National Institute of Environmental Health Sciences (NIEHS)
Comments on the Interagency Science Consultation (Step 3)
Draft IRIS Toxicological Review of Perfluorononanoic Acid (PFNA) and Related Salts
Dated July 2023
(Date Received August 25, 2023)

---

1. Overview of Background Information and Assessment Methods

1.1 Background Information on PFNA

General comments:

- Page xxi, line 26: "population," – please change to "population."
- The word "coherence" is used often. Consider defining at first use. Coherence means "systematic or logical connection or consistency." Please consider whether this is the appropriate word or perhaps, synonyms or a mix of words with similar meaning could be utilized.

1.2 Summary of Assessment Methods

General comments:

- EPA's Toxicological Review (TR) of PFNA is another assessment of a specific PFAS that EPA developed as part of a coordinated set of systematic reviews of select PFAS (PFBA, PFHxA, PFHxS, PFNA and PFDA) under a single protocol published by EPA in 2019. The protocol was updated in 2020 and again in 2021 in response to public comments and the updates were tracked and documented.
- EPA is following the highest level of transparency, review and rigor in publishing the protocol, linking the related assessments, and following best practices for systematic reviews as well as EPA's systematic review methods in alignment with their standard IRIS Systematic Review Handbook.
- Page 1-10 through 1-12: The use of populations, exposures, comparators, and outcomes (PECO) and inclusion/exclusion criteria are followed for best practices and clearly described.
- Page 1-13: As noted for other EPA PFAS assessments, EPA does a nice job of transparently describing the challenge of addressing potential confounding across PFAS exposures in section 1.2.2 under additional epidemiological considerations. The approach taken on addressing confounding is both reasonable and clearly presented.
- Page 1-15: Methods for evidence synthesis are adequately described here with reference to detailed methods in the protocol /Appendix A.

Suggested revisions:

- Page 1-9 line 27: EPA states that a literature search was first conducted in 2017, which is 2 years before the protocol was published. Suggest clarifying how that took place when the protocol had not yet been developed or at least not released.

2. Literature Search and Screening Results

General comments:

- Page 1-19: The literature search and screening are clearly presented, and the results of evidence categorization shown in an evidence map that EPA to as literature inventories/literature Tagtrees. The interactive nature of figures such as the Tagtree is useful to the reader and EPA should continue to find ways to use similar interactive approaches for other figures.

3. Pharmacokinetics, Evidence Synthesis, and Evidence Integration

General comments:

- The section is comprehensive and well-organized. Appropriate studies reported in the literature were utilized with justification for inclusion or omission. The approach used to generate average PK values for species and sexes is sound.

Suggested revisions:

- Page 3-1, paragraph 2 states that ADME characteristics of PFNA include rapid absorption from the GI tract, yet subsequent sections note that the T_{max} of PFNA is >12 h suggesting slow absorption. Please clarify as appropriate.

Minor editorial comments:

- Both level of quantitation and limit of quantitation are used; sometimes the acronym LOQ is used for both and sometimes no acronym is used. Recommend using limit of quantitation consistently with the acronym LOQ.
- Page 3-10, line 26: Please provide the LOQ value as context for the statement made.
- In tables where population summaries for Bayesian analysis were given, it is not clear whether it is for a given route or all routes combined. Please clarify.

3.1 Pharmacokinetics

3.2. Noncancer Evidence Synthesis and Integration

3.2.1. General Toxicity

- No comments

3.2.2. Developmental Effects

General comments:

- Page 3-47, line 37: The evidence mentioned does not clearly support an inverse association between PFNA and birth length. There were 13 studies with seven showing no associations and six showing associations. There were more high confidence studies that showed no effect than there were that showed an effect (three high confidence studies showing no effect vs. two high confidence studies showing an effect). There is too much unexplained inconsistency across the studies to support an inverse association. Suggest revising the final conclusion for birth length and update the language to reflect this change on page 3-80, line 27 and 3-83, line 28.

