TASK ORDER 68HERH21F0090 UNDER
CONTRACT EP-C-17-017

EXTERNAL PEER REVIEW OF EPA'S DRAFT AQUATIC LIFE AMBIENT WATER QUALITY CRITERIA FOR PERFLUOROOCTANOIC ACID (PFOA)

FINAL PEER REVIEW REPORT

August 26, 2021

Submitted to:
U.S. Environmental Protection Agency
Office of Water, Office of Science and Technology
1200 Pennsylvania Avenue, NW
Washington, DC 20460
Attn: James Justice
Justice.JamesR@epa.gov

Submitted by:
Eastern Research Group, Inc.
110 Hartwell Avenue
Lexington, MA 02421

ERG
www.erg.com

CONTENTS

1.0 INTRODUCTION

The U.S. Environmental Protective Agency (EPA) Office of Water (OW) is charged with protecting ecological integrity and human health from adverse anthropogenic, water-mediated effects, under the purview of the Clean Water Act (CWA). In support of this mission, EPA has developed draft water quality criteria to protect aquatic life and aquatic-dependent wildlife from the presence of Perfluorooctanoic Acid (PFOA) and Perfluorooctane Sulfonate (PFOS) in freshwater and saltwater environments. Derivation of these criteria is described in two draft documents: *Aquatic Life Ambient Water Quality Criteria for Perfluorooctanoic Acid (PFOA)* and *Aquatic Life Ambient Water Quality Criteria for Perfluorooctane Sulfonate (PFOS)*.

This report documents the results of an independent letter peer review of the EPA's draft *Aquatic Life Ambient Water Quality Criteria for Perfluorooctanoic Acid (PFOA)*. Eastern Research Group, Inc. (ERG), a contractor to EPA, organized this external peer review for EPA OW and developed this report. Independent peer review of the draft *Aquatic Life Ambient Water Quality Criteria* for *Perfluorooctane Sulfonate (PFOS)* document is covered in a separate report.

Section 2.0 of this report presents the individual reviewer comments organized by charge question. Appendix A provides EPA's charge to reviewers and Appendix B presents the reviewer comments organized by reviewer.

1.1 Development of the Draft Documents

Toxicity studies used to derive the PFOA and PFOS criteria were carefully evaluated and thoroughly reviewed to ensure studies were of sufficient data quality to use in criteria derivation. Scientists from EPA OW and Office of Research and Development (ORD) conducted an extensive review of the PFOA and PFOS toxicity studies. Additionally, EPA obtained replicate-level (or treatment-level, when replicates were unavailable) concentration-response (C-R) data from publications, supplemental materials, or via contacting authors so that EPA could independently fit C-R models to estimate acute LC_{50} and chronic EC_{10} values that were used to derive the criteria to ensure endpoints used were statistically sound. Individual C-R models and resultant point estimates were also reviewed and discussed between OW and ORD to ensure the most statistically robust models informed the derivation of the PFOA and PFOS criteria. In addition to contacting study authors for C-R data (when not reported in the open literature), EPA also consulted primary authors for methods clarifications in many instances during the data quality review phase to ensure that the studies used to derive criteria were of high quality.

Overall, due to the paucity of measured freshwater toxicity data, EPA included a number of tests with unmeasured treatments to derive criteria to ensure the dataset was representative of a range of taxa and there were sufficient data to develop criteria. EPA also conducted meta-analyses to evaluate the relationship between nominal and measured test concentrations using tests with measured treatment concentrations. These meta-analyses (described in detail as Appendix L of the PFOA criteria document and Appendix O of the PFOS criteria document) suggested measured concentrations were similar to nominal concentrations and that the use of unmeasured tests, in light of data limitations, was appropriate. Additionally, estuarine/marine toxicity data limitations did not allow for the direct derivation of acute or chronic estuarine/marine criteria for PFOA or PFOS. Therefore, to develop recommendations that states and tribes could use in adopting protective values for estuarine/marine waters, EPA developed acute PFOA and PFOS protective benchmarks using a New Approach Methodology (detailed in Appendix K of the PFOA criteria document and Appendix L of the PFOS criteria document).

Addressing data limitations to derive robust criteria/benchmarks, extensively reviewing studies, and calculating point estimates meant that the derivation of the PFOA and PFOS aquatic life criteria were developed via comprehensive, rigorous process that included collaborations across EPA scientists in OW and

ORD. Beyond detailed discussions between OW and ORD, the PFOA and PFOS drafts also underwent two rounds of review with the EPA Scoping Workgroup (consisting of additional scientists from both OW and ORD) and one round of review with a group of internal EPA reviewers that included representatives from the OW, ORD, other EPA Program Offices, and EPA Regions.

Subsequently, EPA contracted with ERG to organize an independent external peer review of both draft documents. Results of the PFOA review are described in this report. Results of the PFOS review are documented in a separate report.

1.2 Peer Reviewers

ERG identified, screened, and selected the following five experts who met technical selection criteria provided by EPA and had no conflict of interest in performing this review:

- **Jason Conder, Ph.D**.; Principal, Geosyntec Consultants
- **Anu Kumar, Ph.D.**; Principal Research Scientist, Environment Protection and Technologies, Commonwealth Scientific and Industrial Research Organization (CSIRO)
- **Ryan Prosser, Ph.D.**; Associate Professor, University of Guelph
- **Christopher J. Salice, Ph.D.**; Director, Environmental Science and Studies Program, Towson University
- **Jamie G. Suski, Ph.D.**; Senior Scientist, EA Engineering, Science, and Technology, Inc.

ERG provided reviewers with instructions, the draft *Aquatic Life Ambient Water Quality Criteria for Perfluorooctanoic Acid (PFOA)*, and the charge to reviewers (Appendix A of this report) prepared by EPA. Reviewers worked individually to develop written comments in response to the charge questions. After receiving reviewer comments, ERG compiled responses by charge question (see Section 2.0) and included the responses organized by reviewer (Appendix B of this report).

2.0 REVIEWER COMMENTS ORGANIZED BY CHARGE QUESTION

This section organizes reviewer comments by charge question (see Appendix B for reviewer comments organized by reviewer).

2.1 Please comment on the overall clarity of the document as it relates to the derivation of each criterion.

2.1. Clarity of Document as it Relates to the Derivation of Each Criterion	
Reviewer	**Comments**
Reviewer 1	Overall, the document is clear and the reader can follow the logic of criteria derivation, and track the values used back to the cited research articles or values calculated by EPA.
Reviewer 2	I thought that the document was well written and laid out. I thought that the document clearly laid out the approach that the EPA used to derive each criterion. I thought it clearly outlined the approach that the EPA chose in deciding which data to use in their derivation and how these data would be used in derivation.

2.1. Clarity of Document as it Relates to the Derivation of Each Criterion	
Reviewer	**Comments**
	The appendices are very useful in providing added detail and the data that were used in the derivation of the criteria. The appendices allow for a high level of transparency around how the criteria were generated. In Table 3-1, the acronym "GMAV" was used in the caption, but I could not locate where this acronym was defined earlier in the document. The captions of figures and tables are not sufficiently detailed. Figures and tables should be able to stand on their own. Also, the use of acronyms in the caption of tables and figures decreases clarity, e.g., Fig. 3-5. The use of acronyms in the figure or table is fine, as long as they are defined in the caption of the figure or table.
Reviewer 3	I have confidence is the PFOA draft criteria, these are more in-line with thresholds put forth by other agencies. Importantly, a discussion on the difference in derivation of the thresholds would be welcome; for instance, others are derived using Species Sensitivity Distributions of the complete dataset. In comparison, EPA uses a subset of data on the genus identified as most sensitive. Table 1 – does not list units of thresholds. Define CMC at first use Table 3-6 reverses order of sensitive taxa compared to the previous tables.
Reviewer 4	EPA has drafted the PFOA aquatic life criteria to be consistent with methods described in EPA's "*Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses*" (U.S. EPA 1985). I congratulate the EPA Team for a very thorough, comprehensive analysis of toxicological data to derive each criterion. • The report is technically sound and is very clearly written. • The criteria have been derived using strong science-based evidence. • Sub-sections on overview of PFAS, PFAS nomenclature, problem formulation, exposure pathways, transformation and degradation of PFOA precursors in the aquatic environment sources, concentration reported in environment and existing criteria (both nationally and internationally) help to set the scene before toxicological data is presented and assessed for developing various criterion. • The freshwater acute water column-based criterion, the chronic water column-based chronic criterion, the chronic fish whole-body tissue criterion, the chronic fish muscle tissue criterion and the chronic invertebrate whole-body tissue criterion have been developed and reported in this report. • Acute and chronic MDRs for PFOA estuarine/marine criteria derivation were not met due to fewer empirical PFOA toxicity data. EPA Team developed an acute aquatic life benchmark for estuarine/marine environments based on Interspecies Correlation Estimation (ICE) model.

2.1. Clarity of Document as it Relates to the Derivation of Each Criterion	
Reviewer	**Comments**
	• The relative sensitivity of freshwater plants to PFOA exposures indicated plants are less sensitive than aquatic vertebrates and invertebrates so plant criteria were not considered. • EPA Team has provided extensive background information on toxicity data assessment and collated this information in various appendices such as Appendix A-Summary Table of Acceptable Quantitative Freshwater Acute PFOA Toxicity Studies; Appendix B-Detailed PFOA Acute Toxicity Study Summaries and Corresponding Concentration-Response Curves (when calculated; Appendix C: Acceptable Freshwater Chronic PFOA Toxicity Studies); Appendix D- Acceptable Estuarine/Marine Chronic PFOA Toxicity Studies; Appendix E-Acceptable Freshwater Plant PFOA Toxicity Studies; Appendix F-Acceptable Estuarine/Marine Plant PFOA Toxicity Studies; Appendix G-Summary Table of Qualitative Freshwater PFOA Toxicity Studies; Appendix H-Other Estuarine/Marine PFOA Toxicity Studies Unused PFOA Toxicity Studies; Appendix I-Unused PFOA Toxicity Studies; Appendix J- EPA Methodology for Fitting Concentration-Response Data and Calculating Effect Concentration-Fitting Concentration Response Data in R; Appendix K-Derivation of Acute Protective PFOA Benchmarks for Estuarine/Marine Waters through a New Approach Method (NAM); and Appendix L Meta-Analysis of Nominal Test Concentrations Compared to Corresponding Measured Test Concentrations.
Reviewer 5	Overall, similar to PFOS, the document for PFOA is very well written, generally free of grammatical errors and clear. It is long but not as long as PFOS and, therefore, easier to digest. I think the background material for both chemicals is especially good and provides an excellent overview and summary for readers less familiar with PFAS. In particular, the summary of PFOA concentrations in water bodies and other environmental media is wonderfully useful. To me, the derivation of the criteria for PFOA is easier to follow than PFOS, mostly because there are fewer data and, hence, fewer assumptions, calculations, and evaluations that need to be made. Unlike for PFOS, the criteria developed by EPA for PFOA are more similar to values derived by other jurisdictions for both acute and chronic values. To me, this lends strength to EPA's criteria values and is also consistent with the scientific method, in general. If we are all following the same threads of logic, using similar analyses, and looking at the same data, it makes sense that criteria values would be similar. So, again, in the case of PFOA the congruence with other published criteria values is a strength. I will use this as an argument for EPA to reconsider the criteria (and the data supporting it) for PFOS as, for that chemical, EPA's numbers are higher than other jurisdictions.

2.2 Please comment on the approach used to derive the draft criterion for PFOA. Please provide detailed comments.

- **Is the technical approach used to derive the criterion logical?**
- **Does the science support the conclusions?**

- **Is it consistent with the protection of freshwater aquatic life from acute, chronic, and bioaccumulative effects?**

2.2. The Technical Approach Used to Derive the Draft Criterion for PFOA	
Reviewer	**Comments**
Reviewer 1	• **Is the technical approach used to derive the criterion logical?** Yes, the technical approach used to derive the criteria elements is generally logical. I disagree with some of the elements of the analyses, as noted in my detailed comments (see below, responses to charge question 8) • **Does the science support the conclusions?** In general, the science is supportive of the general conclusions. As noted in my below detailed responses to other charge questions, I believe the science is not supportive of the work in a few key instances including: 1. I believe the Criterion Continuous Concentration (CCC) should be potentially re-calculated considering my comments provided in response to charge question 5a. 2. The science does not support the assumption of a 10-year recovery time for PFOA in aquatic systems. 3. The generation of tissue criteria is weakly supported, and the uncertainty associated with these criteria should be emphasized. 4. The NAM-generated marine Final Acute Value (FAV) and FAV/2 values (Appendix K) are highly uncertain. 5. It is unclear if the EPA-calculated Effective Concentration 10% (EC_{10}) values are supported; additional details on the modeling and the variability and fit of each EC_{10} model need to be provided. • **Is it consistent with the protection of freshwater aquatic life from acute, chronic, and bioaccumulative effects?** The criteria derived are aimed at protecting aquatic life (e.g., fish, invertebrates) from the direct acute and chronic toxicity of PFOA in water. Generally, the values applied are protective and are generally similar to protective values derived by other regulatory organizations and independent (i.e., academic, private sector) scientists. Although, as based on my comments, I believe there is room for improvement. The criteria derived for tissues attempt to provide criteria that take into account bioaccumulation so that measurements in tissue can be interpreted with respect to the potential for potential effects; however, the uncertainty with the tissue criteria is high. The water and tissue criteria are not intended protective of bioaccumulative effects that may affect higher trophic levels, such as wildlife that may consume aquatic life.
Reviewer 2	Yes, the technical approach used by the EPA to derive the criterion is logical and defensible. The approach is also clearly laid out in the document. Dividing the 5th centile of the acute GSD by 2 is sufficiently conservative to ensure the protection of 95% of species, based on the data

2.2. The Technical Approach Used to Derive the Draft Criterion for PFOA	
Reviewer	**Comments**
	currently available. The use of the ACR from Brachionus calyciflorus to construct a chronic GSD is a valid approach, along with being the most conservative option. Yes, I think the science supports the EPA's conclusions. However, there appears to be several studies that were not considered by the EPA. I have listed these studies below. Yes, I think the approach taken by the EPA is sufficiently conservative to be protective of freshwater aquatic life from acute, chronic, and bioaccumulative effects based on the data that was available at the time. It was a good idea to evaluate the influence on non-North American species on the derivation of the criteria.
Reviewer 3	• **Is the technical approach used to derive the criterion logical?** Why did EPA derive the LC_{50} for Chydorus when an EC_{50} was provided by the authors; however, accepted the EC50s for the two mussel species? This is an inconsistency. • **Does the science support the conclusions?** See response immediately below • **Is it consistent with the protection of freshwater aquatic life from acute, chronic, and bioaccumulative effects?** Overall, the draft criteria are in agreement with other thresholds generated using the species/genus sensitivity distributions; which seems like a more robust approach given data are not restricted to a subset of studies (albeit most sensitive). Although new data should be evaluated and potentially incorporated into these criteria calculations it is unclear if those data would alter these currently drafted thresholds.
Reviewer 4	This EPA report provides a critical review of toxicity data identified in EPA's literature search for PFOA, including the anionic form (CAS No. 45285-51-6), the acid form (CAS No. 335-67-1), and the ammonium salt (CAS No. 3825-26-1). It quantifies the toxicity of PFOA to aquatic life, and provides criteria intended to protect aquatic life from the acute and chronic toxic effects of PFOA. The detailed assessment is as follows: • These criteria have been derived using robust methods and the best available toxicity data on aquatic life. • The approach used to derive the draft criterion for PFOA is very logical and consistent with the protection offered by acute and chronic aquatic life criteria derived using empirical data, as prescribed in the 1985 *Guidelines.* • Exclusion and inclusion criteria are appropriately discussed in the context of the toxicological data reported in the literature and provide additional evidence on the selection of toxicity data criteria development. • With limited toxicity datasets to North American resident species, non-North American resident species served as taxonomically-related surrogate test organisms. For example, *Oryzias latipes* is a common ecotoxicity test species that served as a surrogate for untested fish species residing in North America.

2.2. The Technical Approach Used to Derive the Draft Criterion for PFOA	
Reviewer	**Comments**
	• The acute measures of effect on aquatic organisms selected included the lethal concentration (LC_{50}), effect concentration (EC_{50}), or inhibitory concentration (IC_{50}) estimated to produce a specific effect in 50 percent of the test organisms • The endpoint for chronic exposures incorporated the effect concentration estimated to produce a chronic effect on survival, growth, or reproduction in 10 percent of the test organisms (EC_{10}). This approach has been also consistent with the harmonized guidelines from OECD and the generally preferred effect level for countries such as Canada, Australia, and New Zealand. • Reported (No Observed Effect Concentrations) (NOECs) and (Lowest Observed Effect Concentrations) (LOECs) were only used for the derivation of a chronic criterion when a robust EC_{10} could not be calculated for the genus. • Furthermore, EPA independently calculated these toxicity values if sufficient raw data were available for EPA to conduct statistical analyses. EPA's independently-calculated toxicity values were used preferentially, where available. • EPA developed protective tissue-based criteria through a bioaccumulation factor approach. This was based on the application of evaluation criteria for screening bioaccumulation factors (BAFs). • The freshwater Final Acute Value (FAV) for PFOA was calculated as 91.34 mg/L and freshwater acute criterion water column magnitude (criterion maximum concentration, CMC), as 46 mg/L PFOA, using the procedures described in the 1985 Guidelines. This values is expected to be protective of 95% of freshwater genera potentially exposed to PFOA under short-term conditions of one-hour of duration, if the one-hour average magnitude is not exceeded more than once in three years. • Toxicity data were available for only two families, an estuarine/marine FAV could not be calculated to derive an estuarine/marine acute criterion. Further benchmark was developed using predictive approach and discussed later in this document. • Tissue-based criteria were also developed using comprehensive methods and assessment is provided as response to charge Question 5.
Reviewer 5	The overall approach to derive criteria for PFOA is logical…except for the use of only the 4 most sensitive endpoints and then a model was fitted (unspecified, I believe) to obtain the 5% most sensitive species (in general). I am not familiar with this as an approach as I have not seen other scientists use this and, instead, I have more commonly seen the application of a species sensitivity distribution based on more data and usually following an s-shape. Indeed, many well-cited papers on toxicity thresholds and criteria have used this approach. I suspect EPA has justification for the approach used in the document and that it is well-supported. I would suggest adding any details, beyond just citing the 1985 guidelines, that supports a focus on just the 4 most sensitive toxicity endpoints for the criteria development. My apologies if this information is in the document and I missed it. . One point of clarification is needed in the explanation of the regression analysis (p. 64). The document states: "When LOECs and NOECs were used, a Max Aceept. Tox Conc. (MATC) was calculated, with is the geometric mean of the NOEC and LOEC. For the calculation of chronic criterion, **point estimates** were selected for use as the measur of effect in favor of the

2.2. The Technical Approach Used to Derive the Draft Criterion for PFOA	
Reviewer	**Comments**
	MATCs…" – WHAT IS MEANT BYT "POINT ESTIMATES"? Isn't the LOEC or NOEC a point estimate? This needs to be clarified here (and for PFOS). On p. 66 – I also recall that in some cases EPA obtained data from plots using web plot digitizer software. If I am correct, this should be explained on this page. One point worth mentioning with regard to the technical approach is that I think EPA was correct to consider non-North American resident species in developing the criteria. While I can understand why some scientists feel strongly about focusing on native species, I also cannot think of a clear example of widely different chemical tolerances among species from different countries. Sure, organisms from contaminated environments are likely to differ compared to the same organisms (species) from uncontaminated areas but barring this, it seem sensitivity to chemicals is not geographically determinated (again, barring extremes). Additionally, I think using the EC10 makes sense for PFAS chronic criteria and perhaps other chemicals as well. That said, I can't say that using a 50% effect level for acute toxicity/exposures makes sense. To me, using an EC20 or LC20 for acute would be more reasonable. As I understand it, the current approach divides the calculated FAVs by 2.0 to further ensure protection. Here again, it would seem easier and more straight-forward to use the EC20 or LC20. Perhaps dividing the FAV by 2.0, however, commonly results in a low threshold (like an LC20, for example). If this is the case, it would be worth EPA mentioning to give some sense of magnitude to what could be considered an arbitrary "safety factor" of 2.0. What highlighted this issue for me was the estuarine mussel species where there was a 27% effect (malformations) at 0.0001 mg PFOA/L but because an EC50 could not be determined, EPA chose to use the highest concentration of 1 mg/L. If I were in charge of managing resources in an estuary, I can't say I'd have much confidence in that 1 mg/L value. Or at least, I would be very uncomfortable. Having said all this, I understand that the frequency and duration of environmental exposures would make is such that the acute and chronic criteria would be protective but please see my comments below with regard to frequency and duration of exposure.

2.3 Please comment on the approach used to derive the draft acute estuarine/marine benchmark for PFOA. Given the limited estuarine/marine test data available, a new approach method was used to support the derivation of an acute estuarine/marine benchmark to provide states and tribes with a protective value. Please provide detailed comments.

