United States Environmental Protection Agency | Office of Water 4304T | EPA-822-R-24-010 September 2024

EPA Response to External Peer Review of U.S. EPA's

"Draft Acute Freshwater Aquatic Life Benchmarks for Eight Data-Limited PFAS:

PFBA, PFHxA, PFNA, PFDA, PFBS, PFHxS, 8:2 FTUCA, and 7:3 FTCA"

U.S. Environmental Protection Agency
Office of Water, Office of Science and Technology
Health and Ecological Criteria Division
Ecological Risk Assessment Branch

September 2024

TABLE OF CONTENTS

I. INTRODUCTION

An independent contractor, Versar Global Solutions, for the Environmental Protection Agency (EPA), coordinated an external letter peer review of the *Derivation of Acute Protective Freshwater Benchmarks for Selected PFAS through a New Approach Method* report (January, 2024). The peer review was conducted for the EPA's Office of Water, Office of Science and Technology.

Background

The U.S. Environmental Protection Agency (EPA) Office of Water is charged with protecting human health and the environment from chemicals in water, under the purview of the Clean Water Act (CWA). In accordance with this mission, the EPA developed acute aquatic life freshwater benchmark values for eight PFAS: Perfluorobutanoic acid (PFBA), Perfluorobutanesulfonic acid (PFBS), Perfluorononanoic acid (PFNA), Perfluorodecanoic acid (PFDA), Perfluorohexanoic acid (PFHxA), Perfluorohexanesulfonic acid (PFHxS), Hexadecafluoro-2-decenoic acid (8:2 FTUCA), and Pentadecafluorodecanoic acid (7:3 FTCA). Aquatic life benchmarks, developed under 304(a)(2) of the CWA, are informational values that the EPA generates when there are limited high quality toxicity data available and data gaps exist for several families of aquatic organisms. The EPA developed acute benchmarks for these eight PFAS using available freshwater species empirical test data in conjunction with the application of a New Approach Method (NAM), specifically the EPA's Office of Research and Development's (ORD) peer-reviewed web-based Interspecies Correlation Estimate tool (Web-ICE; Version 4.0; https://www.epa.gov/webice/) (Raimondo et al. 2010). The EPA additionally investigated the approach described in Giddings et al. (2019) to determine whether a data binning approach based on similar chemical structure could be applied to the above PFAS to derive protective values for carboxylic acid PFAS and sulfonic acid PFAS. The EPA also conducted a comparison of its acute benchmark values to acute data and interim values used by the Department of Defense in developing ecological screening values (Grippo et al, 2021).

Versar conducted an independent search for scientific experts with expertise in one or more of the following disciplines: a) application of NAMs to the derivation of protective aquatic life benchmark values; b) toxicity of PFAS to aquatic life; c) aquatic ecotoxicology; and d) the acceptability of methods, statistical analyses and data interpretation applied to the determination of data and methods acceptability.

As a result of this search, the contractor identified and contacted 25 experts and received eight positive responses expressing interest and availability to participate. The remaining 17 experts were not available during the peer review timeframe or did not respond to the invitation. For each interested and available peer reviewer, the contractor evaluated their qualifications and conducted conflict of interest (COI) screening to ensure that the experts had no COI.

Versar selected the following five scientific experts to serve as peer reviewers:

David Buchwalter, Ph.D.
North Carolina State University

Anupama Kumar, Ph.D.
CSIRO Environment, Australia

Håkon Austad Langberg, Ph.D.
Norwegian Geotechnical Institute, Norway

Ryan Prosser, Ph.D.
University of Guelph, Canada

William Stubblefield, Ph.D.
Oregon State University

II. CHARGE TO PEER REVIEWERS

1) Please comment on the overall clarity of the documents and construction as it relates to assessing the effects and derivation of acute benchmarks for the eight selected PFAS.

2) Please comment on **each of the technical approaches** used to derive the draft benchmark values for the eight selected PFAS presented in EPA's ***Derivation of Acute Protective Freshwater Benchmarks for Selected PFAS through a New Approach Method*** (ICE-based approach using extrapolation [Sections 3-5], ICE-based approach using scaled data [Appendix F], and PFAS benchmark calculations using a data binning approach [Appendix G]).

    a. Are the technical approaches used to derive the benchmark values logical?
    b. Does the science support the conclusions?
    c. Are the approaches and resulting values consistent with the protection of aquatic life?
    d. For the ICE-based models in particular, please compare and contrast the strengths and weaknesses of the extrapolation approach [Sections 3-5] with those of the approach using scaled data [Appendix F].

3) Please comment on the empirical (direct test) and ICE-generated toxicity data used to derive the benchmark values presented in the draft document.

    a. Were the data adequately used and sufficiently comprehensive to represent risks to sensitive aquatic life?
    b. Were the data selected and/or excluded from the benchmark values derivation appropriately utilized?
    c. Are there relevant data that you are aware of that should be included? If so, please provide for derivation of benchmark values.

4) Are the derived benchmark values appropriately protective of sensitive aquatic life?

III. PEER REVIEWER COMMENTS TABLE

I. General Impressions		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
1	This report describes the process and results for acquiring draft acute recommended freshwater aquatic life benchmarks for eight PFAS: PFBA, PFHxA, PFNA, PFDA, PFBS, PFHxS, 7:3 FTCA, and 8:2 FTUCA. Acute benchmark values were derived using the procedure described in EPA's "Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses" where sensitivity distributions based on toxicity databases are used to derive acute freshwater criteria. Due to the lack of acceptable empirical data for constructing the sensitivity distributions, Interspecies Correlation Estimation (ICE) models were applied to complete the toxicity databases. The applied methods are transparent and the methods have undergone previous peer review. The approaches and data used are reported in a clear manner, making it possible to examine the whole process of acquiring the benchmarks.	Thank you for your comment.
1	Some clarification would, however, strengthen the report and make it less likely that readers will misunderstand what is the appropriate use of these benchmarks: 1) It should be more clearly expressed in the summary that these benchmarks are for acute exposure, i.e., to protect aquatic life in freshwater from acute toxic effects of PFAS. For example, these benchmarks are not sufficient to prevent pollution of the environment, adverse effects of chronic exposure to aquatic animals, human exposure, etc. This is important information as the problems with PFAS pollution are often due to the combination of their persistency, mobility, potential for bioaccumulation, and long-term toxicological effects.	Thank you for your comment. Text was added to further clarify that the values expressed are for acute exposure and effects, as well as to clarify that the derived values do not address potential chronic and/or bioaccumulative effects of these PFAS chemicals.
1	2) The assumptions for the use of the ICE models should be summarized in the main report.	The ICE model construct and assumptions associated with the ICE model application are already discussed in Sections 2 and 3, and in the cited literature documenting the application of the ICE models. The application of ICE models to the PFAS values does not inherently change or expand beyond these already-documented model assumptions.

I. General Impressions		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
1	3) The resulting benchmark values should be discussed in more detail. Especially trends for acute toxicity depending on PFAS group and chain length.	Thank you for your comment. Text was added to Section 5.10 to further discuss and compare the resulting benchmark values, including a discussion of the benchmark values in relation to PFAS group and chain length.
1	4) There is a lack of clear conclusions regarding the different methods explored for deriving the benchmarks.	Thank you for your comment. Text was added to Section 5.10 to further discuss and compare the resulting benchmark values.
2	Establishing environmental standards for the protection human health and the environment is critical. Unfortunately, our laws and regulatory approaches regarding the release of chemicals into the environment have allowed for the generation of thousands of chemicals used in commerce without the requirement that they be tested for safety. This leads us to our current situation where the pace of new chemistries being introduced to the world is far outpacing our ability to evaluate their toxicity to all forms of life including aquatic life. This document reflects this current state-of-affairs. PFAS as a chemical class are ubiquitous with new chemistries being introduced a rapid pace. Few environmental standards for their concentrations in surface waters have been established and toxicity data are limited. Methods to extrapolate toxicity data are unfortunately necessary in light of this data-limited situation. Here, the authors use a "New Approach Method" of Interspecies Correlation Estimation developed by Raimondo and others as a way of generating predictions of toxicity to aquatic organisms such that acute benchmarks for selected PFAS could be established based on the 1985 *Guidelines for the Protection of Aquatic Life*. As these compounds are generally not acutely toxic, and the benchmarks proposed in the document appear to be orders of magnitude higher than expected environmental concentrations, it is unclear what the rationale is for proposing these benchmarks in the absence of chronic benchmarks. By releasing these benchmarks as the only protective values available for the compounds in question, it is possible that the discovery of environmental concentrations that are well below these benchmarks (but could be chronically toxic) might not receive an appropriate response by states and tribes. I generally like the ICE approach to fill data gaps such that environmental standards can be set. It just does not feel like these acute standards are particularly relevant in light of the expected environmental concentrations.	Thank you for your comment. The acute toxicity concentrations of these compounds is generally expected to be higher than chronic, as is typical of an acute value. However, characterization of what is a protective acute concentration for chemicals of concern still remains important. Text was added to the document to further clarify the values provided are for acute effects and do not account for potential chronic and/or bioaccumulative effects. Importantly, the EPA is working towards approaches for the development of chronic values for data-limited chemicals, including PFAS, which we agree is important, but the NAMs tools to develop chronic benchmarks were not available to apply at this time. Additionally, there is a general lack of chronic data, especially for these PFAS, upon which to base chronic analyses. Additional research on the chronic toxicity of a range of PFAS would support development of additional values.

I. General Impressions		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
2	In general, the document requires the reader to consult some of the original literature on ICE models and there is not enough explanation of these approaches contained within the document itself. I would encourage the authors to add more technical information about how ICE models work, what their limitations are, and perhaps build in some uncertainty factors given the language in the Forward section of the document.	Thank you for your suggestion. Additional details have been added to the summary description of ICE models (Section 2), and the description of the application of ICE models to the derivation of PFAS benchmarks was expanded (Section 3). The underlying ICE models have been comprehensively documented in previous publications and are available on a public-facing website, and direct references to these resources have been provided in the document. Additional references have been added to the document to facilitate access to these materials; however, inclusion of these materials would greatly expand the size of this document, and in our opinion, would make it more difficult for the reader to identify and interpret the specific application and approach used for benchmark derivation.
3	The document is generally well-written and easy to follow, especially in the discussion of the data used and derivation of the benchmark. Greater detail and perhaps some examples would help in the discussion of the difference between “extrapolated” and “scaled” ICE-based data. It is not clear exactly what the difference is between these, or at least the implications of the two approaches other than to increase the number of species represented. Reviewing both the Raimondo et al (2010) and Willming et al (2016) papers did not address this topic. Raimondo et al (202?) is “in review” and perhaps will address the issues; however, the manuscript was not included in the review materials. The following discussion applies to the document and approach and was not specifically addressed in the charge questions.	Thank you for your comment. Text was added to the summary section to further differentiate between the extrapolated and scaled ICE-based models and to provide the reason for increased test acceptability with the scaled approach. The paper identified as “in review” has now been published and the full reference (Raimondo et al. 2024) has been added to the reference list. Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr

I. General Impressions		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
		Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884
3	To sum up the issue, the problem is that there is insufficient high-quality empirical data available to derive AWQC for the "selected" PFAS compounds. This issue is not new, data limitations in deriving Ambient Water Quality Criteria (AWQC) have been an issue since shortly after the implementation of the Clean Water Act (Kimerle et. al., 1985[1]) and became a greater concern with the reduction of AWQC data development at the EPA-ORD research laboratories. EPA previously proposed an approach to address this issue in the Type II standards methodology developed as part of the Great Lakes Initiative (GLI) published in 1995 (USEPA 40 CFR 9, 122, 123, 131, and 132, Final Water Quality Guidance for the Great Lakes System; Final Rule, March 23 1995). EPA presented a method to develop Secondary Maximum Concentrations (SMC) and the Secondary Continuous Concentrations (SCC) based on data sets that were insufficient to satisfy the eight minimum data requirements (MDR) to derive a national AWQC. Briefly, a secondary acute value (SAV) is calculated by dividing the lowest GMAV in the database by a Secondary Acute Factor (SAF) that is designated in Table A-1 in the document (ranging from 4.3 to 21.9) based on the number of satisfied MDRs available for the compound. This approach is somewhat crude and certainly lacks a great deal of technical basis; nonetheless, it probably should be discussed in the current document. In addition, application of the GLI technique to the PFAS compounds in this document has been previously conducted and presented in Grippo et al (2021)[2]. Resulting values between the Grippo et al (2021) report and application of the method to EPA's data result in slightly differing values, likely due to the acceptance and availability of different empirical data. In Table 1, a comparison of the benchmarks reported in this document (using the ICE approach) is made with the values calculated using the GLI approach (based on the empirical data reported in this document).	Thank you for the comment. Reference to the GLI methodology and comparison to values presented in Grippo et al. (2021) were added to the document summary. The EPA compares the two sets of values, the EPA's and Grippo et al. (2021)(Section 5.10) and discusses the use of ICE vs. assessment factors in developing benchmarks or screening values (Section 5.1.2). The EPA added additional discussion comparing these approaches in the final document, including text indicating that ICE uses PFAS data for the derivation of data to fill the data gaps, while the Grippo et al. (2021) approach uses assessment factors that did not incorporate data for PFAS. A further note is that the EPA's final values were revised to include consideration of new, low toxicity values from 2 species, the insect mayfly, *Neocloeon traingulifer* (Soucek et al. 2023) and the cladoceran *Moina micrura* (Razak et al. 2023), while the Grippo et al. (2021) publication did not include the consideration of these data for the derivation of protective values for these PFAS. The EPA final acute freshwater benchmark values

[1] *Kimerle, R. A., Werner, A. F., and Adams, W. J., "Aquatic Hazard Evaluation Principles Applied to the Development of Water Quality Criteria," Aquatic Toxicology and Hazard Assessment: Seventh Symposium, ASTMSTP854, R. D. Cardwell, R. Purdy, and R. C. Bahner, Eds., American Society for Testing and Materials, Philadelphia, 1985, pp. 538-547*

[2] *M. Grippo, J. Hayse, I. Hlohowskyj, and K. Picel. 2021. Derivation of PFAS Ecological Screening Values. Environmental Science Division. Argonne National Laboratory. September 2021.*

I. General Impressions

REVIEWER	REVIEWER COMMENT	EPA RESPONSE

Table 1. Comparison of benchmark values using the GLI and ICE-based derivation methods.

Chemical	Lowest empirical value	GLI factor	Tier II GLI calculated value (mg/L)	EPA Benchmark (Extrapolation) FAV/2 (mg/L)
Perfluorobutanoic acid (PFBA)	110	8	13.75	83
Perfluorobutanesulfonic acid (PFBS)	1938	13	149.1	183
Perfluorononanoic acid (PFNA)	27.84	13	2.14	10.3
Perfluorodecanoic acid (PFDA)	32	8	4	7.9
Perfluorohexanoic acid (PFHxA)	140	8	17.5	75
Perfluorohexanesulfonic acid (PFHxS)	22.5	13	1.7	9.1
Hexadecafluoro-2-decenoic acid (8:2 FTUCA)	3.2	13	0.24	0.58
Pentadecafluorodecanoic acid (7:3 FTCA)	0.959	13	0.074	0.18

are compared to the Grippo et al. (2021) values in Table 5-30 (Reproduced as Table 1, below).

Table 1. Comparison of Benchmark Outcomes with Water Quality-Based Ecological Screening Values Calculated by Grippo et al. (2021) Using the Great Lakes Initiative (GLI) Approach.

Chemical	EPA Recommended Acute Benchmark (mg/L)	Grippo et al. (2021) Calculated Tier II GLI Values (mg/L)	Factor Difference Between EPA Acute Benchmark and Tier II GLI Values
Carboxylic Acids			
PFBA	5.3	13.75	2.6
PFHxA	4.8	17.5	3.6
PFNA	0.65	2.14	3.3
PFDA	0.50	4	8
Sulfonic Acids			
PFBS	5.0	149.1	30
PFHxS	0.21	1.7	8

Grippo, M., J. Hayse, I. Hlohowskyj and K. Picel. 2021. Derivation of PFAS ecological screening values. Final. September 2021. Environmental Science Division, Argonne National Laboratory.