Suggested revisions:

- Page 3-37, Figure 3-3, and line 5: The Lee 2018 paper has the same deficiencies as Cao 2018, Marks 2019, and Shi 2017 (participant selection, confounding, sensitivity) but was given an overall rating of uninformative, while the other papers had an overall rating of low confidence. It might be helpful to explain why the Lee paper received a lower rating than the other papers with the same deficiencies. Also, if the figure is correct, the Lee paper did not have "critical study deficiencies" but rather "deficiencies" (line 5).
- Page 3-39, Figure 3-4: Based on previous comment, if the overall rating is updated (i.e., Lee overall rating is low confidence instead of uninformative), then this figure will need to be updated to reflect that.
- Page 3-59, Figure 3-11: Heatmap appears to be cut off at the top (one study is missing).
- Page 3-45, line 7: Please provide links to the meta-analysis results on birth weight.
- Page 3-63, line 17: Heatmap for anogenital distance seems to be missing.

Minor editorial comments:

- Page 3-35, line 5: The close parenthesis appears to be in the incorrect location: should read, "(and early childhood) weight and height."
- Page 3-35, line 9: Should read, "While some uncertainty and potentially reduced sensitivity are associated…"

3.2.3. Immune Effects

Suggested revisions:

- Page 3-91, line 18: Change “least” to less (e.g., leukopenia/lymphopenia can be informative).
- Page 3-92, line 15: Should it be 4 (not 6) medium confidence studies (5 not 7 publications) for PFNA and antibody responses, or does the 6/7 also include infectious disease?
- Figure 3-20: Should Kvalem et al. 2020 be indicated as medium (green) and not low (yellow)? Should Stein et al. 2016a and Stein et al. 2016b be indicated as low and medium, respectively (i.e., reversed)? Also, “et al.” should be added to the Figure for many of the studies.
- Page 3-94, line 4: Should it be noted that the association was significant only for diphtheria?
- Page 3-98, line 17: Is Kvalem et al. medium not low confidence? Also, line 34: Should it be 3 not 2 (medium) with Kvalem included?
- Page 3-100, Table 3-18: Should the results for ear infection (age 6-7) be highlighted?
- Figure 3-21: Shouldn’t the Stein et al. 2016 study be “b” not “a”? Also, “et al.” should be added to the Figure for many of the studies.
- Page 3-103, line 11-14: It is not clear what the difference is between “asthma incidences (i.e., diagnosis within the past year)” and “current asthma (generally experiencing symptoms in the past year with asthma diagnosis)”?
- Page 3-105, Table 3-19: Dong et al. (2015) should be 2013.
- Page 3-115, line 11 and Page 3-122, Line 34: splenocyte not “spleenocyte”
- Page 3-124 to 3-125: It was not clear how/why “slight” translates into “suggests” for the data in humans, unless the rationale is defined elsewhere in the document, or is the slight/suggests due to footnote #12? Perhaps clarify better.

3.2.4. Hepatic Effects

Suggested revisions:

- Page 3-131, line 4: Lin et al. 2010 is cited twice, is this correct? Please clarify whether Lin should be included in the first grouping.
- Page 3-150, Table 3-25: Consider putting incidence first with asterisks, then severity scores. In some cases, severity score was not provided (female 25 mg, necrosis) or asterisk is with severity scores (female cytoplasmic alteration). Consider putting cytoplasmic alteration (CA) at the top, and necrosis, the most severe lesion at the bottom, or put in alphabetical order (CA at top).

Minor editorial comments:

- Page 3-164, line 13: Suggest rewording using “observe” instead of “evaluate”; NTP studies did not evaluate hepatocellular necrosis.

3.2.5. Male Reproductive Effects

3.2.6. Endocrine Effects

General comments:

- Authors did a good job of reporting on, at times, very heterogeneous data.

Suggested revisions:

- Page 3-209, line 5: The overall effect found by Aimuzi et al. 2019 and Shah-Kulkarni et al. 2016 is unclear. Did they find an overall effect prior to assessing a sex-specific effect? What were the results from the remaining three studies as they are not included in this paragraph?
- Page 3-209, line 6: This title sentence mentions T3, but the paragraph only includes data on T4. Please address.
- Page 3-209, line 16: Please cite the remaining two studies evaluating T3. In this paragraph only 5 of 7 studies that evaluated T3 in infants are cited. Also, please cite the two studies that evaluated T4 and clarify if those data were also from infants?

