

Expert Panel Status Update and Interim Results

Prepared for:
USEPA Science Advisory Board
PFAS Review Panel

On behalf of:
The American Chemistry Council



March 28, 2022

Executive Summary

An expert panel has been assembled to provide interpretation and guidance on the topic of per- and polyfluoroalkyl substances (PFAS) and potential effects on immune system effects. This panel was defined to have multiple subpanels to ensure sufficient coverage of the expertise areas for this topic area. Interim feedback from the dose-response subpanel identified several concerns with the study of Budtz-Jorgensen and Grandjean (2018) and its potential application to risk assessment, including those associated with the shape of the dose response relationship, confidence limit calculation, PFAS adjustments and correlations, validation (models used, internal, and external), and overall generalizability of the results. This review is ongoing, and additional panel input will become available in the near future.

1. Introduction

Concerns have been raised regarding a potential link between PFAS exposures and a variety of health effects, including those on the immune system. The National Toxicology Program (NTP) concluded that both PFOA and PFOS should be “presumed to be an immune hazard to humans” based on evidence that the two compounds suppressed the antibody response from animal studies and a moderate level of evidence from studies in humans. More recently, USEPA has relied upon immune system effects to derive oral reference dose (RfD) values for PFOA and PFOS (USEPA, 2021a,b).

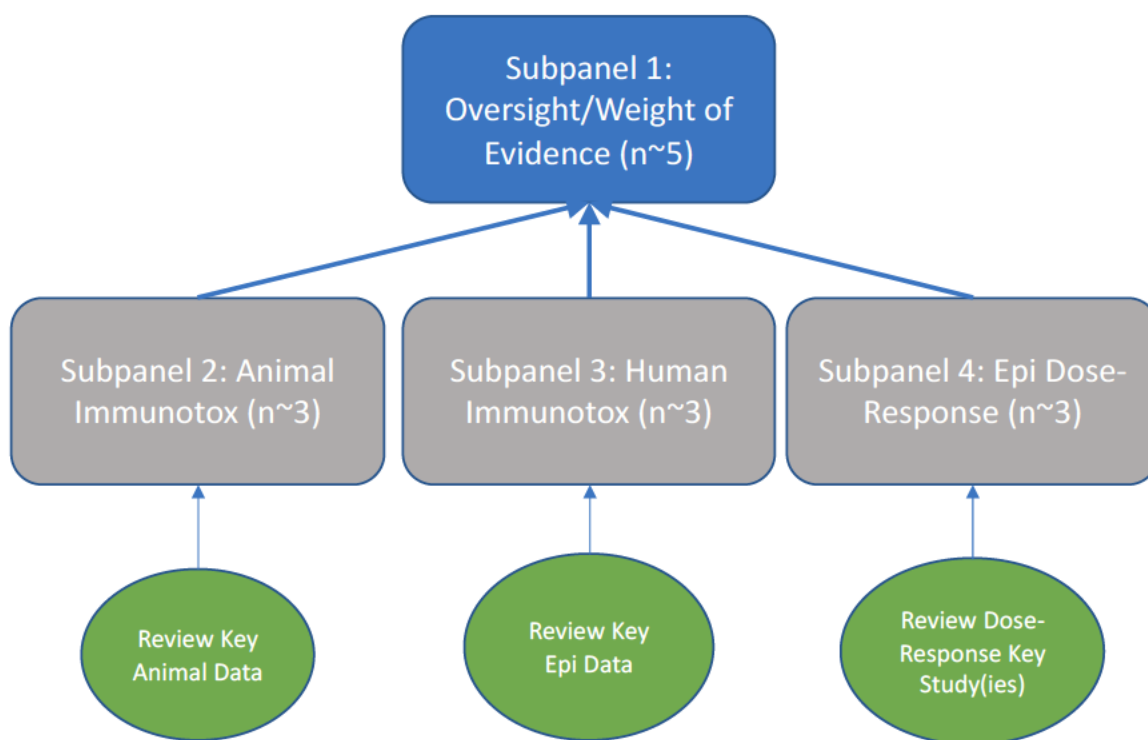
SciPinion LLC was retained by the American Chemistry Council to conduct an independent review of current state of the science for per and polyfluoroalkyl substances (PFAS) and its potential effects on the immune system. This review is ongoing. The text below provides a summary of the methods and interim input for this expert panel engagement.

2. Methods

Panel Design

A nested panel design was adopted for this review, which consisted of four subpanels to ensure adequate coverage of the expertise required for this topic (**Figure 1**).

Figure 1. Nested Panel Design



Panel Assembly and Engagement

A panel of 14 experts was assembled for this review, along with 2 topic leads, using the following steps: (1) Panel Recruitment; (2) Panel Selection; and (3) Panel Engagement. Each of these steps is summarized below.

Panel Recruitment

The goal of the panel recruitment is to cast as wide a net to reach out to as many potential candidates as is feasible. Potential candidates were identified as having relevant experience PFAS, immune system effects, and/or weight of evidence reviews from multiple sources: (1) SciPinion's internal database; (2) searches for authors of recent publications on the topic of interest in online databases (e.g., Pubmed; Google Scholar); (3) searches of profiles on social media databases (e.g., LinkedIn); (4) general internet searches; and (5) referrals. Email addresses were obtained for as many potential candidates as possible with the exception of those who were considered conflicted by working in the industry of the sponsor. An email invitation was sent to all potential candidates, requesting interested candidates to volunteer on <https://app.scipinion.com> and upload a copy of their CV. A total of 138 applications were received.