- **Is the technical approach used to derive the benchmark logical?**
- **Does the science support the conclusions?**
- **Is it consistent with the protection offered by acute estuarine/marine aquatic life criteria derived using empirical data, as prescribed in the 1985 *Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses*?**

2.3. The Technical Approach used to Derive the Draft Acute Estuarine/Marine Benchmark for PFOA	
Reviewer	**Comments**
Reviewer 1	• **Is the technical approach used to derive the benchmark logical?** The derivation of the acute marine benchmarks (FAV and Criterion Maximum Concentration (CMC)) using the New Approach Method (NAM) is highly uncertain, and I would recommend this analysis not be included as in this document. I do not feel that the analysis and subsequent criteria have high confidence for use in a regulatory application. I understand that similar analyses with other chemicals have about a 90% probability of the predicted effect value being within a factor of 5 of the actual value (Raimondo et al., 2010 – cited in document). Given the calculated CMC (3.4 mg/L), this implies the CMC has about a 90% probability of being within 0.68 to 17 mg/L. If the NAM approach stays in the document, this uncertainty and range of values should be acknowledged in the discussion. I would rather see tentative or provisional acute criterion developed from the limited empirical marine acute data highlighted in Appendix B and other recently published marine acute data. I place higher confidence in empirical data (even if limited to a few studies) and would suggest EPA emphasize it in addition to or in place of the values calculated by the NAM. I am hopeful that as new toxicity information on marine species are developed, these values can be supplanted with a proper and robust criteria calculation. If such a future analysis is possible, it should be noted. • **Does the science support the conclusions?** See above comment. • **Is it consistent with the protection offered by acute estuarine/marine aquatic life criteria derived using empirical data, as prescribed in the 1985 Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses?** The approach seems to be consistent with the approach in the 1985 guidelines. As noted above, the uncertainty with regards to the predictive capability of the interspecies correlations should be acknowledged quantitatively.
Reviewer 2	The technical approach using Web-ICE to determine an acute benchmark for estuarine/marine species is logical. The science has shown that Web-ICE can effectively be used to derive effect measures for additional species using species for which data is available. I think the approach taken by EPA has included sufficient conservatism to address the relatively large amount of uncertainty around the acute toxicity of PFOA to estuarine and marine species. The proposed acute benchmark for estuarine and marine species is an order of magnitude lower than the acute benchmark for freshwater species, which I think underscores the conservatism used by EPA in determining an acute benchmark for estuarine and marine species. That said, the benchmark should be used cautiously due to the relatively large amount of uncertainty and effort should be made to generate acute and chronic toxicity data for PFOA on estuarine and marine species, particularly fish.

2.3. The Technical Approach used to Derive the Draft Acute Estuarine/Marine Benchmark for PFOA	
Reviewer	**Comments**
Reviewer 3	• **Is the technical approach used to derive the benchmark logical?** See comment immediately below • **Does the science support the conclusions?** The PFOA - LOEC reported for *Mytilus* in Fabbri et al. 2014 is 0.1µg/L; in the draft criteria this is listed as >1 mg/L. The justification provided for dismissing this effect given 50% of the test organisms did not experience is not compelling. • **Is it consistent with the protection offered by acute estuarine/marine aquatic life criteria derived using empirical data, as prescribed in the 1985** *Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses***?** No, this is a new approach; however, it follows the spirit of the 1985 guidelines.
Reviewer 4	• EPA applied The ICE model predictions to supplement the available test dataset to help fill missing MDRs and allow the derivation of acute estuarine/marine benchmark recommendations for aquatic life using procedures consistent with those in the 1985 Guidelines**.** Total of 3104 datapoints from 398 models were evaluated. • The draft acute benchmark for estuarine/marine aquatic life is lower than the recommended acute freshwater criterion (46 mg/L), suggesting that estuarine/marine species may be more acutely sensitive to PFOA. According to Hayman et al., 2021, marine species, compared to freshwater, may have a higher sensitivity to PFOA. • There are two more studies published, the toxicity values for marine/estuarine species. ○ Stuart L. Simpson, Yawen Liu, David A. Spadaro, Xinhong Wang; Rai S. Kookana and Graeme E. Batley Chronic effects and thresholds for estuarine and marine benthic organism exposure to perfluorooctane sulfonic acid (PFOS)-contaminated sediments: Influence of organic carbon and exposure routes https://doi.org/10.1016/j.scitotenv.2021.146008 ○ Nicholas T Hayman , Gunther Rosen , Marienne A Colvin , Jason Conder , Jennifer A Arblaster Aquatic toxicity evaluations of PFOS and PFOA for five standard marine endpoints. https://doi.org/10.1016/j.chemosphere.2021.129699 **It is recommended to assess the quality of the toxicity data on marine/estuarine species and recalculate estuarine criteria based on this recently available information.**
Reviewer 5	The most sensitive estuarine/marine species was Mytilus g. in the study by Fabbri et al. (2014). EPA chose to use 1 mg/L PFOA because there was not a 50% effect level even at the highest tested concentration (1 mg/L). HOWEVER, there was 27% effect at the LOEC (0.0001 mg/L). Is this not problematic? If I were concerned about mussels or other bivalves in an estuary, I don't think I would hang my hat on a 1 mg/L PFOA concentration given there was a 27% decrease in normal D-larvae at a concentration several orders of magnitude below that. This raises the issue of why EPA is using a 50% effect level for acute criteria – this seems excessively high,

2.3. The Technical Approach used to Derive the Draft Acute Estuarine/Marine Benchmark for PFOA	
Reviewer	**Comments**
	doesn't it? EPA justified using the EC_{10} for chronic criteria so it seems reasonable to use the EC/LC20 for acute. A 50% effect, if it occurs, is likely to manifest as ecologically relevant for any species…20% may not be protective depending on the species and endpoint. As mentioned before, I understand that there is a frequency/duration element to all the criteria but from a functional standpoint, I don't see how the ferequency and duration elements are helpful because nobody collects or reports environmental data on a relevant temporal scale (every hour; 4-day running average??). See below for more on frequency and duration. Given the data, I believe the new approach methods based on WEB-ICE are appropriate. The estimation tool has been in development and used for a considerable length of time and several publications have supported it's use. Of course it would be better to have more data but, again, given the lack of data for estuarine/marine species, the WEB-ICE approach is likely the best available.

2.4 Please comment on the use of measured and unmeasured toxicity tests to derive the respective criterion. In particular please comment on the supporting justification for using unmeasured toxicity tests in Appendix L.

2.4. The Use of Measured and Unmeasured Toxicity Tests to Derive Respective Criterion	
Reviewer	**Comments**
Reviewer 1	The consideration of toxicity data from experiments in which PFOA measurements were not made seems appropriate. The Appendix L analysis is supportive of the general observation that actual concentrations in the toxicity test waters approximated nominal values for the freshwater studies.
Reviewer 2	I am concerned with the approach of using the agreement of measured and nominal concentrations from studies that measured the concentration of PFOA in their tests to determine whether to use toxicity data from studies that did not measure the concentration PFOA in their tests. My concern stems from this approach having to assume that studies that did not measure the concentration of PFOA in their experiments performed the dosing of PFOA with the same care and skill as those studies that did measure the concentration of PFOA in their experiments and measured concentrations within 20% of nominal. My concern is compound by 79.5% and 60% of the acute and chronic tests, respectively, only reporting nominal test concentrations. The EPA's approach uses the agreement of measured and nominal concentration in a minority of studies to determine whether to include the majority of studies on their assessment. I am assuming that there wouldn't be sufficient data to determine a criterion without using data from studies that did not measure the concentrations of PFOA in their experiment?

2.4. The Use of Measured and Unmeasured Toxicity Tests to Derive Respective Criterion	
Reviewer	**Comments**
	I think the approach that the EPA has used to determine the level of agreement between the nominal and measured concentration of PFOA in the studies that measured the concentration is logical and valid. It is encouraging that the agreement on average is high. Again, my largest concern is assuming this agreement in a minority of studies is present in all studies.
Reviewer 3	*Reviewer 3's response to this question was provided in this person's review of the draft PFOS document, with a statement that the comment also pertained to PFOA; therefore, it is copied here.* This seems acceptable for the time being. Having worked in the laboratory with PFOS, I can make a first-hand testament that mixing PFOS into exposures solutions does not guarantee a homogenous mixture despite working at solutions well below the solubility limit. There are nuances associated with achieving homogeneity of the exposure solution, we have developed a PFAS mixing protocol to reduce chemical clumping and this increases uniformity of the solutions. Furthermore, there is approximately 30% variability of PFOS quantitatively (see…Rewerts et al. 2020); so, the best measurement still has significant variability.
Reviewer 4	PFOA is a highly stable compound, resistant to hydrolysis, photolysis, volatilization, and biodegradation (as described in Section 1.1.1 of the Report) and, therefore, expected to vary only minimally in the course of a toxicity test. To determine if nominal and measured PFOA concentrations were typically in close agreement, pairs of nominal and corresponding measured PFOA concentrations were compared to one another through (1) linear correlation analysis and (2) an assessment of measured concentrations as a percent of its paired nominal concentration. The analysis conducted by EPA Team showed strong correlation (correlation = 0.9995) of the 79 pairs of nominal and measured concentrations from freshwater studies, and similar strong correlation (correlation = 0.9999) of the 11 pairs of nominal and measured concentrations from saltwater studies (Figure L1 in the Report). In addition, the experimental conditions did not influence the correlation between nominal and measured concentrations. ***This confirms inclusion of unmeasured PFOA toxicity tests for quantitative use in criteria derivation.*** Personal experience on analyzing PFOA in ecotoxicological studies using freshwater and saltwater species have also exhinited strong correlation between nominal and measured concentrations. **Additional information for L.I.4 summary section** could include additional information based on the two additional published papers and the key points from these studies are listed below. *Despite the concerns and avoidance of glass, few studies have presented data describing the sorption of perfluorooctanoate (PFOA) to glass and other container materials (**Lath et al. 2019**). Lath et al. (2019), who had reported that sorptive losses of PFOA for lower concentration (~20 µg/L) aqueous solutions were highest in polypropylene centrifuge tubes. However, the container type did not influence the measured concentration as reported in section – of this report. According to **Rewerts et al., 2021** additional handling steps, which*

2.4. The Use of Measured and Unmeasured Toxicity Tests to Derive Respective Criterion	
Reviewer	**Comments**
	are not typically reported for ecotoxicological studies but may contribute to variability, include solution homogenization, subsampling procedures, and the container materials selected for storage. https://doi.org/10.1002/etc.4667 Lath S, Knight ER, Navarro DA, Kookana RS, McLaughlin MJ. 2019. Sorption of PFOA onto different laboratory materials: Filter membranes and centrifuge tubes. Chemosphere 222: 671– 678. DOI: 10.1016/j.chemosphere.2019.01.096
Reviewer 5	Similar to PFOS, the approach EPA used for PFOA was to consider studies in which the chemical was not measured. This was justified based on an analysis EPA did comparing nominal and measured concentrations and finding close agreement. This generally seems reasonable to me especially considering the stability of the chemical. However, EPA used a criteria of 20% (p. 61) which is not consistent with the analytical precision of most methods used to analyze PFAS. According to several very prolific environmental chemists that have made a career of measuring PFAS, they have communicated to me that the methods are accurate to within 30%. That means that if the measured were within 30% of nominal, we basically have concluded these were nominal. I would encourage EPA to explore their 20% acceptability threshold and perhaps offer an explanation as to why this is appropriate.

2.5 Please comment on the toxicity data used to derive the draft criteria.

- **Were the data selected and/or excluded from the derivation of the criteria derivation appropriately utilized?**
- **Are there relevant data that you are aware of that should be added to the analyses (note that EPA is working on updating the toxicity data to reflect the data in ECOTOX between Sept. 2019 through the latest update)? If so, please provide references for consideration.**

In particular, please comment on:

2.5.a. The toxicity values used to derive the PFOA criteria, with a particular emphasis on:

2.5.a.i. the selection and the Acute to Chronic Ratio (ACR) to serve as the Final Acute to Chronic Ratio (FACR) and its application to derive the Final Chronic Value (FCV).

2.5.a.ii. the use of the qualitatively acceptable acute midge (*Chironomus plumosus*) data from Yang et al. (2014) to suggest aquatic insects are relatively tolerant to acute PFOA exposures. Specifically, Yang et al. (2014) conducted a 96-hour renewal, measured PFOA acute test with the midge, *Chironomus plumosus*. This study was not acceptable for quantitative use due to the potentially problematic source of the organisms. The reported LC_{50} was 402.24 mg/L PFOA indicating that these insects may not be one of the more sensitive taxonomic groups. Therefore, this test was excluded from the criterion calculation, but used to waive the missing insect MDR.

2.5.b. EPA's approach for fitting concentration-response (C-R) data (described in Appendix J) as well as the specific acute LC_{50} values (Appendix A.2) and chronic EC_{10} values (Appendix C.2) that were estimated (for sensitive genera when C-R data were available) and used to derive criteria.

2.5. The Toxicity Data to Derive the Draft Criterion	
Reviewer	**Comments**
Reviewer 1	• **Were the data selected and/or excluded from the derivation of the criteria derivation appropriately utilized?** In most cases, yes. Please see detailed comments on particular studies and interpretations in response to other charge questions. • **Are there relevant data that you are aware of that should be added to the analyses (note that EPA is working on updating the toxicity data to reflect the data in ECOTOX between Sept. 2019 through the latest update)? If so, please provide references for consideration.** Hayman, N.T., Rosen, G., Colvin, M.A., Conder, J., Arblaster, J.A. 2021. Aquatic toxicity evaluations of PFOS and PFOA for five standard marine endpoints. Chemosphere 273:129699. **2.5.a** **2.5.a.i.** As stated in the 1985 guidelines, "the Final Acute-Chronic Ratio should be calculated as the geometric mean of the acute-chronic ratios for species whose SMAVs are close to the Final Acute Value." The guidance does not quantify "close", but it also does not specify that the Acute to Chronic Ratio (ACR) should only be derived from a single ACR from the study with the acute value that is closest to the Final Acute Value (FAV). In the case of the PFOA document, EPA is using only one ACR value, which is derived from the Zhang et al. (2013) rotifer study in which the LC_{50} is 150 mg/L. I would agree 150 mg/L relatively close to the FAV of 94 mg/L. However, I would also argue that the LC_{50} of 166 mg/L from the study with *Moina macrocopa* is also "close" to 94 mg/L (it is only 11% higher than 150 mg/L). The Species Mean Acute Value (SMAV) for *D. magna* is 253.7 mg/L is well within a factor of 3 of 94 mg/L and could also be considered "close". These three data points tend to cluster, as shown in Figure 3-5. Taking the geometric mean of the three ACRs from these three studies, which are studies that have acute values all within a factor of 2-3 of the 94 mg/L FAV, seems more consistent with the 1985 guidance to select several ACRs that are "close" to the FAV, rather than simply selecting a single ACR. The resulting recalculated ACR would be approximately 90, which provides a good, but conservative measure of central tendency and is higher than all but the highest ACR (299) shown in Table 3-8. Applying an ACR of 90 to the FAV would result in a FCV of approximately 1 mg/L, which is protective of chronic values shown in Figure 3-6 (note the Zhang et al.-derived rotifer SMCV is 0.8 mg/L, which is only slightly lower 1 mg/L). Such as recalculation would better incorporate variability in the ACRs and avoid both water-based criteria being completely driven by the results from a single organism (rotifers) as reported in a single study (Zhang et al., 2013).

2.5. The Toxicity Data to Derive the Draft Criterion	
Reviewer	**Comments**
	2.5.a.ii. I disagree with excluding this data point from the acute criteria calculations. I assume this data has been removed under the assumption that these animals may have been pre-exposed to PFOA and may have been more tolerant of PFOA exposures, which would result in biased-high median lethal concentration (LC_{50}) values. If so, this should be explicitly stated. Assuming these *Chironomus* can develop tolerance to PFOA, it seems that they would have to be exposed to rather high mg/L ranges of PFOA in water given the reported 96-hour LC_{50} of 402 mg/L. Based on published literature, I am unaware of natural ecosystems in China (where the animals may have been originally harvested) with concentrations of PFOA that approach this order of magnitude range (in which they could build up a tolerance). The animals were obtained from a local market, so it is also possible that they were cultured for several generations, presumably using uncontaminated water (which would further reduce the chance that multiple generations were exposed at these levels). Overall, I think it is more reasonable to assume that the animals used in the experiment have not built up an acute lethal tolerance to PFOA, and the that LC_{50} result is unbiased. It does seem clearly show that insects may be less sensitive to acute lethality effects of PFOA. As such, I think it should be included as a quantitative endpoint. Additionally, it seems inconsistent to exclude this Yang et al (2014) study, when acute data from a study by Yuan et al. (2015) were included for quantitative consideration. As noted on page A-10, the animals in the Yuan et al. (2015) study were "collected from a fountain in Quanhetou, Boshan, China, and acclimated in the laboratory for an unspecified time period before use". The source of the animals is just as uncertain as the Yang et al (2014) animals, and it is unclear (if PFOA tolerance at lethal levels is possible) how many generations would be needed to shed adaptive tolerance and how this time period would compare to an "unspecified time period." Simply put, if data from experiments like Yuan et al. (2015) are quantitatively included, those from Yang et al. (2014) should also be quantitatively included (with some notes on the uncertainty of the animal sources). **2.5.b.** More details need to be provided on the dose response modeling using R. Appendix J is helpful for providing the reader with details on the general approach, but where EC_{10}s are modeled by EPA, the model being used (out of the 22 available in the R software package) needs to be specified. Providing some indication of variability (such as a 95% confidence interval) for the model-generated EC_{10}s is standard practice for dose response modeling, and this information should be provided somewhere in the document. Showing the R package output of the goodness of fit statistics (or equivalent) for the modeling in an Appendix would be helpful; since this was used to select the model used in each instance of an EC_{10} calculation, it must be available, so I would recommend including it for full transparency and to aid future efforts in understanding the aquatic toxicology of this chemical. Additionally, it would be helpful to show the selected model fits for all calculated EC_{10}s (as shown for the most sensitive EC_{10}s estimated). These steps would be helpful to ensure and demonstrate quality of the model fits and reproducibility of the modeling work.
Reviewer 2	• **Were the data selected and/or excluded from the derivation of the criteria derivation appropriately utilized?**

2.5. The Toxicity Data to Derive the Draft Criterion	
Reviewer	**Comments**
	I think the data used in the derivation of the criteria were appropriate. As mentioned above, I am a little concerned about the use of toxicity data from studies that did not measure the concentration of PFOA in their experiments, especially considering the proportion of studies that did not measure the concentrations. The confirmation of exposure concentrations is an important principle of sound ecotoxicology. • **Are there relevant data that you are aware of that should be added to the analyses (note that EPA is working on updating the toxicity data to reflect the data in ECOTOX between Sept. 2019 through the latest update)? If so, please provide references for consideration.** Hayman, N.T., Rosen, G., Colvin, M.A., Conder, J., Arblaster, J.A., 2021. Aquatic toxicity evaluations of PFOS and PFOA for five standard marine endpoints. Chemosphere 273, 129699.. doi:10.1016/j.chemosphere.2021.129699 Logeshwaran, P., Sivaram, A.K., Surapaneni, A., Kannan, K., Naidu, R., Megharaj, M., 2021. Exposure to perfluorooctanesulfonate (PFOS) but not perflurorooctanoic acid (PFOA) at ppb concentration induces chronic toxicity in Daphnia carinata. Science of The Total Environment 769, 144577.. doi:10.1016/j.scitotenv.2020.144577 Bartlett, A.J., De Silva, A.O., Schissler, D.M., Hedges, A.M., Brown, L.R., Shires, K., Miller, J., Sullivan, C., Spencer, C., Parrott, J.L., 2021. Lethal and sublethal toxicity of perfluorooctanoic acid (PFOA) in chronic tests with Hyalella azteca (amphipod) and early-life stage tests with Pimephales promelas (fathead minnow). Ecotoxicology and Environmental Safety 207, 111250.. doi:10.1016/j.ecoenv.2020.111250 Tornabene, B.J., Chislock, M.F., Gannon, M.E., Sepúlveda, M.S., Hoverman, J.T., 2021. Relative acute toxicity of three per- and polyfluoroalkyl substances on nine species of larval amphibians. Integrated Environmental Assessment and Management 17, 684–690.. doi:10.1002/ieam.4391 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.. doi:10.1002/etc.4690 BAF data provided in the supplementary information of Prosser et al. 2016 study for three freshwater species does not appear to have been considered. Prosser, R.S., Mahon, K., Sibley, P.K., Poirier, D., Watson-Leung, T., 2016. Bioaccumulation of perfluorinated carboxylates and sulfonates and polychlorinated biphenyls in laboratory-cultured Hexagenia spp., Lumbriculus variegatus and Pimephales promelas from field-collected sediments. Science of The Total Environment 543, 715–726. doi:10.1016/j.scitotenv.2015.11.062

2.5. The Toxicity Data to Derive the Draft Criterion	
Reviewer	**Comments**
	2.5.a. **2.5.a.i.** I think the approach that led to the selection of the ACR for Brachionus calyciflorus was appropriate. It is also the most conservative approach with the ACR for this species being the largest of the four species with ACRs. **2.5.a.ii.** I think the EPA's decision that the data from Yang et al. (2014) was not acceptable for quantitative use was appropriate. The source of the larvae is problematic. The conclusion that insects may not be one of the most sensitive taxa is valid. The NOEC for Chironomus tentans of 100 mg/L reported by MacDonald et al. (2004) also supports this conclusion. **2.5.b.** I think the approach that the EPA used to determine effect measure from concentration-response data was appropriate. The use of the drc package in R to fit 22 different models to the empirical data and then using several criteria (e.g., AIC, residual standard errors, confidence intervals) to evaluate the fit of the different models is robust. It would have been useful if the EPA reported the 22 different models in Appendix J. I think the LC_{50} and EC_{10} values determined by the EPA using the approach mentioned in the previous paragraph was appropriate. It is valid for these effect measures to be determined when the concentration-response data has been provided by the authors of the study. The EPA has also made is clear in Appendix A.2 and C.2 how they determined these effect measures using the concentration-response data provide in the studies. This generates a high level of transparency in the derivation of the criterion.
Reviewer 3	• **Were the data selected and/or excluded from the derivation of the criteria derivation appropriately utilized?** With the exception of the Fabbri et al. 2014, data currently evaluated, and the associated decision matrix seem appropriate. However, as noted below, there are new data available following this draft. • **Are there relevant data that you are aware of that should be added to the analyses (note that EPA is working on updating the toxicity data to reflect the data in ECOTOX between Sept. 2019 through the latest update)? If so, please provide references for consideration.** o McCarthy et al. 2021 - freshwater o Hayman et al. 2021 – marine o Logeshwaran et al. 2021 – freshwater o Li et al. 2021 – freshwater/plant o Etc. **2.5.a.** **2.5.a.i.** The use of ACR is appropriate given data limitations. However, either I am missing it or it is unclear as to why Chydorus is not included in the chronic data set when it is the most sensitive in the acute.