- Page 3-214, lines 12-18: Three papers are cited but not included in the risk of bias figure: Laws et al. 2007, O'Connor et al. 1999, and Stoker et al., 2000. Please clarify if these papers evaluated PFNA or whether they provide some context for the association between body weight changes and thyroid hormone levels.
- Page 3-215, line 14: Please clarify whether both free and total T4 were reduced by 36 to 53% or are the reduction ranging from 36% to 53%, respectively, for free and total T4?
- Page 3-216, lines 5-6: Please clarify the meaning of the text in parentheses (large TSH decreases of >30% at 0.625 mg/kg-day were not reported as statistically significant). Does this mean that the reductions in serum TSH are not significant, but serum T4 are?
- Page 3-223, line 20: The abbreviation T-screen is not defined. What does the "T" stand for?

Minor editorial comments:

- Page 3-214, line 17: Please add a comma after (NTP, 2018).
- Page 3-216, line 4 and 3-222, lines 5-6: Please double-check how the U.S. EPA reference is cited. If you are using Endnote, a comma is needed after U.S. EPA in order for the full abbreviation to display in the citation.
- Page 3-221, line 22-24: Remove extra spaces and lines.

3.2.7. Nervous System Effects (Developmental Neurotoxicity)

General comments:

- Throughout the developmental neurotoxicity section of the document, the authors consistently and incorrectly equate evidence of precision/imprecision with statistical significance. Statistical significance of an effect estimate is not a proxy for precision, nor does it guarantee that an effect estimate is precise. Similarly, non-statistical significance of an effect estimate is not a proxy for imprecision. Precision is the degree of certainty surrounding an effect estimate with respect to a given outcome and should be based on the wideness of the confidence intervals. For example, a ratio measure (e.g., odds ratio or OR) may be considered imprecise when the ratio of the upper to lower 95% CI for most studies is ≥ 10. Adequate power of the study to detect an effect is another factor to consider, especially when interpreting findings that do not provide support for an association.
- The protocol for this assessment provides a correct description of how precision should be considered, so it is unclear whether (1) the reviewers' assessment of precision/imprecision did not adhere to the protocol or (2) the reviewers inaccurately described the assessment of precision/imprecision in their write-up of the section.
- The primary basis for the evidence judgement is justified as "generally consistent, but imprecise, associations between PFNA and ADHD and potentially related behaviors." Given that an incorrect definition of precision was used throughout the sections that precede the evidence judgement, especially in the ADHD section, the authors should revisit whether (1) the protocol was not followed with respect to judgements about precision and/or (2) the authors of the section have incorrectly described how precision was determined. If statistical significance was used to determine that the associations in the ADHD studies were imprecise, imprecision may not be as large a concern as is currently indicated.

Suggested revisions:

- Page 3-237 to Page 3-238: Consider revising the text so that precision is not equated with statistical significance. Examples of this include, but are not limited to, the following:

  "ADHD diagnosis plus medication is considered a more specific outcome than diagnosis alone and may focus on the more severe forms of ADHD, which may explain the inconsistency in results in (Stein and Savitz, 2011), but the estimates in both studies are imprecise (i.e., not statistically significant)."

  "They observed no increase in ADHD when exposure was modeled as a continuous variable but did observe imprecise indications of higher risk of ADHD in the second and third quartiles ($p > 0.05$) in 19 categorical analysis, and when adjusted for other PFAS exposures, observed a monotonic increase of 20 across quartiles (data not extracted, $p < 0.05$)."

"However, results for ADHD diagnosis (and diagnosis plus medication), the most specific and reliable outcomes, were imprecise. It is possible that exposure levels/contrast were not adequate to reach statistical significance, but there is not a clear pattern between exposure levels/contrast and detecting an association."

- Page 3-243 to Page 3-244: As stated under the general comments, the primary basis for the evidence judgement is given as "generally consistent, but imprecise, associations between PFNA and ADHD and potentially related behaviors." Given use of an incorrect definition of precision throughout the sections that precede the evidence judgement, especially in the ADHD section, the authors should revisit the evidence integration summary judgement. If statistical significance was used to determine that the associations in the ADHD studies were imprecise, imprecision may not be as large a concern as is currently described.