Panel Selection

14 panel members were selected from the available applicants based upon a consideration of their expertise based upon objective metrics for expertise (e.g., years of experience, number of publications, number of first/last author publications, key

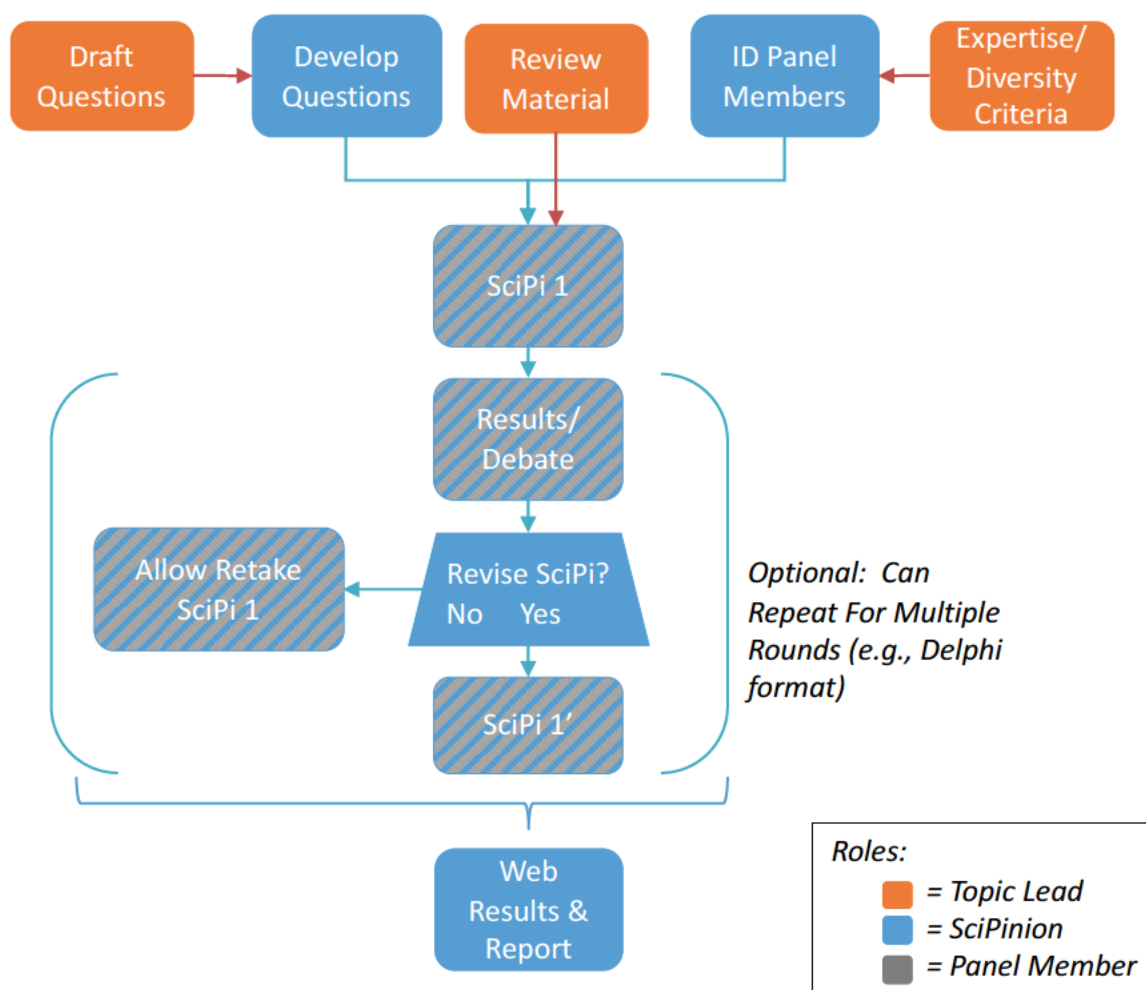
word counts), topic-specific expertise, and experience in making weight-of-evidence decisions for chemical risk assessment.

Panel Engagement

The expert engagement is structured to have multiple rounds, using a modified Delphi format (**Figure 2**) that includes the following design elements:

- During the initial round, all experts will review their assigned review materials and answer initial charge questions independently.
- A comment and debate round is included so that all responses to the charge questions are viewable by the panel members. Panel members are invited to comment and debate each others' responses anonymously (panel members identified as "*Expert 1, Expert 2, ..., Expert 14*"; numbers assigned randomly).
- Subsequent rounds are provided so that experts can revisit their initial responses (in case comments and debate with fellow panel members results in a change in their position) as well as to answer follow-up and clarifying questions.
- Oversight subpanel members are given the opportunity to provide their own charge questions to each of the subpanels to ensure the review scope/charge is not limited or biased.
- The review Sponsor remains blinded to the identities of the panel members until the entire review is complete.

Figure 2. Panel Engagement Process



3. Results

Expert Panel Members

The expert panel consisted of 14 scientists from all over the world, for which the following information was obtained from their CVs:

- Geographic Region: Australia(1), Canada(1), India(1), Italy(1), Mexico(1), United States(9)
- Gender: Male (10), Female (4)
- Degrees: PhD (13), MD (2), DVM (1)
- Total person-years of experience: >450 years
- Total number of publications: >2,500

Interim Results

Interim input is available from **Subpanel 4**, which was initially tasked with reviewing the study of Budtz-Jorgensen and Grandjean (2018) and its application to dose-response assessment (i.e., serves as the basis of USEPA's proposed RfD values for PFOA and PFOS). Interim input from the panel is available in **Appendix A**, and summarized below.