2.5. The Toxicity Data to Derive the Draft Criterion	
Reviewer	**Comments**
	2.5.a.ii. McCarthy et al. 2021 reported midge data for PFOA with an EC50 of 192 mg PFOA/L following a 20-day exposure. Following EPA review of this publication, this may fulfill the missing insect MDR; however, still supports the conclusion of likely not a sensitive taxa. **2.5.b.** This is a sound scientific approach, what is unclear is when EPA employs this vs other times when it is not used. i.e. chydorus vs two mussle spp in the acute studies
Reviewer 4	• **Were the data selected and/or excluded from the derivation of the criteria derivation appropriately utilized?** The data selected to derive PFOA criteria are appropriate. Studies that did not fully meet the data quality objectives outlined in the 1985 Guidelines were not considered for inclusion in the criteria derivation, including some studies with other PFAS exposures, but were considered qualitatively as supporting information. A brief summary of each study describing the experimental conditions and summary tables providing all the relevant information such as strengths and limitations of each study, end points selected for deriving criteria are well documented by the EPA team. The key acceptable exclusion/inclusion criteria used to derive draft criteria are listed below: • Only single chemical toxicity tests with PFOA were considered for possible inclusion in criteria derivation, studies that tested chemical mixtures, including mixtures with PFAS compounds were excluded from criteria derivation. • Both controlled laboratory experiments and field observations/studies were included. • PFOA toxicity tests were not excluded from quantitative use in criteria derivation on the basis of unmeasured test concentrations alone. • Only single chemical toxicity tests with PFOA were considered for possible inclusion in criteria derivation, studies that tested chemical mixtures, including mixtures with PFAS compounds were excluded from criteria derivation. • Due to lower sensitivity, insect MDR was excluded from the criterion calculation, but were used to waive the missing insect MDR. • Further supporting information on acceptable and unused studies for acute and chronic endpoints and for freshwater and marine studies are documented and summarized as appendices in this report. • **Are there relevant data that you are aware of that should be added to the analyses (note that EPA is working on updating the toxicity data to reflect the data in ECOTOX between Sept. 2019 through the latest update)? If so, please provide references for consideration.** Additional toxicity data published over the last six months is listed below:

2.5. The Toxicity Data to Derive the Draft Criterion	
Reviewer	**Comments**
	Marine/estuarine Stuart L. Simpson, Yawen Liu, David A. Spadaro, Xinhong Wang; Rai S. Kookana and Graeme E. Batley Chronic effects and thresholds for estuarine and marine benthic organism exposure to perfluorooctane sulfonic acid (PFOS)-contaminated sediments: Influence of organic carbon and exposure routes https://doi.org/10.1016/j.scitotenv.2021.146008 Nicholas T Hayman , Gunther Rosen , Marienne A Colvin , Jason Conder , Jennifer A Arblaster Aquatic toxicity evaluations of PFOS and PFOA for five standard marine endpoints. https://doi.org/10.1016/j.chemosphere.2021.129699 **Fresh water** Christopher J. McCarthy, Shaun A. Roark, Demitria Wright, Kelly O'Neal, Brett Muckey, Mike Stanaway, Justin N. Rewerts, Jennifer A. Field, Todd A. Anderson, Christopher J. Salice, Toxicological Response of Chironomus dilutus in Single-Chemical and Binary Mixture Exposure Experiments with 6 Perfluoralkyl Substances, Environmental Toxicology and Chemistry, 10.1002/etc.5066, **40**, 8, (2319-2333), (2021). **2.5.a.** **2.5.a.i.** The selection and the Acute to Chronic Ratio (ACR) to serve as the Final Acute to Chronic Ratio (FACR) and its application to derive the Final Chronic Value (FCV) for PFOA is acceptable. The 1985 Guidelines allow the use of a Final Acute-Chronic Ratio (FACR) to convert the FAV to the FCV (i.e., FAV/FACR=FCV), which is equivalent to the chronic criterion (Criterion Continuous Concentration, CCC), intended to protect 95 percent of the taxa in aquatic ecosystems. For PFOA, the 8-family MDR requirement was not met for the chronic dataset, as acceptable chronic studies for species representing three MDR groups were not available (benthic crustacean and third phylum or second insect order not already represented). Therefore, the Final Chronic Value (FCV) was calculated with the use of an ACR (acute-chronic ratio). When more than a single ACR was calculated for the same species, the Species Mean Acute-Chronic Ratio (SMACR) was calculated as the geometric mean value of all ACRs for that species. . The specifications for derivation of a FACR for aquatic animals was met for PFOA based on 1985 Guidelines: ACRs for at least three different families provided that at least one was a fish, at least one was an invertebrate, and at least one was an acutely sensitive freshwater species. The 1985 Guidelines provides recommendations to calculate the FACR when SMACRs are dissimilar. The 1985 Guidelines states that if SMACRs tend to increase or decrease as the SMAV increases, the FACR should be calculated as the geometric mean of the ACRs for species whose SMAVs are close to the FAV. For PFOA, relationship between SMAV and SMACR showed that SMACRs decreased as the SMAVs increased. SMAV of the rotifer *B. calyciflorus* was closest to the FAV. The CCC was calculated by dividing the FAV by the FACR to determine the FCV (91.34/299.1=0.3054). *The PFOA FCV derived by this method is lower than the all of the quantitatively-acceptable chronic values (ranges between 0.76- 40 mg/L) and listed in listed in Table C.1*. **2.5.a.ii.** The acute data set for PFOA contained 14 genera representing seven of the eight taxonomic MDR groups. The missing MDR was a representative from an insect family. There was no evidence to suggest aquatic insect are among the four most sensitive genera. EPA

2.5. The Toxicity Data to Derive the Draft Criterion	
Reviewer	**Comments**
	calculated the PFOA CMC using all acceptable quantitative studies, but did not include the insect data in the criterion calculation (i.e., the relatively tolerant insect LC_{50} value was not included in the total count ("n") of Genus Mean Acute Values in the criterion calculation). In addition, waiving an unfulfilled MDR when available data suggest it is not among the four most sensitive genera is consistent with previous EPA criteria documents, including U.S. EPA (2016). In addition, Stefani et al. (2014), Macdonald et al. (2004), and Marziali et al. (2019) conducted chronic toxicity tests with *Chironomus* spp. and reported apical endpoints. *Results of these studies, taken together, also suggest that insects may not be among the most sensitive taxa to chronic PFOA exposures. Therefore, these tests were excluded from the criterion calculation, but were used to waive the missing insect MDR.* **2.5.b.** • This is an excellent approach utilized by the EPA Team. EPA's approach for fitting concentration-response (C-R) data resulted in consistent approach across various ecotoxicological studies. The R drc package was used to fit 22 different models to each individual C-R dataset. A single model was then selected from the 22 models to serve as the representative C-R model. The selected model represented the most statistically-robust model available. In certain cases, this approach even improved and helped to select most sensitive toxicological endpoint, for example, o Page 85- Ji et al. (2008) conducted a chronic life-cycle test on the effects of PFOA (with *Moina macrocopa*). The *M. macrocopa* 7-day NOEC (reproduction: number of young per adult) reported by authors was 3.125 mg/L, the LOEC was 6.25 mg/L, and the MATC is 4.419 mg/L. EPA performed C-R analysis for this study and determined the number of young per starting female as the most sensitive endpoint with an acceptable C-R curve. The EPA-calculated EC_{10} was 2.194 mg/L PFOA for *M. macrocopa* and used it directly as the *Moina* GMCV. o Page 88 Yang et al., 2014 Chronic survival using EPA's method was more sensitive than reproduction endpoint related EC_{10} values reported by the authors. • In depth analyses and associated dose-response graphs in Appendix A.2 and Appendix C.2 provides further in-depth information on the EPA's approach for fitting concentration-response (C-R) data.
Reviewer 5	In general, the data selected or excluded for criteria development were appropriately used. A study published by McCarthy et al. in 2021 reports toxicity of PFOA to Chironomus dilutus – if I recall correctly, I believe these data (on an insect) indicate that PFOA is not very toxic to C. dilutus…even for exposures greater than 96 hours. This agrees with EPA's assessment of available acute toxicity data for freshwater insects; that PFOA is generally not toxic to insects. I would, however, still urge more studies on aquatic insects as this is an obviously diverse taxa with many sensitive species. It is very possible that the available test species of insects are not sensitive but other insects such as mayflies or damselflies, etc. may, in fact, be quite sensitive.

2.5. The Toxicity Data to Derive the Draft Criterion	
Reviewer	**Comments**
	The write up of the Zhang et al. (2013 and 2014) papers has a few grammatical errors and could benefit from additional editing for clarity. I agree that using the intrinsic rate of increase (or similar) is a very relevant endpoint but it reads as though this was based on a 4 day observation period for the 2014 paper. This does not quite make sense to me…offhand, I would hypothesize that the reported/calculated effect level of 1.166 mg/L would decrease with a longer observation period. I think I may be misunderstanding the experimental design – some editing would help clarify. This is an important series of studies (see below regatding resting egg production) so clarity is critical. NOTE: the EC_{10} of 0.076 mg/L for resting egg production observed by Zhang et al. (2014b) is potentially a big deal. EPA appears justified in not using this because it was only one replicate, etc. but these data clearly point to a potentially relevant effect at a relatively low concentration. **2.5.a.** **2.5.a.i.** The EPA followed the 1985 guidelines allowing them to calculate the FCV using the ACR approach as outlined around p. 101. Given the lack of data, this seems like a reasonable approach. The actual calculation of the FCV (based on the data from B. calyciflorus) is appropriate and the use of this FACR to determine the FCV of 0.3054 is appropriate. Morevoer, as stated above, this value is also generally in line with other criteria for PFOA published by other jurisdictions. **2.5.a.ii.** I commented on this above but will mention again. Overall, I think EPA is correct that the available data on chironimids (Yange et al. 2014 and McCarthy et al. 2021) indicate that chironomids are not sensitive to PFOA. That said, it is probably the case that other insects such as mayflies or damselflies (or other species?) are more sensitive than chironomids. In contrast to this statement, however, is that chironomids are among the most sensitive to PFOS. Bottom line: for PFOA and available insect toxicity data, it appears PFOA is not toxic to insect but, clearly, more data are needed to improve confidence in this estimate. **2.5.b.** EPA's approach to fitting C-R data using the drc package in R is, in my opinion, state of the art. The method can easily test a variety of curves and the fit criteria can be used to select the best fitting curves. Comments related to specific studies and LC_{50} and EC_{10} estimates are elsewhere in these comments. In general, my opinion is that the PFOA criteria are slightly more defensible than the PFOS criteria; this is explained more in the review of PFOS.

2.6 Please comment on the translation of the chronic water column criterion elements for aquatic life to derive the tissue-based criterion elements, considering the bioaccumulation of PFOA and PFOS. In particular, please comment on:

2.6.a. Uncertainty surrounding the bioaccumulation factors (BAFs) used to translate of the chronic water column criterion elements into tissue-based criterion elements.

2.6.b. EPA's determination of appropriate BAFs and the tissue types that the tissue criterion elements were based.

2.6. The Translation of the Chronic Water Column Criterion Elements for Aquatic Life to Derive the Tissue-Based Criterion Elements Considering Bioaccumulation	
Reviewer	**Comments**
Reviewer 1	The derivation of the tissue criteria in this manner is highly uncertain. To my knowledge this is the first time EPA has applied ambient water quality criteria protective of aquatic life direct toxicity with uptake factors (bioaccumulation factors (BAFs), bioconcentration factors (BCFs)) in this manner to calculate tissue criteria. References are made to the selenium tissue criteria, but those are used in the reverse (i.e., criteria based on measured concentrations in tissue used to calculate water criteria). The use of criteria for water with a assumed uptake factor carries a large amount of uncertainty, and in general, the use of measured concentrations in tissue linked to adverse effects is a more straightforward approach since it does not involve uptake model predictions. This needs to be noted in the text. Also, are the predicted tissue criteria meant to be a temporary stop-gap until tissue effect data become available? This should be discussed and clarified. **2.6.a.** The use of BAFs derived from field studies is inherently uncertain due to the wide variety of techniques used in the compiled studies, their analytical data quality, the differences in species and ecosystems, experimental designs, spatial uncertainties for mobile animals like fish, etc. That being said, the use of a BAF value (or BCF) in criteria derivation is consistent with other criteria developed by EPA. As noted above, the use of the tissue criteria needs to be considered carefully, and I think empirical tissue data from toxicity experiments should form the basis of a next iteration of a tissue criteria. **2.6.b.** The development of BAFs for invertebrates, fish (whole body), and fish (muscle) seems reasonable for the application in estimating a draft or interim tissue criteria until empirical tissue data can be used to calculate tissue criteria directly
Reviewer 2	**2.6.a.** There appears to be an error in sub-section 2.11.3; "The resulting tissue-based criteria magnitudes correspond to the tissue type from the geometric mean BAF used in the equation (see Section 2.12.3.1)." I cannot locate sub-section 2.12.3.1. I assume the authors meant sub-section 2.11.3.1? Sub-section 2.12.3 and sub-section 2.12.3.1 are also referenced on pages 105 & 106, respectively. I think the EPA has sufficiently addressed the uncertainty around the use of BAFs and the chronic water column criterion in the derivation of tissue-based criterion. They have indicated that tissue-based criterion should only be observed once in 10 years. The use of the geometric mean of the reported BAFs incorporates the range of BAFs that may be present for different invertebrate and fish species. The use of the chronic water column criterion also builds in added conservatism to the tissue-based criterion. Prosser et al. (2016) reported BAFs for PFOA in three freshwater species (two invertebrates and one fish) (See Tables S29-31 in Supplementary Information), but it was not considered in this assessment. It is not clear why it was not considered.

2.6. The Translation of the Chronic Water Column Criterion Elements for Aquatic Life to Derive the Tissue-Based Criterion Elements Considering Bioaccumulation	
Reviewer	**Comments**
	Prosser, R.S., Mahon, K., Sibley, P.K., Poirier, D., Watson-Leung, T., 2016. Bioaccumulation of perfluorinated carboxylates and sulfonates and polychlorinated biphenyls in laboratory-cultured Hexagenia spp., Lumbriculus variegatus and Pimephales promelas from field-collected sediments. Science of The Total Environment 543, 715–726. doi:10.1016/j.scitotenv.2015.11.062 **2.6.b.** The evaluation criteria for BAFs outline in Table 2-3 are appropriate and the decision to only use high and medium quality BAFs is justified based on the criteria that would make a BAF low quality. It was a good idea to use fish BAFs based on the concentration in muscle and whole body. Muscle tissue is usually exclusively sampled in large fish, especially as part of fish consumption guidelines. The whole body is more appropriate for small fish and invertebrate species, e.g., minnows, benthic macroinvertebrates.
Reviewer 3	This sentence is very confusing: "EPA examined the potential for criteria using only those studies in which test organisms were exposed to PFOA in their diet, because such studies would most closely replicate real-world exposures (diet and/or diet plus water)." The tissue criteria are based on water exposures, the relevance of this statement and evaluation is lost on me. A table summarizing the species used to derive the BAFs would be helpful to evaluate the comprehensiveness, i.e., pelagic vs sediment feeders.
Reviewer 4	**2.6.a.** Tissue criteria derived from the chronic water column concentration (CCC) with the use of bioaccumulation factors were developed by EPA. The chronic fish whole-body tissue criterion is 54.1 mg/kg wet weight, the chronic fish muscle tissue criterion is 9.37 mg/kg wet weight and the chronic invertebrate whole-body tissue criterion is 23.9 mg/kg wet weight. The freshwater chronic PFOA toxicity data with measured tissue concentrations was limited, with no quantitatively acceptable tissue-based tests. Qualitatively acceptable tissue-based tests were reported for four species (three fish species and one amphibian) across five publications. Therefore, there were insufficient data to derive tissue-based criteria using a GSD approach from empirical tissue data from toxicity studies. EPA thus developed protective tissue-based criteria through a bioaccumulation factor approach (Burkhard 2021). Only BAFs of high and medium quality were used to derive the tissue criteria. BAFs used in the derivation of the PFOA tissue-based criteria consisted of > 2 water and organism samples each and were collected within one year and 2 km distance of one another. Criteria for protection of aquatic life and wildlife will need to use whole- organism BAFs because the criteria are based on whole-body toxicology for aquatic life (e.g., fish) and for wildlife (e.g., birds), both of which consume the whole fish (Stephan et al. 1985). **2.6.b.** Within the body, PFOA tends to bioaccumulate within protein-rich tissues, such as the blood serum proteins and liver. BAFs are different for muscle/fillet and whole-body tissues. Humans consume muscle/fillets from fish and soft tissues from bivalves, therefore the water quality criteria recommended by EPA used BAFs based on these tissues. EPA calculated

2.6. The Translation of the Chronic Water Column Criterion Elements for Aquatic Life to Derive the Tissue-Based Criterion Elements Considering Bioaccumulation	
Reviewer	**Comments**
	additional tissue values for liver, blood, and reproductive tissues by transforming the freshwater chronic water column criterion (i.e., 0.31 µg/L) into representative tissue concentrations using tissue-specific bioaccumulation factors (BAFs). Furthermore, EPA team justified to use female reproductive tissues due to its relevance for potential maternal transfer to offspring. There additional tissue-based values were calculated for comparative purposes and were not proposed as recommended criteria.
Reviewer 5	**2.6.a.** Overall, this seems like a reasonable approach – to estimate tissue-based criteria using the water column criteria multiplied by the bioaccumulation factors. The difficulty arises when we consider the accuracy or robustness of the BAFs. I agree with the criteria in Table 2-3 and especially emphasize the importance of concurrent collections in space AND time for tissues and environmental media. **2.6.b.** I am very familiar with the Burkhard (2021) paper which the PFOA document follows closely in terms of BAFs. The BAFs used by EPA are appropriate given the data. I also agree that the most useful/appropriate tissues for BAFs are invertebrates, fish muscle and fish whole body – these are the most commonly analyzed and most abundant in the literature. For what it's worth, in my own research in which we collected and analyzed fish tissues and co-located water samples, our calculated BAFs for PFOA (and PFOS) were close to the central tendency BAFs reported by Burkhard (2021). I also agree that co-located and sampled at the same time yield the most defensible BAFs as PFAS concentrations can vary considerably in space and time (not often shown in the literature).

2.7 Please comment on the frequency and duration of the criterion elements, in particular please comment on the frequency and duration components of the tissue-based criterion elements.