  "Five of six *medium* confidence studies reported higher ADHD diagnosis or ADHD-related behaviors with higher PFNA exposure. However, there is considerable uncertainty in this association, including imprecision in all the estimates from the three studies evaluating ADHD diagnosis, the most specific outcome, and some unexplained inconsistency (no association found in one of the three *medium* confidence study of behaviors potentially related to ADHD with adequate sensitivity)."

Minor editorial comments:

- In Tables 3-33 (pages 3-233 to 3-236) and 3-34 (pages 3-239 to 3-241, check that all acronyms are clarified.

3.2.8. Cardiometabolic Effects

General comments:

- It could be helpful to comment on the overall lack of *high confidence* studies for this health effect area – is this a common or uncommon occurrence?
- Page 3-248, lines 10-12: When referencing Dong et al. 2019, it is interesting that a positive association with total cholesterol was seen with all but one NHANES wave. Consider commenting on the 2009-2010 wave that did not show a positive association.
- Pages 3-254 through 3-256: While much of the data for cardiometabolic outcomes were inconsistent, blood pressure and hypertension endpoints stand out as outcomes that were particularly inconsistent. The corresponding summary and the Evidence Integration section capture these inconsistencies well, but it could be worth elaborating a bit more.
- Page 3-269, lines 2-3: It would be useful to the reader to further define "elevated" when describing "elevated waist circumference" in the context of metabolic syndrome.

Suggested revisions:

- Page 3-282, line 19: It might be helpful to include a source supporting that 28 days is not expected to be a sufficient exposure time to develop lesions.

Minor editorial comments:

- Page 3-249, line 23: There is an extra comma after "The medium confidence studies in adults, support…"
- Page 3-252: The font sizing looks inconsistent in Table 3-36, in the row for Manzano-Salgado et al. (2017b).

3.2.9. Female Reproductive Effects

General comments:

- It is reasonable that parous and nulliparous results may differ. The infant "inherits" much of mom's body burden. Parity is an important variable in the statistical models that should be included.
- Pages 3-299 and 3-300: Attention is needed to the text on gestational weight gain. If understood correctly, studies reported increased weight gain or no weight loss in pregnant female mice that had reduced litter sizes or decreased pregnancy success. Pre-eclampsia induced edema in mice/rats is rarely evaluated and will lead to increased weight gain in female rodents. It has been reported in Blake et al., 2020 EHP for PFOA and GenX.

Suggested revisions:

- The effects of PFNA on shortened duration of lactation in women seems to have been missed and should be added, as there seems to be consistency across several of the existing studies. There is a recent systematic review that should be cited and the text on PFNA directly from that paper is here: "PFNA was likewise examined in four studies, and while the hazard of terminating breastfeeding at 3 months was reduced by 23% (95% CI: 7; 37%) with an interquartile range increase in PFNA among Norwegian women [42], a doubling in PFNA was associated with 1.3 (95% CI: 0.7; 2.0) months shorter duration of breastfeed among Faroese women [44], a 17% (95% CI: 4; 31%) increased hazard of terminating breastfeeding at any given time among Danish women [43], and a 12% (95% CI: −19; 53%) increased risk of terminating breastfeeding before 3 months, among American women [45]." Another paper was published after this one that can be included, although they evaluated effects differently for the other papers (in general): Rokoff LB, Wallenborn JT, Harris MH, Rifas-Shiman SL, Criswell R, Romano ME, Young JG, Calafat AM, Oken E, Sagiv SK, Fleisch AF. Plasma concentrations of per- and polyfluoroalkyl substances in pregnancy and breastfeeding duration in Project Viva. Sci Total Environ. 2023 Sep 15;891:164724. doi: 10.1016/j.scitotenv.2023.164724. Epub 2023 Jun 7. PMID: 37290653.
- Page 3-300, line 34-36: Testosterone levels and all other hormones reported in female rodents are only informative if analyzed and reported by cycle stage. If the females were not estrous cycle staged, then the information on hormones is not useful and the study value should be decreased for those endpoints. (This idea holds in rodents just like the statement made for the human data in the female reproductive section). Please clarify if the data reported in this section are adequate in this respect or not.

Minor editorial comments:

- Page 3-300, line 4-5: Please clarify if in pregnant mice or all mice.
- Page 3-300, line 10-13: Please clarify if the pups were smaller or litters were smaller.