- Strengths of Budtz-Jorgensen and Grandjean (2018) noted by the panel include:
 - The authors' analysis considered and adjusted for exposure to multiple PFAS
 - Consideration of IgG data for two vaccines
 - Use of BMD does not depend on the covariates
 - Use of sensitivity analysis to explore implications of model uncertainty
 - Treatment of continuous variables, without converting to dichotomous variables, is good.
 - Some limitations were recognized by the study authors
 - Collection of serum samples at two times points (birth and 5 years of age) is advantageous
 - The cohort is reasonably representative of the population (although those from caesarean deliveries and obstetric complications were excluded), homogenous, and the study had relatively high retention rates.
- Weakness of Budtz-Jorgensen and Grandjean (2018) noted by the panel include:
 - The shape of the dose response was not sufficiently examined, and does not appear to be biologically sound
 - Confidence limit calculations would have been better based on an approximate method such as the profile likelihood that does not assume an exact normal distribution.
 - A BMD analysis for a weighted combination/relative potency of all PFAS is needed
 - The lack of adjustment for all five PNASs was an interesting decision. It would have been helpful to also see an analysis with all PFASs treated equally. It would also be helpful to know whether this decision was part of the original planning or whether it was made only after seeing the results of the analyses.
 - The models used are not justified. There are no regression diagnostics and no model validation results are presented. Modeling assumptions (e.g., normal distributions) were not tested
 - External validity is not addressed using valid methods. The authors state that ""Because the PFAS exposure levels are similar to those of other populations [7], our results should be applicable beyond the Faroes."" This conclusion does not follow from its premise unless

confounders, modifiers, etc. are also similar across populations, which has not been shown.

- Internal validity is not addressed using valid methods. It is unclear to what extent confounders, model specification errors, latent variables, etc. are responsible for the results. The paper dismisses confounders in a sentence with no supporting data or analyses. Other studies in different populations have reached different conclusions (e.g., "Meta-analysis showed that a higher income is associated with a higher internal exposure to PFAS in adults (Buekers et al. 2018). With few, often methodologically limited studies of any particular health condition, generally inconsistent results, and an inability to exclude confounding, bias, or chance as an explanation for observed associations, the available epidemiologic evidence is insufficient to reach a conclusion about a causal relationship between exposure to PFOA and PFOS and any immune-related health condition in humans." (Chang et al. 2016). Given this past attention to confounding, careful analysis of confounding, residual confounding, and potential latent confounders would have been advisable.
- More information about the correlations between the five major PFAS at birth and at 5 years is needed to have a better understanding about possible sources of exposure, and relevance of timing.
- The generalizability of these findings is limited. The dominant source of PFAS is from marine food contamination. The population of the island is largely homogenous (in terms of ethno-racial characteristics).
- Panelist's confidence (rated on a scale of 1-10 each; 1=lowest, 10=highest) in the study methods, results, and conclusions were highly varied, ranging from 1-10 (mean=4.7), 1-9 (mean=5.7), and 1-9 (mean=4.7), respectively. Based on an overall mean rating of 5, overall confidence is rated medium (i.e., in the middle of the mid-tertile).
- Similarly, panelists' confidence (rated on a 1-10 scale) in application of the results of Budtz-Jorgensen and Grandjean (2018) to human health risk assessment for individual PFAS was also highly varied, ranging from 1-7 (mean=4). Based on this mean, overall confidence in the application of these results rated as medium-low (at the low end of the mid-tertile).

Efforts for other subpanels, which will cover weight of evidence in humans and in animals, the generalizability of the available evidence to other PFAS, are either ongoing or scheduled to start soon. These data will be made available to interested parties in the very near future.

Appendix A: Interim Results

Peer review of dose-response modeling study relating
human biomarkers of exposure to adverse health
outcomes

Paid peer review panel opportunity

Generated: 2022-01-30 18:22:50 +0000

URL: <https://app.scipinion.com/scipis/391/report>

ROUND 1 CHARGE QUESTIONS

Result 1.1 (ID: 5309)

Question 1.1 (ID: 4330)

Please identify the top strengths (up to 3) of Budtz-Jorgensen et al. (2018)

Expert 1

The analysis for each PFAS is adjusted for exposure to other PFASs. The BMD method has the advantage that the BMD does not depend on the covariates. The analysis seems to have been well thought-out and implemented

Expert 3

1. Consideration of multiple PFASs (correlated co-exposures) 2. Use of sensitivity analysis to explore implications of model uncertainty 3. Treating the continuous variables as continuous (not dichotomizing) is good. The authors also deserve credit for recognizing and stating some key limitations (e.g., "some results depended on a part of the dose-response curve for which the data do not hold any information", lack of interaction terms)

Expert 2

1. The collection of serum samples at two times points (birth and 5 years of age) is a distinct advantage of this birth cohort and provides a longitudinal measure of exposure 2. The cohort is reasonably representative of the population (although those from caesarean deliveries and obstetric complications were excluded), homogenous, and the study had relatively high retention rates. 3. Consideration of 5 PFASs and IgG from two vaccines, and the ability of modelling to include multiple PFAs measures (as adjustment factors)

Comments 5

SCORE
-1

Expert 3 01/11/2022 22:16

My main concern is that the regression models used may not describe the data very well. The "adjustment" of each PFAS for the others depends on the model used, as does the estimated BMDL. I would like to see some model diagnostics and/or results from non-parametric modeling.

SCORE
1

Expert 1 01/13/2022 10:36

I agree mostly with the strengths of the study mentioned by the other reviewers, particularly having data multiple PFASs, having samples both at birth and five years of age and use of continuous variables and not dichotomizing them. However use of the continuous data increases the risk that highly influential points (e.g., outliers) could be having a untoward effect on the results. Therefore, use of diagnostics that check for such data points would have been very useful.