Razak, M.R., A.Z. Aris, A.H. Zainuddin, F.M. Yusoff, Z.N.B. Yusof, S.D. Kim, and K.W. Kim. 2023. Acute toxicity and risk assessment of

I. General Impressions		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
		perfluorooctanoic acid (PFOA) and perfluorooctanesulfonate (PFOS) in tropical cladocerans *Moina micrura*. Chemosphere 313: 9 p. Soucek, D.J., R.A. Dorman, E.L. Pulster, B.G. Perrotta, D.M. Walters and J.A. Steevens. 2023. Perfluorooctanesulfonate adversely affects a mayfly (*Neocloeon triangulifer*) at environmentally realistic concentrations. Environ. Sci. Technol. Lett. DOI: DOI: 10.1021/acs.estlett.3c00056 <3March2023>
3	Other comments not specifically addressed in the charge questions are provided below: • In at least three cases (i.e., PFBS, PFNA, and PFHxS) the derived criteria, did not comply with the MDR minimum of n=8, it appears that EPA disregarded the MDR minimum for these materials and calculated the benchmark with fewer MDRs, i.e., 7. This is not keeping with the 1985 guidance that states: "*Similarly. if all required data are not available, a numerical criterion should not be derived except in special cases."* This should be acknowledged in the text and some statement regarding the minimum number of MDRs to calculate a benchmark addressed.	The EPA is aware that missing the final minimum data requirement (MDR) could mean additional uncertainty with the derived value. However, these values are benchmarks, not criteria, and are expected to be less certain due to the limited empirical data on PFAS. Text was added to further note the uncertainties and discuss the specific MDRs for which data are missing. Further, the goal of the EPA here is to provide the best available scientific information for states, Tribes and others to consider in their water quality protection programs, rather than not providing any information to support environmental protection.
3	• Table 2 summarizes the MDRs available for each of the 8 PFAS compounds. In all cases 62.5 to 75% of the MDR data used in deriving the SSD-based benchmark was estimated using the ICE model method. Thus, most of the data used for MDRs and to develop SSDs, are estimated values. It is interesting that in the best case, only 1 empirical data point was available among the 4 most	Thank you for the comment. As noted, the entire purpose of developing benchmarks is to provide information to states, Tribes and other interested stakeholders on the toxicity of these chemicals given that there are very limited direct empirical

<table>
<tr><th colspan="3">I. General Impressions</th></tr>
<tr><th>REVIEWER</th><th>REVIEWER COMMENT</th><th>EPA RESPONSE</th></tr>
<tr><td></td><td>sensitive species for 50% of the materials. The other 4 materials had no empirical data represented among the most sensitive species. Although there may be good correlations between species making estimation for one species based on data from another possible, questions remain regarding the extent of the role that estimated values should play in the derivation of water quality criteria, standards, or benchmarks.
Table 2. Numbers of MDRs used in derivation of benchmarks
<table><tr><th>Chemical</th><th># of empirical MDRs</th><th># of estimated MDRs</th><th>Total MDRs met</th><th># quantitively accepted empirical data in 4 lowest species</th></tr><tr><td>Perfluorobutanoic acid (PFBA)</td><td>3</td><td>5</td><td>8</td><td>1</td></tr><tr><td>Perfluorobutanesulfonic acid (PFBS)</td><td>2</td><td>5</td><td>7</td><td>0</td></tr><tr><td>Perfluorononanoic acid (PFNA)</td><td>2</td><td>5</td><td>7</td><td>1</td></tr><tr><td>Perfluorodecanoic acid (PFDA)</td><td>3</td><td>5</td><td>8</td><td>0</td></tr><tr><td>Perfluorohexanoic acid (PFHxA)</td><td>3</td><td>5</td><td>8</td><td>1</td></tr><tr><td>Perfluorohexanesulfonic acid (PFHxS)</td><td>2</td><td>3</td><td>5</td><td>1</td></tr><tr><td>Hexadecafluoro-2-decenoic acid (8:2 FTUCA)</td><td>2</td><td>6</td><td>8</td><td>0</td></tr><tr><td>Pentadecafluorodecanoic acid (7:3 FTCA)</td><td>2</td><td>6</td><td>8</td><td>0</td></tr></table></td><td>test data, including data for the most sensitive test species.

After the draft was peer reviewed, the EPA developed application factors to account for two genera (the mayfly, Neocloeon and the cladoceran, Moina) that were indicated by direct empirical test data to have markedly greater sensitivity to PFOA and PFOS. These application factors were used to lower the PFAS values to account for these highly sensitive genera, for which PFAS toxicity data are not available. The pattern noted for the other direct empirical test data does not invalidate the outcome, but instead simply reflects the relative sensitivity of these highly sensitive species for which these direct test data were available. The EPA acknowledges there is greater uncertainty associated with the calculated benchmark values, and that the presence of direct empirical test data amongst the most sensitive test species would provide a greater degree of certainty. However, the absence of sensitive species test does not undermine the validity of the resulting benchmarks. These values were calculated based on methods presented in the peer-reviewed scientific literature and provide useful information that states and Tribes can choose to use for the protection of their surface waters.</td></tr>
<tr><td>3</td><td>• If EPA is going to revise the AWQC minimum MDR data requirement (i.e., 8) for the purpose of “Benchmark” derivation, then EPA should develop guidance regarding the minimum quantity</td><td>Thank you for your comment.
The EPA is not revising the AWQC MDRs for the purposes of benchmark derivation, but is using a</td></tr>
</table>

I. General Impressions		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
	and quality of **empirical** data required before a benchmark can be derived; this should potentially include: ○ a minimum number of empirical data required to be contained among the 4 lowest species (should benchmarks should be derived based solely on "estimated" values?). Is there a minimum percent of empirical to estimated data that should be met to establish a benchmark? ○ a requirement for empirical data with a representation among a base set of organisms that have historically been shown to be sensitive to a range of toxicants, e.g., *Ceriodaphnia dubia*, fathead minnow, would be useful. Kimerle et al (1985)[3] suggested that a minimum base data set composed of an algae, daphnid, and fish could consistently predict the most sensitive species based on available data at the time. Use of the ICE database could help to identify species consistently shown to be among those predicted to be most sensitive among chemical groups with a common mode-of-action. The table below, composed from the ICE-modeled data for PFAS compounds, suggests that a base data set composed of a freshwater mussel, cladoceran, and fish, would cover the majority of predicted most-sensitive species for PFAS compounds. ○ a minimum number of empirically derived MDR data points should be established, and a maximum number of ICE-estimated values that are allowed to be considered in an SSD should be established.	publicly available, data-rich, peer-reviewed modeling approach to fill in data gaps, in consonance with the EPA's NAMs workplan, which was created to support reduction of animal testing while continuing to protect human health and the environment, as noted by another peer reviewer. Further, it is noteworthy that web-ICE is based on empirical data for many species and substances. Although the presence of direct empirical test data for the chemical of concern amongst the most sensitive test species could potentially increase the level of certainty in the resulting values, the absence of a direct test data for a chemical within the four most sensitive test values does not invalidate the development of a benchmark values. Instead, it may simply reflect an absence of direct test data for the most sensitive test species for that chemical. Accordingly, the absence of direct empirical test data for a chemical amongst the most sensitive species does not invalidate the derivation of a benchmark value. The EPA has also been conducting a separate evaluation to characterize patterns of relative taxa sensitivity across chemicals. This ongoing evaluation of taxa relative sensitivity will further

[3] *Kimerle, R. A., Werner, A. F., and Adams, W. J., "Aquatic Hazard Evaluation Principles Applied to the Development of Water Quality Criteria," Aquatic Toxicology and Hazard Assessment: Seventh Symposium, ASTMSTP854, R. D. Cardwell, R. Purdy, and R. C. Bahner, Eds., American Society for Testing and Materials, Philadelphia, 1985, pp. 538-547.*

I. General Impressions

REVIEWER	REVIEWER COMMENT	EPA RESPONSE

Table 3. Comparison of the species sensitivity ranking from the ICE model for each of the PFAS compounds

Chemical	Species Sensitivity Rank from ICE model			
	1	2	3	4
Perfluorobutanoic acid (PFBA)	*Brachionus calyciflorus* (rotifer)	*Oncorhynchus mykiss* (rainbow trout)	*Gammarus fasciatus* (amphipod)	*Amblema plicata* (mussel)
Perfluorobutanesulfonic acid (PFBS)	*Amblema plicata* (mussel)	*Gammarus fasciatus* (amphipod)	*Hyalella azteca* (amphipod)	*Ceriodaphnia dubia* (cladoceran)
Perfluorononanoic acid (PFNA)	*Amblema plicata* (mussel)	*Chydorus sphaericus* (cladoceran)	*Megalonias nervosa* (mussel)	*Oncorhynchus mykiss* (rainbow trout)
Perfluorodecanoic acid (PFDA)	*Caecidotea brevicauda* (isopod)	*Micropterus salmoides* (bass)	*Perca flavescens* (yellow perch)	*Salvelinus fontinalis* (brook trout)
Perfluorohexanoic acid (PFHxA)	*Brachionus calyciflorus* (rotifer)	*Amblema plicata* (mussel)	*Gammarus fasciatus* (amphipod)	*Chydorus sphaericus* (cladoceran)
Perfluorohexanesulfonic acid (PFHxS)	*Danio rerio* (zebrafish)	*Jordanella floridae* (flagfish)	*Daphnia magna* (cladoceran)	*Limnodrilus hoffmeisteri* (oligochaete)
Hexadecafluoro-2-decenoic acid (8:2 FTUCA)	*Amblema plicata* (mussel)	*Palaemonetes kadiakensis* (grass shrimp)	*Chydorus sphaericus* (cladoceran)	*Megalonias nervosa* (mussel)
Pentadecafluorodecanoic acid (7:3 FTCA)	*Amblema plicata* (mussel)	*Macrobrachium nipponense* (river shrimp)	*Chydorus sphaericus* (cladoceran)	*Megalonias nervosa* (mussel)

inform data considerations for the derivation of future benchmark and criteria values.

In considering the peer reviewers comment that "Kimerle et al (1985)[4] suggested that a minimum base data set composed of an algae, daphnid, and fish". The EPA notes that an insect (the mayfly *Neocloeon triangulifer*) is by far the most acutely sensitive species for PFOS and among the most acutely sensitive for PFOA (Soucek at al 2023), while one cladoceran species (*Moina micrura*) was the most acutely sensitive species to PFOA and amongst the most sensitive species to PFOS (Razak et al 2023) and was much more sensitive than any other cladoceran species. Although data on these very sensitive species were not available for the 8 chemicals for which these acute PFAS benchmarks were developed, the sensitivities of these taxa were accounted for through the use of an application factor that was derived following the peer review. The calculated PFAS values were divided by the application factor, to lower the benchmark values and account for these highly sensitive species. Additionally, although in most cases not falling amongst the four most sensitive species, empirical toxicity text data were available for daphnids for seven of the eight chemicals evaluated (PFBA, PFHxA, PFNA, PFDA, PFBS,

[4] *Kimerle, R. A., Werner, A. F., and Adams, W. J., "Aquatic Hazard Evaluation Principles Applied to the Development of Water Quality Criteria," Aquatic Toxicology and Hazard Assessment: Seventh Symposium, ASTMSTP854, R. D. Cardwell, R. Purdy, and R. C. Bahner, Eds., American Society for Testing and Materials, Philadelphia, 1985, pp. 538-547.*

I. General Impressions		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
		8:2 FTUCA, 7:3 FTCA) and for fish for six of the eight chemicals evaluated (PFBA, PFDA, PFBS, PFHxS, 8:2 FTUCA, 7:3 FTCA), in addition to web-ICE models more broadly representing these taxa. For these analyses, the EPA used the most current data set and models available. The EPA considers it useful to present the best available science, using publicly available, peer-reviewed models, to develop benchmarks as information for states, Tribes, stakeholders and the public to consider, instead of providing no information whatsoever. Soucek, D.J., R.A. Dorman, E.L. Pulster, B.G. Perrotta, D.M. Walters and J.A. Steevens. 2023. Perfluorooctanesulfonate adversely affects a mayfly (*Neocloeon triangulifer*) at environmentally realistic concentrations. Environ. Sci. Technol. Lett. DOI: DOI: 10.1021/acs.estlett.3c00056 <3March2023>. Razak, M.R., A.Z. Aris, A.H. Zainuddin, F.M. Yusoff, Z.N.B. Yusof, S.D. Kim, and K.W. Kim. 2023. Acute toxicity and risk assessment of perfluorooctanoic acid (PFOA) and perfluorooctanesulfonate (PFOS) in tropical cladocerans *Moina micrura*. Chemosphere 313: 9 p.

I. General Impressions		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
3	• If EPA chooses to develop criteria/benchmarks for materials that have limited empirical data, then a two-tiered approach should be adopted much like that previously adopted by EPA for the GLI program. It is unclear from the document what the long-term intent and regulatory status will be for "benchmarks." Those materials that do not have sufficient empirical data to permit derivation of a "Tier 1" criteria, could be addressed by a "Tier II" benchmark, as suggested. The Tier II benchmark could be derived using the proposed ICE-based methodology or a method like that used for the GLI Secondary Acute Values. That said, will benchmarks serve the same purpose as the current AWQC? They do not have the same scientific basis as "Tier I" AWQC, but if adopted as "standards" by states and tribes, they will have the same regulatory/legal status. This is briefly addressed in the document's forward; however, greater clarity regarding the "scientific confidence" and "regulatory validity" could be provided. Questions regarding use of the approach should be considered, for example, if a state developed a proposed standard for a chemical based on limited empirical data and relying predominately on ICE-estimated data (e.g., 7:3 FTCA), would EPA approve it?	The EPA is not developing a new two-tiered approach for criteria. Benchmarks were derived for these chemicals instead of criteria because there were insufficient data to develop criteria, recognizing that the resulting values have a greater degree of uncertainty than criteria, as stated in the document. Text has been added to the document to further clarify the difference between aquatic life criteria and benchmarks. The EPA is providing these benchmark values as a source of information that states can choose to consider for use in the protection of their surface waters. The EPA also provides 304(a) ambient water quality criteria as recommendations for states and Tribes to consider in their water quality protection programs. Additionally, states and Tribes can develop their own scientifically defensible values and/or develop site-specific values and submit them for approval as state water quality standards. The EPA considers all values states and Tribes develop and propose for use in their water quality standards.
4	The draft *Derivation of Acute Protective Freshwater Benchmarks for Selected PFAS Compounds through a New Approach Method (NAM)* was well organized and well written. The accuracy of each element of the derivation process was satisfactory. The structure and writing of the draft document clearly communicated the rationale and the process of deriving the benchmarks. I think that the proposed benchmarks are reasonable and protective of aquatic life based on the acceptable empirical acute toxicity data available for the eight PFAS. The major source of uncertainty for the derived benchmarks is the lack of acceptable empirical acute toxicity data on freshwater primary producers and freshwater invertebrates for the eight PFAS. I think the "New Approach Method" is reasonable when there is a lack	Thank you for your comment. The use of a NAM in calculating protective values is not intended to replace existing empirical data but is instead intended to supplement the dataset when limited direct test data are available and incorporate the consideration of new approaches going beyond those developed in EPA's Aquatic Life Criteria Guidelines. The EPA's NAMs workplan supports

I. General Impressions		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
	of acceptable empirical data, and a benchmark needs to be derived. However, the NAM should not replace the derivation of benchmarks with empirical data.	reduction of animal testing while continuing to protect human health and the environment.
5	This document provides draft Acute Protective Freshwater Benchmarks for the following eight PFAS Compounds using New Approach Method (NAM): 1. Perfluorobutanoic acid (PFBA) 2. Perfluorobutanesulfonic acid (PFBS) 3. Perfluorononanoic acid (PFNA) 4. Perfluorodecanoic acid (PFDA) 5. Perfluorohexanoic acid (PFHxA) 6. Perfluorohexanesulfonic acid (PFHxS) 7. Hexadecafluoro-2-decenoic acid (8:2 FTUCA), and 8. Pentadecafluorodecanoic acid (7:3 FTCA) The detailed methodology used for the derivation of benchmarks has been thoroughly explained. The process used and results of a systematic review of available empirical toxicity data for aquatic organisms identified via EPA's literature search for the eight PFAS has been adequately addressed.	Thank you for your comment.
5	The aquatic life benchmarks for the eight PFAS compounds have been developed using the empirical and Web-ICE data for these chemicals and were calculated by applying statistical methods. This method aligns with the EPA's objective to decrease reliance on animal testing by employing NAMS in toxicity assessment. In addition, the EPA applied 'binning' approach to calculate protective benchmark values for six PFAS, utilizing combined carboxylic acid (PFBA, PFNA, PFDA, PFHxA) and sulfonic acid (PFBS, PFHxS) groupings to facilitate value derivation.	Thank you for your comment.
5	The use of estimated data suggests a proactive approach in addressing gaps in empirical data. It also demonstrates agencies' commitment to methodological rigor and adaptability in the face of data challenges when deriving protective values for PFAS compounds.	Thank you for your comment.

