carcinogenicity of PFBS or K^+PFBS in humans or animals by any route of exposure. Therefore, an oral (oral slope factor [OSF]) or inhalation (inhalation unit risk [IUR]) cannot be derived. I agree with the conclusion that there is inadequate evidence to assess the carcinogenic potential of PFBS.

9. **Editorial or Additional Comments: Please provide any editorial or additional comments you would like to make here. These should be any comments that are not in direct response to the technical charge questions above.**

   In general, the document was very thorough and well written.

   P32, line 16 – a word appears to be missing from this sentence.

   P69, line 26 – a word appears to be missing from this sentence.

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 Peer Review of EPA's Draft
Human Health Toxicity Values for Perfluorobutane Sulfonic Acid and
Related Compound Potassium Perfluorobutane Sulfonate

1. **The document describes and applies a systematic review process for identifying and selecting pertinent studies. Please comment on whether the literature search strategy, study selection considerations, and study evaluation are appropriate and clearly described. Please identify additional peer-reviewed studies that the assessment should consider.**

   The draft report describes the systematic approach taken by the EPA toward the identification and selection of pertinent studies on this topic. The search strategy is overall easy to understand and transparent. The amount of detail provided in describing the screening process of potentially useful studies, the number of reviewers who completed each step, as well as the availability of each study's associated details in HAWC, are particularly appreciated. Overall, the strategy is appropriately comprehensive, and there do not appear to be any other peer-reviewed studies which need to be considered.

   However, two epidemiologic studies (Bao 2017 and Kim 2016) were excluded due to their large number of samples (96% and 72%, respectively) below their limits of detection (Table 4 in the draft report). Although this appears reasonable for these particular studies, it would be beneficial to describe what a reasonable threshold for inclusion might be for potential future assessments.

2. **For PFBS, the key study chosen for determining the subchronic and chronic RfDs is the Lieder et al. (2009) 90-day rat study and the critical effect is increased incidence of kidney hyperplasia in female rats. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and defensible?**

   a. **If so, please explain your justification.**
   b. **If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the subchronic and chronic RfDs.**
   c. **In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described.**

   The section summarizing the included studies and Table 7 are well-organized and thoughtfully prepared. Separation of the human vs animal studies in each organ system, as well as presentation of their findings via interactive graphics, greatly enhances the readability of the document. From the available data, the two organ systems demonstrating adverse effects from PFBS exposure with the highest level of confidence are the kidney and the thyroid gland. Reproductive, liver, lipid, and immune effects were all considered to be equivocal, which appears appropriate, given the paucity of data in these areas.

   Thyroid: There are no human studies regarding thyroidal effects from PFBS exposure, but the available animal studies support a consistent trend of associated hypothyroidism, which notably includes an up to 97% reduction in serum thyroid hormone concentrations among groups exposed to very high PFBS doses.

Kidney: Animal data also show that PFBS exposure is associated with increased renal weight and abnormal histopathologic findings (mostly renal hyperplasia). The one available human renal study suggesting that uric acid levels may be elevated in exposed boys is not particularly robust.

Overall, the selection of Lieder et al 2009 as the key study for determining subchronic and chronic RfDs (corresponding to the critical effect of abnormal renal histopathology in primarily female rats) is reasonably supported by these data.

However, the adverse thyroidal effects from exposure are substantial, and hypothyroidism can arguably be considered also as another critical effect. Specifically, the NTP 2018 study was notable for the following points:

1. In the high-dose PFBS group (500mg/kg/day) after 28 days, striking reductions of approximately 92% in serum total thyroxine (TT4), 85% in free thyroxine (FT4), and 56% in total triiodothyronine (TT3) concentrations among male rats, and 69% in serum TT4, 65% in FT4, and 43% in TT3 among female rats were observed. The one female rat which was exposed to 1000 mcg/kg/day was found to have even more profound hypothyroidism, in accordance with a monotonic dose-response effect. The associated standard error of measurement (SEM) values in all groups were reasonably small to support these overall trends.

   [I was not able to access the NTP 2011 study to examine it greater detail, and there is no associated link in the references, but it is also described in the draft report to have produced similarly dramatic reductions in serum thyroid hormone levels.]

2. In this timeframe of exposure (28 days), the lack of significant abnormalities related to thyroid weight and histopathology is not unexpected. However, one would have expected a greater elevation of serum TSH than what was observed for this degree of hypothyroidism within this timeframe.

3. In mild hypothyroidism, individuals with an intact hypothalamic-pituitary-thyroid axis can compensate by the action of increased TSH production to stimulate thyroid hormone production at the thyroid gland. However, for the greatest severity of hypothyroidism seen in the NTP 2011 and 2018 studies, normal physiologic adaptive processes would be insufficient to restore normal thyroid function. Thus, given the importance of normal thyroid hormone levels in multiple systems, the significant reduction of thyroid hormone availability can justifiably be considered a critical effect of PFBS exposure.