SCORE
0

Expert 3 01/13/2022 19:55

I agree: identifying influential observations and performing other regression diagnostics and checks of modeling assumptions would have been useful. Without those checks (or cross-validation etc.) is hard to guess how robust the conclusions are.

SCORE
0

Expert 2 01/14/2022 13:38

I agree with many of these comments. While Expert 1 notes that PFAS is adjusted for exposure to other PFASs it would have been helpful to see more data on how they were all related to each other. Generally speaking, I thought it was a well conducted study but would have liked to see more sensitivity analyses, and an examination of other functional forms of the exposure response relationship. Modelling the continuous data is good, but given this why not explore other shapes of the exposure response curve. Good point about diagnostics and looking at outliers (by Expert 1) above

SCORE
0

Expert 3 01/16/2022 23:04

Agree on all points.

Please identify the top weaknesses (up to 3) of Budtz-Jorgensen et al. (2018)

Expert 1

The shape of the dose response was not sufficiently examined to my satisfaction. Only log, linear and piecewise linear models were studied. The piecewise linear model, with an abrupt change of slope, doesn't appear to be biologically sound. I did not understand the argument that the confidence limits for the BMD were exact because they were "based on the fact that regression coefficients follow normal distributions." I checked the stated reference for this conclusion and found nothing that supported this conclusion. It appears to me that the confidence limit calculations would have been better based on an approximate method such as the profile likelihood that doesn't assume an exact normal distribution. The analysis was directed at determining benchmark results for each PFAS individually. Even each PFAS individually is present in sufficiently low blood concentrations the total of their individual contributions could still be problematic. It would seemingly greatly aid in the the regulation of these chemicals to be able to regulate them as a group. Assuming the individual PFASs have a similar mode of action, could these data support an analysis that provides a BMD analysis for a weighted combination of all PFASs? Or could relative potencies derived from other data be used with these data to develop such a BMD analysis? The lack of adjustment for all five PNASs was an interesting decision. It was supported by the fact that "that PFOS and PFOA are considered the PFASs with the best documentation for immunotoxicity." It would have been helpful to see an analysis with all PFASs treated equally. It would also be helpful to know whether this decision was part of the original planning or whether it was made only after seeing the results of the analyses.

Expert 3

1. Models used are not justified. There are no regression diagnostics and no model validation results are presented. Modeling assumptions (e.g., normal distributions) were not tested (or at least no results from such tests are shown). Thus, the conclusions rest on assumptions and models of unknown validity. 2. External validity is not addressed using valid methods. The authors state that "Because the PFAS exposure levels are similar to those of other populations [7], our results should be applicable beyond the Faroes." This conclusion does not follow from its premise unless confounders, modifiers, etc. are also similar across populations, which has not been shown. 3. Internal validity is not addressed using valid methods. To what extent are confounders, model specification errors, latent variables, etc. responsible for the results? The analysis leaves the answer unclear. The paper dismisses confounders in a sentence with no supporting data or analyses ("No important confounders were identified among a wide range of social and demographic parameters, and adjustments therefore included only sex and age and, for the age-7 data, the type of booster vaccination at age 5.") Other studies in different populations have reached different conclusions (e.g., "Meta-analysis showed that a higher income is associated with a higher internal exposure to PFASs (PFOS or perfluorooctanesulfonic acid, PFOA or perfluorooctanoic acid, PFNA or perfluorononanoic acid, PFHxS or perfluorohexane sulfonate)" in adults (Buekers et al. (2018) <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6313392/>); ""With few, often methodologically limited studies of any particular health condition, generally inconsistent results, and an inability to exclude confounding, bias, or chance as an explanation for observed associations, the available epidemiologic evidence is insufficient to reach a conclusion about a causal relationship between exposure to PFOA and PFOS and any immune-related health condition in humans." (Chang et al. 2016). Given this past attention to confounding, careful analysis of confounding, residual confounding, and potential latent confounders would have been advisable.

Expert 2

1. An objective of their analyses was to try and characterize shape of exposure response curve. In my view, they did not fully explore the shape as curve as they could have. They only consider linear, and piecewise linear. It would have been helpful to explore this in more sophisticated models (e.g., spline functions). Moreover, the impacts of the adjustment factors would likely vary across models, and this could have been described in greater detail. There were substantial differences in the BMD across statistical models - the focus on the paper was on the mutual adjustment of other PFASs and I felt differences by type of model was done discussed as in depth as it could have. In fact, there was no mention whatsoever of different models or results in the abstract (linear dose response vs piecewise exponential) 2. I would have liked to have seen more information about the correlations between the five majors PFAS at birth and at 5 years to have a better understanding about possible sources of exposure, and relevance of timing. I recognize there may be some constraints in sample size, but there seems to be little efforts looking at differences in the two cohorts. Have there been any changes in the types of vaccines used over the two different cohorts? 3. Limited generalizability of these finding as the dominant source of PFASs is from marine food contamination. The population of the island is largely homogenous (in terms of ethno-racial characteristics) and I can't help but wonder if this may also it difficult to generalize findings from this study population to other regions.

Comments 2

SCORE -1	Expert 3 I agree with these insightful comments from both of you. My concerns about potential model form misspecification errors, lack of model diagnostics, and lack of model validation are similar. Non-parametric analyses (e.g., MARS for splines) could be useful here. I agree specifically that the comments on normal distributions do not seem well justified.	01/11/2022 22:21
SCORE 0	Expert 2 These are solid comments, and I agree with nearly all of them. All of us drew attention to the fact the authors could have considered other exposure-response curves (shapes). I strongly agree with Reviewer 3's concern about the issue of external validity. I too am skeptical that these findings would be applicable to other populations.	01/14/2022 13:52

Please rate your confidence in Budtz-Jorgensen based on consideration of the following study components (1=lowest confidence; 10=highest confidence). Please explain your rating in the space below.