II. Response to Charge Questions

Charge Question 1: ***Please comment on the overall clarity of the documents and construction as it relates to assessing the effects and derivation of acute benchmarks for the eight selected PFAS.***

REVIEWER	REVIEWER COMMENT	EPA RESPONSE
1	The overall methodology and data used is clearly presented. The conclusions for the explored approaches, i.e., the use of extrapolation versus the "scaled" approach and the use of the binning approach should be made clearer. These conclusions should be included in the summary.	Thank you for your comment. Text was added to the summary (Section 5.10) to further discuss attributes and conclusions made about each approach.
1	The report would be strengthened by including an introduction to PFAS and its uses (and emissions). The brief introduction in the summary is very good, however a version with some more details in the main report would be good to include. Further, I would encourage the authors to include a brief discussion on the trends for the benchmarks for the different PFAS (including PFOA and PFOS) compared to the scientific literature (i.e., trends for toxicity depending on group and chain length).	Thank you for your comment. A discussion of PFAS (including PFOA and PFOS) toxicity based on chain length was added to the document. Additionally, references to the final PFOA and PFOS criteria documents were added. These documents provide detailed discussions of PFAS sources, fate, and transport in the environment.
2	It would be useful to better explain what a Benchmark is related to a Water Quality Criterion. The forward states that benchmarks are "less certain than Water Quality Criteria", but the reader should also be informed about enforcement differences between benchmarks and WQC.	Thank you for your suggestion. Text has been added to the document to further define benchmarks and their intended application. Recognizing there is greater uncertainty associated with these values, as indicated by the "benchmark" designation, the EPA provides these values as a source of information. Benchmarks, provide information that states and Tribes can choose to use for the protection of their surface waters.

Charge Question 1: ***Please comment on the overall clarity of the documents and construction as it relates to assessing the effects and derivation of acute benchmarks for the eight selected PFAS.***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
3	In general, the document is clear and well-organized. The sections of the document follow a "template" making the subsections parallel to each other making it easy to compare. As previously suggested, some additional detail regarding the extrapolation and scaled estimation technique would be helpful. This may be contained in the Raimondo et al (in review) document but it was not provided for this review. Also, a brief discussion of the "binning" technique (Giddings et al 2019) was made in Section 5.10 and in Appendix G and a comparison of the benchmarks derived using the binning vs ICE-based techniques is provided, but discussion or assessment of the utility/advantages/disadvantages of the technique is not provided.	The Raimondo et al. paper identified as "in review" has now been published and the full reference (Raimondo et al. 2024) has been added to the reference list. Text was also added to the summary (Section 5.10) to further discuss attributes and conclusions made about each of the approaches. Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884
4	I thought the overall clarity of the writing and construction of the document were good. I found the structure logical and easy to follow in the progression of the derivation process. I would not recommend any changes to the overall writing or structure of the document. There were a few spelling and grammatical errors but all very minor.	Thank you for your comment.
5	Great job on the overall structure and organization of the document! The logical flow and seamless transitions between sections significantly enhance the readability and understanding of the content. Information has been laid out in appendices with detailed information on the approaches and examples for deriving benchmarks. The list of Tables and Figures provides information on all the empirical data and acceptable ICE models used in deriving aquatic life acute benchmarks of all eight compounds. The lowest quantitatively acceptable empirical toxicity studies used to derive aquatic life benchmarks for eight PFAS compounds were detailed in the appendix. Ranked GMAVs and FAVs have been provided for all eight PFAS compounds. Data incorporated in SSDs have been listed and all figures are self-explanatory.	Thank you for your comment.
5	It is apparent that considerable thought and effort were invested in crafting a document with a well-considered and smooth progression.	Thank you for your comment.

Charge Question 2: ***Please comment on each of the technical approaches used to derive the draft benchmark values for the eight selected PFAS presented in EPA's Derivation of Acute Protective Freshwater Benchmarks for Selected PFAS through a New Approach Method (ICE-based approach using extrapolation [Sections 3-5], ICE-based approach using scaled data [Appendix F], and PFAS benchmark calculations using a data binning approach [Appendix G]).***

2.a. Are the technical approaches used to derive the benchmark values logical?		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
1	All these 3 methods are logical. However, I am missing a discussion comparing these methods. It is mentioned that the "scaled" approach is undergoing evaluation by Raimondo et al. (in review), however no information from this work is reported. The lack of a discussion and clear conclusions for the comparisons of these methods makes it difficult for the reader.	The Raimondo paper identified as "in review" has now been published and the full reference (Raimondo et al. 2024) has been added to the reference list. Text was also added to the summary to further discuss attributes and conclusions made about each of the approaches. Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884 Text was also added to the summary (Section 5.10) to further discuss attributes and conclusions made about each of the approaches.
2	I think the reader does not have enough information to evaluate ICE-based approaches without consulting the original literature. The introduction to Web-ICE (p. 17) is remarkably brief in explaining the technical approach and this section could be expanded significantly. It is helpful that the reader is given references to read that point to the successful applications of the approach, but this document should be self-contained with respect to describing the technical approach in detail.	Thank you for your comment. Because the underlying approaches represent publicly-available published methods, the EPA has chosen to focus the document on describing the specific application of these methods, rather than restating a description of the tools being used for the evaluation. References were provided to facilitate access to these documents.
2	The same criticism can be applied to the scaled data and data binning approaches. These technical approaches should be explained in more detail particularly in terms of the mechanisms by which these different approaches could yield different toxicity estimates.	Text was added to the summary (Section 5.10) to further discuss attributes and conclusions made about each of the approaches. However, as noted above, because the underlying approaches represent publicly-available published methods, the EPA has chosen to focus the document on describing the specific application of these methods, rather than restating a description of the tools used for the evaluation.

2.a. Are the technical approaches used to derive the benchmark values logical?		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
2	In general, I favor action on the creation of environmental standards, even when data are limiting. The approach of making toxicity predictions is logical and the process follows the 1985 Guidelines. However, these guidelines are in need of modernization, and it is unclear to me how aggregated, individual species toxicity tests (that ignore dietary exposure pathways and species interactions) provide compelling evidence for protecting aquatic communities in nature.	The EPA is continuing work on revising EPA's Aquatic Life Criteria Guidelines. Consideration of NAMs-based approaches represents a component of this revision process. This document focuses on the consideration of short-term exposures based on the derivation of acute toxicity values. Dietary exposure pathways and species interactions are both more relevant to chronic exposures.
3	The approaches used by the three methods are logical and creative methods to address the issue of data limitations. Although the calculated data are provided in the document and in Appendix F, it is difficult for this reviewer to fully understand the technical differences in the ICE-based approach using extrapolation or scaled data. Perhaps hands on evaluation of the models or review of the Raimondo et al (in review) manuscript would help. At the least, an example showing calculations both ways would help the reader.	Figure 3-1 and supporting text were added to the benchmarks document (Section 3) to depict side-by-side the extrapolation and scaled approaches based on a selected example. This example figure is shown in Appendix A of this document. The Raimondo paper identified as "in review" has now been published and the full reference (Raimondo et al. 2024) has been added to the reference list. Text was also added to the summary (Section 5.10) to further discuss attributes, conclusions made, and differences between each of the approaches. Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884
3	PFAS benchmark calculations using the data binning approach presented in Giddings et al (2019) are logical and provide a method to expand the quantity of empirical data provided that the assumption of a common mode of action (MOA) is valid. Raimondo et al (2010) notes that MOA-specific models are more robust and improve the fit of the ICE model approach.	Thank you for your comment.
3	Approaching the problem of missing data using the ICE model and binning techniques are more elegant than the previous GLI technique and are more generally scientifically supportable.	Thank you for your comment.
4	I think the technical approaches taken to derive the benchmark values were logical. In the absence of acceptable empirical data, the use of the ICE models to generate a data-	Thank you for your comment. Correct, the intent would be to use the acceptable direct test data for each chemical, as available. In addition to

2.a. Are the technical approaches used to derive the benchmark values logical?		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
	rich SSD is a logical approach. Obviously, the benchmarks could be re-evaluated if and when acceptable empirical data is available to regulators.	the use of ICE models, the EPA has since developed application factors to account for two genera (the mayfly, *Neocloeon* and the cladoceran, *Moina*) that were indicated by direct empirical test data on PFOA and PFOS to have markedly greater sensitivity. These application factors were used to lower the PFAS values to account for these highly sensitive genera, for which PFAS toxicity data for these eight benchmark chemicals are not available.
5	The methodologies employed to determine the benchmark values are rational and sound from a technical perspective. The available empirical data for the eight PFAS under consideration fulfill only 2-3 Minimum Data Requirements (MDRs). Consequently, the EPA opted to employ the peer-reviewed Interspecies Correlation Estimation (ICE) models developed by Raimondo et al. in 2010. The primary objective of this application was to provide acute toxicity data to fulfill MDRs in instances where direct toxicity data were not at hand. The ICE models underwent rigorous evaluation based on acceptance parameters, including mean square error (MSE), R^2, and slope, as delineated in Box 1. Only models meeting these predefined acceptance criteria were utilized in the derivation of species-specific toxicity data. This data, when integrated with empirical toxicity data, served to strengthen the process of establishing benchmark values.	Thank you for your comment. Following the peer review, model selection has been further refined. Only ICE models based on freshwater species were utilized. This is based on information indicating that saltwater species may differ in sensitivity to PFAS from freshwater species because of their ionic composition. Additionally, when more than one ICE model was available for a predicted species, only the model derived using the species with the closest taxonomic relationship was used, based on the greater accuracy of prediction of models with closer taxonomic relationships. The EPA considers these two improvements to further support the scientific rogor of the benchmarks.

2.a. Are the technical approaches used to derive the benchmark values logical?		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
5	The EPA's investigation into the "binning" approach for establishing protective values for grouped carboxylic acid PFASs and grouped sulfonic acid PFAS is grounded in the precedent established by Giddings et al. in 2019. A similar methodology proved successful for pyrethroids in that study. This strategic approach involved consolidating chemicals with shared modes of action, offering advantages in scenarios where limitations in available data present challenges to value determination. The calculated values were based on the amalgamation of carboxylic acid compounds (PFBA, PFNA, PFDA, PFHxA) and sulfonic acid compounds (PFBS, PFHxS), thereby substantiating the derivation of these values. Calculated benchmark values for carboxylic acids and sulfonic acids consistently demonstrated higher values when utilizing the SSD generator in comparison to the Guidelines-based approach. However, these benchmarks displayed variability in magnitude when contrasted with the ICE-based benchmark values. This might be influenced by the constrained empirical datasets for certain PFAS and the restricted number of data points available for the species employed in normalization.	Thank you for your comment. We concur, most notably, the “binning” approach incorporated the much larger empirical datasets provided by the inclusion of PFOA and PFOS than were available for the other PFAS compounds, which were evaluated using the ICE-based approaches. This is likely to have influenced the outcome of the evaluation, as noted by the reviewer. The EPA also updated the acute benchmarks to reflect improvements in use of the ICE models (see response immediately above) and also the use of information indicting expected high sensitivity of two invertebrate species to these eight data-limited PFAS, based on empirical data for these sensitive species (*Neocloeon triangulifer* and *Moina micrura*) with PFOA and PFOS. These refinements to the acute benchmarks following peer review resulted in the benchmarks being much closer in magnitude to the calculated binned values using EPA’s Aquatic Life Criteria Guidelines-based calculations. Table 2, provided below and which is also included in the revised document, shows a comparison of the benchmark outcomes compared with values that EPA calculated using the approach by Giddings et al. (2019), using both EPA’s Aquatic Life Criteria Guidelines and SSD Generator-based calculation methods.

2.a. Are the technical approaches used to derive the benchmark values logical?

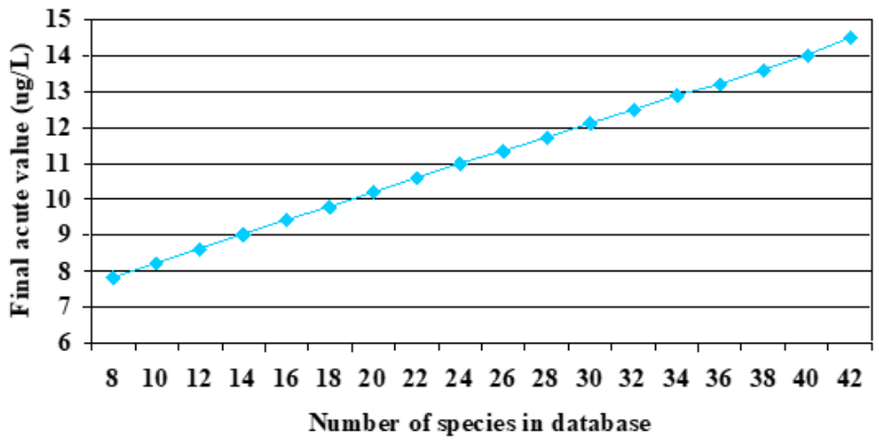
REVIEWER	REVIEWER COMMENT	EPA RESPONSE
		Table 2. Comparison of Benchmark Outcomes with Acute Benchmarks Calculated Based on Approach by Giddings et al. (2019) (Aquatic Life Criteria Guidelines and SSD Generator-Based Values).

Chemical	EPA Recommended Acute Benchmark (mg/L)	Acute Benchmarks Calculated Using Binning Approaches		Factor Difference Between Acute Benchmark and Binning-based Benchmark Values	
		Aquatic Life Criteria Guidelines-Based Values (mg/L)	SSD Generator-Based Values (mg/L)	Aquatic Life Criteria Guidelines-Based Values	SSD Generator-Based Values
Carboxylic Acids					
PFBA	5.3	19	110	3.6	21
PFHxA	4.8	4.3	24	0.9	5.1
PFNA	0.65	0.35	2.0	0.5	3.0
PFDA	0.50	0.49	2.8	1.0	5.6
Sulfonic Acids					
PFBS	5.0	15	75	3.0	15
PFHxS	0.21	0.11	0.56	0.5	2.7

Giddings, J.M., J. Wirtz, D. Campana and M. Dobbs. 2019. Derivation of combined species sensitivity distributions for acute toxicity of pyrethroids to aquatic animals. Ecotoxicol. 28: 242-250.

5	By stating that the derived benchmarks are considered less certain than ambient water quality criteria, the authors acknowledge a level of uncertainty. This acknowledgment is crucial in providing a realistic assessment of the reliability of the benchmarks.	Thank you for your comment.