2.7. The Frequency and Duration of the Criterion Elements	
Reviewer	**Comments**
Reviewer 1	The 4-day duration seems to be supported by the time scale of toxicity discussed for the limited chronic studies selected by EPA. This assumption should be revisited as more data become available (please note in the document). For the tissue-based criterion (page 107), there is no clear support for assuming a 10-year exceedance frequency. Given the uncertainty with the BAF-predicted tissue criteria, and how little is known regarding the recalcitrance of PFOA in aquatic ecosystems and recovery time if PFOA inputs in water were halted, the assignment of a 10-year exceedance frequency at this stage seems completely arbitrary. We simply do not yet know the time frame over which aquatic ecosystems recover from PFOA. It is not technically supported to cite recovery times for selenium to support a 10-year recovery time for PFOA, these are completely different toxicants that have their own unique fate and behavior. USEPA (1985) guidance suggests

2.7. The Frequency and Duration of the Criterion Elements	
Reviewer	**Comments**
	assuming a 3-year frequency as a default, and the discussion on page 17-108 is not scientifically convincing enough to modify it to 10 years. Additionally, it should be noted that the exceedance frequency for another organic chemical, Tributyltin (TBT) was set at 3 years by EPA in derivation of that criteria. TBT exhibits uptake factors similar to PFOA (i.e., BCF of approximately 2,000 L/kg, wet weight for goldfish, as noted in the EPA TBT criteria document, which is higher than the PFOA BAFs of 30-175 L/kg, wet weight being used to calculate the fish tissue criteria). TBT is also persistent in aquatic ecosystems, as noted by EPA. Given TBT is at least an organic chemical, it is a closer analog than selenium, which is an element. As such, the exceedance frequency for the PFOA tissue criterion should be set at the default of 3 years unless EPA can provide convincing technical information specific to recovery times for PFOA. Additionally, on page 108, the paragraph that begins with "Metals and other chemical pollutants such as PFOA..." is not convincing as any quantitative support for EPA's 10-year exceedance frequency for the chronic tissue-based criteria. The text as written may give the reader the conclusion that PFOA recovery may be "on the order of decades", as EPA notes for selenium. There is no support for the conjecture that PFOA recovery may be "relatively slow" or require decades, as noted in my above comment.
Reviewer 2	As per Table 0-1, I think the chosen durations and frequencies for the acute and chronic criteria are appropriate. They will ensure protection of aquatic life. The duration of the tissue-based criterion is appropriate as the concentration will be measured when biota is collected. The 10 year frequency is appropriate considering that for biota to reach the tissue-based criteria, they would likely to have be exposed to concentrations at or above the chronic criteria for an extended period of time.
Reviewer 3	This is a not an easy statement to comment on, as it may be unlikely that the aquatic receptors will exceed or reach these tissue concentrations prior to exceedances from the CCC. Importantly, PFOA is not particularly bioaccumulative compared to PFOS and this likely a less sensitive threshold.
Reviewer 4	PFOA concentrations in tissues are generally expected to change only gradually over time in response to environmental fluctuations. The chronic tissue-based criteria averaging periods, or duration components, were therefore specified as instantaneous, because tissue data provide point, or instantaneous, measurements that reflect integrative accumulation of PFOA over time and space in population(s) at a given site. It was appropriate for EPA to inform the recommended ten-year exceedance frequencies for the chronic tissue-based criteria given the large variation in possible biological and physical variable influencing ecological recovery.
Reviewer 5	Conceptually, the frequency and duration of the criterion elements seem reasonable – the acute water column criterion can't be exceeded for more than one-hour of duration which is then not to be exceeded more than once in three years. Even though many of the effect levels are 50%, this is likely protective given the duration of most acute toxicity studies is certainly

2.7. The Frequency and Duration of the Criterion Elements	
Reviewer	**Comments**
	and convincingly more than one hour. This applies to the chronic water column criteria as well which is based on not exceeding a 4-day duration; most chronic studies are much longer than 4 days. The only real issue with this is that, for all intents and purposes, these frequency and duration elements are not measured in practice. In other words, I am unaware of hourly measurements of PFAS in water or 4-day running averages. So, conceptually, I believe the frequency and duration elements are protective. However, in practice it is not clear to me how these would be useful or would help with regulation because the data related to frequency and magnitude of PFAS in water is not at a fine enough temporal resolution. The frequency and duration for tissue-based criteria is a little different, however. Tissue concentrations (as mentioned in the document) represent an integration through time and so a measure of fish tissues, for example, provides some insight to exposures that have occurred over longer than an hour or 4 days. In this case, not exceeding the tissue-based criteria more than once in 10 years is likely protective. It is also more likely that monitoring programs would sample fish at least yearly which means this criteria is likely the most useful from a monitoring and clean-up perspective. It is still possible that high concentrations in tissues will be "missed" with only sampling once a year but this is far better than the frequency and duration elements for the water column criteria.

2.8 Please provide any additional technical comments that you believe should be considered.

2.8. Additional Technical Comments to Consider	
Reviewer	**Comments**
Reviewer 1	I have the additional detailed comments: a) Please note that the comments provided in this file reflect a focus on of key portions of the "Draft of the Aquatic Life Water Quality Criterion…" document as directed by the above charge questions provided to me. Given time and resource constraints and the scope of my review, it was not feasible to provide a detailed review of the entire document and all of the supporting references and their associated results and conclusions. As such, I reserve my right to supplement or amend my comments in future, pending additional review or new information. Thank you for the opportunity to assist EPA in its work on this very important matter, and I was honored to be selected as a reviewer. b) Page 7: Please note in Figure 1-1 that this is the linear isomer of PFOA. It would be helpful to note that the PFOA data in this study are likely from experiments with water spiked with the linear PFOA isomer. It is hypothesized that toxicity and bioaccumulation may differ between branched and linear forms of PFCAs and PFASs. Linear PFOA is thought to be more accumulative (as noted on Page 61) and potentially more toxic to aquatic life when the dose is expressed as an external water concentration. At some sites, a portion of the concentrations of PFOA in water (which are reported as the sum

2.8. Additional Technical Comments to Consider	
Reviewer	**Comments**
	of branched and linear PFOA) can be branched PFOA, so criteria derived from linear PFOA could be overly protective. Please include this uncertainty in the discussion in the document. c) Page 47: With regard to the discussion of biomagnification factors (BMFs), please acknowledge the Martin et al. 2003 paper where the BMF for fish was noted to be less than one. This paper is very helpful for understanding biomagnification from the diet in fish, as it was a controlled PFOA-spiked food study. d) Page 51: Starting here on this page and in the rest of this section, most of the units need to be specified for dry weight or wet weight for concentrations of PFOA in tissue. There were other instances of this error in the document as well. For units of every concentration of PFOA in tissue, please be sure to specify dry weight or wet weight. e) Page 61: Regarding "nearly 90% of measured concentrations fell within 20% of paired nominal concentrations", I believe that applies to the freshwater studies reviewed (please note). f) Page 64: The use of EC_{10} values instead of effective concentration 20% (EC20) values for chronic values is inconsistent with EPA's general practice for developing aquatic life values. The selection of EC_{10}s for the selenium criteria (EPA, 2016) was associated with the derivation of tissue guidelines. In the EPA (2016) document, EPA noted "EC20s have historically been used in the derivation of EPA criteria applicable to the water medium". As noted in the EPA (2016) selenium guidance EC_{10}s were selected over EC20s "given the nature of exposure and effects for this bioaccumulative chemical." Additionally EPA (2016) selected EC_{10} for selenium because "it was found that the dose-response curves for selenium across a broad range of fish genera are very steep, such that a small change in selenium tissue concentration yielded a large increase in observed adverse effect." g) First, all the derivation of aquatic life criteria for "bioaccumulative chemicals" have not followed the process used for selenium, and there is no quantitative discussion in the current document that compares the bioaccumulation values for selenium to those of PFOA in a manner than justifies the use of EC_{10}s. For example, EPA in its 2016 aquatic life criteria for cadmium noted that cadmium "can bioaccumulate in aquatic organisms", but EC20s (not EC_{10}s) were used as chronic values in the derivation of aquatic life criteria in that document. Fundamentally, there is a logical disconnect between adding additional conservativism (i.e., using EC_{10}s instead of EC_{20}s) simply because a chemical has a higher bioaccumulative potential than another chemical or exceeds a BCF or BAF criteria used to determine a chemical has "bioaccumulative" status by typical chemical registration guidelines. The use of chronic exposure toxicology data generally assumes that concentrations in the organisms have reached steady state and, and thus, any bioaccumulation that has occurred is accounted for and manifests in toxic action. Coincidentally, the general assumption is that toxic responses have plateaued as well and that effective doses (measured via external concentrations in water or concentrations in the organism) will not change significantly with additional exposure

2.8. Additional Technical Comments to Consider	
Reviewer	**Comments**
	time. The bioaccumulative nature of the toxicant at that point is a moot point with regards to toxic effects in an aquatic organism, so there seems no need to add additional conservatism in the estimation of a threshold for potential ecologically-significant effects on aquatic life. Adding additional conservatism to the aquatic life criteria to protect other trophic levels (i.e. wildlife that consume aquatic life) or human consumers of aquatic life, which does involve bioaccumulation of chemicals in aquatic organisms, is not justified. Criteria to protect wildlife and humans exposed via exposure pathways involving bioaccumulation of chemicals in aquatic life are handled via separate approaches, and are completely disconnected from the acute and chronic toxicity data developed to evaluate the risks to aquatic invertebrates and lower trophic level vertebrates like fish and amphibians. h) Second, EPA has not provided any analysis of the dose response curves that demonstrates the need for EC_{10}s versus EC20s (as was mentioned for selenium). Additionally, justification of the use of EC_{10}s by simply referencing the regulatory policies of other countries seems to be insufficient as the basis for a US policy, and is unsatisfying from a scientific perspective. i) More discussion is needed to support the poorly-supported move from EC20s to EC_{10}s, or alternately, EC20s need to be used in throughout the document, as consistent with past EPA practice in aquatic life criteria derivation. EC_{10}s are more conservative than EC20s, but there is often greater variability and uncertainty associated with EC_{10} values given the typical 50% effect ranges that are generally targeted in the experimental designs of typical toxicological studies. Additionally, as noted in EPA's 2016 aquatic life criteria document for cadmium, EC_{10}s are "rarely statistically significantly different from the control treatment." A 20% effect has often been discussed as a point of departure of ecologically-significant population- and community-level effects (e.g., Suter, 2000: Suter, G.W., Efroymson, R.A., Sample, B.E., & Jones, D.S. (2000). Ecological Risk Assessment for Contaminated Sites. CRC Press. April). j) Overall, the adoption of a more conservative 10% effect level (i.e., EC_{10}) for chronic values used in criteria calculation carries large environmental management and policy implications. As noted above, clarification and careful justification is needed. EPA needs to clearly articulate (ideally with ample scientific support) why the additional conservatism is needed. This important potential policy matter deserves an open and earnest discourse among the scientific, stakeholder, and regulated communities. k) Page 67: It appears that only studies in which organisms exposed via diet were included for evaluation of tissue criteria. Is this correct? It is questionable to exclude effect concentrations in tissue from experiments in which exposure of PFAS was only via water. EPA (2016) took the "dietary exposure only" approach with selenium because the primary exposure route for selenium has been shown to be via the diet in natural ecosystems. In contrast, for many aquatic animals (especially lower trophic level fish and invertebrates), a significant portion of the exposure to PFOA is via non-dietary pathways. Part of this is due to the fact that controlled studies (e.g., Martin et al., 2003 studies cited in the document) have found that water-to-organism BCFs for aquatic life

2.8. Additional Technical Comments to Consider	
Reviewer	**Comments**
	such as fish are generally larger than diet-to-organism biomagnification factors (BMFs). Additionally, there is no reason to expect dietary or non-dietary exposure pathways would affect toxic responses given the relatively rapid internal kinetics of PFAS in aquatic life (i.e., half-life of hours or days), especially for small invertebrates and fish that are in relative equilibrium with their surrounding exposure water. l) Page 82: The intrinsic rate of increase (basis of the chronic value from the Zhang et al. (2013) study is not a typical endpoint used in aquatic life criteria derivation. This chronic value has a large influence over the criteria calculated in this document, as noted below. More detail is needed on the calculation of the endpoints used in the two Zhang et al. papers, and some additional information should be provided on what this endpoint means ecologically in comparison to typical reproductive endpoints used in criteria derivation. Note that for stable populations, the intrinsic rate is 0, so achieving a maximal rate of increase is not always ecologically sustainable. Some additional explanation would be helpful to the reader. m) Page 98: Both the Criterion Maximum Concentration (CMC) and Criterion Continuous Concentration (CCC) are heavily influenced by toxicity tests on a single organism (rotifers, Brachionus calyciflorus) conducted by Zhang et al. (2013) and Zhang et al. (2014). For example, the rotifer acute value is the second lowest value of the four values used to calculate the FAV and CMC. The ACR used to calculate the CCC (using the with the FAV) is derived entirely from the Zhang et al. (2013) study (Page 101). I would recommend the ACR be recalculated to reflect more of a central tendency estimate that incorporates other ACRs, rather than relying solely on the Zhang et al. (2013) result (see above 5a comment). Regardless, I think a few sentences should be added to note the strong influence of these rotifer studies on PFOA criteria. Given this, additional experimentation and verification of the Zhang et al. results seems would seem to be useful and this is worth mentioning in the document as well. n) Page G-1: Please explain the acceptable duration acceptable for tests for which "Duration too short" is noted. It would be good to provide the acceptable durations that would be considered acceptable for these species. Perhaps a summary table for acceptable durations for particular endpoints could be provided in this document. o) Appendix G: The endpoints such as "decrease mRNA expression levels of neural genes DjFoxD, DjotxA and DjotxB" and "decrease in inflammatory cytokines (IL-1ß and IL-21) in spleen" are atypical. These should be noted under "Deficiencies". p) Page H-2: First use of "atypical duration" in the table. This entry is inconsistent with other entries (e.g., "duration too short") and does not clearly describe why the experiment is not considered. Please explain this table entry.
Reviewer 2	I think the EPA's criteria for PFOA are very defensible based on the science and data available. I think they did a great job clearly laying out how they derived the criteria and providing all of the data that was used in the derivation.

<table>
<tr><th colspan="2">2.8. Additional Technical Comments to Consider</th></tr>
<tr><th>Reviewer</th><th>Comments</th></tr>
<tr><td>Reviewer 3</td><td>All technical comments have been previously mentioned</td></tr>
<tr><td>Reviewer 4</td><td>
Some additional edits/suggestions are listed below:

1- Appendix list in Table of Contents in missing

2- List of Tables not matching with Tables listed in the text
Table 2-4. is Table 2.1 Measured Perfluorooctanoic acid (PFOA) Concentrations in Surface Waters Across the United States. 24

Table 2-5. is Table 2.2 Summary of Assessment Endpoints and Measures of Effect Used in the Criteria Derivation for PFOA. 30

Table 2-6 is Table 2.3. Evaluation Criteria for Screening Bioaccumulation Factors (BAFs) in the Public Literature. 68

3- Page xiv Table 0-1. Recommended Freshwater Perfluorooctanoic acid (PFOA) Aquatic Life Ambient Water Quality Criteria-
Superscript 3 [3] listed as footnote in the Table but not referenced in the table

4- Table 0-2. Summary of Assessment Endpoints and Measures of Effect Used in the Criteria Derivation for PFOA

<table>
<tr><th>Assessment Endpoints for the Aquatic Community</th><th>Measures of Effect</th></tr>
<tr><td>Aquatic Life: Survival, growth, and reproduction of freshwater and estuarine/marine aquatic life (i.e., fish, amphibians, aquatic invertebrates)</td><td>For effects from acute exposure:
1. LC_{50} concentrations in water, diet, and/or tissue (e.g., muscle, blood, egg)
2. NOEC and LOEC concentrations in water, diet, and/or tissue (e.g., muscle, blood, egg)

For effects from chronic exposure:
1. EC_{10} concentrations in water, diet, and/or tissue (e.g., muscle, blood, egg)
2. NOEC and LOEC concentrations in water, diet, and/or tissue (e.g., muscle, blood, egg); Only used when an EC_{10} could not be calculated for a genus.</td></tr>
</table>
</td></tr>
</table>

2.8. Additional Technical Comments to Consider

Reviewer	Comments

Please review if the highlighted muscle, blood and egg would be relevant to this section.

5- Section 1.1 Previously Derived PFOA Toxicity Values and Thresholds
Table 1.1 to be updated by current information for Australia and new Zealand.

Freshwater values are to be used on an interim basis until final marine guideline values can be set using the nationally-agreed process under the Australian and New Zealand Guidelines for Fresh and Marine Water Quality. Marine guideline values developed by CRC CARE are under consideration through the nationally-agreed water quality guideline development process. The Australian and New Zealand Guidelines for Fresh and Marine Water Quality (ANZG, 2018; ANZWQG, 2000) and Australia's PFAS Environmental Management Plan (HEPA, 2020) both recommend that a 99% species protection level be used as a guideline for bioaccumulating substances such as PFOS, PFOA, and PFHxS (as given below) as a precautionary approach.

Exposure scenario	PFOA	Exposure scenario	Comments and source
Freshwater	19 µg/L	99% species protection - high conservation value systems	Australian and New Zealand Guidelines for Fresh and Marine Water Quality - technical draft default guideline values for PFOA. The draft guidelines do not account for effects which result from the biomagnification of toxicants in air-breathing animals or in animals which prey on aquatic organisms. The WQGs advise P40FP40F[1]PP that the 99% level of protection be used for slightly to moderately disturbed systems. This approach is generally adopted for chemicals that bioaccumulate and
	220 µg/L	95% species protection - slightly to moderately disturbed systems	
	632 µg/L	90% species protection - highly disturbed systems	
	1824 µg/L	80% species protection - highly disturbed systems	

[1] https://www.waterquality.gov.au/anz-guidelines/guideline-values/default/water-quality-toxicants/local-conditions#bioaccumulation.

2.8. Additional Technical Comments to Consider	
Reviewer	**Comments**

			biomagnify in wildlife. Regulators may specify or environmental legislation may prescribe the level of species protection. required, rather than allowing for case-by-case assessments.
Exposure scenario	**PFOA**	**Exposure scenario**	**Comments and source**
Interim marine	19 µg/L	99% species protection - high conservation value systems	As above. Freshwater values are to be used on an interim basis until final marine guideline values can be set using the nationally-agreed process under the Australian and New Zealand Guidelines for Fresh and Marine Water Quality. The WQG advise that in the case of estuaries, the most stringent of freshwater and marine criteria apply, taking account of any available salinity correction. Marine guideline values developed by CRC CARE are under consideration through the nationally-agreed water quality guideline development process.
	220 µg/L	95% species protection - slightly to moderately disturbed systems	
	632 µg/L	90% species protection - highly disturbed systems	
	1824 µg/L	80% species protection - highly disturbed systems	

6- Table 1 2. Two Primary Categories of PFAS

2.8. Additional Technical Comments to Consider	
Reviewer	**Comments**
	Please refer to OECD 2021 to be consistent with PFAS terminology/nomenclature OECD (2021), Reconciling Terminology of the Universe of Per- and Polyfluoroalkyl Substances: Recommendations and Practical Guidance, OECD Series on Risk Management, No. 61, OECD Publishing, Paris **7- Table 1.3 page 5** Please review Figure 9 OECD 2021 (also attached as PDF) **8- Conceptual Model of PFOA in the Aquatic Environment and Effects** Figure 2.7 page 56- Growth as an endpoint missing in the endpoints – first pentagon **9-** Page 69 last paragraph- **Error! Reference source not found.** outlines the screening criteria for study evaluation and ranking. Cross reference to be updated to Table 2.3
Reviewer 5	Overall, I think the PFOA criteria document and the reported criteria are robust given the constraints imposed by data availability. I agree with most assumptions made by EPA and the actual criteria values reported. I would, however, suggest that EPA consider revising the 1985 Guidelines. Re-evaluating the guidelines and publishing and update, even if changes are minimal would be a valued effort. I am sure that EPA has more than enough to do but given the importance of AWQC, I think it would be better to base the science on a more recent effort. There has been a lot of solid science in the last 35 years with a focus on criteria, species sensitivity distributions, etc. Although not used heavily for PFOA the use of the 4 most sensitive taxa and then a linear regression to estimate the criteria seems less robust than other methods. I could be wrong but here is where an updated Guidelines would be helpful – that may be an effort worth exploring quantitatively.

APPENDIX A

CHARGE TO REVIEWERS

Technical Charge to External Peer Reviewers

Contract No. EP-C-17-017
Task Order 68HERH21F0090 (ERG Task 49)

External Peer Review of EPA's Draft Aquatic Life Water Quality Criterion for Perfluorooctanoic Acid (PFOA)

BACKGROUND

PFOA and PFOS toxicity studies have been conducted on a limited number of aquatic organisms, including species of fish and invertebrates, indicating that exposure to elevated concentrations of PFOA and PFOS can cause effects on survival, growth, and reproduction. In these draft criteria documents EPA is proposing two separate water quality criteria to ensure the protection of aquatic life species from the exposure to PFOA and PFOS individually.

Background on the PFOA and PFOS Aquatic Life Criteria Development Process:

Toxicity studies used to derive the PFOA and PFOS criteria were carefully evaluated and thoroughly reviewed to ensure studies were of sufficient data quality to use in criteria derivation. Scientists from the Office of Water (OW) and Office of Research and Development (ORD) conducted an extensive review of the PFOA and PFOS toxicity studies. Additionally, EPA obtained replicate-level (or treatment-level, when replicates were unavailable) concentration-response (C-R) data from publications, supplemental materials, or via contacting authors so that EPA could independently fit C-R models to estimate acute LC_{50} and chronic EC_{10} values that were used to derive the criteria to ensure endpoints used were statistically sound. Individual C-R models and resultant point estimates were also reviewed and discussed between OW and ORD to ensure the most statistically robust models informed the derivation of the PFOA and PFOS criteria. In addition to contacting study authors for C-R data (when not reported in the open literature), EPA also consulted primary authors for methods clarifications in many instances during the data quality review phase to ensure that the studies used to derive criteria were of high quality.

Overall, due to the paucity of measured freshwater toxicity data, EPA included a number of tests with unmeasured treatments to derive criteria to ensure the dataset was representative of a range of taxa and there were sufficient data to develop criteria. EPA also conducted meta-analyses to evaluate the relationship between nominal and measured test concentrations using tests with measured treatment concentrations. These meta-analyses (described in detail as Appendix L of the PFOA criteria document and Appendix O of the PFOS criteria document) suggested measured concentrations were similar to nominal concentrations and that the use of unmeasured tests, in light of data limitations, was appropriate. Additionally, estuarine/marine toxicity data limitations did not allow for the direct derivation of acute or chronic estuarine/marine criteria for PFOA or PFOS. Therefore, to develop of recommendations that states and tribes could use in adopting protective values for estuarine/marine waters, EPA developed acute PFOA and PFOS protective benchmarks using a New Approach Methodology (detailed in Appendix K of the PFOA criteria document and Appendix L of the PFOS criteria document).

Addressing data limitations to derive robust criteria/benchmarks, extensively reviewing studies, and calculating point estimates meant that the derivation of the PFOA and PFOS aquatic life criteria were developed via comprehensive, rigorous process that included collaborations across EPA scientists in OW and ORD. Beyond detailed discussions between OW and ORD, the PFOA and PFOS drafts also underwent two

rounds of review with the EPA Scoping Workgroup (consisting of additional scientists from both OW and ORD) and one round of review with a group of internal EPA Reviewers that included representatives from the Office of Water, Office of Research and Development, other EPA Program Offices, and EPA Regions. In this peer review EPA is seeking to obtain a focused, objective evaluation of the two separate draft criteria documents, one for PFOA and the other for PFOS. Generally, the charge questions below are the same for EPA's PFOA and PFOS draft aquatic life water quality criteria. However, there are some unique questions specific to the individual drafts and therefore, there are two separate sets of charge questions.