Component	Confidence Rating	Total
Methods: Consideration of and adjustments for covariates, model form, modeling decisions, statistical tests and assumptions	1 33.33% 1 5 33.33% 1 8 33.33% 1	3
Results: Were all results reported with transparency and without bias? Are the results accurately reported?	1 33.33% 1 10 33.33% 1 6 33.33% 1	3
Conclusions: Are the conclusions drawn supportable based on the methods used and results reported? Are there alternate conclusions that could/should be presented?	1 33.33% 1 4 33.33% 1 9 33.33% 1	3

Answer Explanations

Expert 3

Component	Confidence Rating
Methods: Consideration of and adjustments for covariates, model form, modeling decisions, statistical tests and assumptions	1
Results: Were all results reported with transparency and without bias? Are the results accurately reported?	1
Conclusions: Are the conclusions drawn supportable based on the methods used and results reported? Are there alternate conclusions that could/should be presented?	1

Methods: No regression diagnostics or model validation results are presented. Model assumptions are not tested. The model form is assumed but not justified. Exclusion of interaction terms and nonlinearities (other than a permitted change in slope at the median) is not justified. The model form is ad hoc and the modeling assumptions are not critically evaluated or validated. Confounding is dismissed with not careful analysis. Sensitivity analyses do not address the full range of what other plausible model forms might show. Results: No results are presented for tests of model assumptions. Conclusions: The conclusion that the results can be expected to hold in other populations (external validity) is indefensible based on the reasoning given (similar exposures). Neither internal nor external validity of conclusions has been established.

Expert 1

Component	Confidence Rating
Methods: Consideration of and adjustments for covariates, model form, modeling decisions, statistical tests and assumptions	8
Results: Were all results reported with transparency and without bias? Are the results accurately reported?	10
Conclusions: Are the conclusions drawn supportable based on the methods used and results reported? Are there alternative conclusions that could/shoud be presented?	9

The methods were well described, and were appropriate for the analysis goals. As mentioned earlier, I questioned the assumption of an exact normal distribution. However, I doubt that this would have a big impact on the results. I detected no bias in how the results were presented. No strong conclusions were presented, which is appropriate, I believe.

Expert 2

Component	Confidence Rating
Methods: Consideration of and adjustments for covariates, model form, modeling decisions, statistical tests and assumptions	5
Results: Were all results reported with transparency and without bias? Are the results accurately reported?	6
Conclusions: Are the conclusions drawn supportable based on the methods used and results reported? Are there alternative conclusions that could/shoud be presented?	4

As per my comments above, I think the authors could have provided more data and analysis that allowed for; (i) different patterns in exposure response; (ii) assessed differences between the two cohorts, and iii) caution that their findings may not be readily generalizable to other populations. The authors could be more cautious about extrapolating findings reported in their paper to other populations

Comments 2

SCORE	Expert 2	01/14/2022 13:48
0	<p>After reviewing comments in the two previous pages of discussion, if I had to fill this out again, I would probably assign a 4 (for methods), a 6 for results, and a 5 for conclusions. For methods, I do agree that additional sensitivity analyses and model checks could have been done. That said, for what they did, I still think they were presenting results in a transparent way or with bias. Conclusions, authors could have been a little more cautious on generalizing findings elsewhere.</p>	

Very fair. In light of our discussion, I would be comfortable coming up some (my ultra-low scores were mainly reflecting uncertainty), although I do not think I would go above 5 on any criterion even now. I agree that the authors were transparent, but I what that clarity reveals seems to me to be possibly model-dependent conclusions of uncertain validity. (BTW, did you mean "without bias" instead of "or with bias"? It seems to me that the authors were good on transparency and lack of any obvious deliberate bias (other than over-generalization and perhaps unwarranted causal interpretation), but weaker on explicit tests of assumptions and internal and external validity tests.

On a scale of 1-10 (1=low; 10=high), how confident are you in applying the study of Budtz-Jorgensen et al. (2018) to human health risk assessment for each PFAS considered (i.e., as the basis for an oral reference dose) ? Please explain your answer below

	Rating	Total
PFOA	1 33.33% 1 4 33.33% 1 7 33.33% 1	3
PFOS	1 33.33% 1 4 33.33% 1 7 33.33% 1	3
PFHxS	1 33.33% 1 4 33.33% 1 7 33.33% 1	3
PFNA	1 33.33% 1 4 33.33% 1 7 33.33% 1	3

PFDA	1	3
	33.33%	
	1	
	4	
	33.33%	
	1	
	7	
	33.33%	
	1	

Answer Explanations

Expert 3

	Rating
PFOA	1
PFOS	1
PF xS	1
PFNA	1
PFDA	1

The results presented have no known internal or external validity. They are driven by ad hoc modeling assumptions for which no tests or validation results are presented. Choosing different, equally good modeling assumptions could lead to opposite conclusions. Roles of confounders, latent variables, and model specification errors are unknown. Any interpretation of the findings is therefore purely speculative; that analysis presented does not justify any clear conclusions about causal exposure-response relations in the real world.

Expert 1

	Rating
PFOA	7
PFOS	7
PF xS	7
PFNA	7
PFDA	7

This question is difficult to answer without knowing what other information is available. Assuming these results are all that is available, and an oral reference dose is needed, I would recommend using the results of this paper for determining an oral reference dose, at least on an interim basis, until more definitive information becomes available.