3. In addition, for PFBS, a RfD associated specifically with a developmental lifestage is derived. The gestational exposure mouse study by Feng et al. (2017) is chosen as the key study and the critical effect is decreased total thyroxine (T4) in offspring. Is the selection of the key study and critical effect for the derivation of this developmental RfD for PFBS scientifically justified and defensible?

   a. If so, please explain your justification.
   b. If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the developmental RfD.
   c. In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described and whether the critical effect is associated with clinically relevant health outcome(s).

Although the statement "The selection of total T4 as the critical effect is based on the consideration that this represents the aggregate of potential thyroid endocrine signaling (i.e., free T4 + protein bound T4) at any given time" (page 75, lines 18-20) is accurate, potential thyroid hormone availability is less relevant for the present issue. Free T4 is the form of thyroid hormone which is biologically active, and thus is the relevant measure for assessing the effects of thyroid toxicant exposures. In addition to the interspecies differences in free vs total T4 physiology as described in the draft report, total T4 is a less optimal surrogate for free FT4 and would be valid only if all serum thyroid binding protein concentrations (which were not measured in the Feng 2017 study) are within their normal ranges.

The summary of early gestational thyroid hormone levels as a determinant of later health outcomes is accurately described in the draft report. It is well-accepted that low serum free T4 levels (hypothyroxinemia) during development is associated with impaired somatic growth and neurocognitive deficits. The use of gestational serum free T4 concentrations (particularly during early gestation) as a critical effect would be a well-supported measure of clinically relevant health outcomes.

4. **EPA employed benchmark dose modeling (USEPA, 2012) in the identification of points-of-departure (PODs) for PFBS, one based on kidney hyperplasia and the other based on a decrease in total T4 levels in offspring.**

    a. **Are the modeling approaches used, selection of benchmark response levels, and the selected models used to identify each POD for RfD derivation scientifically justified and defensible?**
    b. **Specifically, considering species and/or lifestage specific differences in thyroid economy (e.g., differential reserve capacities for thyroid hormone in infants compared to adults and mice compared to humans), comment on how EPA addresses these factors in the choice of a biologically based benchmark response level (i.e., level of change that characterizes the lower limit of biological significance compared with normal background responses).**

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

Regarding the differences in thyroid physiology between species and lifestages, the draft report (pages 52 and 74) accurately describes the critical role of thyroid hormone in early gestation for brain and overall development, the known differences in rodent and human thyroidal physiology, and that rodents serve as a reliable model for human thyroid physiology. The references supplied in these sections are scientifically robust and appropriate. Overall, the available animal data are sufficient to support adverse thyroidal effects from PFBS exposures. I defer to the other reviewers on the appropriateness of the uncertainty factors chosen.

5. **Given what is known and not known about the interspecies differences in toxicokinetics of PFBS, EPA applied body weight to the ¾ allometric scaling to adjust the POD to estimate a human equivalent dose (HED) in the derivation of the respective RfDs (USEPA, 2011).**

    a. **Is applying the body weight to the ¾ for PFBS scientifically justified and defensible? If not, please provide your rationale and detail the alternative approach you would use.**

b. 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?

This is not my area of expertise, thus I defer to the other reviewers.

6. 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.

a. Has uncertainty been adequately accounted for in the derivation of the RfDs? Please describe and provide suggestions, if needed.
b. Does the provided scientific rationale support the application of the selected uncertainty factors? Please explain.

This is not my area of expertise, thus I defer to the other reviewers.

7. The draft assessment for PFBS identifies several potential human hazards. Please comment on whether the available data have been clearly and appropriately synthesized for each toxicological hazard. Please comment on whether the weight of evidence for hazard identification has been clearly described and scientifically justified.

Available data of PFBS exposures in human studies is overall scarce, as described in the draft report. Table 7 provides a clear summary of available data as separated by human vs animal studies. There are no human studies regarding thyroid, developmental, and hepatic effects, and the available human studies describing reproductive, renal, lipid, asthma, and atopic dermatitis effects are appropriately categorized as generally low confidence with insufficient ability to assess the current exposures of interest.

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

This is an appropriate statement; the currently available data have not assessed carcinogenicity of PFBS exposure and are thus insufficient to make a determination regarding this potential effect.

9. Editorial or Additional Comments: Please provide any editorial or additional comments you would like to make here. These should be any comments that are not in direct response to the technical charge questions above.

Page 74, lines 25-26: In addition to TBG and transthyretin, serum albumin is also considered to be an important thyroid binding protein and is usually included in this list.

Page 75, lines 21-23: Since free (or unbound) T4 is the form of thyroid hormone which is biologically active, it would be the driver of toxicant-mediated effects on the thyroid, rather than total T4.