2.b. Does the science support the conclusions?		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
1	The applied methods have undergone peer review in previous publications. Previous testing of ICE model performance indicate validity for the assumptions of this approach: 1) that the relationship of inherent sensitivity between two species is conserved across chemicals, mechanisms of action, and ranges of toxicity; and 2) that the nature of a contaminant that was tested on the surrogate reflects the nature of the contaminant in the predicted species.	Thank you for your comment.
1	A deeper scientific understanding would require in-depth knowledge of the toxicological mechanisms. Modes of action (MOA) specific models have previously been reported to be more robust. However, such mechanistic understanding, including data for MOA specific toxic effects, is not available. Therefore, the applied methodology can be considered as the best available option.	Thank you for your comment.
2	The science tells us that these compounds are not acutely toxic and the benchmark values could give states and tribes a false sense of safety if they encounter high concentrations that are below the benchmark. The language in the document is explicit about the application as a 1 hour maximum every 3 years, but in the absence of more environmentally relevant standards, I'm not sure what these benchmarks do for environmental protection.	Thank you for your comment. The acute toxicity concentrations of these compounds is expected to be higher than chronic, as is typically the case for acute vs chronic toxicity. However, characterization of what is a protective acute concentration for chemicals still remains important, especially in the context of PFAS releases. Text was added to the document to further clarify the values provided are for acute effects and do not account for potential chronic effects and/or bioaccumulation. The EPA is working towards approaches for the development of chronic values for such data-limited chemicals, but the NAMs tools to develop chronic benchmarks were not available to apply at this time. Additionally, there is a general lack of chronic data, especially for these PFAS, upon which to base chronic analyses. Additional research on the chronic toxicity of a range of PFAS would support development of additional values.

2.b. Does the science support the conclusions?		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
3	I am not sure what is being asked since there are no specific conclusions called out in the document. The approach of estimating species sensitivities to toxins is logical and scientifically supportable, provided that sufficient underpinning data are available to support the models. Derivation of 1985 AWQC compliant values (FAV/FCVs) are dependent on 5 datapoints; the four lowest Genus Mean Values (GMAV/GMCV) and the total number of species represented in the database. The calculation is more sensitive to the relationship of the 4 lowest values than to the total number of species represented. As an example, Figure 1 provides an example of increasing the size of the database for a compound (related to the number of available ICE models); increasing from 8 GMAVs to 42 results in a slightly less than a doubling in the calculated FAV (8 to 14.5). So, the choice of extrapolated vs scaled ICE models may result in a slight increase in the calculated benchmark due to acceptance of more models. Far more important is the validity and relationship of the 4 lowest GMAVs. It is critical that these values be valid and as accurate as possible. Relying on estimated values can introduce a large degree of uncertainty in the resulting benchmark value.  *Figure 1. Final Acute and Chronic values are sensitive to the number of species in the database*	In this case, the conclusions are the benchmark values. As noted by the reviewer, the lowest four values play a significant role in determining the final benchmark value. As previously discussed, since the peer review the EPA has developed and included the use of application factors to account for two genera (the mayfly, *Neocloeon* and the cladoceran, *Moina*) that were indicated by recent direct empirical test data to have much greater sensitivity to PFOA and PFOS. These application factors were used to lower the benchmarks for the 8 data-limited PFAS to account for these particularly sensitive genera, for which toxicity data are not available.
4	Overall, yes, the conclusions are supported by the available science.	Thank you for your comment.

2.b. Does the science support the conclusions?		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
4	An important element of the science supporting the conclusions is the validation of the ICE models to predict the acute toxicity of PFAS. In the benchmark document on page 17, an unpublished work by Raimondo et al. is cited to support the validation of the ICE models to predict acute toxicity of PFAS. The documents states, "*ICE models have been developed from a broad range of chemicals (e.g., metals and other inorganics, pesticides, solvents, and reactive chemicals) and across a wide range of toxicity values and have been validated as accurate predictors of PFAS acute toxicity when model criteria parameters are followed (Raimondo et al., in review).*" As the unpublished manuscript by Raimondo et al. is not available as part of this review, I am left to assume that the statement made in the benchmark document about the validation of the ICE models to predict the acute toxicity of PFAS to be accurate.	The paper identified as "in review" has now been published and the full reference (Raimondo et al. 2024) has been added to the reference list. Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884
4	Another question is whether the statement "…*have been validated as accurate predictors of PFAS acute toxicity…"* is solely based on data with PFOS and PFOA?	This statement is referring to the validation of ICE models across a broad range of chemicals.
5	This is a valuable contribution to the current scientific knowledge on the toxicity of PFAS compounds to aquatic life, even while acknowledging the inherent uncertainties associated with using estimated data in the derivation process.	Thank you for your comment.

2.c. Are the approaches and resulting values consistent with the protection of aquatic life?		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
1	In my opinion, the approaches and resulting values are protective for acute toxic effects on aquatic life as defined in the report (i.e., to be protective of 95% of freshwater genera potentially exposed to these specific PFAS under short-term conditions of one-hour duration, if the one-hour average magnitude is not exceeded more than once in three years).	Thank you for your comment.
1	However, it is important to note that the most problematic properties of PFAS are not the acute toxicological effects but rather the combination of their persistency, mobility, potential for bioaccumulation (for some PFAS), and long-term toxicological effects. Emissions of PFAS resulting in recipient concentrations comparable to the threshold values in this report may (depending on the volume of water in the recipient) be substantial sources of PFAS pollution to the environment. As many PFAS (PFCA and PFSA) are both persistent and relatively mobile, such emissions may result in problematic pollution of drinking water and wildlife with adverse long-term consequences for both human health and the environment.	Thank you for your comment. The EPA agrees with the reviewer. However, characterization of what is a protective acute concentration for chemicals of concern still remains important. Text was added to the document to further clarify the values provided are for acute effects and do not account for potential chronic effects and/or bioaccumulation. The EPA is working towards approaches for the development of chronic values for data-limited chemicals. but the NAMs tools to develop chronic benchmarks were not available to apply at this time. There is a general lack of chronic data, especially for these PFAS, upon which to base chronic analyses. Additional research on the chronic toxicity of a range of PFAS would support development of additional values.
2	See comments above. Adhering to an outdated (1985) understanding of toxicology and species sensitivity differences remains an unfortunate state-of-affairs at EPA.	The EPA disagrees with the commenter's vague and negative characterization of the current work. The NAMs-based approach presented in this document, although incorporating elements of the EPA's Aquatic Life Criteria Guidelines (e.g., using sensitivity distributions and fulfilling MDRs), represents the application of an innovative approach to addressing data-limited chemicals. The approach used in this evaluation does not strictly "adhere" to the EPA's Aquatic Life Criteria Guidelines, as doing so would have meant that none of the protective aquatic life values presented within the document could have been derived.

2.c. Are the approaches and resulting values consistent with the protection of aquatic life?		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
		The use of empirical data on chemical toxicity, consideration of intra and interspecies (and genus) sensitivities, and use of new scientific tools to leverage data are all aspects of this document's analysis and are also common and current approaches to evaluating the effects of chemicals on both ecological species and humans that is applied across the globe. The understanding of the toxicological activity of PFAS is an ongoing international research endeavor, both for human and ecological health. As previously discussed, the EPA has also been conducting a separate evaluation to characterize patterns of relative taxa sensitivity across chemicals. This ongoing evaluation of taxa relative sensitivity will further inform data considerations for the derivation of benchmark and criteria values.
3	The table below summarizes the benchmarks calculated using the various methods described in the reviewed document and the values calculated using the GLI method. Values vary substantially, in some cases as much as an order of magnitude. The Tier II GLI value frequently provided the lowest calculated value; however, the values were not inconsistent with the other methods. Given the minimal amount of empirical data available for these materials, it is difficult to identify if the values are *"consistent with the protection of aquatic life."* Perhaps conducting an analysis with a data rich compound (e.g., copper or a pesticide), using only a limited portion of the available data followed by a comparison to the full AWQC database would give some insight into the comparability of the benchmark and the standard AWQC approach.	Thank you for your comment. The values presented in the reviewer's comment have been updated in the revised document and a summary of these revised values is presented in the table below. The values do vary between those derived with the benchmark extrapolation approach and the other approaches (up to a factor difference of approximately 13 for all chemicals except for PFBS using the Tier II approach, which differed by a factor of approximately 30); however, most of the differences were less than a factor of 10. Text has been added to the summary section (Section 5.10) to compare the values derived using the ICE-based, binning, and Tier II GLI approaches and to discuss in greater detail

2.c. Are the approaches and resulting values consistent with the protection of aquatic life?		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
	Table 5. Summary of calculated benchmark values Chemical | EPA Benchmark (Extrapolation) (mg/L) | EPA Benchmark (Scaled) (mg/L) | Binning approach (Guidelines-based) (mg/L) | Binning approach (SSD-based) (mg/L) | Tier II GLI calculated value (mg/L) Perfluorobutanoic acid (PFBA) | 83 | 174 | 194 | 467 | 13.75 Perfluorobutanesulfonic acid (PFBS) | 183 | 237 | 24 | 102 | 149.1 Perfluorononanoic acid (PFNA) | 10.3 | 12 | 3.4 | 8.3 | 2.14 Perfluorodecanoic acid (PFDA) | 7.9 | 10 | 4.9 | 12 | 4 Perfluorohexanoic acid (PFHxA) | 75 | 95 | 43 | 103 | 17.5 Perfluorohexanesulfonic acid (PFHxS) | 9.1 | 9.4 | 0.18 | 0.76 | 1.7 Hexadecafluoro-2-decenoic acid (8:2 FTUCA) | 0.58 | 0.65 | | | 0.24 Pentadecafluorodecanoic acid (7:3 FTCA) | 0.18 | 0.23 | | | 0.074	the differences between these values. As noted, values did vary between the approaches used. However, all the methodologies tested rely on relatively robust *underlying empirical* datasets for their development. The ICE models rely on underlying empirical test data used to derive each model, the binning approach bases value development on direct empirical datapoints for PFAS that were grouped to determine the chemical-specific benchmarks, while the Tier II approach is directly based on available empirical data. The approach used to derive the ICE-based benchmark values additionally used empirical data for mayfly (*Neocloeon triangulifer*) and a cladoceran (*Moina micrura*) to account for these two highly sensitive species that were not otherwise represented in the underlying direct empirical or ICE datasets. All approaches, however, do rely on the extrapolation of these empirical data to derive the benchmark values, which is likely to be leading to some of the differences observed for these values. Consistent with the suggestion by the reviewer, the EPA has been exploring the evaluation of these approaches using chemicals with more data-rich direct toxicity datasets.

2.c. Are the approaches and resulting values consistent with the protection of aquatic life?		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
		<table><tr><th>Chemical</th><th>EPA Benchmark (Extrapolation; mg/L)</th><th>EPA Benchmark (Scaled; mg/L)</th><th>Binning Approach (Guidelines-based; mg/L)</th><th>Binning Approach (SSD-based; mg/L)</th><th>Tier II GLI Value (mg/L)</th></tr><tr><td>PFBA</td><td>5.3</td><td>11</td><td>19</td><td>110</td><td>13.75</td></tr><tr><td>PFHxA</td><td>4.8</td><td>6.0</td><td>4.3</td><td>24</td><td>17.5</td></tr><tr><td>PFNA</td><td>0.65</td><td>0.73</td><td>0.35</td><td>2.0</td><td>2.14</td></tr><tr><td>PFDA</td><td>0.50</td><td>0.65</td><td>0.49</td><td>2.8</td><td>4</td></tr><tr><td>PFBS</td><td>5.0</td><td>7.6</td><td>15</td><td>75</td><td>149.1</td></tr><tr><td>PFHxS</td><td>0.21</td><td>0.21</td><td>0.11</td><td>0.56</td><td>1.7</td></tr><tr><td>8:2 FTUCA</td><td>0.037</td><td>0.041</td><td>-</td><td>-</td><td>-</td></tr><tr><td>7:3 FTCA</td><td>0.012</td><td>0.015</td><td>-</td><td>-</td><td>-</td></tr></table>
4	Yes, I think the approaches and resulting values are consistent with the protection of aquatic life based on the acceptable empirical data that was available to the assessors.	Thank you for your comment.
5	Yes, the approaches and resulting values align with the protection of aquatic life. The aquatic life benchmarks for the eight PFAS compounds were established using empirical and Web-ICE data, employing statistical methods for calculation. This approach aligns with the EPA's goal to reduce reliance on animal testing by incorporating NAMS in toxicity assessment.	Thank you for your comment.
5	Detailed response as in 2a.	Thank you for your comment.
5	*Limitation* ICE models have not been developed for chronic toxicity data and therefore only acute criteria were developed.	Thank you for your comment. That is correct that ICE models have been developed to address acute toxicity only. The EPA is pursuing similar NAMs-based approaches for the derivation of chronic values.

2.d. For the ICE-based models in particular, please compare and contrast the strengths and weaknesses of the extrapolation approach [Sections 3-5] with those of the approach using scaled data [Appendix F].		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
1	The use of extrapolation beyond the model range may result in large confidence intervals and hence, many ICE models will not meet the acceptability parameters. However, in my opinion, this method is the most intuitive. The "scaled" approach produced more models that met the acceptability parameters. Nevertheless, based on the results, both these methods provided similar benchmark values.	Thank you for your comment. As noted, the extrapolation method resulted in fewer models meeting with acceptability parameters (outlined in Box 1) than using the scaled approach (between one and 64 more models accepted for PFHxS and PFBS, respectively. However, there was only a marginal increase in the Minimum Data Requirements met using the scaled approach (one additional MDR met for PFBS and for PFHxS). Further, as noted by the reviewer, both approaches resulted in similar benchmark values. As discussed within the document, EPA selected the extrapolation approach as the primary approach for derivation of the benchmark values.
1	It would strengthen the report to include some of the conclusions from Raimondo et al. (in review) regarding the use of the scaling method.	The paper identified as "in review" has now been published and the full reference (Raimondo et al. 2024) has been added to the reference list. Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884

2.d. For the ICE-based models in particular, please compare and contrast the strengths and weaknesses of the extrapolation approach [Sections 3-5] with those of the approach using scaled data [Appendix F].		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
2	We are not given enough technical information to make this comparison. The scaled approach information we are given seems more like instructions for how to run the model when values fall out of environmental realism than a detailed description of how it differs technically from the normal model.	Figures and supporting text were added to the benchmarks document (Section 3) to compare the extrapolation and scaled approaches side-by-side, using a specific example. This example figure is shown in Appendix A of this document. Text was also added to the summary section (Section 5.10) of the benchmarks document to values derived using the ICE-based (extrapolation and scaled), binning, and GLI approaches. Additional details were provided to further clarify the differences between the extrapolated and scaled models. Finally, the recently published paper by Raimondo et al. (2024) also provides comparison of the scaled vs extrapolation approach. Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884

2.d. For the ICE-based models in particular, please compare and contrast the strengths and weaknesses of the extrapolation approach [Sections 3-5] with those of the approach using scaled data [Appendix F].		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
3	As stated above, the technical approach between these techniques is not clear. The implication of the use of the scaled approach rather that the extrapolation method results in the acceptance of more GMAVs, thus resulting in an increased benchmark.	Figure 3-1 and supporting text were added to the benchmarks document (Section 3) to compare the extrapolation and scaled approaches side-by-side, using a specific example. This example figure is shown in Appendix A of this document. Text was also added to the summary section (Section 5.10) of the benchmarks document to values derived using the ICE-based (extrapolation and scaled), binning, and GLI approaches. Additional details were provided to further clarify the differences between the extrapolated and scaled models. Finally, the recently published paper by Raimondo et al. (2024) also provides comparison of the scaled vs extrapolation approach. As noted by the reviewer, the increase in number of models accepted using the scaled approach does result in slightly higher benchmark value. Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884

2.d. For the ICE-based models in particular, please compare and contrast the strengths and weaknesses of the extrapolation approach [Sections 3-5] with those of the approach using scaled data [Appendix F].		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
4	Based on the benchmarks derived using the two different approaches, the extrapolation approach generated lower benchmarks across the eight PFAS compared to using scaled data. For the purpose of the protection of aquatic life, the extrapolation approach would be more protective than the scaled approach. I don't know if this would be the case for other groups of chemicals, but it appears that for PFAS, the extrapolation approach is a more protective approach. In the absence of acceptable empirical data, the more protective approach should be selected. This is critical to avoiding a type II error (i.e., false negative), which is an important consideration in risk assessment.	Thank you for your comment. Additional details were provided to further clarify the differences between the extrapolated and scaled models. The recently published paper by Raimondo et al. (2024) also provides comparison of the scaled vs extrapolation approach. Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884

2.d. For the ICE-based models in particular, please compare and contrast the strengths and weaknesses of the extrapolation approach [Sections 3-5] with those of the approach using scaled data [Appendix F].