Expert 2

	Rating
PFOA	4
PFOS	4
PF xS	4
PFNA	4
PFDA	4

I am uncomfortable given many of the unique features of this cohort to applying it to other populations. In my view, it would be importance to replicate similar studies elshwere before doing so. The evaluation of mutual adjustment of other PFASs may well vary across different types of models, and this was not explored in as much detail as it could have, Having correlations presented in the paper (between the different PFASs) would have been very helpful for helping me interpret the findings, and critique the methodological approach

Comments 4

SCORE 0	Expert 1 I agree with the short-comings of the analysis that were pointed out by the other two reviewers. However, I don't consider then quite so serious as Reviewer 3, in particular, seems to believe as reflected in his or her ratings..	01/13/2022 12:13
------------	---	------------------

SCORE 0	Expert 3 I take your point that they may not be as serious as I have indicated (i.e., maximal severity). A more nuanced description would be that in the absence of diagnostics, validation tests, or robust (e.g., nonparametric or resampling-based) analyses, I am left very uncertain about the adequacy of the models and the soundness of the conclusions. I agree that this does not necessarily imply that they are wrong, just that there is no good way to tell how trustworthy they are. My low score (high severity) reflects uncertainty about whether the conclusions are sound, not an assertion that they are unsound.	01/14/2022 13:27
------------	--	------------------

SCORE 0	Expert 2 Ultimately, I think the design of the study is fairly strong - with good information and exposure and outcome. I think it is a valid, yet not complete analyses. More could have been done about confounding, and sensitivity analyses. For me though, I am not convinced materially the measures of association. My largest concerns is that I have doubt these findings could be generalized/applied to other populations (external validity). I also can't help but note that 2nd author has published a number of excellent papers over the year, and in my view, at the onset led me to believe likely represented solid research,	01/14/2022 13:58
------------	--	------------------

SCORE 0	Expert 3 Fair points. I was trying to evaluate this paper only on its own methodological strengths and limitations.	01/16/2022 23:13
------------	---	------------------

MISCELLANEOUS

Result 2.1 (ID: 5313)
Question 2.1 (ID: 4334)

Please identify any potential conflicts of interest you may have that might impact your ability to review the assigned paper(s) impartially

Expert 3

None

Expert 1

None

Expert 2

None

Comments 1

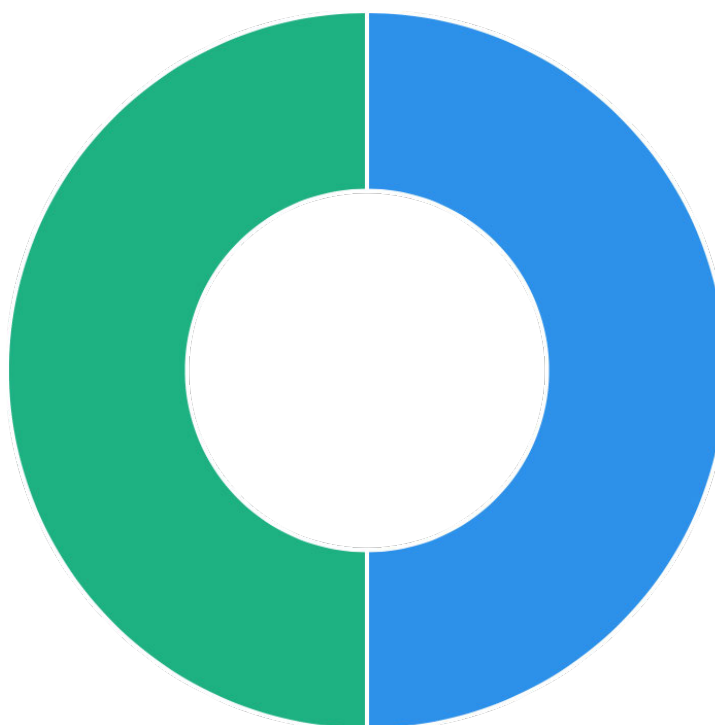
SCORE	Expert 2	01/14/2022 13:59
0	Always nice to see no one is conflicted (lol)	

ROUND 3 QUESTIONS

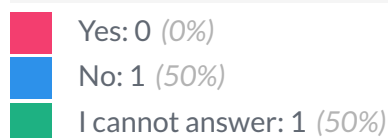
Result 3.1 (ID: 5345)

Question 3.1 (ID: 4361)

The study design involved sampling of serum levels of mothers and young children at various time points. Certain exposure periods were selected in the statistical analysis to account for potential interactions (confounders). Biomarkers of exposure in this study populations have been assessed for other chemicals, including PCBs and methylmercury (e.g., Heilman et al., 2010; Osuna et al., 2014). Was the identification and assessment of potential confounders in Budtz-Jorgensen and Grandjean (2018) sufficient for this study?



Legend



answers: 2

skips: 1

Answer Explanations

I cannot answer

Expert 1

I do not have sufficient understanding of the potential confounders in this study to feel comfortable in answering this question.