COMMENTS SUBMITTED BY

Angela L. Slitt, PhD
Associate Professor
Department of Biomedical and Pharmaceutical Sciences
and
Director, Graduate Programs, College of Pharmacy
University of Rhode Island
Kingston, Rhode Island

External Peer Review of EPA's Draft
Human Health Toxicity Values for Perfluorobutane Sulfonic Acid and
Related Compound Potassium Perfluorobutane Sulfonate

1. **The document describes and applies a systematic review process for identifying and selecting pertinent studies. Please comment on whether the literature search strategy, study selection considerations, and study evaluation are appropriate and clearly described. Please identify additional peer-reviewed studies that the assessment should consider.**

   The literature search strategy was appropriate and thorough, with the overall method being very thorough and objective. The method was described and included clear criteria for the inclusion and exclusion of studies. The databases utilized (i.e., PubMed, Web of Science, Toxline, and TSCATS via Toxline) are appropriate and the search terms were comprehensive in nature. The methods in appendices A-C used to evaluate study quality were systematic and thorough. The metrics and criteria applied for Animal and *in vitro* toxicity studies were exceedingly thorough and well defined. The weighting and relative important used for weighting the criteria was appropriate. Specifically, the approach for evaluating epidemiological and animal toxicology studies was well described and rationalized in several well-organized diagrams (Figures 3-5 and Tables 3 & 4). The overall framework for judging the health effect was systematic, objective, and unbiased.

2. **For PFBS the key study chosen for determining the subchronic and chronic RfDs is the Lieder et al. (2009) 90-day rat study and the critical effect is increased incidence of kidney hyperplasia in female rats. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and defensible?**

   a. **If so, please explain your justification.**

   The Leider et al., 2009a study was considered to meet the criteria to be included. This was a 90-day gavage study for adult male and female rats. The study methods are well described, and the study had a reasonable "n" of 10 animals per treatment group to statistically detect effects. The Leider et al., 2009a study for kidney effects is also supported by NTP, 2018, 2011; Leider et al., 2009; 3M, 2001; 200d. These studies describe alterations in renal weight and some evidence of histopathological changes (i.e., inflammation) that were dose dependent. Thus, the renal effects being considered to be the critical effect is supported.

   b. **If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the subchronic and chronic RfDs.**

   I do not have a suggestion for alternative studies or effects to be selected to support the RfD. The Leider et al., 2009a study was thoroughly designed and met the evaluation criteria to be included. Along with renal, that publication considered multiple other endpoints, such as other organs, clinical chemistry, hematology, and histopathology for four other tissues. Although liver effects were also noted among several studies (incuding Leider et al, 2009a), none noted evidence or cytotoxicity or single-cell necrosis.

c. In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described.

Yes, the conclusions regarding adversity about supported and described. Table 7 on pages 62 and 63 presents the findings from six high confidence studies that there are various kidney effects observed in rats, with histopathology changes in kidney for female rats being the primary effect.

3. In addition, for PFBS, a RfD associated specifically with a developmental lifestage is derived. The gestational exposure mouse study by Feng et al. (2017) is chosen as the key study and the critical effect is decreased total thyroxine (T4) in offspring. Is the selection of the key study and critical effect for the derivation of this developmental RfD for PFBS scientifically justified and defensible?

a. If so, please explain your justification.

After evaluation of the three key developmental studies, it does appear that the gestational exposure mouse study by Feng et al. (2017) is the key study to select for critical developmental effects. All studies seem to have similar limitations in that the PFBS used is not ultrapure and there was no blinding in the studies. However, even with the limitations, the authors did use a sufficient "n" and the study meets the criteria as "high criteria". There are other studies cited (NTP, 2018 and NTP, 2011) in the document that support thyroid hormone changes as an adverse effect – Figure 7 and Table 7 outline other studies that support the effect.

b. If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the developmental RfD.

Both Leider et al., 2009a and York, 2003 meet the criteria described and are considered "high". Both describe other effects at 200-300 mg/kg/day, which is in the general range of Feng et al., 2017. Other effects to consider would be increased liver weight described by York, 2003. However, like thyroid measures it also doesn't appear to be dose responsive. Liver weight without evidence of cytotoxicity is not considered an adverse effect (Hall et al., 2012).

c. In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described and whether the critical effect is associated with clinically relevant health outcome(s).

The document makes a sound argument for thyroid disruption as an adverse endpoint. Feng et al., 2017 is supported by other studies in the document that cite decreased serum T3, T4 and increased TSH (NTP, 2018; NTP 2011). 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.

However, Table 7, page 58, states "rodents are considered good models for human thyroid effects". There needs to be discussion about this along with some support from the literature that this is correct. A quick search on this revealed that it is not obvious. This should be addressed to support thyroid as the critical measure.