<table>
<tr><th>REVIEWER</th><th>REVIEWER COMMENT</th><th>EPA RESPONSE</th></tr>
<tr><td>4</td><td>Benchmarks derived using extrapolation approach:

<table>
<tr><th>Chemical[1]</th><th>PFBA (mg/L)</th><th>PFBS (mg/L)</th><th>PFNA (mg/L)</th><th>PFDA (mg/L)</th><th>PFHxA (mg/L)</th><th>PFHxS (mg/L)</th><th>8:2 FTUCA (mg/L)</th><th>7:3 FTCA (mg/L)</th></tr>
<tr><td>Magnitude</td><td>83</td><td>183</td><td>10.3</td><td>7.9</td><td>75</td><td>9.1</td><td>0.58</td><td>0.18</td></tr>
<tr><td>Duration</td><td colspan="8">One hour average</td></tr>
<tr><td>Frequency</td><td colspan="8">Not to be exceeded more than once in three years on average</td></tr>
</table>
Benchmarks derived using scaled approach:

<table>
<tr><th>Chemical[1]</th><th>PFBA (mg/L)</th><th>PFBS (mg/L)</th><th>PFNA (mg/L)</th><th>PFDA (mg/L)</th><th>PFHxA (mg/L)</th><th>PFHxS (mg/L)</th><th>8:2 FTUCA (mg/L)</th><th>7:3 FTCA (mg/L)</th></tr>
<tr><td>Magnitude</td><td>174</td><td>237</td><td>12</td><td>10</td><td>95</td><td>9.4</td><td>0.65</td><td>0.23</td></tr>
<tr><td>Duration</td><td colspan="8">One hour average</td></tr>
<tr><td>Frequency</td><td colspan="8">Not to be exceeded more than once in three years on average</td></tr>
</table></td><td>Thank you for your comment. Additional details were provided in Section 3 (pgs. 8-10) of the document to further clarify the differences between the extrapolated and scaled models. The recently published paper by Raimondo et al. (2024) also provides comparison of the scaled vs extrapolation approach.

Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884</td></tr>
<tr><td>4</td><td>Based on the reported validation that has been conducted by Raimondo et al., both approaches seem reasonable. It would be nice to be able to see the validation that has been conducted by Raimondo et al., but it appears that this manuscript is currently in review.</td><td>The paper identified as “in review” has now been published and the full reference (Raimondo et al. 2024) has been added to the reference list.</td></tr>
</table>

2.d. For the ICE-based models in particular, please compare and contrast the strengths and weaknesses of the extrapolation approach [Sections 3-5] with those of the approach using scaled data [Appendix F].		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
		Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884
5	The acknowledgment of potential challenges, such as large confidence intervals and potential limitations in accepting ICE models beyond the model range, demonstrates transparency and a thorough understanding of the modeling process.	Thank you for your comment.
5	The decision to select the "scaled" approach as an alternative approach for deriving benchmark values, showcases a proactive and meticulous approach. The alternative scaled approach modifies toxicity values, as needed, to align them with the ICE model range, avoiding the extrapolation of regressions beyond the established model range. There is close agreement between the benchmark values calculated using either approach (as listed in Table 5-26).	Thank you for your comment.
5	This consideration of alternative methods highlights a commitment to rigorous evaluation and continuous improvement, reflecting a commendable scientific rigor in the approach to deriving benchmark values.	Thank you for your comment.

Charge Question 3: ***Please comment on the empirical (direct test) and ICE-generated toxicity data used to derive the benchmark values presented in the draft document.***

3.a. ***Were the data adequately used and sufficiently comprehensive to represent risks to sensitive aquatic life?***

REVIEWER	REVIEWER COMMENT	EPA RESPONSE
1	Yes. The data covers the eight MDRs, and hence a wide range of taxa with different characteristics in aquatic ecosystems. This, combined with the statistical approach which is focusing on the lowest GMAVs ensures that acute toxic effects on sensitive aquatic life are taken into account. For PFBS and PFHxS, only seven and six, respectively, of the eight MDRs were fulfilled and hence, these benchmarks are associated with greater uncertainty. This information should be included in the summary (could be footnotes in Table Ex1-1).	Thank you for your comment. Footnotes and text were added to the noted summary to clarify this point.
2	Probably not. There are too few empirical data to be secure in understanding which species in the real world might be sensitive.	Thank you for your comment. It should be noted that the ICE models are derived from empirical test data, though as discussed, these data are extrapolated to other species. The binning approach is also based on empirical data, with the assumption that the constituent empirical data have similar or the same MOAs. As previously discussed, following the peer review the EPA has developed application factors to account for two species (the mayfly *Neocloeon triangulifer* and the cladoceran species *Moina micrura*) that were indicated by direct empirical test data to have markedly greater sensitivity to PFOA and PFOS. These application factors were used to lower the PFAS values to account for these particularly sensitive species, for which PFAS toxicity data are not available. This adjustment captures the sensitivity of the most sensitive known species to acute effects of PFAS, yielding protective values.

3.a. ***Were the data adequately used and sufficiently comprehensive to represent risks to sensitive aquatic life?***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
		Further, the outcomes of the benchmark approach using ICE data is corroborated by two other methods for calculating protective values in data - limited situations, the Giddings et al (2019) binning approach and the Grippo et al (2021) approach (using EPA's Aquatic Life Criteria Guidelines-based approach for the SSD) to calculate protective values. The protective values calculated by all 3 methods fall within a factor of 30 of each other (except for PFBS), with most falling within a factor of less than 10 of each other, which in the stochastic world of environmental science is remarkably close, providing further support and validation of the EPA's approach using web-ICE. See Table 2 above, and in Section 5.10.3 of the document.
3	As stated, the problem is a lack of empirical data and a reliance on data estimation techniques. The table below provides the data for the 4 most sensitive species used to derive the benchmark values for the PFAS materials. Actual empirical data are highlighted (4 of 32 data points, 12.5%); as you can see, most of the data for the range of species are estimated values. Without additional confirmatory experimental data, it is difficult to say anything about how comprehensive the data represent sensitive aquatic organisms.	Thank you for your comment. The methodologies tested both rely on relatively robust *underlying empirical* datasets for their development. The ICE models rely on underlying empirical test data used to derive each model, while the binning approach bases value development on direct empirical datapoints for PFAS that were grouped to determine the chemical-specific benchmarks. Both approaches, however, do rely on the extrapolation of these empirical data (to other compounds) to derive the benchmark values. The EPA has been exploring the evaluation of these approaches using chemicals with other, more data-rich direct toxicity datasets.

3.a. ***Were the data adequately used and sufficiently comprehensive to represent risks to sensitive aquatic life?***

REVIEWER	REVIEWER COMMENT	EPA RESPONSE
	Table 4. Species sensitivity ranking for PFAS compounds. Empirical data are in bold, all others are ICE-estimated values.	

Chemical	Species Sensitivity Rank from ICE model[1]			
	1	2	3	4
Perfluorobutanoic acid (PFBA)	***Brachionus calyciflorus* (rotifer)**	*Oncorhynchus mykiss* (rainbow trout)	*Gammarus fasciatus* (amphipod)	*Amblema plicata* (mussel)
Perfluorobutanesulfonic acid (PFBS)	*Amblema plicata* (mussel)	*Gammarus fasciatus* (amphipod)	*Hyalella azteca* (amphipod)	*Ceriodaphnia dubia* (cladoceran)
Perfluorononanoic acid (PFNA)	*Amblema plicata* (mussel)	***Chydorus sphaericus* (cladoceran)**	*Megalonias nervosa* (mussel)	*Oncorhynchus mykiss* (rainbow trout)
Perfluorodecanoic acid (PFDA)	*Caecidotea brevicauda* (isopod)	*Micropterus salmoides* (bass)	*Perca flavescens* (yellow perch)	*Salvelinus fontinalis* (brook trout)
Perfluorohexanoic acid (PFHxA)	***Brachionus calyciflorus* (rotifer)**	*Amblema plicata* (mussel)	*Gammarus fasciatus* (amphipod)	*Chydorus sphaericus* (cladoceran)
Perfluorohexanesulfonic acid (PFHxS)	***Danio rerio* (zebrafish)**	*Jordanella floridae* (flagfish)	*Daphnia magna* (cladoceran)	*Limnodrilus hoffmeisteri* (oligochaete)
Hexadecafluoro-2-decenoic acid (8:2 FTUCA)	*Amblema plicata* (mussel)	*Palaemonetes kadiakensis* (grass shrimp)	*Chydorus sphaericus* (cladoceran)	*Megalonias nervosa* (mussel)
Pentadecafluorodecanoic acid (7:3 FTCA)	*Amblema plicata* (mussel)	*Macrobrachium nipponense* (river shrimp)	*Chydorus sphaericus* (cladoceran)	*Megalonias nervosa* (mussel)

[1] It should also be noted that the 4 data points listed in bold are all based on nominal concentrations, so the accuracy of the EC50 values may be questioned.

4	I think the available empirical was adequately used.	Thank you for your comment.
4	However, there was an overall lack of empirical acute toxicity data, particularly for freshwater primary producers and freshwater invertebrates. For example, there was no acute toxicity data on the eight PFAS for freshwater primary producers. Freshwater invertebrates were also under-represented in the empirical data set. For example, there was only three empirical data points for PFHxS and all three were for freshwater vertebrates. When empirical toxicity data on the eight PFAS was available for a freshwater invertebrate species, it was usually *Daphnia magna*. Consequently, I don't think that the data is sufficiently comprehensive to represent risk to sensitive aquatic life.	Thank you for your comment. The entire effort to develop benchmarks for the eight PFAS using web-ICE was pursued precisely because data are limited, yet the need to protect the environment from potential impacts of these chemicals remains. It should be noted that the ICE models are derived from an extremely extensive empirical test dataset (>10,700 toxicity records); these data are extrapolated to other species based on the ICE models. The binning approach is also based on empirical data, with the assumption that the

3.a. ***Were the data adequately used and sufficiently comprehensive to represent risks to sensitive aquatic life?***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
		constituent empirical data all have similar or the same MOAs. As previously discussed, following the peer review, the EPA has developed application factors to account for two species (the mayfly *Neocloeon triangulifer* and the cladoceran species *Moina micrura*) that were indicated by direct empirical test data to have markedly greater sensitivity to PFOA and PFOS. These application factors were used to lower the PFAS values to account for these particularly sensitive species, for which PFAS toxicity data are not available. This adjustment captures the sensitivity of the most sensitive known species to acute effects of PFAS, yielding protective values. The EPA agrees that additional data generation by researchers on the toxicity of PFAS to aquatic life, and other species would be useful to inform development of protective values.
5	The EPA employed both empirical test data and ICE values, derived for missing Minimum Data Requirements (MDRs), to determine acute freshwater benchmark recommendations for aquatic life. The utilization of ICE-predicted values by various independent, international groups to establish protective values for aquatic life confirms that values derived from ICE-generated Species Sensitivity Distributions (SSDs) offer a consistent level of protection comparable to using directly measured laboratory data.	Thank you for your comment.

3.b. ***Were the data selected and/or excluded from the benchmark values derivation appropriately utilized?***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
1	The data that was deemed of sufficient quality to be used (i.e., quantitatively acceptable freshwater acute toxicity studies in appendix A) were appropriately utilized as these data were used to produce the sensitivity distributions. However, the qualitative freshwater acute toxicity studies in appendix C should be included in a discussion about the derived benchmarks.	Thank you for your comment. Text was added to provide a general discussion of the data in Appendix C, and the relationship of the patterns observed with these data relative to the quantitatively-accepted data.
2	I think the authors did the best they could with the available data on hand. It is a shame that more resources are not being deployed to generate more empirical data.	Thank you for your comment.
3	Much of the data accepted would not meet current standards for data acceptability or criteria derivation. The authors have tried to maintain a degree of fidelity to the Stephan et al (1985) AWQC methodology; however, several concerns exist with the data considered: • The benchmark document states: "Toxicity studies accessed from the ECOTOX database were further evaluated by Office of Water. Studies were evaluated for data quality as described by EPA OW's data quality standard operating procedure (SOP), and consistent with OW's data quality review approach U.S. EPA (1985), and EPA's Office of Chemical Safety and Pollution Prevention (OPP)'s Ecological Effects Test Guidelines (U.S. EPA 2016c)." These documents were not included in the reference list and were not provided.	Thank you for your comment. The EPA disagrees with the peer reviewer's comment that the accepted data "would not meet current standards for data acceptability." The toxicity data were thoroughly reviewed, first through the ECOTOX data evaluation process and subsequently by OW's extensive data quality review analyses, which have been consistently applied in all criteria. The data quality review for the PFAS benchmark studies is discussed on pages 11-14 of the PFAS acute benchmark report, and tables identifying which tests were used quantitatively, qualitatively or not used are detailed in Appendices A through D of the report. Appendices C and D describe deficiencies in the studies for the qualitative and unused studies, respectively, regarding their lack of quantitative use. Please also see specific responses to the next five comments about data acceptability. The noted references have been added to the document.

3.b. ***Were the data selected and/or excluded from the benchmark values derivation appropriately utilized?***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
3	• In general, most of the accepted empirical studies are reported based on nominal exposure concentrations rather than analytically measured concentrations; this is not consistent with state-of-the-science standards of acceptability for empirical toxicology data. In fact, one of the studies that measured exposure concentrations (Ding et al 2012) ultimately reported test endpoint data (EC50) based on nominal concentrations rather than measured values. The reported analytical data indicates that test concentrations differed from nominals by 10-20%, so the value reported based on nominals is likely to be 10-20% off.	Toxicity tests used in EPA aquatic life criteria documents are typically based on measured chemical concentrations if such data are available. However, for PFOA and PFOS, a substantial proportion of the available data were reported by the study authors only as nominal test concentrations. The EPA completed a thorough analysis of the data to determine whether nominal (unmeasured) and measured concentrations are in agreement with each other, based on an analysis of available studies reporting both nominal and measured concentrations in the EPA's 2024 Final Aquatic Life Criteria documents. (U.S. EPA 2024a,b). PFOA and PFOS toxicity having pairs of nominal and corresponding measured concentrations were compared through: (1) linear correlation analysis and; (2) an assessment of measured concentrations as a percent of its paired nominal concentration. The EPA's analyses demonstrated that the linear correlation between measured and corresponding nominal concentrations show a high degree of precision between paired concentrations across all test conditions, in most cases within the test acceptability threshold identified by the EPA's OCSPP's Ecological Effects Test Guidelines. These analyses were published in a peer-reviewed journal in 2023 by Jarvis et al. As broadly accepted by the international community, OCSPP's Guidelines recommend measured concentrations be within +/- 20% of nominal

3.b. ***Were the data selected and/or excluded from the benchmark values derivation appropriately utilized?***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
		concentrations to be used. Although a parallel analysis could not be conducted based upon the more limited datasets for these PFAS benchmarks, the PFAS concentrations in test waters are expected to remain relatively constant over the course of acute exposures given its ability to resist breakdown and transformation (Ahrens et al. 2011). Accordingly, PFAS toxicity tests were not excluded from quantitative use in criteria derivation on the basis of unmeasured test concentrations alone based on results of the analysis with PFOA and PFOS. Finally, Ding et al.(2012) appears to have calculated point estimates using nominal data over measured data because only a portion of the test replicates/treatments were actually measured. For example, the PFAS Benchmarks document states, "*Although partial chemical analysis was conducted, only nominal concentrations were reported. To assess the test concentration variability, all samples at a selected nominal concentration were analyzed, with a relative standard deviation of 6.5%. The measured concentrations were between 82 and 91% of the nominal concentration*." Ahrens, L. 2011. Polyfluoroalkyl compounds in the aquatic environment: a review of their occurrence and fate. J. Environ. Monit. 13(1): 20-31.