3.2.10. Urinary System Effects

Suggested revisions:

- Page 3-212, after line 5: It may be worth mentioning that the decreased total protein in males at 0.625 mg/kg did not occur during significant weight loss at this dose (-3.5%).

Minor editorial comments:

- Page 3-316, Table 3-49: Please clarify why is Markers underlined in the Studies, Outcome and Confidence column for Evidence from In vivo animal Studies.

3.2.11. Other Noncancer Health Effects

General comments:

- Page 3-318, paragraph starting on line 18: It does seem like there are hematological effects as there are decreased reticulocytes that are initiated at a lower non-overt tox dose, that become substantial at the higher over toxicity dose. However, given the current dataset and dose evaluations, the overall assessment of inadequate evidence is appropriate.

3.3 Carcinogenicity

3.3.1. Cancer

General comments:

- DNA damage observed under cytotoxic condition, in which DNA damages are expected and do not indicate the exposure agent is genotoxic or non-genotoxic, does not need to be reported. In Comet assays, the increased DNA damage at non-cytotoxic condition is, on the other hand, worth mentioning, as it suggests the substance

is genotoxic. Several places citing the Eriksen et al. 2010 study need revision. The focus should be at the low concentration result, not the high concentration result. Additionally, page 3-321, line 12, is it ready for a revision?

- In a single study, oxidative stress is often measured by the presence of one or two reactive oxygen species (ROS) or reactive nitrogen species (RNS) among a long list of possible ROS and RNS species. Negative results from one ROS species and positive results from another ROS species are not considered contradictory. Consider providing more details on the oxidative stress measurement in Table 3-50 (page 3-322 to 3-323) so the information mentioned on page 3-320, lines 34-36 is clear.

Suggested revisions:

- Page 3-319, line 4-6: ***Suggested edits****:* due to ~~self-reporting and~~ critical deficiencies in exposure measurement, outcome ascertainment (e.g., outcomes were only self-reported), and participant selection, and are not discussed further.
- Page 3-320, Lines 35-36: It states, ***"***while the other reported damage without ROS production but was concurrent with cytotoxicity (Eriksen et al., 2010)." ➔ ***Suggested edits:*** while the other reported increases in DNA damages (at 100 µM) were observed without increase in ROS production (at either 40 or 200 µM, as 100 µM was not tested for ROS production) (Eriksen et al., 2010).
- Page 3-321, line 10: It states, "significance of these results to human biology is difficult to interpret", could the authors add a sentence or two about why? Is it because the activation of DNA damage-associated genes may lead to DNA repair and no DNA damage? Is it because these genes are unique to the prokaryotic system used (i.e., don't have functional counterparts in human cells)?
- Page 3-323, Result column for Eriksen et al, 2010 (in Table 3-50): It states, "Comet assay indicated PFNA increased DNA strand breaks at both concentrations (400 µM was associated with cytotoxicity measured by LDH release)." ➔ ***Suggested edits:*** Comet assay indicates PFNA increased DNA strand breaks (100 µM, a non-cytotoxic concentration as measured by LDH release)
- Page 3-323, Conclusion column for Eriksen et al., 2010 (in Table 3-50): ***Suggested edits:*** Non-cytotoxic ~~and cytotoxic~~ concentration~~s~~ of PFNA induced DNA damage without significant changes in ROS production. [Also, what ROS was measured?]

Minor editorial comments:

- Page 3-319, line 9-10: a nonsignificant decreased risk ➔ ***Suggested edits:*** "a nonsignificant decrease in risk"
- Page 3-320, line 4: The second *low* confidence study by (Hardell et al., 2014) reported….➔ ***Suggested edits:*** The second *low* confidence study (Hardell et al., 2014) reported… [otherwise, it sounded like Hardell et al., 2014 has two low confidence studies in it]
- Page 3-321, line 5: were consistently negative ➔ ***Suggested edits:*** were all negative [because each assay was detecting a different type of mutation].
- Page 3-322, Table 3-50, Conclusion column for Fang et al., 2010: Consider rewording "PFNA exposure is associated with DNA damage associated with apoptosis and oxidative stress." Are all three endpoints related to each other? Please clarify if "associated" means "positively" or "negatively" associated. Would it be correct to say "PFNA exposure increased DNA damage, apoptosis, and oxidative stress"?