CHARGE QUESTIONS

PFOA

1) Please comment on the overall clarity of the document as it relates to the derivation of each criterion.

2) Please comment on the approach used to derive the draft criterion for PFOA. Please provide detailed comments.
   - Is the technical approach used to derive the criterion logical?
   - Does the science support the conclusions?
   - Is it consistent with the protection of freshwater aquatic life from acute, chronic, and bioaccumulative effects?

3) Please comment on the approach used to derive the draft acute estuarine/marine benchmark for PFOA. Given the limited estuarine/marine test data available, a new approach method was used to support the derivation of an acute estuarine/marine benchmark to provide states and tribes with a protective value. Please provide detailed comments.
   - Is the technical approach used to derive the benchmark logical?
   - Does the science support the conclusions?
   - Is it consistent with the protection offered by acute estuarine/marine aquatic life criteria derived using empirical data, as prescribed in the 1985 *Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses*?

4) Please comment on the use of measured and unmeasured toxicity tests to derive the respective criterion. In particular please comment on the supporting justification for using unmeasured toxicity tests in Appendix L.

5) Please comment on the toxicity data used to derive the draft criteria.
   - Were the data selected and/or excluded from the derivation of the criteria derivation appropriately utilized?
   - Are there relevant data that you are aware of that should be added to the analyses (note that EPA is working on updating the toxicity data to reflect the data in ECOTOX between Sept. 2019 through the latest update)? If so, please provide references for consideration.

In particular, please comment on:

5a. The toxicity values used to derive the PFOA criteria, with a particular emphasis on:
   i. the selection and the Acute to Chronic Ratio (ACR) to serve as the Final Acute to Chronic Ratio (FACR) and its application to derive the Final Chronic Value (FCV).

ii. the use of the qualitatively acceptable acute midge (*Chironomus plumosus*) data from Yang et al. (2014) to suggest aquatic insects are relatively tolerant to acute PFOA exposures. Specifically, Yang et al. (2014) conducted a 96-hour renewal, measured PFOA acute test with the midge, *Chironomus plumosus*. This study was not acceptable for quantitative use due to the potentially problematic source of the organisms. The reported LC_{50} was 402.24 mg/L PFOA indicating that these insects may not be one of the more sensitive taxonomic groups. Therefore, this test was excluded from the criterion calculation, but used to waive the missing insect MDR.

5b. EPA's approach for fitting concentration-response (C-R) data (described in Appendix J) as well as the specific acute LC_{50} values (Appendix A.2) and chronic EC_{10} values (Appendix C.2) that were estimated (for sensitive genera when C-R data were available) and used to derive criteria.

6) Please comment on the translation of the chronic water column criterion elements for aquatic life to derive the tissue-based criterion elements, considering the bioaccumulation of PFOA and PFOS. In particular, please comment on:

6a. Uncertainty surrounding the bioaccumulation factors (BAFs) used to translate of the chronic water column criterion elements into tissue-based criterion elements.

6b. EPA's determination of appropriate BAFs and the tissue types that the tissue criterion elements were based.

7) Please comment on the frequency and duration of the criterion elements, in particular please comment on the frequency and duration components of the tissue-based criterion elements.

8) Please provide any additional technical comments that you believe should be considered.

APPENDIX B

INDIVIDUAL REVIEWER COMMENTS

COMMENTS SUBMITTED BY

REVIEWER 1

External Peer Review of EPA's Draft Aquatic Life Water Quality Criterion for Perfluorooctanoic Acid (PFOA)

1. Please comment on the overall clarity of the document as it relates to the derivation of each criterion.

Overall, the document is clear and the reader can follow the logic of criteria derivation, and track the values used back to the cited research articles or values calculated by EPA.

2. Please comment on the approach used to derive the draft criterion for PFOA. Please provide detailed comments.

- **Is the technical approach used to derive the criterion logical?**

  Yes, the technical approach used to derive the criteria elements is generally logical. I disagree with some of the elements of the analyses, as noted in my detailed comments (see below, responses to charge question 8)

- **Does the science support the conclusions?**

  In general, the science is supportive of the general conclusions. As noted in my below detailed responses to other charge questions, I believe the science is not supportive of the work in a few key instances including:

  1. I believe the Criterion Continuous Concentration (CCC) should be potentially re-calculated considering my comments provided in response to charge question 5a.
  2. The science does not support the assumption of a 10-year recovery time for PFOA in aquatic systems.
  3. The generation of tissue criteria is weakly supported, and the uncertainty associated with these criteria should be emphasized.
  4. The NAM-generated marine Final Acute Value (FAV) and FAV/2 values (Appendix K) are highly uncertain.
  5. It is unclear if the EPA-calculated Effective Concentration 10% (EC10) values are supported; additional details on the modeling and the variability and fit of each EC10 model need to be provided.

- **Is it consistent with the protection of freshwater aquatic life from acute, chronic, and bioaccumulative effects?**

  The criteria derived are aimed at protecting aquatic life (e.g., fish, invertebrates) from the direct acute and chronic toxicity of PFOA in water. Generally, the values applied are protective and are generally similar to protective values derived by other regulatory organizations and independent (i.e., academic, private sector) scientists. Although, as based on my comments, I believe there is room for improvement. The criteria derived for tissues attempt to provide criteria that take into account bioaccumulation so that measurements in tissue can be interpreted with respect to the potential for potential effects; however, the uncertainty with the tissue criteria is high. The water and tissue criteria are not intended protective of bioaccumulative effects that may affect higher trophic levels, such as wildlife that may consume aquatic life.

3. **Please comment on the approach used to derive the draft acute estuarine/marine benchmark for PFOA. Given the limited estuarine/marine test data available, a new approach method was used to support the derivation of an acute estuarine/marine benchmark to provide states and tribes with a protective value. Please provide detailed comments.**

    - **Is the technical approach used to derive the benchmark logical?**

      The derivation of the acute marine benchmarks (FAV and Criterion Maximum Concentration (CMC)) using the New Approach Method (NAM) is highly uncertain, and I would recommend this analysis not be included as in this document. I do not feel that the analysis and subsequent criteria have high confidence for use in a regulatory application. I understand that similar analyses with other chemicals have about a 90% probability of the predicted effect value being within a factor of 5 of the actual value (Raimondo et al., 2010 – cited in document). Given the calculated CMC (3.4 mg/L), this implies the CMC has about a 90% probability of being within 0.68 to 17 mg/L. If the NAM approach stays in the document, this uncertainty and range of values should be acknowledged in the discussion.

      I would rather see tentative or provisional acute criterion developed from the limited empirical marine acute data highlighted in Appendix B and other recently published marine acute data. I place higher confidence in empirical data (even if limited to a few studies) and would suggest EPA emphasize it in addition to or in place of the values calculated by the NAM.

      I am hopeful that as new toxicity information on marine species are developed, these values can be supplanted with a proper and robust criteria calculation. If such a future analysis is possible, it should be noted.

    - **Does the science support the conclusions?**

      See above comment.

    - **Is it consistent with the protection offered by acute estuarine/marine aquatic life criteria derived using empirical data, as prescribed in the 1985 Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses?**

      The approach seems to be consistent with the approach in the 1985 guidelines. As noted above, the uncertainty with regards to the predictive capability of the interspecies correlations should be acknowledged quantitatively.

4. **Please comment on the use of measured and unmeasured toxicity tests to derive the respective criterion. In particular please comment on the supporting justification for using unmeasured toxicity tests in Appendix L.**

The consideration of toxicity data from experiments in which PFOA measurements were not made seems appropriate. The Appendix L analysis is supportive of the general observation that actual concentrations in the toxicity test waters approximated nominal values for the freshwater studies.

5. **Please comment on the toxicity data used to derive the draft criteria.**

    - **Were the data selected and/or excluded from the derivation of the criteria derivation appropriately utilized?**

In most cases, yes. Please see detailed comments on particular studies and interpretations in response to other charge questions.

- **Are there relevant data that you are aware of that should be added to the analyses (note that EPA is working on updating the toxicity data to reflect the data in ECOTOX between Sept. 2019 through the latest update)? If so, please provide references for consideration.**

  Hayman, N.T., Rosen, G., Colvin, M.A., Conder, J., Arblaster, J.A. 2021. Aquatic toxicity evaluations of PFOS and PFOA for five standard marine endpoints. Chemosphere 273:129699.

In particular, please comment on:

5a. The toxicity values used to derive the PFOA criteria, with a particular emphasis on:

i. the selection and the Acute to Chronic Ratio (ACR) to serve as the Final Acute to Chronic Ratio (FACR) and its application to derive the Final Chronic Value (FCV).

As stated in the 1985 guidelines, "the Final Acute-Chronic Ratio should be calculated as the geometric mean of the acute-chronic ratios for species whose SMAVs are close to the Final Acute Value." The guidance does not quantify "close", but it also does not specify that the Acute to Chronic Ratio (ACR) should only be derived from a single ACR from the study with the acute value that is closest to the Final Acute Value (FAV). In the case of the PFOA document, EPA is using only one ACR value, which is derived from the Zhang et al. (2013) rotifer study in which the LC_{50} is 150 mg/L. I would agree 150 mg/L relatively close to the FAV of 94 mg/L. However, I would also argue that the LC_{50} of 166 mg/L from the study with *Moina macrocopa* is also "close" to 94 mg/L (it is only 11% higher than 150 mg/L). The Species Mean Acute Value (SMAV) for *D. magna* is 253.7 mg/L is well within a factor of 3 of 94 mg/L and could also be considered "close". These three data points tend to cluster, as shown in Figure 3-5. Taking the geometric mean of the three ACRs from these three studies, which are studies that have acute values all within a factor of 2-3 of the 94 mg/L FAV, seems more consistent with the 1985 guidance to select several ACRs that are "close" to the FAV, rather than simply selecting a single ACR. The resulting recalculated ACR would be approximately 90, which provides a good, but conservative measure of central tendency and is higher than all but the highest ACR (299) shown in Table 3-8. Applying an ACR of 90 to the FAV would result in a FCV of approximately 1 mg/L, which is protective of chronic values shown in Figure 3-6 (note the Zhang et al.-derived rotifer SMCV is 0.8 mg/L, which is only slightly lower 1 mg/L). Such as recalculation would better incorporate variability in the ACRs and avoid both water-based criteria being completely driven by the results from a single organism (rotifers) as reported in a single study (Zhang et al., 2013).

ii. the use of the qualitatively acceptable acute midge (*Chironomus plumosus*) data from Yang et al. (2014) to suggest aquatic insects are relatively tolerant to acute PFOA exposures. Specifically, Yang et al. (2014) conducted a 96-hour renewal, measured PFOA acute test with the midge, *Chironomus plumosus*. This study was not acceptable for quantitative use due to the potentially problematic source of the organisms. The reported LC_{50} was 402.24 mg/L PFOA indicating that these insects may not be one of the more sensitive taxonomic groups. Therefore, this test was excluded from the criterion calculation, but used to waive the missing insect MDR.

I disagree with excluding this data point from the acute criteria calculations. I assume this data has been removed under the assumption that these animals may have been pre-exposed to PFOA and may have been more tolerant of PFOA exposures, which would result in biased-high median lethal concentration (LC50) values. If so, this should be explicitly stated. Assuming these *Chironomus* can

develop tolerance to PFOA, it seems that they would have to be exposed to rather high mg/L ranges of PFOA in water given the reported 96-hour LC50 of 402 mg/L. Based on published literature, I am unaware of natural ecosystems in China (where the animals may have been originally harvested) with concentrations of PFOA that approach this order of magnitude range (in which they could build up a tolerance). The animals were obtained from a local market, so it is also possible that they were cultured for several generations, presumably using uncontaminated water (which would further reduce the chance that multiple generations were exposed at these levels). Overall, I think it is more reasonable to assume that the animals used in the experiment have not built up an acute lethal tolerance to PFOA, and the that LC50 result is unbiased. It does seem clearly show that insects may be less sensitive to acute lethality effects of PFOA. As such, I think it should be included as a quantitative endpoint.

Additionally, it seems inconsistent to exclude this Yang et al (2014) study, when acute data from a study by Yuan et al. (2015) were included for quantitative consideration. As noted on page A-10, the animals in the Yuan et al. (2015) study were "collected from a fountain in Quanhetou, Boshan, China, and acclimated in the laboratory for an unspecified time period before use". The source of the animals is just as uncertain as the Yang et al (2014) animals, and it is unclear (if PFOA tolerance at lethal levels is possible) how many generations would be needed to shed adaptive tolerance and how this time period would compare to an "unspecified time period." Simply put, if data from experiments like Yuan et al. (2015) are quantitatively included, those from Yang et al. (2014) should also be quantitatively included (with some notes on the uncertainty of the animal sources).

5b. EPA's approach for fitting concentration-response (C-R) data (described in Appendix J) as well as the specific acute LC50 values (Appendix A.2) and chronic EC10 values (Appendix C.2) that were estimated (for sensitive genera when C-R data were available) and used to derive criteria.

More details need to be provided on the dose response modeling using R. Appendix J is helpful for providing the reader with details on the general approach, but where EC10s are modeled by EPA, the model being used (out of the 22 available in the R software package) needs to be specified. Providing some indication of variability (such as a 95% confidence interval) for the model-generated EC10s is standard practice for dose response modeling, and this information should be provided somewhere in the document. Showing the R package output of the goodness of fit statistics (or equivalent) for the modeling in an Appendix would be helpful; since this was used to select the model used in each instance of an EC10 calculation, it must be available, so I would recommend including it for full transparency and to aid future efforts in understanding the aquatic toxicology of this chemical. Additionally, it would be helpful to show the selected model fits for all calculated EC10s (as shown for the most sensitive EC10s estimated). These steps would be helpful to ensure and demonstrate quality of the model fits and reproducibility of the modeling work.

6. Please comment on the translation of the chronic water column criterion elements for aquatic life to derive the tissue-based criterion elements, considering the bioaccumulation of PFOA and PFOS.

The derivation of the tissue criteria in this manner is highly uncertain. To my knowledge this is the first time EPA has applied ambient water quality criteria protective of aquatic life direct toxicity with uptake factors (bioaccumulation factors (BAFs), bioconcentration factors (BCFs)) in this manner to calculate tissue criteria. References are made to the selenium tissue criteria, but those are used in the reverse (i.e., criteria based on measured concentrations in tissue used to calculate water criteria). The use of criteria for water with a assumed uptake factor carries a large amount of uncertainty, and in general, the use of measured concentrations in tissue linked to adverse effects is a more straightforward approach since it does not involve uptake model predictions. This needs to be noted in the text. Also, are the predicted tissue criteria

meant to be a temporary stop-gap until tissue effect data become available? This should be discussed and clarified.

In particular, please comment on:

6a. Uncertainty surrounding the bioaccumulation factors (BAFs) used to translate of the chronic water column criterion elements into tissue-based criterion elements.

The use of BAFs derived from field studies is inherently uncertain due to the wide variety of techniques used in the compiled studies, their analytical data quality, the differences in species and ecosystems, experimental designs, spatial uncertainties for mobile animals like fish, etc. That being said, the use of a BAF value (or BCF) in criteria derivation is consistent with other criteria developed by EPA. As noted above, the use of the tissue criteria needs to be considered carefully, and I think empirical tissue data from toxicity experiments should form the basis of a next iteration of a tissue criteria.

6b. EPA's determination of appropriate BAFs and the tissue types that the tissue criterion elements were based.

The development of BAFs for invertebrates, fish (whole body), and fish (muscle) seems reasonable for the application in estimating a draft or interim tissue criteria until empirical tissue data can be used to calculate tissue criteria directly.

7. Please comment on the frequency and duration of the criterion elements, in particular please comment on the frequency and duration components of the tissue-based criterion elements.

The 4-day duration seems to be supported by the time scale of toxicity discussed for the limited chronic studies selected by EPA. This assumption should be revisited as more data become available (please note in the document).

For the tissue-based criterion (page 107), there is no clear support for assuming a 10-year exceedance frequency. Given the uncertainty with the BAF-predicted tissue criteria, and how little is known regarding the recalcitrance of PFOA in aquatic ecosystems and recovery time if PFOA inputs in water were halted, the assignment of a 10-year exceedance frequency at this stage seems completely arbitrary. We simply do not yet know the time frame over which aquatic ecosystems recover from PFOA. It is not technically supported to cite recovery times for selenium to support a 10-year recovery time for PFOA, these are completely different toxicants that have their own unique fate and behavior. USEPA (1985) guidance suggests assuming a 3-year frequency as a default, and the discussion on page 17-108 is not scientifically convincing enough to modify it to 10 years.

Additionally, it should be noted that the exceedance frequency for another organic chemical, Tributyltin (TBT) was set at 3 years by EPA in derivation of that criteria. TBT exhibits uptake factors similar to PFOA (i.e., BCF of approximately 2,000 L/kg, wet weight for goldfish, as noted in the EPA TBT criteria document, which is higher than the PFOA BAFs of 30-175 L/kg, wet weight being used to calculate the fish tissue criteria). TBT is also persistent in aquatic ecosystems, as noted by EPA. Given TBT is at least an organic chemical, it is a closer analog than selenium, which is an element. As such, the exceedance frequency for the PFOA tissue criterion should be set at the default of 3 years unless EPA can provide convincing technical information specific to recovery times for PFOA.

Additionally, on page 108, the paragraph that begins with "Metals and other chemical pollutants such as PFOA..." is not convincing as any quantitative support for EPA's 10-year exceedance frequency for the chronic tissue-based criteria. The text as written may give the reader the conclusion that PFOA recovery may be "on

the order of decades", as EPA notes for selenium. There is no support for the conjecture that PFOA recovery may be "relatively slow" or require decades, as noted in my above comment.

8. Please provide any additional technical comments that you believe should be considered.

I have the additional detailed comments:

a) Please note that the comments provided in this file reflect a focus on of key portions of the "Draft of the Aquatic Life Water Quality Criterion…" document as directed by the above charge questions provided to me. Given time and resource constraints and the scope of my review, it was not feasible to provide a detailed review of the entire document and all of the supporting references and their associated results and conclusions. As such, I reserve my right to supplement or amend my comments in future, pending additional review or new information. Thank you for the opportunity to assist EPA in its work on this very important matter, and I was honored to be selected as a reviewer.

b) Page 7: Please note in Figure 1-1 that this is the linear isomer of PFOA. It would be helpful to note that the PFOA data in this study are likely from experiments with water spiked with the linear PFOA isomer. It is hypothesized that toxicity and bioaccumulation may differ between branched and linear forms of PFCAs and PFASs. Linear PFOA is thought to be more accumulative (as noted on Page 61) and potentially more toxic to aquatic life when the dose is expressed as an external water concentration. At some sites, a portion of the concentrations of PFOA in water (which are reported as the sum of branched and linear PFOA) can be branched PFOA, so criteria derived from linear PFOA could be overly protective. Please include this uncertainty in the discussion in the document.

c) Page 47: With regard to the discussion of biomagnification factors (BMFs), please acknowledge the Martin et al. 2003 paper where the BMF for fish was noted to be less than one. This paper is very helpful for understanding biomagnification from the diet in fish, as it was a controlled PFOA-spiked food study.

d) Page 51: Starting here on this page and in the rest of this section, most of the units need to be specified for dry weight or wet weight for concentrations of PFOA in tissue. There were other instances of this error in the document as well. For units of every concentration of PFOA in tissue, please be sure to specify dry weight or wet weight.

e) Page 61: Regarding "nearly 90% of measured concentrations fell within 20% of paired nominal concentrations", I believe that applies to the freshwater studies reviewed (please note).

f) Page 64: The use of EC10 values instead of effective concentration 20% (EC20) values for chronic values is inconsistent with EPA's general practice for developing aquatic life values. The selection of EC10s for the selenium criteria (EPA, 2016) was associated with the derivation of tissue guidelines. In the EPA (2016) document, EPA noted "EC20s have historically been used in the derivation of EPA criteria applicable to the water medium". As noted in the EPA (2016) selenium guidance EC10s were selected over EC20s "given the nature of exposure and effects for this bioaccumulative chemical." Additionally EPA (2016) selected EC10 for selenium because "it was found that the dose-response curves for selenium across a broad range of fish genera are very steep, such that a small change in selenium tissue concentration yielded a large increase in observed adverse effect."

g) First, all the derivation of aquatic life criteria for "bioaccumulative chemicals" have not followed the process used for selenium, and there is no quantitative discussion in the current document that compares the bioaccumulation values for selenium to those of PFOA in a manner than justifies the

use of EC10s. For example, EPA in its 2016 aquatic life criteria for cadmium noted that cadmium "can bioaccumulate in aquatic organisms", but EC20s (not EC10s) were used as chronic values in the derivation of aquatic life criteria in that document. Fundamentally, there is a logical disconnect between adding additional conservativism (i.e., using EC10s instead of EC20s) simply because a chemical has a higher bioaccumulative potential than another chemical or exceeds a BCF or BAF criteria used to determine a chemical has "bioaccumulative" status by typical chemical registration guidelines. The use of chronic exposure toxicology data generally assumes that concentrations in the organisms have reached steady state and, and thus, any bioaccumulation that has occurred is accounted for and manifests in toxic action. Coincidentally, the general assumption is that toxic responses have plateaued as well and that effective doses (measured via external concentrations in water or concentrations in the organism) will not change significantly with additional exposure time. The bioaccumulative nature of the toxicant at that point is a moot point with regards to toxic effects in an aquatic organism, so there seems no need to add additional conservatism in the estimation of a threshold for potential ecologically-significant effects on aquatic life. Adding additional conservatism to the aquatic life criteria to protect other trophic levels (i.e. wildlife that consume aquatic life) or human consumers of aquatic life, which does involve bioaccumulation of chemicals in aquatic organisms, is not justified. Criteria to protect wildlife and humans exposed via exposure pathways involving bioaccumulation of chemicals in aquatic life are handled via separate approaches, and are completely disconnected from the acute and chronic toxicity data developed to evaluate the risks to aquatic invertebrates and lower trophic level vertebrates like fish and amphibians.