3.b. ***Were the data selected and/or excluded from the benchmark values derivation appropriately utilized?***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
		Ding, G.H., T. Fromel, E.J. Van den Brandhof, R. Baerselman and W.J.G.M. Peijnenburg. 2012. Acute toxicity of poly- and perfluorinated compounds to two cladocerans, *Daphnia magna* and *Chydorus sphaericus*. Environ. Toxicol. Chem.31(3): 605-610. Jarvis, A.L., J.R. Justice, B. Schnitker and K. Gallagher. 2023. Meta-analysis comparing nominal and measured concentrations of perfluorooctanoic acid and perfluorooctane sulfonate in aquatic toxicity studies across various experimental conditions. Environ. Toxicol. Chem. 42(11): 2289-2301. USEPA (U.S. Environmental Protection Agency). 2024a. Aquatic life ambient water quality criterion for PFOA-freshwater. EPA-842-R-24-002. Office of Water. Washington, D.C. USEPA (U.S. Environmental Protection Agency). 2024b. Aquatic life ambient water quality criterion for PFOS-freshwater. EPA-842-R-24-003. Office of Water. Washington, DC.
3	• Some of the toxicity data used in the derivation of the aquatic benchmarks comes from studies that used non-native species (i.e., zebrafish, *Danio rerio*), which adds uncertainty associated with the representativeness of such species to native North American aquatic fauna. Stephan et al (1985) states: II. G. "Questionable data, data on formulated mixtures and emulsifiable concentrates, and data obtained with non-resident species in North America or previously	While it remains uncertain if there are established resident zebrafish (*Danio rerio*) populations in the conterminous United States (USFWS 2018), zebrafish are common ecotoxicity test organisms that serve as taxonomic surrogates for untested

3.b. *Were the data selected and/or excluded from the benchmark values derivation appropriately utilized?*		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
	exposed organisms may be used to provide auxiliary information **but should not be used in the derivation of criteria**." Appendix I Resident North American Species of Aquatic Animals Used in Toxicity and Bioconcentration Tests defines zebrafish (*Danio rerio*) as "Non-resident" species and therefore should not be included for criteria derivation. Use of Non-resident species is briefly discussed in the report and a reference to US EPA 2018b is cited; however, this reference is not included in the reference list. The zebrafish is in the family *Cyprinidae*, which all North American native minnows (including the fathead minnow), shiners, and dace belong. Although not native to North America, EPA seems to have decided that in the absence of suitable data on native cyprinids, the zebrafish is an acceptable representative. However, given that zebrafish are frequently among the more sensitive species and at least some studies with PFAS compounds have suggested that fathead minnows may be more sensitive[5], it would be good to have some comparative additional data with NA species.	fish species and are also considered in effects assessments conducted under the Toxic Substances Control Act (TSCA) and the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA). Moreover, zebrafish data were used in deriving the final aquatic life criteria for PFOA and PFOS (USEPA 2024a, b), and for the derivation of other aquatic life criteria (e.g., aluminum). USEPA (2024a, b) specifically showed that inclusion of non-North American species had a negligible to slight impact on the final PFOA and PFOS aquatic life criteria. The noted reference (i.e., USEPA 2018) was added to the document. U.S. EPA (United States Environmental Protection Agency). 2018. Final aquatic life ambient water quality criteria for aluminum - 2018 (EPA-822-R-18-001). Office of Water. Washington, D.C. USEPA (U.S. Environmental Protection Agency). 2024a. Aquatic life ambient water quality criterion for PFOA-freshwater. EPA-842-R-24-002. Office of Water. Washington, D.C. USEPA (U.S. Environmental Protection Agency). 2024b. Aquatic life ambient water quality criterion

[5] Suski et al. 2023. Ecotoxicity and Accumulation of Perfluorononanoic Acid in the Fathead Minnow (*Pimephales promelas*) and an Approach to Developing Protective Thresholds in the Aquatic Environment Through Species Sensitivity Distribution. Environ Toxicol Chem. https://doi.org/10.1002/etc.5692

3.b. ***Were the data selected and/or excluded from the benchmark values derivation appropriately utilized?***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
		for PFOS-freshwater. EPA-842-R-24-003. Office of Water. Washington, DC. U.S. FWS. (U.S. Fish and Wildlife Service). 2018. Zebra Danio (Danio rerio) Ecological Risk Screening Summary.
3	• Some of the test methods used are not consistent with the 1985 guidance. EPA 1985 states that "Acute EC50s that are based on effects that are not severe, such as reduction in shell deposition and reduction in growth, are not used in calculating the Final Acute Value." The zebrafish tests included in the benchmark document (Annunziato et al 2020) followed the OECD 236 method and reported results based on a growth rather than survival endpoint. These data would not be acceptable for derivation of an FAV based on the 1985 guideline.	The ~96 hr. growth-based results were not used. Despite the authors measuring growth, this test provided information to identify an acute value as a "< LC50" from a relevant acute exposure duration. For example, the PFAS Benchmarks document states, "*Test concentrations of PFHxS were too low to derive an LC50 value. Larval growth (length) was assessed at test termination. The author-reported a growth (weight) NOEC of 22.5 mg/L, and a 93-hr LC_{50} >45 mg/L. This latter value was acceptable for quantitative use in deriving the recommended acute freshwater PFHxS benchmark."*
3	• In addition to the above concerns, it was noted that at least two studies (Ding et. al. 2012, Annunziato et al 2020) that reported tests with PFBA, PFBS, PFNA, PFDA, and PFHxS conducted their studies using dimethylsulfoxide (DMSO) as a carrier solvent. In EPA's current test guidelines, it is recommended that if a carrier solvent must be used, *"Preferred solvents are dimethyl formamide, triethylene glycol, methanol, acetone, or ethanol. Solvent use should be avoided if possible."* DMSO is known to transport nonionized molecules thorough many biological membranes (Jacob and Herschler 1985[6]). Although the authors of the lab tests conducted a "solvent control" this does not control for possible synergistic interactions of DMSO acting as a membrane carrier, thus potentially increasing observed toxicity. Because these tests represent a large portion of the quantitatively acceptable freshwater toxicity tests (20%, 7 of 36),	OECD (2019) indicated that when a solvent must be used, the final concentration of the solvent used should be minimized and not exceed 100 mg/L or 0.1 mL/L and should be the same in all test vessels, excluding the dilution water control. OCSPP 850.1000 and OCSPP 850.1075 similarly report 0.1 mL/L as a level that should not be exceeded when a solvent is used. These Test Guidelines do not recommend excluding tests with DMSO. Studies conducted by Ding et. al. (2012)

[6] *Jacob, S. W., & Herschler, R. (1986). Pharmacology of DMSO. Cryobiology, 23(1), 14-27.*

3.b. ***Were the data selected and/or excluded from the benchmark values derivation appropriately utilized?***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
	EPA should consider the potential for inclusion of these data resulting in lower than desired criteria, due to an overestimation of toxicity due to DMSO synergy.	and Annunziato et. al. (2020) are consistent with these guidelines. These data were accordingly considered acceptable for quantitative use. OECD. 2019. Guidance Document on Aqueous-phase Aquatic Toxicity Testing of Difficult Test Chemicals. Report ENV/JM/MONO(2000)6/REV1. Organisation for Economic Co-operation and Development. Series on Testing and Assessment. No 23 (Second Edition).
3	EPA should reassess the quality and acceptability of the available data for regulatory purposes.	The EPA disagrees with the peer reviewer's comment and wishes to clarify the purpose of the benchmarks. Importantly, these benchmarks are not regulatory values; rather, they provide information to states, Tribes and other stakeholders. The toxicity data were thoroughly reviewed, first through the ECOTOX data evaluation process, and subsequently by OW's extensive data quality review analyses, which have been consistently applied in all criteria. The data quality review for the PFAS benchmark studies is discussed on pages 11-14 of the PFAS acute benchmark report, and tables identifying which tests were used quantitatively, qualitatively or not used are detailed in Appendices A through D of the report. Appendices C and D describe deficiencies in the studies for the qualitative and unused studies,

3.b. ***Were the data selected and/or excluded from the benchmark values derivation appropriately utilized?***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
		respectively, regarding their lack of quantitative use. Notably, other peer reviewers agreed the EPA's data review and selection process was clear, logical and well-established.
4	Yes, the process of including and excluding empirical data for derivation of the benchmark values was clearly explained, logical, and well established.	Thank you for your comment.
5	Quantitatively acceptable empirical acute toxicity data available for each of the eight PFAS was tabulated for each individual study. All toxicity values, including LC values, EC values, NOECs, LOECs, and species- and genus-mean values, were presented with four significant figures. This practice avoided round-off errors in subsequent calculations. Studies that were determined to be qualitatively acceptable as supporting information, but not acceptable for quantitative use were listed with deficiencies in each study. Furthermore, studies that were deemed unsuitable for either quantitative or qualitative were also cited. I endorse the choice to incorporate toxicity data for studies solely based on unmeasured test concentrations. This decision is rooted in findings for PFOA and PFOS (U.S. EPA 2022a, b), leading the EPA to determine that nominal test concentrations effectively represent real PFAS exposures in standard acute laboratory-based toxicity tests. In addition, Hoke et al.,2012 (https://doi.org/10.1016/j.chemosphere.2011.12.066) also reported mean measured test concentrations were similar (within 80–120% of nominal) to the targeted nominal test concentrations for fluorinated acids with the exception of the 5:3 acid.	Thank you for your comment.
5	The authors applied the criteria recommended by Willming et al., 2016 to enhance models' reliability and robustness (Box 1). Models adhering to these acceptance parameters were employed to generate species toxicity data, which were then combined with empirical toxicity data to strengthen the derivation of benchmark values. This approach demonstrated logical and consistent application of standard criteria across all eight PFAS compounds.	Thank you for your comment.

3.c. ***Are there relevant data that you are aware of that should be included? If so, please provide for derivation of benchmark values.***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
1	I am not aware of data not already included in the study.	Thank you for your comment.
2	There is some mayfly data from the Soucek lab that does not seem to be acknowledged here. The general lack of insect data is a systemic problem – particularly when a single midge is used to represent the toxicity of an entire class or organisms that is likely close to 10,000 species in N. America.	Thank you for your comment. The EPA concurs with the importance of additional insect data and uses acceptable empirical toxicity test data whenever they are available in the scientific literature. As discussed in earlier responses, following the peer review the EPA developed application factors to account for two species (the mayfly *Neocloeon triangulifer* and the cladoceran species *Moina micrura*) that were indicated by recent direct empirical test data to have markedly greater sensitivity to PFOA and PFOS. These application factors were used to lower the PFAS values to account for these particularly sensitive species, for which PFAS toxicity data are not available, thus addressing the peer reviewer's concern there is a lack of insect data, including the sensitive mayfly *N. triangulifer*.
3	Several recent publications have critically reviewed the available data for PFAS compounds, e.g., Pandelides et al. 2023[7]. The references below are just an example of amphibian references included in one of the review articles, these include acute and chronic endpoints: Abercrombie, S. A., de Perre, C., Choi, Y. J., Tornabene, B. J., Sepúlveda, M. S., Lee, L. S., & Hoverman, J. T. (2019). Larval amphibians rapidly bioaccumulate poly- and perfluoroalkyl substances. *Ecotoxicology and Environmental Safety*, **178**, 137–145. https://doi.org/10.1016/j.ecoenv.2019.04.022;	Thank you for your comment. The cited studies and those within the cited review paper focus on the evaluation of bioaccumulation, subacute, subchronic, chronic responses, non-apical endpoints, and/or mixtures and the reported data are not directly applicable to the acute endpoints evaluated in this document.

[7] *Pandelides Z, J Conder, Y Choi, E Allmon, T Hoskins, L Lee, J Hoverman, M Sepúlveda. 2023. A Critical Review of Amphibian Per- and Polyfluoroalkyl Substance Ecotoxicity Research Studies: Identification of Screening Levels in Water and Other Useful Resources for Site-Specific Ecological Risk Assessments. Environ Toxicol Chem. https://doi.org/10.1002/etc.5695*

3.c. ***Are there relevant data that you are aware of that should be included? If so, please provide for derivation of benchmark values.***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
	Ankley, G. T., Kuehl, D. W., Kahl, M. D., Jensen, K. M., Butterworth, B. C., & Nichols, J. W. (2004). Partial life-cycle toxicity and bioconcentration modeling of perfluorooctane sulfonate in the northern leopard frog (*Rana pipiens*). *Environmental Toxicology and Chemistry*, **23**, 2745. https://doi.org/10.1897/03-667 Brown, S. R., Flynn, R. W., & Hoverman, J. T. (2021). Perfluoroalkyl substances increase susceptibility of northern leopard frog tadpoles to trematode infection. *Environmental Toxicology and Chemistry*, **40**, 689–694. https://doi.org/10.1002/etc.4678 Flynn, R. W., Chislock, M. F., Gannon, M. E., Bauer, S. J., Tornabene, B. J., Hoverman, J. T., & Sepúlveda, M. S. (2019). Acute and chronic effects of perfluoroalkyl substance mixtures on larval American bullfrogs (*Rana catesbeiana*). *Chemosphere*, **236**, 124350. https://doi.org/10.1016/j.chemosphere.2019.124350 Flynn, R. W., Hoover, G., Iacchetta, M., Guffey, S., de Perre, C., Huerta, B., Li, W., Hoverman, J. T., Lee, L., & Sepúlveda, M. S. (2022). Comparative toxicity of aquatic per- and polyfluoroalkyl substance exposure in three species of amphibians. *Environmental Toxicology and Chemistry*, **41**, 1407–1415. https://doi.org/10.1002/etc.5319 Flynn, R. W., Iacchetta, M., Perre, C., Lee, L., Sepúlveda, M. S., & Hoverman, J. T. (2021). Chronic per-/polyfluoroalkyl substance exposure under environmentally relevant conditions delays development in northern leopard frog (*Rana pipiens*) larvae. *Environmental Toxicology and Chemistry*, **40**, 711–716. https://doi.org/10.1002/etc.4690 Foguth, R. M., Hoskins, T. D., Clark, G. C., Nelson, M., Flynn, R. W., de Perre, C., Hoverman, J. T., Lee, L. S., Sepúlveda, M. S., & Cannon, J. R. (2020). Single and mixture per- and polyfluoroalkyl substances accumulate in developing northern leopard frog brains and produce complex neurotransmission alterations. *Neurotoxicology and Teratology*, **81**, 106907. https://doi.org/10.1016/j.ntt.2020.106907 Fort, D. J., Mathis, M. B., Guiney, P. D., & Weeks, J. A. (2019). Evaluation of the developmental toxicity of perfluorooctane sulfonate in the Anuran, *Silurana*	

3.c. ***Are there relevant data that you are aware of that should be included? If so, please provide for derivation of benchmark values.***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
	tropicalis. *Journal of Applied Toxicology*, **39**, 365–374. https://doi.org/10.1002/jat.3727 Hoover, G. M., Chislock, M. F., Tornabene, B. J., Guffey, S. C., Choi, Y. J., De Perre, C., Hoverman, J. T., Lee, L. S., & Sepúlveda, M. S. (2017). Uptake and depuration of four per/polyfluoroalkyl substances (PFAS) in northern leopard frog *Rana pipiens* tadpoles. *Environmental Science and Technology Letters*, **4**, 399–403. https://doi.org/10.1021/acs.estlett.7b00339 Hoskins, T. D., Allmon, E. B., Flynn, R. W., Lee, L. S., Choi, Y., Hoverman, J. T., & Sepúlveda, M. S.(2022). An environmentally relevant mixture of perfluorooctane sulfonic acid and perfluorohexane sulfonic acid does not conform to additivity in northern leopard frogs exposed through metamorphosis. *Environmental Toxicology and Chemistry*, **41**, 3007–3016. https://doi.org/10.1002/etc.5486 Lech, M. E., Choi, Y. J., Lee, L. S., Sepúlveda, M. S., & Hoverman, J. T. (2022). Effects of per- and polyfluoroalkyl substance mixtures on the susceptibility of larval American bullfrogs to parasites. *Environmental Science & Technology*, **56**, 15953–15959. https://doi.org/10.1021/acs.est.2c04574	
4	I am not aware of relevant data that should be included in this process. An extensive and complete review of available data has been conducted in preparation for this process of deriving benchmarks.	Thank you for your comment.
5	Below, a recent chronic study by Kadlec et al., 20203 has been listed as an additional reference, some aspects may be relevant. **https://doi.org/10.1002/etc.5784** Sarah M. Kadlec, Will J. Backe, Russell J. Erickson, J. Russell Hockett, Sarah E. Howe, Ian D. Mundy, Edward Piasecki, Henry Sluka, Lauren K. Votava, David R. Mount (2023) Sublethal Toxicity of 17 Per- and Polyfluoroalkyl Substances with Diverse Structures to *Ceriodaphnia dubia*, *Hyalella azteca*, and *Chironomus dilutus*	Thank you for your comment. The cited paper contains subchronic studies which exceed the test duration considered applicable for inclusion within the acute dataset.