4. **EPA employed benchmark dose modeling (USEPA, 2012) in the identification of points-of-departure (PODs) for PFBS, one based on kidney hyperplasia and the other based on a decrease in total T4 levels in offspring.**

    a. **Are the modeling approaches used, selection of benchmark response levels, and the selected models used to identify each POD for RfD derivation scientifically justified and defensible?**

    I do not have the expertise to comment on modeling approaches or selection of benchmark response levels.

    b. **Specifically, considering species and/or lifestage specific differences in thyroid economy (e.g., differential reserve capacities for thyroid hormone in infants compared to adults and mice compared to humans), comment on how EPA addresses these factors in the choice of a biologically based benchmark response level (i.e., level of change that characterizes the lower limit of biological significance compared with normal background responses).**

    This is again out of my area of expertise, but it the documents reads as appropriate and provides a solid rationale for the selection of life stage with regard to thyroid function as being the biological benchmark response level. Some aspects of the document read well regarding the influence of thyroid hormone impact on aspects of development. The document should add additional citations to address the "differential reserve capacities" in infants compared to adults. But, with these assumptions in place, the selection of this characterizing the lower limit of biological significance is justifiable.

5. **Given what is known and not known about the interspecies differences in toxicokinetics of PFBS, EPA applied body weight to the ¾ allometric scaling to adjust the POD to estimate a human equivalent dose (HED) in the derivation of the respective RfDs (USEPA, 2011).**

    a. **Is applying the body weight to the ¾ for PFBS scientifically justified and defensible? If not, please provide your rationale and detail the alternative approach you would use.**

    This is not within my expertise to respond. However, the guidance supplied (US EPA, 2011b) outlines a strong scientific rationale for the appropriateness of body weight to the ¾. There is no clear reason to not apply this criteria for PFBS. I have no suggestions for alternatives.

    b. **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?**

    Yes, the methods used do account for the appropriate uncertainties.

6. **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.**

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

Yes. The use of uncertainty factors was used according to EPA guidance (US EPA, 2011b and US EPA, 1991). Based on guidance from "Recommended Use of Body Weight3/4 as the Default Method in Derivation of the Oral Reference Dose, US EPA, 2011b" the application of UF3 is appropriate based on guidance on pages ix and x in that guidance. The application of a UF of 3 for interspecies differences is acceptable although the interspecies differences listed in the document were fairly similar regarding dose. With NOEALs ranging from 50 mg/kd-day (Thyroid hormone changes; Feng et al., 2017) to 100 mg/kg-day in rats (kidney effects in P0 and F1 generation), UF of 3 is in line with the variation observed within the species for NOAEL. A UFH of 10 is appropriate given the lack of data for K^+PFBS or PFBS in humans. UFL of 1 is appropriate because a NOAEL was used and and a UFs of 1 because the POD was from Feng et al., 2017 which was a developmental study in mice and represents a susceptible life stage.

b. Does the provided scientific rationale support the application of the selected uncertainty factors? Please explain.

Yes, Table 10 outlines and provide sufficient rationales for the application of the selected uncertainty factors. See response in part a.

7. The draft assessment for PFBS identifies several potential human hazards. Please comment on whether the available data have been clearly and appropriately synthesized for each toxicological hazard. Please comment on whether the weight of evidence for hazard identification has been clearly described and scientifically justified.

The document is based on literature recovered from very thorough searches and then applied rigorous criteria from which the recovered literature was allowed to be used to identify potential hazardous effects. From this process, several key studies emerged that pointed to developmental and chronic effects for thyroid hormone and kidney. In addition to those studies being utilized each study (Feng et al., 2017 and Leider et al., 2009a) were supported by additional studies supporting related or similar adverse effect. Overall the weight of the evidence provided does support thyroid hormone disruption and renal effects in well conducted rodent studies.

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

Yes, this assessment is correct given the literature search did not reveal any studies that assessed the carcinogenic potential of PFBS. Several studies (i.e., mutagenicity test, Ames, DNA damage, CHO chromosomal aberration, micronucleus assay) failed to detect mutagenicity or genotoxicity. There were no mechanistic studies that met the criteria or recovered from the literature search performed that were included, so no mechanisms were proposed. There are a couple of studies by EPA scientists that do warrant being included that would address mechanistic aspects of PFBS signaling. The following study "Activation of mouse and human peroxisome proliferator-activated receptor alpha by perfluoroalkyl acids of different functional groups and chain lengths, Wolf CJ, Takacs ML, Schmid JE, Lau C, Abbott BD" was not included in the report but does

seem to meet the criteria. This study does describe PFBS activation of both human and mouse PPAR-α at 30 μM and 150 μM, respectively and compares PFBS activity to other PFASs known to induce PPAR-α activity. A second paper by EPA scientists, entitled "Evaluation of perfluoroalkyl acid activity using primary mouse and human hepatocytes" by Rosen et al., 2013 evaluated transcriptional pathways induced in human and mouse hepatocytes by 12 different PFASs, including PFBS. This publication should be reviewed to determine if it meets the criteria and the findings could be presented in the report.