h) Second, EPA has not provided any analysis of the dose response curves that demonstrates the need for EC10s versus EC20s (as was mentioned for selenium). Additionally, justification of the use of EC10s by simply referencing the regulatory policies of other countries seems to be insufficient as the basis for a US policy, and is unsatisfying from a scientific perspective.

i) More discussion is needed to support the poorly-supported move from EC20s to EC10s, or alternately, EC20s need to be used in throughout the document, as consistent with past EPA practice in aquatic life criteria derivation. EC10s are more conservative than EC20s, but there is often greater variability and uncertainty associated with EC10 values given the typical 50% effect ranges that are generally targeted in the experimental designs of typical toxicological studies. Additionally, as noted in EPA's 2016 aquatic life criteria document for cadmium, EC10s are "rarely statistically significantly different from the control treatment." A 20% effect has often been discussed as a point of departure of ecologically-significant population- and community-level effects (e.g., Suter, 2000: Suter, G.W., Efroymson, R.A., Sample, B.E., & Jones, D.S. (2000). Ecological Risk Assessment for Contaminated Sites. CRC Press. April).

j) Overall, the adoption of a more conservative 10% effect level (i.e., EC10) for chronic values used in criteria calculation carries large environmental management and policy implications. As noted above, clarification and careful justification is needed. EPA needs to clearly articulate (ideally with ample scientific support) why the additional conservatism is needed. This important potential policy matter deserves an open and earnest discourse among the scientific, stakeholder, and regulated communities.

k) Page 67: It appears that only studies in which organisms exposed via diet were included for evaluation of tissue criteria. Is this correct? It is questionable to exclude effect concentrations in tissue from experiments in which exposure of PFAS was only via water. EPA (2016) took the "dietary exposure only" approach with selenium because the primary exposure route for selenium has been shown to be via the diet in natural ecosystems. In contrast, for many aquatic animals (especially

lower trophic level fish and invertebrates), a significant portion of the exposure to PFOA is via non-dietary pathways. Part of this is due to the fact that controlled studies (e.g., Martin et al., 2003 studies cited in the document) have found that water-to-organism BCFs for aquatic life such as fish are generally larger than diet-to-organism biomagnification factors (BMFs). Additionally, there is no reason to expect dietary or non-dietary exposure pathways would affect toxic responses given the relatively rapid internal kinetics of PFAS in aquatic life (i.e., half-life of hours or days), especially for small invertebrates and fish that are in relative equilibrium with their surrounding exposure water.

l) Page 82: The intrinsic rate of increase (basis of the chronic value from the Zhang et al. (2013) study is not a typical endpoint used in aquatic life criteria derivation. This chronic value has a large influence over the criteria calculated in this document, as noted below. More detail is needed on the calculation of the endpoints used in the two Zhang et al. papers, and some additional information should be provided on what this endpoint means ecologically in comparison to typical reproductive endpoints used in criteria derivation. Note that for stable populations, the intrinsic rate is 0, so achieving a maximal rate of increase is not always ecologically sustainable. Some additional explanation would be helpful to the reader.

m) Page 98: Both the Criterion Maximum Concentration (CMC) and Criterion Continuous Concentration (CCC) are heavily influenced by toxicity tests on a single organism (rotifers, Brachionus calyciflorus) conducted by Zhang et al. (2013) and Zhang et al. (2014). For example, the rotifer acute value is the second lowest value of the four values used to calculate the FAV and CMC. The ACR used to calculate the CCC (using the with the FAV) is derived entirely from the Zhang et al. (2013) study (Page 101). I would recommend the ACR be recalculated to reflect more of a central tendency estimate that incorporates other ACRs, rather than relying solely on the Zhang et al. (2013) result (see above 5a comment). Regardless, I think a few sentences should be added to note the strong influence of these rotifer studies on PFOA criteria. Given this, additional experimentation and verification of the Zhang et al. results seems would seem to be useful and this is worth mentioning in the document as well.

n) Page G-1: Please explain the acceptable duration acceptable for tests for which "Duration too short" is noted. It would be good to provide the acceptable durations that would be considered acceptable for these species. Perhaps a summary table for acceptable durations for particular endpoints could be provided in this document.

o) Appendix G: The endpoints such as "decrease mRNA expression levels of neural genes DjFoxD, DjotxA and DjotxB" and "decrease in inflammatory cytokines (IL-1ß and IL-21) in spleen" are atypical. These should be noted under "Deficiencies".

p) Page H-2: First use of "atypical duration" in the table. This entry is inconsistent with other entries (e.g., "duration too short") and does not clearly describe why the experiment is not considered. Please explain this table entry.

COMMENTS SUBMITTED BY

REVIEWER 2

External Peer Review of EPA's Draft Aquatic Life Water Quality Criterion for Perfluorooctanoic Acid (PFOA)

1. Please comment on the overall clarity of the document as it relates to the derivation of each criterion.

I thought that the document was well written and laid out. I thought that the document clearly laid out the approach that the EPA used to derive each criterion. I thought it clearly outlined the approach that the EPA chose in deciding which data to use in their derivation and how these data would be used in derivation.

The appendices are very useful in providing added detail and the data that were used in the derivation of the criteria. The appendices allow for a high level of transparency around how the criteria were generated.

In Table 3-1, the acronym "GMAV" was used in the caption, but I could not locate where this acronym was defined earlier in the document.

The captions of figures and tables are not sufficiently detailed. Figures and tables should be able to stand on their own. Also, the use of acronyms in the caption of tables and figures decreases clarity, e.g., Fig. 3-5. The use of acronyms in the figure or table is fine, as long as they are defined in the caption of the figure or table.

2. Please comment on the approach used to derive the draft criterion for PFOA. Please provide detailed comments.

- **Is the technical approach used to derive the criterion logical?**
- **Does the science support the conclusions?**
- **Is it consistent with the protection of freshwater aquatic life from acute, chronic, and bioaccumulative effects?**

Yes, the technical approach used by the EPA to derive the criterion is logical and defensible. The approach is also clearly laid out in the document. Dividing the 5th centile of the acute GSD by 2 is sufficiently conservative to ensure the protection of 95% of species, based on the data currently available. The use of the ACR from Brachionus calyciflorus to construct a chronic GSD is a valid approach, along with being the most conservative option.

Yes, I think the science supports the EPA's conclusions. However, there appears to be several studies that were not considered by the EPA. I have listed these studies below.

Yes, I think the approach taken by the EPA is sufficiently conservative to be protective of freshwater aquatic life from acute, chronic, and bioaccumulative effects based on the data that was available at the time. It was a good idea to evaluate the influence on non-North American species on the derivation of the criteria.

3. Please comment on the approach used to derive the draft acute estuarine/marine benchmark for PFOA. Given the limited estuarine/marine test data available, a new approach method was used to support the derivation of an acute estuarine/marine benchmark to provide states and tribes with a protective value. Please provide detailed comments.

- **Is the technical approach used to derive the benchmark logical?**
- **Does the science support the conclusions?**
- **Is it consistent with the protection offered by acute estuarine/marine aquatic life criteria derived using empirical data, as prescribed in the 1985 Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses?**

The technical approach using Web-ICE to determine an acute benchmark for estuarine/marine species is logical. The science has shown that Web-ICE can effectively be used to derive effect measures for additional

species using species for which data is available. I think the approach taken by EPA has included sufficient conservatism to address the relatively large amount of uncertainty around the acute toxicity of PFOA to estuarine and marine species. The proposed acute benchmark for estuarine and marine species is an order of magnitude lower than the acute benchmark for freshwater species, which I think underscores the conservatism used by EPA in determining an acute benchmark for estuarine and marine species. That said, the benchmark should be used cautiously due to the relatively large amount of uncertainty and effort should be made to generate acute and chronic toxicity data for PFOA on estuarine and marine species, particularly fish.

4. Please comment on the use of measured and unmeasured toxicity tests to derive the respective criterion. In particular please comment on the supporting justification for using unmeasured toxicity tests in Appendix L.

I am concerned with the approach of using the agreement of measured and nominal concentrations from studies that measured the concentration of PFOA in their tests to determine whether to use toxicity data from studies that did not measure the concentration PFOA in their tests. My concern stems from this approach having to assume that studies that did not measure the concentration of PFOA in their experiments performed the dosing of PFOA with the same care and skill as those studies that did measure the concentration of PFOA in their experiments and measured concentrations within 20% of nominal. My concern is compound by 79.5% and 60% of the acute and chronic tests, respectively, only reporting nominal test concentrations. The EPA's approach uses the agreement of measured and nominal concentration in a minority of studies to determine whether to include the majority of studies on their assessment.

I am assuming that there wouldn't be sufficient data to determine a criterion without using data from studies that did not measure the concentrations of PFOA in their experiment?

I think the approach that the EPA has used to determine the level of agreement between the nominal and measured concentration of PFOA in the studies that measured the concentration is logical and valid. It is encouraging that the agreement on average is high. Again, my largest concern is assuming this agreement in a minority of studies is present in all studies.

5. Please comment on the toxicity data used to derive the draft criteria.

- **Were the data selected and/or excluded from the derivation of the criteria derivation appropriately utilized?**

I think the data used in the derivation of the criteria were appropriate. As mentioned above, I am a little concerned about the use of toxicity data from studies that did not measure the concentration of PFOA in their experiments, especially considering the proportion of studies that did not measure the concentrations. The confirmation of exposure concentrations is an important principle of sound ecotoxicology.

- **Are there relevant data that you are aware of that should be added to the analyses (note that EPA is working on updating the toxicity data to reflect the data in ECOTOX between Sept. 2019 through the latest update)? If so, please provide references for consideration.**

Hayman, N.T., Rosen, G., Colvin, M.A., Conder, J., Arblaster, J.A., 2021. Aquatic toxicity evaluations of PFOS and PFOA for five standard marine endpoints. Chemosphere 273, 129699.. doi:10.1016/j.chemosphere.2021.129699

Logeshwaran, P., Sivaram, A.K., Surapaneni, A., Kannan, K., Naidu, R., Megharaj, M., 2021. Exposure to perfluorooctanesulfonate (PFOS) but not perflurorooctanoic acid (PFOA) at ppb concentration

induces chronic toxicity in Daphnia carinata. Science of The Total Environment 769, 144577.. doi:10.1016/j.scitotenv.2020.144577

Bartlett, A.J., De Silva, A.O., Schissler, D.M., Hedges, A.M., Brown, L.R., Shires, K., Miller, J., Sullivan, C., Spencer, C., Parrott, J.L., 2021. Lethal and sublethal toxicity of perfluorooctanoic acid (PFOA) in chronic tests with Hyalella azteca (amphipod) and early-life stage tests with Pimephales promelas (fathead minnow). Ecotoxicology and Environmental Safety 207, 111250.. doi:10.1016/j.ecoenv.2020.111250

Tornabene, B.J., Chislock, M.F., Gannon, M.E., Sepúlveda, M.S., Hoverman, J.T., 2021. Relative acute toxicity of three per- and polyfluoroalkyl substances on nine species of larval amphibians. Integrated Environmental Assessment and Management 17, 684–690.. doi:10.1002/ieam.4391

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.. doi:10.1002/etc.4690

BAF data provided in the supplementary information of Prosser et al. 2016 study for three freshwater species does not appear to have been considered.

Prosser, R.S., Mahon, K., Sibley, P.K., Poirier, D., Watson-Leung, T., 2016. Bioaccumulation of perfluorinated carboxylates and sulfonates and polychlorinated biphenyls in laboratory-cultured Hexagenia spp., Lumbriculus variegatus and Pimephales promelas from field-collected sediments. Science of The Total Environment 543, 715–726. doi:10.1016/j.scitotenv.2015.11.062

In particular, please comment on:

5a. The toxicity values used to derive the PFOA criteria, with a particular emphasis on:

i. the selection and the Acute to Chronic Ratio (ACR) to serve as the Final Acute to Chronic Ratio (FACR) and its application to derive the Final Chronic Value (FCV).

I think the approach that led to the selection of the ACR for Brachionus calyciflorus was appropriate. It is also the most conservative approach with the ACR for this species being the largest of the four species with ACRs.

ii. the use of the qualitatively acceptable acute midge (Chironomus plumosus) data from Yang et al. (2014) to suggest aquatic insects are relatively tolerant to acute PFOA exposures. Specifically, Yang et al. (2014) conducted a 96-hour renewal, measured PFOA acute test with the midge, Chironomus plumosus. This study was not acceptable for quantitative use due to the potentially problematic source of the organisms. The reported LC50 was 402.24 mg/L PFOA indicating that these insects may not be one of the more sensitive taxonomic groups. Therefore, this test was excluded from the criterion calculation, but used to waive the missing insect MDR.

I think the EPA's decision that the data from Yang et al. (2014) was not acceptable for quantitative use was appropriate. The source of the larvae is problematic. The conclusion that insects may not be one of the most sensitive taxa is valid. The NOEC for Chironomus tentans of 100 mg/L reported by MacDonald et al. (2004) also supports this conclusion.

5b. EPA's approach for fitting concentration-response (C-R) data (described in Appendix J) as well as the specific acute LC_{50} values (Appendix A.2) and chronic EC_{10} values (Appendix C.2) that were estimated (for sensitive genera when C-R data were available) and used to derive criteria.

I think the approach that the EPA used to determine effect measure from concentration-response data was appropriate. The use of the drc package in R to fit 22 different models to the empirical data and then using several criteria (e.g., AIC, residual standard errors, confidence intervals) to evaluate the fit of the different models is robust. It would have been useful if the EPA reported the 22 different models in Appendix J.

I think the LC50 and EC10 values determined by the EPA using the approach mentioned in the previous paragraph was appropriate. It is valid for these effect measures to be determined when the concentration-response data has been provided by the authors of the study. The EPA has also made is clear in Appendix A.2 and C.2 how they determined these effect measures using the concentration-response data provide in the studies. This generates a high level of transparency in the derivation of the criterion.

6. Please comment on the translation of the chronic water column criterion elements for aquatic life to derive the tissue-based criterion elements, considering the bioaccumulation of PFOA and PFOS. In particular, please comment on:

6a. Uncertainty surrounding the bioaccumulation factors (BAFs) used to translate of the chronic water column criterion elements into tissue-based criterion elements.

There appears to be an error in sub-section 2.11.3; "The resulting tissue-based criteria magnitudes correspond to the tissue type from the geometric mean BAF used in the equation (see Section 2.12.3.1)." I cannot locate sub-section 2.12.3.1. I assume the authors meant sub-section 2.11.3.1? Sub-section 2.12.3 and sub-section 2.12.3.1 are also referenced on pages 105 & 106, respectively.

I think the EPA has sufficiently addressed the uncertainty around the use of BAFs and the chronic water column criterion in the derivation of tissue-based criterion. They have indicated that tissue-based criterion should only be observed once in 10 years. The use of the geometric mean of the reported BAFs incorporates the range of BAFs that may be present for different invertebrate and fish species. The use of the chronic water column criterion also builds in added conservatism to the tissue-based criterion.

Prosser et al. (2016) reported BAFs for PFOA in three freshwater species (two invertebrates and one fish) (See Tables S29-31 in Supplementary Information), but it was not considered in this assessment. It is not clear why it was not considered.

Prosser, R.S., Mahon, K., Sibley, P.K., Poirier, D., Watson-Leung, T., 2016. Bioaccumulation of perfluorinated carboxylates and sulfonates and polychlorinated biphenyls in laboratory-cultured Hexagenia spp., Lumbriculus variegatus and Pimephales promelas from field-collected sediments. Science of The Total Environment 543, 715–726. doi:10.1016/j.scitotenv.2015.11.062

6b. EPA's determination of appropriate BAFs and the tissue types that the tissue criterion elements were based.

The evaluation criteria for BAFs outline in Table 2-3 are appropriate and the decision to only use high and medium quality BAFs is justified based on the criteria that would make a BAF low quality. It was a good idea to use fish BAFs based on the concentration in muscle and whole body. Muscle tissue is usually exclusively sampled in large fish, especially as part of fish consumption guidelines. The whole body is more appropriate for small fish and invertebrate species, e.g., minnows, benthic macroinvertebrates.

7. Please comment on the frequency and duration of the criterion elements, in particular please comment on the frequency and duration components of the tissue-based criterion elements.

As per Table 0-1, I think the chosen durations and frequencies for the acute and chronic criteria are appropriate. They will ensure protection of aquatic life. The duration of the tissue-based criterion is appropriate as the concentration will be measured when biota is collected. The 10 year frequency is appropriate considering that for biota to reach the tissue-based criteria, they would likely to have be exposed to concentrations at or above the chronic criteria for an extended period of time.

8. Please provide any additional technical comments that you believe should be considered.

I think the EPA's criteria for PFOA are very defensible based on the science and data available. I think they did a great job clearly laying out how they derived the criteria and providing all of the data that was used in the derivation.

COMMENTS SUBMITTED BY

REVIEWER 3

External Peer Review of EPA's Draft Aquatic Life Water Quality Criterion for Perfluorooctanoic Acid (PFOA)

1. **Please comment on the overall clarity of the document as it relates to the derivation of each criterion.**

   - I have confidence is the PFOA draft criteria, these are more in-line with thresholds put forth by other agencies. Importantly, a discussion on the difference in derivation of the thresholds would be welcome; for instance, others are derived using Species Sensitivity Distributions of the complete dataset. In comparison, EPA uses a subset of data on the genus identified as most sensitive.
   - Table 1 – does not list units of thresholds.
   - Define CMC at first use
   - Table 3-6 reverses order of sensitive taxa compared to the previous tables.

2. **Please comment on the approach used to derive the draft criterion for PFOA. Please provide detailed comments.**

   - **Is the technical approach used to derive the criterion logical?**
     - Why did EPA derive the LC50 for Chydorus when an EC50 was provided by the authors; however, accepted the EC50s for the two mussel species? This is an inconsistency.
   - **Does the science support the conclusions?**
     - See response immediately below
   - **Is it consistent with the protection of freshwater aquatic life from acute, chronic, and bioaccumulative effects?**
     - Overall, the draft criteria are in agreement with other thresholds generated using the species/genus sensitivity distributions; which seems like a more robust approach given data are not restricted to a subset of studies (albeit most sensitive). Although new data should be evaluated and potentially incorporated into these criteria calculations it is unclear if those data would alter these currently drafted thresholds.

3. **Please comment on the approach used to derive the draft acute estuarine/marine benchmark for PFOA. Given the limited estuarine/marine test data available, a new approach method was used to support the derivation of an acute estuarine/marine benchmark to provide states and tribes with a protective value. Please provide detailed comments.**

   - **Is the technical approach used to derive the benchmark logical?**
     - See comment immediately below
   - **Does the science support the conclusions?**
     - The PFOA - LOEC reported for *Mytilus* in Fabbri et al. 2014 is 0.1µg/L; in the draft criteria this is listed as >1 mg/L. The justification provided for dismissing this effect given 50% of the test organisms did not experience is not compelling.
   - **Is it consistent with the protection offered by acute estuarine/marine aquatic life criteria derived using empirical data, as prescribed in the 1985** *Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses***?**
     - No, this is a new approach; however, it follows the spirit of the 1985 guidelines.

4. Please comment on the use of measured and unmeasured toxicity tests to derive the respective criterion. In particular please comment on the supporting justification for using unmeasured toxicity tests in Appendix L.

Reviewer 3's response to this question was provided in this person's review of the draft PFOS document, with a statement that the comment also pertained to PFOA; therefore, it is copied here.

i. This seems acceptable for the time being. Having worked in the laboratory with PFOS, I can make a first-hand testament that mixing PFOS into exposures solutions does not guarantee a homogenous mixture despite working at solutions well below the solubility limit. There are nuances associated with achieving homogeneity of the exposure solution, we have developed a PFAS mixing protocol to reduce chemical clumping and this increases uniformity of the solutions. Furthermore, there is approximately 30% variability of PFOS quantitatively (see…Rewerts et al. 2020); so, the best measurement still has significant variability."

5. Please comment on the toxicity data used to derive the draft criteria.

- **Were the data selected and/or excluded from the derivation of the criteria derivation appropriately utilized?**
  - With the exception of the Fabbri et al. 2014, data currently evaluated, and the associated decision matrix seem appropriate. However, as noted below, there are new data available following this draft.
- **Are there relevant data that you are aware of that should be added to the analyses (note that EPA is working on updating the toxicity data to reflect the data in ECOTOX between Sept. 2019 through the latest update)? If so, please provide references for consideration.**
  - McCarthy et al. 2021 - freshwater
  - Hayman et al. 2021 – marine
  - Logeshwaran et al. 2021 – freshwater
  - Li et al. 2021 – freshwater/plant
  - Etc.

In particular, please comment on:

5a. The toxicity values used to derive the PFOA criteria, with a particular emphasis on:

i. the selection and the Acute to Chronic Ratio (ACR) to serve as the Final Acute to Chronic Ratio (FACR) and its application to derive the Final Chronic Value (FCV).