Charge Question 4: ***Are the derived benchmark values appropriately protective of sensitive aquatic life?***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
1	In my opinion, the derived benchmark values are protective for acute toxic effects on aquatic life as defined in the report (i.e., to be protective of 95% of freshwater genera potentially exposed to these specific PFAS under short-term conditions of one-hour duration, if the one-hour average magnitude is not exceeded more than once in three years).	Thank you for your comment.
1	However, it is important to note that the most problematic properties of PFAS are not the acute toxicological effects but rather the combination of their persistency, mobility, potential for bioaccumulation (for some PFAS), and long-term toxicological effects. Emissions of PFAS resulting in recipient concentrations comparable to the threshold values in this report may (depending on the volume of water in the recipient) be substantial sources of PFAS pollution to the environment. As many PFAS (PFCA and PFSA) are both persistent and relatively mobile, such emissions may result in problematic pollution of drinking water and wildlife with adverse long-term consequences for both human health and the environment.	Thank you for your comment. The EPA agrees with the reviewer that the chronic effects of PFAS are a concern. However, characterization of what is a protective acute concentration for chemicals of concern still remains important, particularly for discharges. Text was added to the document to further clarify the values provided are for acute effects and do not account for potential chronic effects and/or bioaccumulation. The EPA is working towards approaches for the development of chronic values for data-limited chemicals, but the NAMs tools to develop chronic benchmarks were not available to apply at this time. Additionally, there is a general lack of chronic data, especially for these PFAS, upon which to base chronic analyses. Additional research on the chronic toxicity of a range of PFAS would support development of additional values.
2	As applied to a one-hour maximum concentration not to be exceeded every 3 years, the values are likely protective to most species.	Thank you for your comment.
3	The benchmark methods attempt to maintain compliance with the EPA's 1985 method for derivation of AWQC and to the extent that the 1985 method was "*appropriately protective of sensitive aquatic life"* the benchmark approach should be as well. However, one major difference between the 1985 guidance and the new benchmark approach is that the requirements for high-quality empirical data for a minimal range of aquatic species have been reduced or eliminated. The benchmark approach seems to rely on existing data or extrapolation of limited data from similar compounds. ICE methods are extremely useful	The intent of deriving benchmark values is not to undermine the importance of using empirical test data for the derivation of protective aquatic life values (benchmarks or criteria). The EPA is incorporating the use of direct test data to the greatest extent possible when they are available.

Charge Question 4: ***Are the derived benchmark values appropriately protective of sensitive aquatic life?***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
	and important in estimating values for species where we cannot generate empirical data, e.g., T&E species (Willming et al 2016). However, the benchmark approach proposed seems to minimize the utility and need for empirical data. Modelling techniques that are based on robust empirical data are extremely useful for **supplementing** extant data for species-of-concern that cannot be easily or cost-effectively tested, or tested due to regulatory restrictions, but they should not suppliant the need for chemical-specific empirical data. Is there a minimum amount of empirical data that are needed to derive a benchmark? In theory, a single acute toxicity test may be sufficient, using ICE-models, to derive regulatory benchmarks. To gain confidence in the proposed approach, EPA should conduct testing to further confirm the accuracy of the ICE estimates, especially for the most sensitive species in the benchmark data sets.	Instead, the objective is to employ methodologies that would allow for the development of protective aquatic life values for chemicals that might otherwise be precluded from the derivation of protective values based on traditional derivation approaches and "minimum data requirements", which would exclude the EPA's ability to derive protective values. Evaluation and the appropriate incorporation of evolving science and NAMs is consistent with the good science clause in the EPA's Aquatic Life Criteria Guidelines. There are no specific direct test requirements that have been established for deriving benchmarks. However, as previously discussed, all methods used incorporate underlying empirical test data, albeit with extrapolation to derive benchmark values. Consistent with the suggestion by the reviewer, the EPA has been exploring the evaluation of these approaches using chemicals with more data-rich direct toxicity datasets.
4	I do have concerns about the lack of acceptable empirical acute toxicity data for freshwater primary producers and freshwater invertebrates. I think the process of deriving benchmarks for the eight PFAS described by the USEPA is appropriate for the empirical data that is available. I think they have done their best with the data that is available to them.	Thank you for your comment.
5	The establishment of aquatic life benchmarks for the eight PFAS compounds involved the utilization of empirical and Web-ICE data, incorporating statistical methods for calculation. This strategy is in accordance with the EPA's objective of minimizing dependence on animal testing by integrating NAMS into toxicity assessments.	Thank you for your comment.

Charge Question 4: ***Are the derived benchmark values appropriately protective of sensitive aquatic life?***		
REVIEWER	**REVIEWER COMMENT**	**EPA RESPONSE**
5	In stating that the benchmarks derived are regarded as less certain than ambient water quality criteria, the authors are acknowledging a degree of uncertainty. This recognition is essential for offering a logical evaluation of the benchmarks' reliability.	Thank you for your comment.

III. Specific Observations				
Reviewer	**Page**	**Paragraph**	**Comment or Question**	**EPA Response**
1	Front page	Title	As “PFAS” is an abbreviation for “per- and polyfluorinated substances”, the term “PFAS Compounds” would mean “per- and polyfluorinated substances Compounds”. Here and throughout the document, I suggest to just use “PFAS” instead of “PFAS Compounds”, “PFAS substances”, or “PFAS chemicals”	The edit was made as suggested.
1	viii	Acronym list	Several of the acronyms used in the report are missing. Examples of missing acronyms are listed in the following, however, there are likely more missing acronyms. This should be reviewed and corrected before publication. Examples of missing acronyms: SMAV, all the eight PFAS in focus (PFBA, PFBS, PFHxA, PFHxS, etc.), EPA, DOD, MSE, SMAV	The missing acronyms were added as suggested.
1	x	1	Please define what is meant by the term Water Quality Criteria in this context	The text is referring to 304(a)(1) Water Quality Criteria. This text first identifies these criteria in order to contrast them with 304(a)(2) benchmark values.
1	x	1	Please define the difference between “draft ambient water quality benchmarks” and “Water Quality Criteria”	Benchmarks refer to the values derived in this document, while “Water Quality Criteria” refers to 304(a)(1) Water Quality Criteria.
1	xi	1	It is somewhat confusing for the reader to understand what was done in the present study and what has been done previously. I suggest making this clearer by using terms such as “in the present study” or similar.	The edit was made as suggested.
1	xi	1	Here, and throughout the document, it would be more logical to present the eight PFAS sorted according to group and number of perfluorinated carbon atoms in the structures. I.e., PFBA, PFHxA, PFNA, PFDA, PFBS, PFHxS, 7:3 FTCA, 8:2 FTUCA.	The edit was made as suggested.
1	xi	1	It should be made clearer in the summary that the benchmarks reported here are for acute exposure. I.e.: “...protective of 95% of freshwater genera potentially exposed to the specific PFAS	Text was added clarifying parameters for frequency and duration.

III. Specific Observations				
Reviewer	**Page**	**Paragraph**	**Comment or Question**	**EPA Response**
			under short-term conditions of one-hour duration, if the one-hour average magnitude is not exceeded more than once in three years"	
1	xi	2	The sentence that starts with "EPA's "*Guidelines for Deriving Numerical*" needs to be revised as it should likely be at least two sentences.	The edit was made as suggested.
1	xii	Table Ex-1-1	Is the superscript "1" referring to a footnote?	The superscript was corrected as suggested.
1	xii	Table Ex-1-1	It would be more logical to present the eight PFAS sorted according to group and number of perfluorinated carbon atoms in the structures. I.e., PFBA, PFHxA, PFNA, PFDA, PFBS, PFHxS, 7:3 FTCA, 8:2 FTUCA.	The edit was made as suggested.
1	xiii	1	Here, and throughout the document, I suggest using the terms perfluoroalkyl carboxylic acids (PFCA) and perfluoroalkyl sulfonic acids (PFSA) as this is the common terminology.	The edit was made as suggested.
1	xiii	1	A conclusion for the binning approach should be included in the summary.	A conclusion about the binning approach was added to the summary.
1	14 (or 1?)	Page numbers	The first page in the section "Background" should probably be 1 (it has page number 14 in the version I received)	The edit was made as suggested.
1	14	1	Here, and throughout the document, it would be more logical to present the eight PFAS sorted according to group and number of perfluorinated carbon atoms in the structures. I.e., PFBA, PFHxA, PFNA, PFDA, PFBS, PFHxS, 7:3 FTCA, 8:2 FTUCA.	The edit was made as suggested.
1	14	2	The acute water quality benchmark concentrations for PFOS and PFOA should be stated here	The PFOS and PFOA final values were added to the summary discussion for the PFAS values.
1	14	2	It would be good to include information on how the selection of these eight PFAS was done. Why these exact substances? Were other PFAS considered, but not included due to limited information available?	The text (now on page 1) states "The above eight chemicals were selected to represent a range of PFAS that are present in aquatic ecosystems and of concern to stakeholders (e.g., states, Tribes, DOD). Another important

III. Specific Observations				
Reviewer	**Page**	**Paragraph**	**Comment or Question**	**EPA Response**
				consideration for selection was the availability of both acute empirical and ICE model toxicity data for these chemicals."
1	14	3	The reference U.S. EPA 1985 is not included in the reference list.	Missing references were added to the document.
1	14	3	It would make it clearer for the reader to state "toxicity data" instead of just "data"	The edit was made as suggested.
1	15	2	It would be good to include a reference to Table 4-2 to show which MDRs are fulfilled for which PFAS.	The edit was made as suggested.
1	15	3	Would it be better to use the term "aquatic life benchmarks" instead of "aquatic life values" ?	The edit was made as suggested.
1	17	2	Raimondo et al. 2023 is not in the References list. Is it the same as "Raimondo et al., in review"?	Correct. This paper has now been published and the full reference (Raimondo et al. 2024) has been added to the reference list. Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884
1	17	3	Wilming et al 2016 does not mention PFOS and PFOA	The text was edited to remove reference to PFOA and PFOS.
1	17	3	Wilming et al 2016 is not included in the References list	Missing references were added to the document.
1	17	3	Wilming et al 2016 defines parameters for listed species. The use of these as general parameters is likely unproblematic, but it should be stated that these parameters are used in a slightly different context here compared to in Wilming et al	The designation as a listed species does not materially affect the analysis conducted by Willming et al. (2016), and reference to this difference may introduce uncertainty regarding why it is being noted.
1	18	1	Bejarano and Wheeler, 2020 is not in the References list	Missing references were added to the document.
1	18	Figure 2-1	The figure is not referred to in the text. A reference to the figure should be included where appropriate.	The edit was made as suggested.
1	18	Figure 2-1	The last sentence in the figure text is not written in bold and appears to be incomplete. Should it be "...develop a log-linear model"?	The edit was made as noted.

III. Specific Observations				
Reviewer	**Page**	**Paragraph**	**Comment or Question**	**EPA Response**
1	21	Table 3-1	It would be good to include references in this table, alternatively to refer to appendix A in the table text.	The edit was made as suggested.
1	21	Table 3-1	The heading “Toxicity” is a bit ambiguous. It would be more intuitive to write e.g., EC50/LC50	The table heading was edited to indicate “Acute Toxicity”.
1	23	1	The sentence “The Office of Water completed a Data Evaluation Record (DER) for each species by chemical combination from the studies identified by ECOTOX for the eight PFAS compounds undergoing evaluation.” is not clear to me. What does “chemical combination” refer to in this context?	The text was edited to clarify that the EPA was referring to chemical mixtures (as opposed to tests with single compounds).
1	23	1	The sentence “Further, only single chemical toxicity tests with PFOA were considered for possible inclusion in benchmark derivation” is confusing. Is this statement correct?	The reference to PFOA was deleted. The text is otherwise correct as stated.
1	24	2	The title for Appendix A used here (“Appendix A: Acceptable Freshwater Acute PFOA Toxicity Studies”) is not correct.	The title was corrected.
1	25	2	Were all toxicity values (e.g., LC, EC, NOEC, LOEC) treated as the same? Please explain	The accepted empirical studies were all based on LC50 or EC50 values. These studies are summarized in Appendix B. Data based on NOECs and LOECs were not determined to be acceptable for quantitative use (see Appendix C) and were not used for calculating the benchmark value.
1	25	2	The last sentence is missing a word (“is”?) (i.e., “of the corresponding benchmarks is stated for each study at the end”?)	The edit was made as suggested.
1	27	2	The EPA 1985 approach should be summarized here. At least the calculation procedure (as detailed in EPA 1985) should be stated. It should be clear to the reader why the four most sensitive values are focused on.	Text was added to refer the reader to EPA’s Aquatic Life Criteria Guidelines for a more detailed description, in order to focus the reader on the primary analyses being conducted within this document.
1	28	1	In the first sentence, I suggest reminding the reader that eight MDR groups are required fulfilled.	The edit was made as suggested.
1	28	2	Please explain why the FAV was divided by two (what is the reasoning behind this approach?)	This is a standard procedure for the derivation of acute criteria, as described in EPA’s Aquatic Life Criteria Guidelines, the purpose of which

III. Specific Observations				
Reviewer	**Page**	**Paragraph**	**Comment or Question**	**EPA Response**
				is to estimate from an LC50/EC50 a concentration that would not impact aquatic life. Additional text was added to the document to clarify this point.
1	40	1	The sentence "A Web ICE model was not available to fulfill the MDR for a third family in the phylum chordata." Is somewhat confusing as it is according to the table already fulfilled using empirical data.	The noted sentence was determined to not be necessary and was removed from the document.
1	40	2	Here and for the other PFAS, explain why it is stated that "GMAVs for the four most sensitive genera were within a factor of 2.0 of each other" (i.e., it is according to the criteria in the guidelines)	This statement provides a general indication of the similarity or "agreement" between the four values used for calculating the final value, and such a statement summarizing the similarity of these values is typically provided as a general statement in criteria documents.
1	81	1	A brief discussion on the results would be appropriate. For example: - The long chained PFAS are more acute toxic than shorter (PFBA, PFBS). This is in agreement with scientific literature. Previously published values for PFOS and PFOA indicate that these are the most toxic of the PFSA and PFCA, respectively. - According to the results published here, FTUCA and FTCA are the most toxic. Is this as expected?	A discussion of chain length and toxicity value magnitude was added to the summary section as suggested.
1	82	1	The sentence "The resulting acute benchmarks, although consistently higher, were also small, with each of the benchmarks falling within a factor of < 2.1 of one another, indicating close agreement between values calculated using either approach." Should be revised. What were small, the difference between values calculated using the different approaches?	The text (now on page 87) was edited to state "The *differences between the* resulting acute benchmarks, although consistently higher using the scaled approach, were also small,…"
1	83	1	A conclusion for the use of the binning method should be included.	Further discussion about use of the binning approach was added to the document.
1	86	References	The reference list needs to be updated as several references used in the text are not included here.	Missing references were added to the document.