9. Editorial or Additional Comments: Please provide any editorial or additional comments you would like to make here. These should be any comments that are not in direct response to the technical charge questions above.

The document distills the information well and is comprehensive.

In the writing (for example page 75, lines 23-34), the authors need to state the dose used in the Feng et al., 2017 that is being used for to derive the $BMDL_{20}$(HED). A reader experienced reading these types of documents would be able to go through the tables of the document and find this, but it should be stated. For example, in Table 9, perhaps there could be another column that lists the dose used in the publication that is then used to derive the POD(HED) for each endpoint. Showing the calculation from start to the final RfD would be be helpful to the reader.

There are some details missing from Feng et al., 2017 that should be considered. The authors do not detail the method of blood collection for the adult mice. Multiple ELISA assays were conducted along with using 100 μl for PFBS quantification. The vendor for the ELISA assays is not a globally recognized company in these assays and fact sheets/directions cannot be easily accessed. For this reason, it is hard to determine whether the serum might have been diluted or how much serum was used per assay. The concern is whether there was enough serum collected to run so many assays, especially from young mice.

COMMENTS SUBMITTED BY

David Alan Warren, MPH, PhD
Program Director, Environmental Health Science
University of South Carolina Beaufort
Beaufort, South Carolina

External Peer Review of EPA's Draft
Human Health Toxicity Values for Perfluorobutane Sulfonic Acid and
Related Compound Potassium Perfluorobutane Sulfonate

1. **The document describes and applies a systematic review process for identifying and selecting pertinent studies. Please comment on whether the literature search strategy, study selection considerations, and study evaluation are appropriate and clearly described. Please identify additional peer-reviewed studies that the assessment should consider.**

   Pages 17-31 and Appendices A-D clearly describe an appropriate, but laborious process by which literature, peer-reviewed and otherwise, is identified and evaluated for the purpose of toxicity assessment. Use of the word "process" is clearly appropriate, as the toxicity assessment describes a series of progressive and interdependent steps as a means to an end. The process requires considerable skill if done well, and I refrain from using the phrase "done correctly" in recognition of the subjective judgments that remain inherent to it. The PECO criteria and additional exclusion criteria are straightforward, and were obviously effective at severely reducing all studies to a manageable number of relevant ones. The listed domains for epidemiological and animal studies should make for a comprehensive evaluation, and the domain-specific questions in Appendix C can be of assistance to even the most experienced reviewer. While Figures 5 and 6 are informative, the toxicity assessment fails to disclose the rules by which domain ratings are combined to reach an overall study classification. Such information would be a welcomed addition to the toxicity assessment document. And while HAWC undoubtedly benefits the toxicity assessment process, it is important that the written work product be sufficiently informative to the end user that web access is purely optional. Such is the case for PFBS and as such, I support the study identification and evaluation process used, while encouraging its refinement over time. As for additional peer-reviewed studies, three were located though no attempt was made to evaluate them against PECO criteria.

   1. Xu et al., 2017. Effects of perfluoroalkyl substances on neurosteroid synthetic enzymes in the rat. *Chemico-Biological Interactions*, Vol. 272, pp. 182-187.
   2. Chen et al., 2018. Multigenerational Disruption of the Thyroid Endocrine System in Marine Medaka after a Life-Cycle Exposure to Perfluorobutane sulfonate. *Environ. Sci. Technol*., 52(7), pp. 4432-4439.
   3. Gyllenhammar et al., 2018. Perfluoroalkyl acids (PFAAs) in serum from 2-4 month-old infants: Influence of maternal serum concentration, gestational age, breast-feeding, and contaminated drinking water. *Environ. Sci. Technol*.52 (12), pp. 7101-7110.

2. **For PFBS the key study chosen for determining the subchronic and chronic RfDs is the Lieder et al. (2009) 90-day rat study and the critical effect is increased incidence of kidney hyperplasia in female rats. Is the selection of the key study and critical effect for the derivation of the subchronic and chronic RfDs for PFBS scientifically justified and defensible?**

   a. **If so, please explain your justification.**

b. If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the subchronic and chronic RfDs.
c. In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described.