- The use of ACR is appropriate given data limitations. However, either I am missing it or it is unclear as to why Chydorus is not included in the chronic data set when it is the most sensitive in the acute.

ii. the use of the qualitatively acceptable acute midge (*Chironomus plumosus*) data from Yang et al. (2014) to suggest aquatic insects are relatively tolerant to acute PFOA exposures. Specifically, Yang et al. (2014) conducted a 96-hour renewal, measured PFOA acute test with the midge, *Chironomus plumosus*. This study was not acceptable for quantitative use due to the potentially problematic source of the organisms. The reported LC50 was 402.24 mg/L PFOA indicating that these insects may not be one of the more sensitive taxonomic groups. Therefore, this test was excluded from the criterion calculation, but used to waive the missing insect MDR.

- McCarthy et al. 2021 reported midge data for PFOA with an EC50 of 192 mg PFOA/L following a 20-day exposure. Following EPA review of this publication, this may fulfill the missing insect MDR; however, still supports the conclusion of likely not a sensitive taxa.

5b. EPA's approach for fitting concentration-response (C-R) data (described in Appendix J) as well as the specific acute LC_{50} values (Appendix A.2) and chronic EC_{10} values (Appendix C.2) that were estimated (for sensitive genera when C-R data were available) and used to derive criteria.

- This is a sound scientific approach, what is unclear is when EPA employs this vs other times when it is not used. i.e. chydorus vs two mussle spp in the acute studies

6. Please comment on the translation of the chronic water column criterion elements for aquatic life to derive the tissue-based criterion elements, considering the bioaccumulation of PFOA and PFOS. In particular, please comment on:

6a. Uncertainty surrounding the bioaccumulation factors (BAFs) used to translate of the chronic water column criterion elements into tissue-based criterion elements.

6b. EPA's determination of appropriate BAFs and the tissue types that the tissue criterion elements were based.

- This sentence is very confusing: "EPA examined the potential for criteria using only those studies in which test organisms were exposed to PFOA in their diet, because such studies would most closely replicate real-world exposures (diet and/or diet plus water)." The tissue criteria are based on water exposures, the relevance of this statement and evaluation is lost on me.
- A table summarizing the species used to derive the BAFs would be helpful to evaluate the comprehensiveness, i.e. pelagic vs sediment feeders.

7. Please comment on the frequency and duration of the criterion elements, in particular please comment on the frequency and duration components of the tissue-based criterion elements.

- This is a not an easy statement to comment on, as it may be unlikely that the aquatic receptors will exceed or reach these tissue concentrations prior to exceedances from the CCC. Importantly, PFOA is not particularly bioaccumulative compared to PFOS and this likely a less sensitive threshold.

8. Please provide any additional technical comments that you believe should be considered.

- All technical comments have been previously mentioned

COMMENTS SUBMITTED BY

REVIEWER 4

External Peer Review of EPA's Draft Aquatic Life Water Quality Criterion for Perfluorooctanoic Acid (PFOA)

1. Please comment on the overall clarity of the document as it relates to the derivation of each criterion.

EPA has drafted the PFOA aquatic life criteria to be consistent with methods described in EPA's "*Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses*" (U.S. EPA 1985). I congratulate the EPA Team for a very thorough, comprehensive analysis of toxicological data to derive each criterion.

- The report is technically sound and is very clearly written.
- The criteria have been derived using strong science-based evidence.
- Sub-sections on overview of PFAS, PFAS nomenclature, problem formulation, exposure pathways, transformation and degradation of PFOA precursors in the aquatic environment sources, concentration reported in environment and existing criteria (both nationally and internationally) help to set the scene before toxicological data is presented and assessed for developing various criterion.
- The freshwater acute water column-based criterion, the chronic water column-based chronic criterion, the chronic fish whole-body tissue criterion, the chronic fish muscle tissue criterion and the chronic invertebrate whole-body tissue criterion have been developed and reported in this report.
- Acute and chronic MDRs for PFOA estuarine/marine criteria derivation were not met due to fewer empirical PFOA toxicity data. EPA Team developed an acute aquatic life benchmark for estuarine/marine environments based on Interspecies Correlation Estimation (ICE) model.
- The relative sensitivity of freshwater plants to PFOA exposures indicated plants are less sensitive than aquatic vertebrates and invertebrates so plant criteria were not considered.
- EPA Team has provided extensive background information on toxicity data assessment and collated this information in various appendices such as Appendix A-Summary Table of Acceptable Quantitative Freshwater Acute PFOA Toxicity Studies; Appendix B-Detailed PFOA Acute Toxicity Study Summaries and Corresponding Concentration-Response Curves (when calculated; Appendix C: Acceptable Freshwater Chronic PFOA Toxicity Studies); Appendix D- Acceptable Estuarine/Marine Chronic PFOA Toxicity Studies; Appendix E-Acceptable Freshwater Plant PFOA Toxicity Studies; Appendix F- Acceptable Estuarine/Marine Plant PFOA Toxicity Studies; Appendix G-Summary Table of Qualitative Freshwater PFOA Toxicity Studies; Appendix H-Other Estuarine/Marine PFOA Toxicity Studies Unused PFOA Toxicity Studies; Appendix I-Unused PFOA Toxicity Studies; Appendix J- EPA Methodology for Fitting Concentration-Response Data and Calculating Effect Concentration-Fitting Concentration Response Data in R; Appendix K-Derivation of Acute Protective PFOA Benchmarks for Estuarine/Marine Waters through a New Approach Method (NAM); and Appendix L Meta-Analysis of Nominal Test Concentrations Compared to Corresponding Measured Test Concentrations.

2. Please comment on the approach used to derive the draft criterion for PFOA. Please provide detailed comments.

- **Is the technical approach used to derive the criterion logical?**
- **Does the science support the conclusions?**
- **Is it consistent with the protection of freshwater aquatic life from acute, chronic, and bioaccumulative effects?**

This EPA report provides a critical review of toxicity data identified in EPA's literature search for PFOA, including the anionic form (CAS No. 45285-51-6), the acid form (CAS No. 335-67-1), and the ammonium salt (CAS No. 3825-26-1). It quantifies the toxicity of PFOA to aquatic life, and provides criteria intended to protect aquatic life from the acute and chronic toxic effects of PFOA. The detailed assessment is as follows:

- These criteria have been derived using robust methods and the best available toxicity data on aquatic life.
- The approach used to derive the draft criterion for PFOA is very logical and consistent with the protection offered by acute and chronic aquatic life criteria derived using empirical data, as prescribed in the 1985 *Guidelines.*
- Exclusion and inclusion criteria are appropriately discussed in the context of the toxicological data reported in the literature and provide additional evidence on the selection of toxicity data criteria development.
- With limited toxicity datasets to North American resident species, non-North American resident species served as taxonomically-related surrogate test organisms. For example, *Oryzias latipes* is a common ecotoxicity test species that served as a surrogate for untested fish species residing in North America.
- The acute measures of effect on aquatic organisms selected included the lethal concentration (LC50), effect concentration (EC50), or inhibitory concentration (IC50) estimated to produce a specific effect in 50 percent of the test organisms
- The endpoint for chronic exposures incorporated the effect concentration estimated to produce a chronic effect on survival, growth, or reproduction in 10 percent of the test organisms (EC_{10}). This approach has been also consistent with the harmonized guidelines from OECD and the generally preferred effect level for countries such as Canada, Australia, and New Zealand.
- Reported (No Observed Effect Concentrations) (NOECs) and (Lowest Observed Effect Concentrations) (LOECs) were only used for the derivation of a chronic criterion when a robust EC_{10} could not be calculated for the genus.
- Furthermore, EPA independently calculated these toxicity values if sufficient raw data were available for EPA to conduct statistical analyses. EPA's independently-calculated toxicity values were used preferentially, where available.
- EPA developed protective tissue-based criteria through a bioaccumulation factor approach. This was based on the application of evaluation criteria for screening bioaccumulation factors (BAFs).
- The freshwater Final Acute Value (FAV) for PFOA was calculated as 91.34 mg/L and freshwater acute criterion water column magnitude (criterion maximum concentration, CMC), as 46 mg/L PFOA, using the procedures described in the 1985 Guidelines. This values is expected to be protective of 95% of freshwater genera potentially exposed to PFOA under short-term conditions of one-hour of duration, if the one-hour average magnitude is not exceeded more than once in three years.
- Toxicity data were available for only two families, an estuarine/marine FAV could not be calculated to derive an estuarine/marine acute criterion. Further benchmark was developed using predictive approach and discussed later in this document.
- Tissue-based criteria were also developed using comprehensive methods and assessment is provided as response to charge Question 5.

3. Please comment on the approach used to derive the draft acute estuarine/marine benchmark for PFOA. Given the limited estuarine/marine test data available, a new approach method was used to support the derivation of an acute estuarine/marine benchmark to provide states and tribes with a protective value. Please provide detailed comments.

- **Is the technical approach used to derive the benchmark logical?**

- **Does the science support the conclusions?**
- **Is it consistent with the protection offered by acute estuarine/marine aquatic life criteria derived using empirical data, as prescribed in the 1985 *Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses*?**

    - EPA applied The ICE model predictions to supplement the available test dataset to help fill missing MDRs and allow the derivation of acute estuarine/marine benchmark recommendations for aquatic life using procedures consistent with those in the 1985 Guidelines**.** Total of 3104 datapoints from 398 models were evaluated.
    - The draft acute benchmark for estuarine/marine aquatic life is lower than the recommended acute freshwater criterion (46 mg/L), suggesting that estuarine/marine species may be more acutely sensitive to PFOA. According to Hayman et al., 2021, marine species, compared to freshwater, may have a higher sensitivity to PFOA.
    - There are two more studies published, the toxicity values for marine/estuarine species.
        - Stuart L. Simpson, Yawen Liu, David A. Spadaro, Xinhong Wang; Rai S. Kookana and Graeme E. Batley Chronic effects and thresholds for estuarine and marine benthic organism exposure to perfluorooctane sulfonic acid (PFOS)-contaminated sediments: Influence of organic carbon and exposure routes https://doi.org/10.1016/j.scitotenv.2021.146008
        - Nicholas T Hayman , Gunther Rosen , Marienne A Colvin , Jason Conder , Jennifer A Arblaster Aquatic toxicity evaluations of PFOS and PFOA for five standard marine endpoints. https://doi.org/10.1016/j.chemosphere.2021.129699

It is recommended to assess the quality of the toxicity data on marine/estuarine species and recalculate estuarine criteria based on this recently available information.

4. Please comment on the use of measured and unmeasured toxicity tests to derive the respective criterion. In particular please comment on the supporting justification for using unmeasured toxicity tests in Appendix L.

PFOA is a highly stable compound, resistant to hydrolysis, photolysis, volatilization, and biodegradation (as described in Section 1.1.1 of the Report) and, therefore, expected to vary only minimally in the course of a toxicity test. To determine if nominal and measured PFOA concentrations were typically in close agreement, pairs of nominal and corresponding measured PFOA concentrations were compared to one another through (1) linear correlation analysis and (2) an assessment of measured concentrations as a percent of its paired nominal concentration. The analysis conducted by EPA Team showed strong correlation (correlation = 0.9995) of the 79 pairs of nominal and measured concentrations from freshwater studies, and similar strong correlation (correlation = 0.9999) of the 11 pairs of nominal and measured concentrations from saltwater studies (Figure L1 in the Report). In addition, the experimental conditions did not influence the correlation between nominal and measured concentrations.

This confirms inclusion of unmeasured PFOA toxicity tests for quantitative use in criteria derivation.

Personal experience on analyzing PFOA in ecotoxicological studies using freshwater and saltwater species have also exhinited strong correlation between nominal and measured concentrations.

Additional information for L.I.4 summary section could include additional information based on the two additional published papers and the key points from these studies are listed below.

*Despite the concerns and avoidance of glass, few studies have presented data describing the sorption of perfluorooctanoate (PFOA) to glass and other container materials (**Lath et al. 2019**). Lath et al. (2019), who had reported that sorptive losses of PFOA for lower concentration (~20 µg/L) aqueous solutions were highest in polypropylene centrifuge tubes. However, the container type did not influence the measured concentration as reported in section – of this report. According to **Rewerts et al., 2021** additional handling steps, which are not typically reported for ecotoxicological studies but may contribute to variability, include solution homogenization, subsampling procedures, and the container materials selected for storage*. https://doi.org/10.1002/etc.4667

Lath S, Knight ER, Navarro DA, Kookana RS, McLaughlin MJ. 2019. Sorption of PFOA onto different laboratory materials: Filter membranes and centrifuge tubes. Chemosphere 222: 671– 678. DOI: 10.1016/j.chemosphere.2019.01.096

5. Please comment on the toxicity data used to derive the draft criteria.

- **Were the data selected and/or excluded from the derivation of the criteria derivation appropriately utilized?**

The data selected to derive PFOA criteria are appropriate. Studies that did not fully meet the data quality objectives outlined in the 1985 Guidelines were not considered for inclusion in the criteria derivation, including some studies with other PFAS exposures, but were considered qualitatively as supporting information. A brief summary of each study describing the experimental conditions and summary tables providing all the relevant information such as strengths and limitations of each study, end points selected for deriving criteria are well documented by the EPA team.

The key acceptable exclusion/inclusion criteria used to derive draft criteria are listed below:

- Only single chemical toxicity tests with PFOA were considered for possible inclusion in criteria derivation, studies that tested chemical mixtures, including mixtures with PFAS compounds were excluded from criteria derivation.
- Both controlled laboratory experiments and field observations/studies were included.
- PFOA toxicity tests were not excluded from quantitative use in criteria derivation on the basis of unmeasured test concentrations alone.
- Only single chemical toxicity tests with PFOA were considered for possible inclusion in criteria derivation, studies that tested chemical mixtures, including mixtures with PFAS compounds were excluded from criteria derivation.
- Due to lower sensitivity, insect MDR was excluded from the criterion calculation, but were used to waive the missing insect MDR.
- Further supporting information on acceptable and unused studies for acute and chronic endpoints and for freshwater and marine studies are documented and summarized as appendices in this report.

- **Are there relevant data that you are aware of that should be added to the analyses (note that EPA is working on updating the toxicity data to reflect the data in ECOTOX between Sept. 2019 through the latest update)? If so, please provide references for consideration.**

Additional toxicity data published over the last six months is listed below:

Marine/estuarine

- Stuart L. Simpson, Yawen Liu, David A. Spadaro, Xinhong Wang; Rai S. Kookana and Graeme E. Batley Chronic effects and thresholds for estuarine and marine benthic organism exposure to

perfluorooctane sulfonic acid (PFOS)-contaminated sediments: Influence of organic carbon and exposure routes https://doi.org/10.1016/j.scitotenv.2021.146008

- Nicholas T Hayman , Gunther Rosen , Marienne A Colvin , Jason Conder , Jennifer A Arblaster Aquatic toxicity evaluations of PFOS and PFOA for five standard marine endpoints. https://doi.org/10.1016/j.chemosphere.2021.129699

Fresh water

- Christopher J. McCarthy, Shaun A. Roark, Demitria Wright, Kelly O'Neal, Brett Muckey, Mike Stanaway, Justin N. Rewerts, Jennifer A. Field, Todd A. Anderson, Christopher J. Salice, Toxicological Response of Chironomus dilutus in Single-Chemical and Binary Mixture Exposure Experiments with 6 Perfluoralkyl Substances, Environmental Toxicology and Chemistry, 10.1002/etc.5066, **40**, 8, (2319-2333), (2021).

In particular, please comment on:

5a. The toxicity values used to derive the PFOA criteria, with a particular emphasis on:

i. the selection and the Acute to Chronic Ratio (ACR) to serve as the Final Acute to Chronic Ratio (FACR) and its application to derive the Final Chronic Value (FCV).

ii. the use of the qualitatively acceptable acute midge (*Chironomus plumosus*) data from Yang et al. (2014) to suggest aquatic insects are relatively tolerant to acute PFOA exposures. Specifically, Yang et al. (2014) conducted a 96-hour renewal, measured PFOA acute test with the midge, *Chironomus plumosus*. This study was not acceptable for quantitative use due to the potentially problematic source of the organisms. The reported LC50 was 402.24 mg/L PFOA indicating that these insects may not be one of the more sensitive taxonomic groups. Therefore, this test was excluded from the criterion calculation, but used to waive the missing insect MDR.

i. The selection and the Acute to Chronic Ratio (ACR) to serve as the Final Acute to Chronic Ratio (FACR) and its application to derive the Final Chronic Value (FCV) for PFOA is acceptable. The 1985 Guidelines allow the use of a Final Acute-Chronic Ratio (FACR) to convert the FAV to the FCV (i.e., FAV/FACR=FCV), which is equivalent to the chronic criterion (Criterion Continuous Concentration, CCC), intended to protect 95 percent of the taxa in aquatic ecosystems. For PFOA, the 8-family MDR requirement was not met for the chronic dataset, as acceptable chronic studies for species representing three MDR groups were not available (benthic crustacean and third phylum or second insect order not already represented). Therefore, the Final Chronic Value (FCV) was calculated with the use of an ACR (acute-chronic ratio). When more than a single ACR was calculated for the same species, the Species Mean Acute-Chronic Ratio (SMACR) was calculated as the geometric mean value of all ACRs for that species. . The specifications for derivation of a FACR for aquatic animals was met for PFOA based on 1985 Guidelines: ACRs for at least three different families provided that at least one was a fish, at least one was an invertebrate, and at least one was an acutely sensitive freshwater species. The 1985 Guidelines provides recommendations to calculate the FACR when SMACRs are dissimilar. The 1985 Guidelines states that if SMACRs tend to increase or decrease as the SMAV increases, the FACR should be calculated as the geometric mean of the ACRs for species whose SMAVs are close to the FAV. For PFOA, relationship between SMAV and SMACR showed that SMACRs decreased as the SMAVs increased. SMAV of the rotifer *B. calyciflorus* was closest to the FAV. The CCC was calculated by dividing the FAV by the FACR to determine the FCV (91.34/299.1=0.3054). *The PFOA FCV derived by this method is lower than the all of the quantitatively-acceptable chronic values (ranges between 0.76- 40 mg/L) and listed in listed in Table C.1*.

ii The acute data set for PFOA contained 14 genera representing seven of the eight taxonomic MDR groups. The missing MDR was a representative from an insect family. There was no evidence to suggest aquatic insect are among the four most sensitive genera. EPA calculated the PFOA CMC using all acceptable quantitative studies, but did not include the insect data in the criterion calculation (i.e., the relatively tolerant insect LC50 value was not included in the total count ("n") of Genus Mean Acute Values in the criterion calculation). In addition, waiving an unfulfilled MDR when available data suggest it is not among the four most sensitive genera is consistent with previous EPA criteria documents, including U.S. EPA (2016).

In addition, Stefani et al. (2014), Macdonald et al. (2004), and Marziali et al. (2019) conducted chronic toxicity tests with *Chironomus* spp. and reported apical endpoints. *Results of these studies, taken together, also suggest that insects may not be among the most sensitive taxa to chronic PFOA exposures. Therefore, these tests were excluded from the criterion calculation, but were used to waive the missing insect MDR.*

5b. EPA's approach for fitting concentration-response (C-R) data (described in Appendix J) as well as the specific acute LC_{50} values (Appendix A.2) and chronic EC_{10} values (Appendix C.2) that were estimated (for sensitive genera when C-R data were available) and used to derive criteria.

- This is an excellent approach utilized by the EPA Team. EPA's approach for fitting concentration-response (C-R) data resulted in consistent approach across various ecotoxicological studies. The R drc package was used to fit 22 different models to each individual C-R dataset. A single model was then selected from the 22 models to serve as the representative C-R model. The selected model represented the most statistically-robust model available. In certain cases, this approach even improved and helped to select most sensitive toxicological endpoint, for example,
    - Page 85- Ji et al. (2008) conducted a chronic life-cycle test on the effects of PFOA (with *Moina macrocopa*). The *M. macrocopa* 7-day NOEC (reproduction: number of young per adult) reported by authors was 3.125 mg/L, the LOEC was 6.25 mg/L, and the MATC is 4.419 mg/L. EPA performed C-R analysis for this study and determined the number of young per starting female as the most sensitive endpoint with an acceptable C-R curve. The EPA-calculated EC10 was 2.194 mg/L PFOA for *M. macrocopa* and used it directly as the *Moina* GMCV.
    - Page 88 Yang et al., 2014 Chronic survival using EPA's method was more sensitive than reproduction endpoint related EC10 values reported by the authors.
- In depth analyses and associated dose-response graphs in Appendix A.2 and Appendix C.2 provides further in-depth information on the EPA's approach for fitting concentration-response (C-R) data.

6. Please comment on the translation of the chronic water column criterion elements for aquatic life to derive the tissue-based criterion elements, considering the bioaccumulation of PFOA and PFOS. In particular, please comment on:

6a. Uncertainty surrounding the bioaccumulation factors (BAFs) used to translate of the chronic water column criterion elements into tissue-based criterion elements.

Tissue criteria derived from the chronic water column concentration (CCC) with the use of bioaccumulation factors were developed by EPA. The chronic fish whole-body tissue criterion is 54.1 mg/kg wet weight, the chronic fish muscle tissue criterion is 9.37 mg/kg wet weight and the chronic invertebrate whole-body tissue criterion is 23.9 mg/kg wet weight.