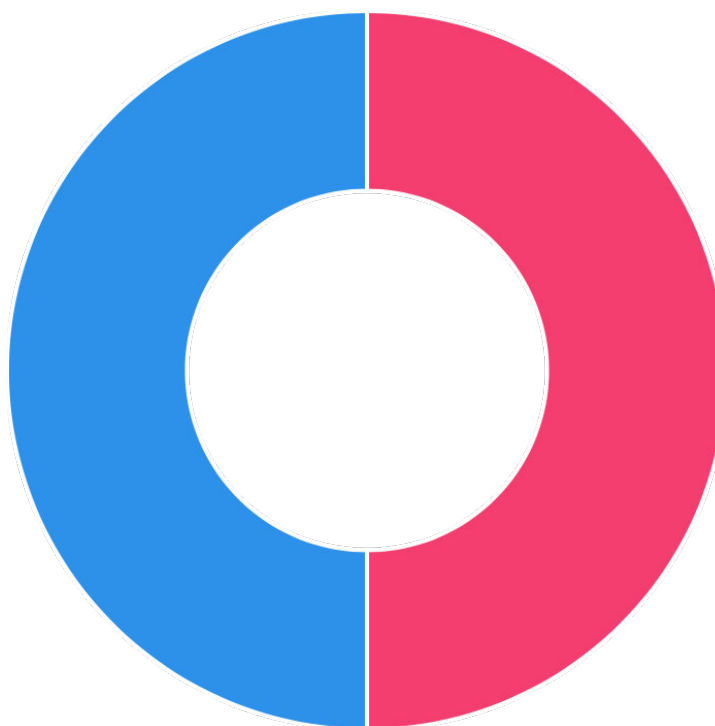
Identification and assessment of potential confounders in the Budtz-Jorgensen and Grafdjean (2018) study were not discussed thoroughly. The authors state that "No important confounders were identified among a wide range of social and demographic parameters," but this may be because the variables examined and the methods used (not discussed in detail) were inadequate. No results of formal tests for latent confounders or residual confounding are included, and no models or results (e.g., DAG models, results of conditional independence) are offered that would allow the adequacy of control for confounding to be evaluated. The authors appear to have dismissed the importance of potential confounders without a careful presentation of relevant analyses and results. (Grandjean was co-author an article by Timmerman et al., 2020, <https://ehp.niehs.nih.gov/doi/10.1289/EHP6517> that examine potential confounding more thoroughly in an African population. Similar analyses might have been useful for the Budtz-Jorgensen and Grafdjean (2018) study.)

No

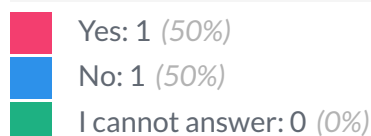
Expert 2

There was little discussion about confounders in the paper, and analyses included the covariates of sex, age and booster type. Previous work by Heilman suggested the relevance of PCBs - perhaps this should have been considered as a possible confounder. Some other studies of antibody response to diphtheria vaccine have identified possible role of a seasonal influence, and perhaps there is a role of factors such as body mass index, nutritional factors, exposure to second hand smoke, etc. I am unsure whether these could be confounders or not, but wish the paper would have commented on this possibility. For all these reasons, I am skeptical that confounding was adequately addressed in the study.

Are you aware of additional studies that should be considered with respect to the characterization and assessment of covariates and potential confounders?



Legend



answers: 2

skips: 1

Answer Explanations

Yes

Expert 3

Several papers such as Bulka et al. 2021 (<https://www.sciencedirect.com/science/article/pii/S0269749121001974?via%3Dihub>) and Timmerman et al. 2020 (<https://ehp.niehs.nih.gov/doi/10.1289/EHP6517>) discuss DAG models for covariates and potential confounders in other PFAS health effects studies. Similar methods may be useful here. The somewhat dated paper of Chang et al. (2016) discusses potential confounders for several measures related to immunotoxicity (Chang ET, Adami HO, Boffetta P, Wedner HJ, Mandel JS. A critical review of perfluorooctanoate and perfluorooctanesulfonate exposure and immunological health conditions in humans. Crit Rev Toxicol. 2016;46(4):279-331. doi: 10.3109/10408444.2015.1122573. Epub 2016 Jan 13. PMID: 26761418; PMCID: PMC4819831.)

Expert 2

Their was the study by Heilman of the same population (EHP, 2010). There was this other study I found whose findings suggest a relevance for seasonality (<https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-021->

11383-7). As part of my own reading to learn more, I also found the following review of some relevance to this question (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6431125/>)

Please comment on the use of a piecewise linear model in this study? What are potential issues with using a piecewise linear model for establishing biological responses? What are the potential advantages and disadvantages of this model? Are there other models that may be more appropriate?

Expert 1

A potential advantage of the piecewise model is that it does provide some information on potential nonlinearity of the dose response. However I feel the shape of the dose response was not studied extensively enough. The piecewise linear shape dose not seem very plausible. I would liked to have see the results of use of a more flexible model.