Considering the database limitations for PFBS, I support selection of Lieder et al. (2009a) as the critical study and kidney hyperplasia in females as the critical effect. However, while these selections can be scientifically justified and defended, an argument can also be made for alternatives (i.e., reduced T4 in mice and rats from Feng et al. (2017) and NTP (2018), respectively) that when subjected to BMD modeling, result in lower BMDLs. While Lieder et al.'s evaluation resulted in its characterization as "good" (Figure 6) or "high confidence" (Table 6), this alone does not distinguish it from several other studies that underwent consideration. The decision ultimately came down to what was considered of most importance - basing RfD derivation on the most sensitive endpoint in adults (thyroid hormone perturbation) or reliance on a peer-reviewed study of subchronic duration. Clearly, the Agency felt as though peer-review and adequate exposure duration trumped sensitivity, despite NTP's reputation for generating high quality data and reliance on Feng et al. (2017) for developmental RfD derivation. As I am unable to make a reasoned choice between the aforementioned alternatives, I support the Agency's decision. It seems clear that given database limitations, the choice of target tissue was limited to the thyroid and kidney. The toxicity assessment does an adequate job of discussing the scientific support for both as hazards.

3. In addition, for PFBS, a RfD associated specifically with a developmental lifestage is derived. The gestational exposure mouse study by Feng et al. (2017) is chosen as the key study and the critical effect is decreased total thyroxine (T4) in offspring. Is the selection of the key study and critical effect for the derivation of this developmental RfD for PFBS scientifically justified and defensible?

a. If so, please explain your justification.
b. If not, please provide your rationale and detail an alternative key study and/or critical effect you would select to support the derivation of the developmental RfD.
c. In addition, please comment on whether the conclusions regarding adversity are scientifically supported and clearly described and whether the critical effect is associated with clinically relevant health outcome(s).

I fully support selection of the "high confidence" study of Feng et al. (2017) as the critical study and decreased total T4 as the critical effect. These selections are scientifically justified and defensible. The text on pp. 69-75 provides a good discussion of the maternal-fetal unit's role in thyroid hormone homeostasis, the importance of thyroid hormone to developmental integrity, and why total T4 is the metric of choice. It also conveys the considerable uncertainties surrounding the degree of thyroid hormone reduction required for developmental insult. However, it does not address the unique sensitivity of the rat thyroid to perturbation by a host of toxicants, a consideration given the use of mice in the Feng et al. (2017) study. Overall, the toxicity assessment does an adequate job of discussing the scientific support for the thyroid as a hazard. However, it is perhaps worthy of mention that numerous epidemiological and clinical studies report that even subclinical maternal hypothyroidism during pregnancy can have neurotoxicological consequences measurable at birth and for years thereafter.

4. EPA employed benchmark dose modeling (USEPA, 2012) in the identification of points-of-departure (PODs) for PFBS, one based on kidney hyperplasia and the other based on a decrease in total T4 levels in offspring.

a. Are the modeling approaches used, selection of benchmark response levels, and the selected models used to identify each POD for RfD derivation scientifically justified and defensible?
b. Specifically, considering species and/or lifestage specific differences in thyroid economy (e.g., differential reserve capacities for thyroid hormone in infants compared to adults and mice compared to humans), comment on how EPA addresses these factors in the choice of a biologically based benchmark response level (i.e., level of change that characterizes the lower limit of biological significance compared with normal background responses).

One aspect of toxicity assessments that is particularly admirable is their presentation of BMD modeling results of multiple datasets or the same dataset using different BMRs. Such is the case with the present assessment that presents numerous BMDLs for consideration in deriving a developmental RfD, and somewhat fewer BMDLs for subchronic and chronic RfD derivation. The presentation allows the end user (and reviewer) to conduct comparative analyses, thereby increasing confidence in the toxicity assessment's end results. For PFBS, the modeling approaches appear consistent with USEPA's *Benchmark Dose Technical Guidance* document based on the text and modeling output in Appendix F. However, while it is recognized that modeled data and output are available via embedded links within the toxicity assessment, summary tables that normally allow for the quick application of model choice decision logic are absent from the document itself. For those unfamiliar with the HAWC interface, this makes an assessment of the scientific justification and defensibility of the modeling effort more difficult (though the modeling results in F.3.1 and F.3.2 are somewhat useful in this regard). Nonetheless, I see no reason to doubt the validity of the modeling effort. As for the BMR level selected as the POD, pp. 69-70 are convincing that a BMR of 20% relative deviation from the control mean is most appropriate (compared to 1 SD) given studies that have examined the relationship between T4 deficiencies and adverse developmental outcomes.

5. Given what is known and not known about the interspecies differences in toxicokinetics of PFBS, EPA applied body weight to the ¾ allometric scaling to adjust the POD to estimate a human equivalent dose (HED) in the derivation of the respective RfDs (USEPA, 2011).

a. Is applying the body weight to the ¾ for PFBS scientifically justified and defensible? If not, please provide your rationale and detail the alternative approach you would use.
b. 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?