III. Specific Observations				
Reviewer	**Page**	**Paragraph**	**Comment or Question**	**EPA Response**
1	Appendix A		The references are not included in any Reference list. This should be corrected.	Missing references were added to the document.
1	Appendix B		Please explain why the four most sensitive values are summarized	Text was added to clarify how the four values are used in the value calculation.
1	C-2	Green alga	The text explaining the Deficiency for this study "Initially identified as Quantitative" is confusing.	The text was edited to clarify the final decision for use of the value.
1	C-3	Green alga	Why is the test duration written in red?	The text was edited to correct to black lettering.
1	E-1		Only the Web-ICE version 3.3 is available online. I assume that the small differences I found between online model parameters and parameters in the Table is due to the updated values in the v4 model?	That is correct. There are very small differences between the version 3.3 and 4.0 model outcomes based on changes (primarily increases) in the underlying data.
1	F-1		The claim "In these situations, a user can either enter the measured toxicity value (LC50/EC50) into the ICE model as µg/L and allow the regression to extrapolate beyond the range of the model or enter a "scaled" toxicity value (i.e. enter the measured LC50 value as mg/L)." should be supported by a literature reference.	The supporting paper has now been published and the full reference (Raimondo et al. 2024) has been added to the reference list. Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884
1	F-73	1	Using the scaled approach, the eight MDRs were fulfilled for seven of the evaluated compounds (not six)?	The edit was made as suggested.
1	G-7	Footnote b	It would be more intuitive for the reader to use "Species x SMAV", similar to what is done for PFAS	The edit was made as suggested.
1	G-9		The table lacks references to the footnotes	References to all footnotes were added to this table.
1	G-12		Figures G-1 and G-2 are too small. Consider showing panel B below panel A	The noted figures were moved to separate pages and the figure sizes were increased.
2	19	Box	Some classes are quite species rich and a single representative is likely not sufficient (see insects)	There was variation in the size of the database used and notable limitations to the data available for some species groups (notably insects); however, all acceptable data were considered for inclusion in the benchmark calculations.

III. Specific Observations				
Reviewer	**Page**	**Paragraph**	**Comment or Question**	**EPA Response**
2	20	3	When is Raimondo et al (in review) going to be released? It seems like this should be available information for the reader	The supporting paper has now been published and the full reference (Raimondo et al. 2024) has been added to the reference list. Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884
2	throughout		Why are no extra uncertainty factors used other than FAV/2 given the uncertainties associated with the process?	Additional uncertainty factors are not used as part of EPA's Aquatic Life Criteria Guidelines criteria derivation process, which was the basic process used to derive the benchmark values, beyond the development of the toxicity database. As previously discussed, following the peer review the EPA has developed application factors to account for two genera (the mayfly, *Neocloeon* and the cladoceran, *Moina*) that were indicated by direct empirical test data to have markedly greater sensitivity to PFOA and PFOS. These application factors were used to lower the PFAS values to account for these particularly sensitive genera, for which PFAS toxicity data are not available.
2	throughout		Within class extrapolation is a whole lot of biodiversity to lump together	Data were grouped according to procedures presented in EPA's Aquatic Life Criteria Guidelines, which was the basic process used to derive the benchmark values, beyond development of the toxicity database.
3	x	2	Are benchmarks expected to carry the same weight as AWQC if adopted as state or tribal standards?	As indicated in the document, the freshwater acute benchmarks are informational values that states and Tribes could choose to consider in their state water quality protection programs; however, the acute benchmarks magnitudes are less certain than Water Quality Criteria since they are based on both empirical and estimated toxicity data. Text has been added to clarify the difference between benchmarks and water quality criteria.
3	14	1	Should there be inclusion of EPA's GLI approach?	This document focuses on consideration of recently-developed NAMs-based methods. Inclusion of the GLI approach is beyond the scope of this document.

III. Specific Observations				
Reviewer	**Page**	**Paragraph**	**Comment or Question**	**EPA Response**
3	17	2	It is difficult to evaluate this statement since the Raimondo et al report is unpublished and not supplied.	The supporting paper has now been published and the full reference (Raimondo et al. 2024) has been added to the reference list. Raimondo, S., Lilavois, C. and Nelson, S.A. (2024), Uncertainty analysis and updated user guidance for Interspecies Correlation Estimation (ICE) models and low toxicity compounds. Integr Environ Assess Manag. Accepted Author Manuscript. https://doi.org/10.1002/ieam.4884
3	23	1	The document references USEPA 2016c for information on how data were evaluated. The reference is not in the reference list and was not provided. It is critical to assess the acceptability of the empirical data accepted in Appendix A. USEPA 1985 was not included in the reference list.	Missing references were added to the document.
3	23	2	The document references USEPA 2018b however, the reference is not in the reference list and was not provided.	Missing references were added to the document.
4	xi	Second	Period missing at the end of "…*minimum data requirements (MDRs) to calculate aquatic life criteria*"	The edit was made as suggested.
4	xii	First	Space needed between "…*(Guidelines)(U.S. EPA 1985*)."	The edit was made as suggested.
4	xii	Table Ex-1-1	Not clear how the "*Duration*" and "*Frequency*" were determined for the recommended benchmarks. Is this standard for USEPA acute freshwater aquatic life benchmarks?	The duration and frequency assumptions are consistent with those presented in EPA's Aquatic Life Criteria Guidelines for acute values and with those used for most criteria values.
4	14	First	A closing bracket missing in "• *Perfluorohezanesulfonic acid (PFHxS (CAS# 355464, 108427538, 3871996, 82382125)"*	The edit was made as suggested.
4	15	First	The first word in f), g), and h) is not capitalized as in a) to e). *a) "insect (e.g., mayfly, dragonfly, damselfly, stonefly, caddisfly, mosquito, midge, etc.)* *b) family in a phylum other than Arthropoda or Chordata (e.g., Rotifera, Annelida, Mollusca, etc.)*	The edit was made as suggested.

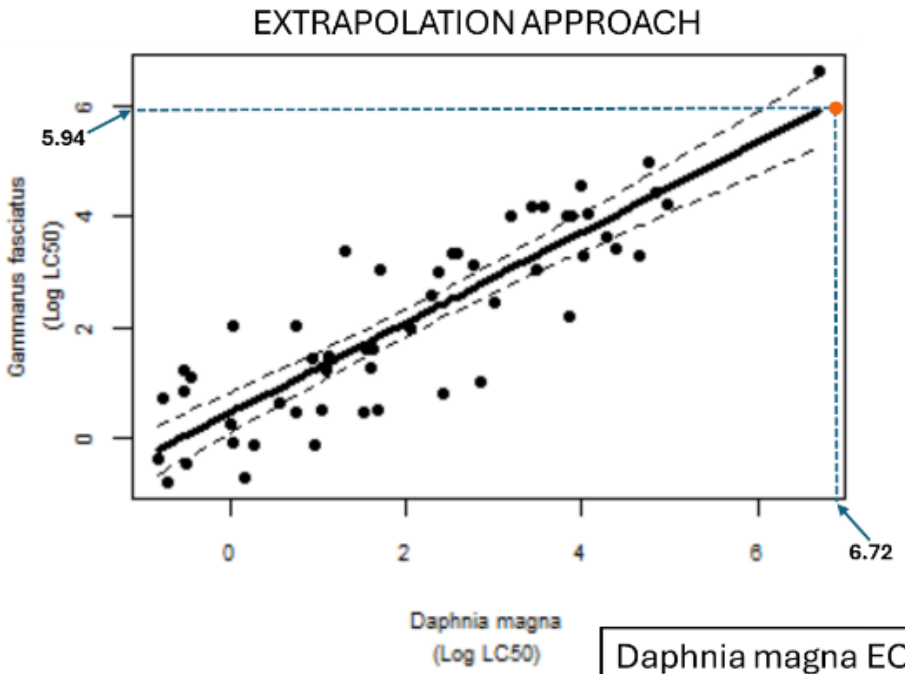
III. Specific Observations				
Reviewer	**Page**	**Paragraph**	**Comment or Question**	**EPA Response**
			c) *family in any order of insect or any phylum not already represented*	
4	20	Final sentence on the page	The final sentence of this section is "*Benchmark values for the eight PFAS using this alternative approach are summarized in* ***Section 5.10****."* I would specifically reference Table 5-26 in section 5.10 at the end of the sentence above. It would make it easier for the reader to find the benchmarks derived using the scaled approach. Or reference section F.9 and/or Table F-29.	A reference to Section F.9 was added to the document.
4	23	First	The first sentence in the paragraph states, "*Empirical studies available for the eight PFAS were identified using the ECOTOXicology Knowledgebase*…" I assume that the "eight PFAS" refers to the compounds for which the benchmarks are being set. However, later in the paragraph, there is a sentence that states, "*Further, only single chemical toxicity tests with PFOA were considered for possible inclusion in benchmark derivation*." PFOA is not one of the eight PFAS for which a benchmark is being set, so it is not clear why acute toxicity data for PFOA is being used. It was stated earlier in section 3 that validation was conducted using measured and predicted values for PFOS and PFOA, but it is not clear at this point in section 4 how acute toxicity data for PFOA will be used in deriving benchmarks for the eight PFAS that are the focus of this document. You may want to make that clear to the reader.	The text was edited and references to PFOA and PFOS were removed from this section of the text.

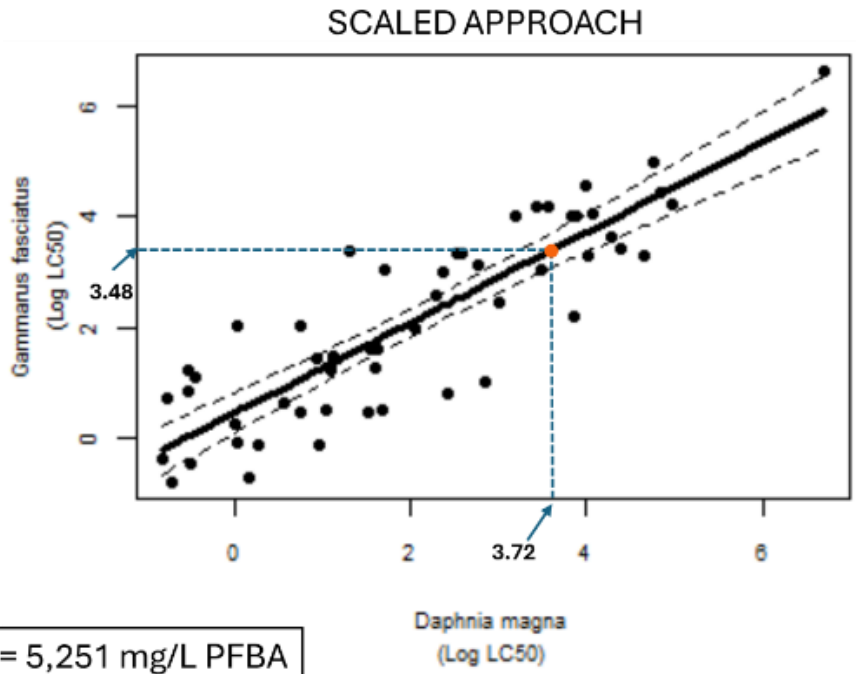
III. Specific Observations				
Reviewer	**Page**	**Paragraph**	**Comment or Question**	**EPA Response**
4	24	First	The paragraph makes references to whether this process should consider studies that only report nominal concentrations of the PFAS in the toxicity study. The rationale given for choosing to consider studies that only report nominal concentrations is a case study that was conducted with measured and nominal concentrations for PFOS and PFOA. While the rationale is sound for considering studies that only include nominal concentrations of PFOS or PFOA, care should be taken to extrapolate to the entire class of chemicals, i.e., PFAS. The eight PFAS that are focus of this process have different physical and chemical properties than PFOS and PFOA, consequently, the probability of the nominal concentrations being with 20% of the measured concentrations for the eight PFAS may be different than PFOS and PFOA, which was 82 and 83%, respectively.	Toxicity tests used in EPA aquatic life criteria documents are typically based on measured chemical concentrations only. For PFOA and PFOS, a substantial proportion of the available data was reported as nominal test concentrations. The EPA examined the issue of whether nominal (unmeasured) and measured concentrations are in close agreement with each other, based on an analysis of available studies reporting both nominal and measured concentrations (U.S. EPA 2024a,b). PFOA and PFOS toxicity having pairs of nominal and corresponding measured concentrations were compared through: (1) linear correlation analysis and; (2) an assessment of measured concentrations as a percent of its paired nominal concentration. Linear correlation between measured and corresponding nominal concentrations suggested a high degree of precision between paired observations across all test conditions, in most cases within the test acceptability threshold identified by the EPA's OCSPP's Ecological Effects Test Guidelines. Similar results were also observed by Jarvis et al. (2023). Although a parallel analysis could not be conducted based upon the more limited datasets for the evaluated PFAS, the PFAS concentrations in test waters are expected to remain relatively constant over the course of acute exposures given its ability to resist breakdown and transformation (Ahrens et al. 2011). Accordingly, PFAS toxicity tests were not excluded from quantitative use in criteria derivation on the basis of unmeasured test concentrations alone based on results of the analysis with PFOA and PFOS. Ahrens, L. 2011. Polyfluoroalkyl compounds in the aquatic environment: a review of their occurrence and fate. J. Environ. Monit. 13(1): 20-31. Jarvis, A.L., J.R. Justice, B. Schnitker and K. Gallagher. 2023. Meta-analysis comparing nominal and measured concentrations of perfluorooctanoic acid and perfluorooctane sulfonate in aquatic toxicity

III. Specific Observations				
Reviewer	**Page**	**Paragraph**	**Comment or Question**	**EPA Response**
				studies across various experimental conditions. Environ. Toxicol. Chem. 42(11): 2289-2301. USEPA (U.S. Environmental Protection Agency). 2024a. Aquatic life ambient water quality criterion for PFOA-freshwater. EPA-842-R-24-002. Office of Water. Washington, D.C. USEPA (U.S. Environmental Protection Agency). 2024b. Aquatic life ambient water quality criterion for PFOS-freshwater. EPA-842-R-24-003. Office of Water. Washington, DC.
4	27	First	Comma needed between "*toxicity database*" and "*benchmark values"* in the first sentence on page 27.	The edit was made as suggested.
4	27	Second	It would be useful to report at some point the number of acceptable empirical LC50 or EC50 used in each SSD that were based on nominal concentrations. The data is available in the document but the reader would have to take a great deal of time to compile these numbers.	This information is summarized in Table 4-2 and all studies used quantitatively are presented in Appendix A.
4	27	Second	Why not use the lowest acute value for a species instead of the mean? Using the lowest acute value for a species would be a more conservative approach in terms of protection of sensitive species. The same question could be asked about the genus mean acute values.	Mean species and genus values are used as part of EPA's Aquatic Life Criteria Guidelines criteria derivation process, which was the basic process used to derive the benchmark values, beyond the development of the toxicity database.
4	Entire document	Entire document	Review the document to ensure that Greek letters are used consistently, e.g., µg/L vs. ug/L	Values were reviewed and changed to µg/L, as appropriate.
4	Appendices A to C		I thought these were very valuable appendices. They clearly laid out the studies that were considered for inclusion in the derivation of the benchmarks and why studies were eventually not included.	Thank you for your comment.

III. Specific Observations				
Reviewer	**Page**	**Paragraph**	**Comment or Question**	**EPA Response**
5	70 and 80	Figure 5-7 and Figure 5-8	Is this a bimodal response- as this is model based SSD, it is challenging to confirm? Invertebrates and fish may have different mode of action for 8:2 FTUCA and 7:3 FTCA	It is possible that results are showing a bimodal distribution for fish and invertebrates, but it may be overstepping to draw any conclusions based on the available data.

Appendix A
Example of Extrapolation vs Scaled Approach Value Calculation: *D. magna* with PFBA





Daphnia magna EC50 = 5,251 mg/L PFBA

EXTRAPOLATION APPROACH

- **WebICE INPUT**: Log of 5,251,000 = 6.72
- 6.72 is beyond the range of values for *D. magna* in WebICE model (prediction is "Extrapolated")
- **WebICE OUTPUT**: 5.94 (Log of 878,785 µg/L)
- Value for *G. fasciatus* used for SSD = **878.8 mg/L PFBA**

SCALED APPROACH

- Value is "Scaled" to fall with the range of values for *D. magna* in the WebICE model
- **WebICE INPUT**: Log of 5,251 = 3.72
- **WebICE OUTPUT**: 3.48 (Log of 3,060 µg/L)
- Prediction is "Scaled" to be consistent with input units
- Value for *G. fasciatus* used for SSD = **3,060 mg/L PFBA**