The freshwater chronic PFOA toxicity data with measured tissue concentrations was limited, with no quantitatively acceptable tissue-based tests. Qualitatively acceptable tissue-based tests were reported for four species (three fish species and one amphibian) across five publications. Therefore, there were insufficient data to derive tissue-based criteria using a GSD approach from empirical tissue data from toxicity studies. EPA thus developed protective tissue-based criteria through a bioaccumulation factor approach (Burkhard 2021). Only BAFs of high and medium quality were used to derive the tissue criteria. BAFs used in the derivation of the PFOA tissue-based criteria consisted of > 2 water and organism samples each and were collected within one year and 2 km distance of one another. Criteria for protection of aquatic life and wildlife will need to use whole- organism BAFs because the criteria are based on whole-body toxicology for aquatic life (e.g., fish) and for wildlife (e.g., birds), both of which consume the whole fish (Stephan et al. 1985).

6b. EPA's determination of appropriate BAFs and the tissue types that the tissue criterion elements were based.

Within the body, PFOA tends to bioaccumulate within protein-rich tissues, such as the blood serum proteins and liver. BAFs are different for muscle/fillet and whole-body tissues. Humans consume muscle/fillets from fish and soft tissues from bivalves, therefore the water quality criteria recommended by EPA used BAFs based on these tissues. EPA calculated additional tissue values for liver, blood, and reproductive tissues by transforming the freshwater chronic water column criterion (i.e., 0.31 µg/L) into representative tissue concentrations using tissue-specific bioaccumulation factors (BAFs). Furthermore, EPA team justified to use female reproductive tissues due to its relevance for potential maternal transfer to offspring. There additional tissue-based values were calculated for comparative purposes and were not proposed as recommended criteria.

7. Please comment on the frequency and duration of the criterion elements, in particular please comment on the frequency and duration components of the tissue-based criterion elements.

PFOA concentrations in tissues are generally expected to change only gradually over time in response to environmental fluctuations. The chronic tissue-based criteria averaging periods, or duration components, were therefore specified as instantaneous, because tissue data provide point, or instantaneous, measurements that reflect integrative accumulation of PFOA over time and space in population(s) at a given site. It was appropriate for EPA to inform the recommended ten-year exceedance frequencies for the chronic tissue-based criteria given the large variation in possible biological and physical variable influencing ecological recovery.

8. Please provide any additional technical comments that you believe should be considered.

Some additional edits/suggestions are listed below:

1. **Appendix list in Table of Contents in missing**

2. **List of Tables not matching with Tables listed in the text**

Table 2-6 is Table 2.3 . Evaluation Criteria for Screening Bioaccumulation Factors (BAFs) in the Public Literature. **Error! Bookmark not defined.**

3. **Page xiv Table 0-1. Recommended Freshwater Perfluorooctanoic acid (PFOA) Aquatic Life Ambient Water Quality Criteria-**

   Superscript 3 [3]listed as footnote in the Table but not referenced in the table

4. **Table 2-2. Summary of Assessment Endpoints and Measures of Effect Used in the Criteria Derivation for PFOA**

Assessment Endpoints for the Aquatic Community	**Measures of Effect**
Aquatic Life: Survival, growth, and reproduction of freshwater and estuarine/marine aquatic life (i.e., fish, amphibians, aquatic invertebrates)	For effects from acute exposure: 1. LC_{50} concentrations in water, diet, and/or tissue (e.g., muscle, blood, egg) 2. NOEC and LOEC concentrations in water, diet, and/or tissue (e.g., muscle, blood, egg) For effects from chronic exposure: 1. EC_{10} concentrations in water, diet, and/or tissue (e.g., muscle, blood, egg) 2. NOEC and LOEC concentrations in water, diet, and/or tissue (e.g., muscle, blood, egg); Only used when an EC_{10} could not be calculated for a genus.

Please review if the highlighted muscle, blood and egg would be relevant to this section.

5. **Section 1.1 Previously Derived PFOA Toxicity Values and Thresholds**

   Table 1.1 to be updated by current information for Australia and new Zealand.

   Freshwater values are to be used on an interim basis until final marine guideline values can be set using the nationally-agreed process under the Australian and New Zealand Guidelines for Fresh and Marine Water Quality. Marine guideline values developed by CRC CARE are under consideration through the nationally-agreed water quality guideline development process. The Australian and New Zealand Guidelines for Fresh and Marine Water Quality (ANZG, 2018; ANZWQG, 2000) and Australia's PFAS Environmental Management Plan (HEPA, 2020) both recommend that a 99% species protection level be used as a guideline for bioaccumulating substances such as PFOS, PFOA, and PFHxS (as given below) as a precautionary approach.

Exposure scenario	PFOA	Exposure scenario	Comments and source
Freshwater	19 µg/L	99% species protection - high conservation value systems	Australian and New Zealand Guidelines for Fresh and Marine Water Quality - technical draft default guideline values for PFOA. The draft guidelines do not account for effects which result from the biomagnification of toxicants in air-breathing animals or in animals which prey on aquatic organisms. The WQGs advise P40FP40F[1]PP that the 99% level of protection be used for slightly to moderately disturbed systems. This approach is generally adopted for chemicals that bioaccumulate and biomagnify in wildlife. Regulators may specify or environmental legislation may prescribe the level of species protection. required, rather than allowing for case-by-case assessments.
	220 µg/L	95% species protection - slightly to moderately disturbed systems	
	632 µg/L	90% species protection - highly disturbed systems	
	1824 µg/L	80% species protection - highly disturbed systems	
Exposure scenario	**PFOA**	**Exposure scenario**	**Comments and source**
Interim marine	19 µg/L	99% species protection - high conservation value systems	As above. Freshwater values are to be used on an interim basis until final marine guideline values can be set using the nationally-agreed process under the Australian and New Zealand Guidelines for Fresh and Marine Water Quality. The WQG advise that in the case of estuaries, the most stringent of freshwater and marine criteria apply, taking account of any available salinity correction. Marine guideline values developed by CRC CARE are under consideration through the
	220 µg/L	95% species protection - slightly to moderately disturbed systems	
	632 µg/L	90% species protection - highly disturbed systems	
	1824 µg/L	80% species protection - highly disturbed systems	

[1] https://www.waterquality.gov.au/anz-guidelines/guideline-values/default/water-quality-toxicants/local-conditions#bioaccumulation.

			nationally-agreed water quality guideline development process.

6. Table 1 2. Two Primary Categories of PFAS

Please refer to OECD 2021 to be consistent with PFAS terminology/nomenclature

OECD (2021), Reconciling Terminology of the Universe of Per- and Polyfluoroalkyl Substances: Recommendations and Practical Guidance, OECD Series on Risk Management, No. 61, OECD Publishing, Paris

7. Table 1.3 page 5

Please review Figure 9 OECD 2021 (also attached as PDF)

8. Conceptual Model of PFOA in the Aquatic Environment and Effects

Figure 2.7 page 56- Growth as an endpoint missing in the endpoints – first pentagon

9. Page 69 last paragraph- Table 2-6 outlines the screening criteria for study evaluation and ranking. Cross reference to be updated to Table 2.3

COMMENTS SUBMITTED BY

REVIEWER 5

External Peer Review of EPA's Draft
Aquatic Life Water Quality Criterion for Perfluorooctanoic Acid (PFOA)

1. Please comment on the overall clarity of the document as it relates to the derivation of each criterion.

RESPONSE: Overall, similar to PFOS, the document for PFOA is very well written, generally free of grammatical errors and clear. It is long but not as long as PFOS and, therefore, easier to digest. I think the background material for both chemicals is especially good and provides an excellent overview and summary for readers less familiar with PFAS. In particular, the summary of PFOA concentrations in water bodies and other environmental media is wonderfully useful.

To me, the derivation of the criteria for PFOA is easier to follow than PFOS, mostly because there are fewer data and, hence, fewer assumptions, calculations, and evaluations that need to be made.

Unlike for PFOS, the criteria developed by EPA for PFOA are more similar to values derived by other jurisdictions for both acute and chronic values. To me, this lends strength to EPA's criteria values and is also consistent with the scientific method, in general. If we are all following the same threads of logic, using similar analyses, and looking at the same data, it makes sense that criteria values would be similar. So, again, in the case of PFOA the congruence with other published criteria values is a strength. I will use this as an argument for EPA to reconsider the criteria (and the data supporting it) for PFOS as, for that chemical, EPA's numbers are higher than other jurisdictions.

2. Please comment on the approach used to derive the draft criterion for PFOA. Please provide detailed comments.
- **Is the technical approach used to derive the criterion logical?**
- **Does the science support the conclusions?**
- **Is it consistent with the protection of freshwater aquatic life from acute, chronic, and bioaccumulative effects?**

RESPONSE: The overall approach to derive criteria for PFOA is logical…except for the use of only the 4 most sensitive endpoints and then a model was fitted (unspecified, I believe) to obtain the 5% most sensitive species (in general). I am not familiar with this as an approach as I have not seen other scientists use this and, instead, I have more commonly seen the application of a species sensitivity distribution based on more data and usually following an s-shape. Indeed, many well-cited papers on toxicity thresholds and criteria have used this approach. I suspect EPA has justification for the approach used in the document and that it is well-supported. I would suggest adding any details, beyond just citing the 1985 guidelines, that supports a focus on just the 4 most sensitive toxicity endpoints for the criteria development. My apologies if this information is in the document and I missed it.

One point of clarification is needed in the explanation of the regression analysis (p. 64). The document states: "When LOECs and NOECs were used, a Max Aceept. Tox Conc. (MATC) was calculated, with is the geometric mean of the NOEC and LOEC. For the calculation of chronic criterion, **point estimates** were selected for use as the measur of effect in favor of the MATCs…" – WHAT IS MEANT BYT "POINT ESTIMATES"? Isn't the LOEC or NOEC a point estimate? This needs to be clarified here (and for PFOS).

On p. 66 – I also recall that in some cases EPA obtained data from plots using web plot digitizer software. If I am correct, this should be explained on this page.

One point worth mentioning with regard to the technical approach is that I think EPA was correct to consider non-North American resident species in developing the criteria. While I can understand why some scientists feel strongly about focusing on native species, I also cannot think of a clear example of widely

different chemical tolerances among species from different countries. Sure, organisms from contaminated environments are likely to differ compared to the same organisms (species) from uncontaminated areas but barring this, it seem sensitivity to chemicals is not geographically determinated (again, barring extremes).

Additionally, I think using the EC10 makes sense for PFAS chronic criteria and perhaps other chemicals as well. That said, I can't say that using a 50% effect level for acute toxicity/exposures makes sense. To me, using an EC20 or LC20 for acute would be more reasonable. As I understand it, the current approach divides the calculated FAVs by 2.0 to further ensure protection. Here again, it would seem easier and more straight-forward to use the EC20 or LC20. Perhaps dividing the FAV by 2.0, however, commonly results in a low threshold (like an LC20, for example). If this is the case, it would be worth EPA mentioning to give some sense of magnitude to what could be considered an arbitrary "safety factor" of 2.0. What highlighted this issue for me was the estuarine mussel species where there was a 27% effect (malformations) at 0.0001 mg PFOA/L but because an EC50 could not be determined, EPA chose to use the highest concentration of 1 mg/L. If I were in charge of managing resources in an estuary, I can't say I'd have much confidence in that 1 mg/L value. Or at least, I would be very uncomfortable. Having said all this, I understand that the frequency and duration of environmental exposures would make is such that the acute and chronic criteria would be protective but please see my comments below with regard to frequency and duration of exposure.

3. **Please comment on the approach used to derive the draft acute estuarine/marine benchmark for PFOA. Given the limited estuarine/marine test data available, a new approach method was used to support the derivation of an acute estuarine/marine benchmark to provide states and tribes with a protective value. Please provide detailed comments.**
    - **Is the technical approach used to derive the benchmark logical?**
    - **Does the science support the conclusions?**
    - **Is it consistent with the protection offered by acute estuarine/marine aquatic life criteria derived using empirical data, as prescribed in the 1985 *Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses*?**

RESPONSE: The most sensitive estuarine/marine species was Mytilus g. in the study by Fabbri et al. (2014). EPA chose to use 1 mg/L PFOA because there was not a 50% effect level even at the highest tested concentration (1 mg/L). HOWEVER, there was 27% effect at the LOEC (0.0001 mg/L). Is this not problematic? If I were concerned about mussels or other bivalves in an estuary, I don't think I would hang my hat on a 1 mg/L PFOA concentration given there was a 27% decrease in normal D-larvae at a concentration several orders of magnitude below that. This raises the issue of why EPA is using a 50% effect level for acute criteria – this seems excessively high, doesn't it? EPA justified using the EC10 for chronic criteria so it seems reasonable to use the EC/LC20 for acute. A 50% effect, if it occurs, is likely to manifest as ecologically relevant for any species…20% may not be protective depending on the species and endpoint. As mentioned before, I understand that there is a frequency/duration element to all the criteria but from a functional standpoint, I don't see how the ferequency and duration elements are helpful because nobody collects or reports environmental data on a relevant temporal scale (every hour; 4-day running average??). See below for more on frequency and duration.

Given the data, I believe the new approach methods based on WEB-ICE are appropriate. The estimation tool has been in development and used for a considerable length of time and several publications have supported it's use. Of course it would be better to have more data but, again, given the lack of data for estuarine/marine species, the WEB-ICE approach is likely the best available.

4. **Please comment on the use of measured and unmeasured toxicity tests to derive the respective criterion. In particular please comment on the supporting justification for using unmeasured toxicity tests in Appendix L.**

RESPONSE: Similar to PFOS, the approach EPA used for PFOA was to consider studies in which the chemical was not measured. This was justified based on an analysis EPA did comparing nominal and measured concentrations and finding close agreement. This generally seems reasonable to me especially considering the stability of the chemical. However, EPA used a criteria of 20% (p. 61) which is not consistent with the analytical precision of most methods used to analyze PFAS. According to several very prolific environmental chemists that have made a career of measuring PFAS, they have communicated to me that the methods are accurate to within 30%. That means that if the measured were within 30% of nominal, we basically have concluded these were nominal. I would encourage EPA to explore their 20% acceptability threshold and perhaps offer an explanation as to why this is appropriate.

5. Please comment on the toxicity data used to derive the draft criteria.

- **Were the data selected and/or excluded from the derivation of the criteria derivation appropriately utilized?**
- **Are there relevant data that you are aware of that should be added to the analyses (note that EPA is working on updating the toxicity data to reflect the data in ECOTOX between Sept. 2019 through the latest update)? If so, please provide references for consideration.**

RESPONSE: In general, the data selected or excluded for criteria development were appropriately used. A study published by McCarthy et al. in 2021 reports toxicity of PFOA to Chironomus dilutus – if I recall correctly, I believe these data (on an insect) indicate that PFOA is not very toxic to C. dilutus…even for exposures greater than 96 hours. This agrees with EPA's assessment of available acute toxicity data for freshwater insects; that PFOA is generally not toxic to insects. I would, however, still urge more studies on aquatic insects as this is an obviously diverse taxa with many sensitive species. It is very possible that the available test species of insects are not sensitive but other insects such as mayflies or damselflies, etc. may, in fact, be quite sensitive.

The write up of the Zhang et al. (2013 and 2014) papers has a few grammatical errors and could benefit from additional editing for clarity. I agree that using the intrinsic rate of increase (or similar) is a very relevant endpoint but it reads as though this was based on a 4 day observation period for the 2014 paper. This does not quite make sense to me…offhand, I would hypothesize that the reported/calculated effect level of 1.166 mg/L would decrease with a longer observation period. I think I may be misunderstanding the experimental design – some editing would help clarify. This is an important series of studies (see below regatding resting egg production) so clarity is critical.

NOTE: the EC10 of 0.076 mg/L for resting egg production observed by Zhang et al. (2014b) is potentially a big deal. EPA appears justified in not using this because it was only one replicate, etc. but these data clearly point to a potentially relevant effect at a relatively low concentration.

In particular, please comment on:

5a. The toxicity values used to derive the PFOA criteria, with a particular emphasis on:

i. the selection and the Acute to Chronic Ratio (ACR) to serve as the Final Acute to Chronic Ratio (FACR) and its application to derive the Final Chronic Value (FCV).

RESPONSE: The EPA followed the 1985 guidelines allowing them to calculate the FCV using the ACR approach as outlined around p. 101. Given the lack of data, this seems like a reasonable approach. The actual calculation of the FCV (based on the data from B. calyciflorus) is appropriate and the use of this FACR to determine the FCV of 0.3054 is appropriate. Morevoer, as stated above, this value is also generally in line with other criteria for PFOA published by other jurisdictions.

ii. the use of the qualitatively acceptable acute midge (*Chironomus plumosus*) data from Yang et al. (2014) to suggest aquatic insects are relatively tolerant to acute PFOA exposures.

Specifically, Yang et al. (2014) conducted a 96-hour renewal, measured PFOA acute test with the midge, *Chironomus plumosus*. This study was not acceptable for quantitative use due to the potentially problematic source of the organisms. The reported LC50 was 402.24 mg/L PFOA indicating that these insects may not be one of the more sensitive taxonomic groups. Therefore, this test was excluded from the criterion calculation, but used to waive the missing insect MDR.

RESPONSE: I commented on this above but will mention again. Overall, I think EPA is correct that the available data on chironimids (Yange et al. 2014 and McCarthy et al. 2021) indicate that chironomids are not sensitive to PFOA. That said, it is probably the case that other insects such as mayflies or damselflies (or other species?) are more sensitive than chironomids. In contrast to this statement, however, is that chironomids are among the most sensitive to PFOS. Bottom line: for PFOA and available insect toxicity data, it appears PFOA is not toxic to insect but, clearly, more data are needed to improve confidence in this estimate.

5b. EPA's approach for fitting concentration-response (C-R) data (described in Appendix J) as well as the specific acute LC_{50} values (Appendix A.2) and chronic EC_{10} values (Appendix C.2) that were estimated (for sensitive genera when C-R data were available) and used to derive criteria.

RESPONSE: EPA's approach to fitting C-R data using the drc package in R is, in my opinion, state of the art. The method can easily test a variety of curves and the fit criteria can be used to select the best fitting curves. Comments related to specific studies and LC50 and EC10 estimates are elsewhere in these comments. In general, my opinion is that the PFOA criteria are slightly more defensible than the PFOS criteria; this is explained more in the review of PFOS.

6. Please comment on the translation of the chronic water column criterion elements for aquatic life to derive the tissue-based criterion elements, considering the bioaccumulation of PFOA and PFOS. In particular, please comment on:

6a. Uncertainty surrounding the bioaccumulation factors (BAFs) used to translate of the chronic water column criterion elements into tissue-based criterion elements.

RESPONSE: Overall, this seems like a reasonable approach – to estimate tissue-based criteria using the water column criteria multiplied by the bioaccumulation factors. The difficulty arises when we consider the accuracy or robustness of the BAFs. I agree with the criteria in Table 2-3 and especially emphasize the importance of concurrent collections in space AND time for tissues and environmental media.

6b. EPA's determination of appropriate BAFs and the tissue types that the tissue criterion elements were based.

RESPONSE: I am very familiar with the Burkhard (2021) paper which the PFOA document follows closely in terms of BAFs. The BAFs used by EPA are appropriate given the data. I also agree that the most useful/appropriate tissues for BAFs are invertebrates, fish muscle and fish whole body – these are the most commonly analyzed and most abundant in the literature. For what it's worth, in my own research in which we collected and analyzed fish tissues and co-located water samples, our calculated BAFs for PFOA (and PFOS) were close to the central tendency BAFs reported by Burkhard (2021). I also agree that co-located and sampled at the same time yield the most defensible BAFs as PFAS concentrations can vary considerably in space and time (not often shown in the literature).

7. Please comment on the frequency and duration of the criterion elements, in particular please comment on the frequency and duration components of the tissue-based criterion elements.

RESPONSE: Conceptually, the frequency and duration of the criterion elements seem reasonable – the acute water column criterion can't be exceeded for more than one-hour of duration which is then not to be exceeded more than once in three years. Even though many of the effect levels are 50%, this is likely protective given the duration of most acute toxicity studies is certainly and convincingly more than one hour. This applies to the chronic water column criteria as well which is based on not exceeding a 4-day duration; most chronic studies are much longer than 4 days. The only real issue with this is that, for all intents and purposes, these frequency and duration elements are not measured in practice. In other words, I am unaware of hourly measurements of PFAS in water or 4-day running averages. So, conceptually, I believe the frequency and duration elements are protective. However, in practice it is not clear to me how these would be useful or would help with regulation because the data related to frequency and magnitude of PFAS in water is not at a fine enough temporal resolution.

The frequency and duration for tissue-based criteria is a little different, however. Tissue concentrations (as mentioned in the document) represent an integration through time and so a measure of fish tissues, for example, provides some insight to exposures that have occurred over longer than an hour or 4 days. In this case, not exceeding the tissue-based criteria more than once in 10 years is likely protective. It is also more likely that monitoring programs would sample fish at least yearly which means this criteria is likely the most useful from a monitoring and clean-up perspective. It is still possible that high concentrations in tissues will be "missed" with only sampling once a year but this is far better than the frequency and duration elements for the water column criteria.

8. Please provide any additional technical comments that you believe should be considered.

RESPONSE: Overall, I think the PFOA criteria document and the reported criteria are robust given the constraints imposed by data availability. I agree with most assumptions made by EPA and the actual criteria values reported. I would, however, suggest that EPA consider revising the 1985 Guidelines. Re-evaluating the guidelines and publishing and update, even if changes are minimal would be a valued effort. I am sure that EPA has more than enough to do but given the importance of AWQC, I think it would be better to base the science on a more recent effort. There has been a lot of solid science in the last 35 years with a focus on criteria, species sensitivity distributions, etc. Although not used heavily for PFOA the use of the 4 most sensitive taxa and then a linear regression to estimate the criteria seems less robust than other methods. I could be wrong but here is where an updated Guidelines would be helpful – that may be an effort worth exploring quantitatively.