Yes, use of the default dosimetric adjustment factor (DAF) of bw to the ¾ is scientifically justified and defensible. Such is the case given the lack of a PBPK model for interspecies dose extrapolation and the dearth of mechanistic information to inform the issue of how internal dose relates to the nature, magnitude, and time-course of biological effects. Despite the existence of some data (pp. 9-14) that allow for interspecies comparisons of elimination half-lives, the

toxicokinetic database is inadequate to derive equivalent human oral exposures to those posing a hazard to the thyroid of PND1 female mice or the kidneys of adult female rats. Furthermore, half-life comparisons between species are based on elimination rates after single dose administration in humans and may not apply to repeated dose conditions that more accurately reflect real-world exposure scenarios. The toxicity assessment provides additional justification for bw to the ¾ allometric scaling by noting PFBS's lack of metabolism and likely parent compound-mediated toxicities. This is consistent with the discussion in Lieder et al. (2009a) that points to PFBS as a strong surfactant with irritant properties, making its contact with renal tubule cells a viable explanation for the resulting histopathological effects. Such a mode of action is supported by the presence of a large percentage of the administered dose of PFBS in urinary output, making elimination half-life a potential determinant of species sensitivity. For example, one might predict that a species with a longer elimination half-life (and lower dose rate to the kidney) would be less susceptible to renal injury, though such a prediction is admittedly, most uncertain. Defaulting to bw to the ¾ allometric scaling is consistent with the hierarchy of approaches for interspecies extrapolation clearly expressed by USEPA in multiple documents, including *Recommended Use of Body Weight3/4 as the Default Method in Derivation of the Oral Reference Dose* and *Harmonization in Interspecies Extrapolation: Use of BW3/4 as Default Method in Derivation of the Oral RfD*. The toxicity assessment's repeated use of a UF of 3 for interspecies differences is also consistent with Agency guidance, as the default DAF appropriately addresses some, but not all, of the considerable cross-species uncertainties in both PFBS toxicokinetics and toxicodynamics. Lastly, some toxicity assessments, including that for GenX, are using an updated bw of 80 kg for adult humans in the DAF equation. While resulting in only a minor change to the DAF and human equivalent doses compared to the use of a 70 kg default, the Agency should consider standardization of this parameter across toxicity assessments.

6. 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.

a. Has uncertainty been adequately accounted for in the derivation of the RfDs? Please describe and provide suggestions, if needed.
b. Does the provided scientific rationale support the application of the selected uncertainty factors? Please explain.

Yes, the considerable uncertainty surrounding the risk of PFBS exposure has been adequately accounted for in the toxicity assessment. Tables 10, 13 and 15 do a good job of explaining the rationale behind each individual UF value and I agree with the composite UFs of 100, 100 and 1,000 used for derivation of the developmental, subchronic and chronic RfDs, respectively. Admittedly, a composite UF of 1,000 is unsettling, but the conduct of a single, well-designed chronic study could result in its substantial reduction. Though peripheral to the question at hand, the thyroid hormone perturbations reported in NTP (2018) are a reminder of the greater sensitivity frequently seen in rats to such effects compared to mice or humans. Such differential sensitivity must be a consideration should NTP (2018) ever find use in a risk assessment context, and I am reminded that some have argued for an interspecies UF less than one to account for it.

7. The draft assessment for PFBS identifies several potential human hazards. Please comment on whether the available data have been clearly and appropriately synthesized for each toxicological hazard. Please comment on whether the weight of evidence for hazard identification has been clearly described and scientifically justified.

The toxicity assessment does an admirable job of synthesizing the available data on PFBS' effects on the thyroid, reproduction, development, kidneys, liver, lipids/lipoproteins, immunity and genotoxicity/mutagenicity. It is skillfully written and its liberal use of tables (e.g., Tables 6) makes a substantial contribution to its organizational clarity. Initially, I thought the Evidence Synthesis and Evidence Integration sections were too repetitious and should be combined for the sake of brevity. I am no longer of that opinion, and realize the existence of both forces the toxicity assessment to logically progress through a series of interdependent steps as intended. The criteria for making a judgment as to whether the evidence supports a hazard are clearly described (Table 3), while Table 7 summarizes the application of these criteria to each study and allows for a transparent process. In my opinion, the most challenging part of the toxicity assessment process is Evidence Integration and Hazard Characterization (pp. 49-66). In the case of PFBS, the challenge was more than met.

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

I agree that the descriptor, "Inadequate evidence to assess carcinogenic potential" is appropriate for PFBS and its potassium salt for both the oral and inhalation exposure routes. As there is a complete absence of data on the carcinogenicity of these compounds, the brief weight-of-evidence narrative afforded the issue is likewise, appropriate. It is noteworthy, however, that the critical effects on which RfD derivations are based are at least, suggestive of carcinogenic potential. Pathways to this end might include high-dose cytotoxicity followed by compensatory renal cell division and an increased incidence of renal cell tumors. Likewise, thyroid overstimulation by TSH in response to a diminution in circulating thyroid hormones may lead to hyperplasia and eventually the development of cancer.

9. Editorial or Additional Comments: Please provide any editorial or additional comments you would like to make here. These should be any comments that are not in direct response to the technical charge questions above.