4. Summary of Hazard Identification Conclusions

General comments:

- Page 4-1 through 4-7: The hazard identification conclusions are clearly presented and consistent with the individual noncancer evidence synthesis and integration chapters.
- Table 4-1 provides a useful comparison of hazard conclusions relative to other EPA PFAS assessments. It is commendable for EPA using the approach and suggest they emphasize the comparison in summary statements, executive summaries, and future communications when other PFAS are completed. In addition, these table should be updated as other assessments are completed in a "live" version of the document if possible.

Suggested revisions:

- Page 4-4, lines 13-14: While the table is excellent (see above comment), it would be very helpful if there were quick summaries of what EPA considers are the reasons behind differences between the individual PFAS (e.g., lack of data or biological response).

5. Derivation of Toxicity Values

5.1. Noncancer and Cancer Health Effect Categories Considered

5.2. Noncancer Toxicity Values

5.2.1 Oral Reference Dose (RfD) Derivation

General comments:

- The foundation for the oral reference dose (RfD) is based on a point of departure (POD) obtained from human epidemiological studies. There is a significant disparity in potency between what was observed in rodent studies compared to human data, with the effects being considerably less potent in rodents. It is worth highlighting that all animal data are graphically represented, offering a comprehensive view of the data's quality and dependability. This level of visual detail is absent in the case of the human epidemiological data. For utmost transparency in presenting the data upon which regulatory decisions are predicated, a graph showcasing all data points along with a fitted from which the POD is derived should also be included.

Suggested revisions:

- Page 5-7, lines 16-18: "Had there been...". This text seems very casual and would not be understood if read in isolation. Please make the point succinctly and cite the studies.
- Page 5-9, line 13-14: "However..." Consider adding this short sentence to the sentence before with a semi-colon. It was not immediately clear which studies were being referred to, but would be clearer if included in the same sentence.

Minor editorial comments:

- Page 5-1, line 29: More U.S. EPA references that are displaying as U.S. only. Please adjust bibliography program to properly display these citations. See Page 5-15, line 13 for a U.S. EPA reference with the correct citation format (U.S. EPA, 2012a). Please review the bibliography software to ensure that the other U.S. EPA references are similarly formatted to this reference.
- Page 5-7, line 15: Consider removing the phrase ", while appropriate for hazard identification,". There are lots of commas in this sentence and this phrase is not critical to the dose response modeling topic.

5.2.2 Subchronic Oral Reference Dose Derivation

No comments

5.2.3. Inhalation Reference Concentration (RfC)

5.3 Cancer Toxicity Values

References

General comments:

- **NTP (National Toxicology Program). (2011). Specifications for the conduct of studies to evaluate the reproductive and developmental toxicity of chemical, biological and physical agents in laboratory animals for the national toxicology program (NTP).**
  - Note that an updated, compiled version of the specifications was published in March 2023 and replaces the 2011 documents (https://doi.org/10.22427/NIEHS-00)

- **NTP (National Toxicology Program). (2019). NTP technical report on the toxicity studies of perfluoroalkyl carboxylates (perfluorohexanoic acid, perfluorooctanoic acid, perfluorononanoic acid, and**

perfluorodecanoic acid) administered by gavage to Sprague Dawley (Hsd:Sprague Dawley SD) rats [NTP]. (Toxicity Report 97). Research Triangle Park, NC. https://ntp.niehs.nih.gov/ntp/htdocs/st_rpts/tox097_508.pdf.

- Note that a revised version of Toxicity Report 97 was published in 2022 and replaces the 2019 reference (https://doi.org/10.22427/NTP-TOX-97)

Other general comments:

- Citations could be integrated (and not just inserted) with consistent format throughout the document. For example:
  - Page 5-9, lines 8-10: "With respect to the offspring liver enlargements, relative liver weight 8 increases at PND 1 and 24 in (Das et al., 2015) and PND 21 in (Wolf et al., 2010) were selected to 9 model effects at birth…". In this sentence the two citations could be integrated and formatted as "…increases at PND1 and 24 in Das et al. (2015) and PND21 in Wolf et al. (2021)…"