Though most USEPA work products that I have reviewed are of high quality, the present effort is truly superior in terms of editorial quality. It is clearly written by author(s) having considerable skill in synthesizing results and integrating appropriate evidence to support judgments of hazard, including selection of critical studies and effects. Nonetheless, as would be expected of any document of such length and complexity, there are minor errors of grammar and technical accuracy, none of which detract significantly from the effort or make its decisions suspect (see recommendations for minor editorial changes below). One additional issue I wish to raise is the treatment of NTP (2018), data from which the Agency included in its toxicity assessment "*because these data had undergone normal NTP quality assurance/control processing and are publicly available"* (p. 17). It is thus puzzling why the Agency would essentially acknowledge

NTP's efforts to ensure research validity, then disqualify these data from serving as the basis of a RfD based, in part, on the lack of their appearance in a peer-reviewed publication (pp. 80-81).

1. line 3, p. 3: change "was" to "were"
2. line 8, p. 8: delete "a" before "groundwater"
3. next to last line, p. 12: pluralize "value"
4. line 5, p. 13: delete the comma following "K^+PFBS"
5. line 15, p. 13: change the comma to a semi-colon
6. lines 22 and 29, p. 13: consider defining PFHxA (perfluorohexanoic acid) in the text or list of abbreviations and acronyms
7. line 17, p. 14: delete "elimination" following "K+PFBS"
8. line 11, p. 20: delete "the" prior to "Appendix C"
9. line 23, p. 25: change "was" to "were"
10. lines 15, 23 and 32, p. 32: "evidence" is grammatically singular, not plural as used
11. line 13, p. 34: delete "able to be"
12. line 10, p. 36: delete the comma
13. line 33, p. 36: add commas following "production/levels" and "progesterone"
14. line 35, p. 40: change "were" to "was"
15. line 35, p. 40: Figure E-4 presents data on developmental effects (eye opening); E-4 should be changed to E-7 and/or E-8.
16. line 8, p. 44: consider defining AEC (absolute eosinophil counts) and ECP (eosinophilic cationic protein) in the text or list of abbreviations and acronyms
17. Table 5, p. 46: In the description of Material and methods for Bomhard and Loser (1996), change "species" to "strain"
18. Table 5, p. 47: In the description of Materials and methods for 3M (2002a), delete "was" before "0.5 g K+PFBS"
19. Table 6, p. 50: Superscript the "b" under Doses tested for NTP (2018); NTP (2011)
20. Table 6, p. 50: For the study of Feng et al. (2017), <u>decreased</u> TSH is listed as an effect at the LOAEL; however, pages 412 and 413 of the Feng et al. publication indicate, as one might predict given reductions in thyroid hormone levels, <u>increased</u> TSH in offspring and dams
21. lines 24 and 31, p. 55: the two-generation study of Lieder et al. should be designated Lieder et al., 2009b rather than Leider et al., 2009a; this mistake likely occurs at other locations within the text as well

22. line 33, p. 55: change “increase” to “increased”
23. lines 31-36, p. 56: though a rare occurrence, the sentence beginning with “While it is recognized…” might be characterized as an error in syntax; as such, consideration should be given to restructuring it
24. Table 7, p. 59: In Summary of findings, delete “either” in the sentence beginning with “Endpoints relating to…”
25. Table 7, pp. 61, 62, 63: there are multiple locations where Lieber et al., 2009a should be changed to Lieber et al., 2009b
26. Table 7, p. 63: The Summary of findings makes reference to the “two longest duration studies”, but only one is cited (i.e., Lieder et al., 2009a)
27. Line 26, p. 69: Change “in myriad physiological” to “in a myriad of physiological”
28. Line 20, p. 70 and line 18, p. 78: Change Appendix G to Appendix F
29. Line 39, p. 70 and line 37, p. 78: Figure 2 is the conceptual model that summarizes data availability; might Figure 3 be the decision process to which the assessment refers?
30. Table 9, Footnote a, p. 73 and Table 12, Footnote a, p. 80 and Table F-1, footnote a, p. 122: a reference human body weight of 70 kg was used to calculate the DAF; why not use a BWh of 80 kg, described by the Agency in its GenX toxicity assessment as an update to the 70 kg default?
31. Table 12, p. 79: Similar to Table 9, consider placing information related to the selected POD in bold-faced type.
32. Table 12, Footnote b, p. 80: Convert animal body weight (BWa) to kg for DAF calculation (i.e., use 0.211 kg rather than 211 g).
33. Line 20, p. 86: change “are” to “is”
34. Line 11, p. 87: singularize “generations”
35. Line 12, p. 93 and line 9, p. 109: change “is” to “are”
36. Table F-1, p. 122: Similar to that done for the developmental POD, place information related to the subchronic and chronic PODs in bold-faced type.