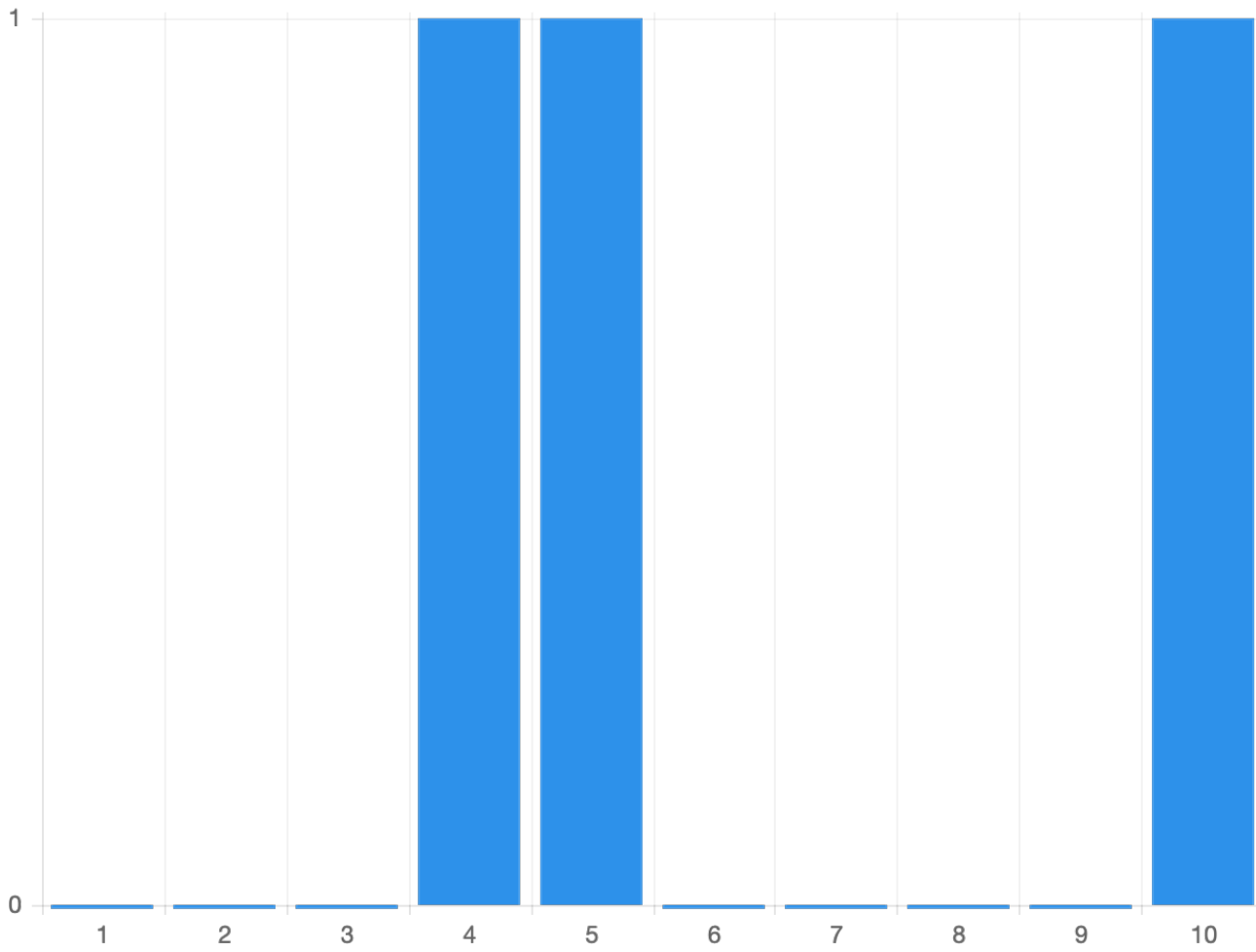
Expert 3

A non-parametric method such as smoothing regression, e.g., using loess or multiple adaptive regression splines (MARS) would avoid the very real dangers of model misspecification inherent in the piecewise linear model. The authors did not present regression model diagnostics or show that the piecewise linear model is appropriate for these data. It is not clear to me to what extent the reported findings result from use of this rather ad hic-seeming model. Using a nonparametric smoothing approach could clarify the extent to which the piecewise linear model is or is not justified for this dataset.

Expert 2

The piecewise model allows for a non-linear association to be modelled. This provides an advantage over the standard linear model. However, there are other types of models that can provide more sophisticated fits of the data to inform on the shape of the exposure response curve (e.g., splines). I often use spline models for assessing the shape of the curve between environmental exposures (e.g., air pollution, diesel exposure) and adverse health outcomes. Another advantage of the piecewise model (even though I might prefer spline type models) is that they may be better equipped to provide insight about the association at low levels of exposure (particularly over the standard linear model) Regardless, fits of models should always be compared - and the authors did so in this paper.

What is your level of concern with the inability to visually inspect model fit (i.e., not depicting response data in Fig. 1) (scale of 1 to 10; 1=no concern, 5 equivocal, 10- serious concern)? In your explanation, please indicate how model fit should be assessed, and if you believe these data should be made available for independent analysis and verification.



Legend

answers: 3
skips: 0

Answer Explanations

5

Expert 1

It is always helpful, in my opinion, to have a visual display of model fit as a supplement to results of statistical analysis. I also feel that it is useful to make data available for independent analysis, particularly if the results of an analysis of the data is to be used to set exposure recommendations or standards.

10

Expert 3

In the absence of model validation results and regression diagnostics, I am left very uncertain about whether the claimed results are model-dependent (driven by model selection and assumptions). My high concern score reflects this uncertainty: it is not meant as an assertion that the reported results are wrong, but as an assertion that there is no way to determine whether they are wrong. Using model ensemble methods (e.g., random forest and MARS) and nonparametric smoothing regression (e.g., MARS) with cross-validation tuning of meta-parameters could reveal whether the piecewise linear approximation gives similar results to more flexible (less assumption-dependent) methods.

4

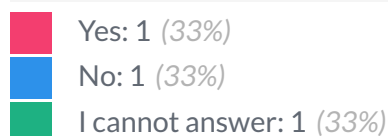
Expert 2

I can understand sometimes that presenting the 'observed' data in such a plot are of limited value as with a large number of observations often there tend to be no clear patterns. So, while ideally, I would like to see this, I can sometimes understand the motivation for not doing so. I would however have liked to see a lot more information on comparisons of model fit, or residual plots (even as an appendix). Graphically presenting confidence intervals for the derived curves could have also been helpful.

Is the use of a benchmark response rate of 5% appropriate if the resulting point of departure values fall outside of the range of observation defined by the data?



Legend



answers: 3

skips: 0

Answer Explanations

Yes

Expert 1

I would not recommend the fact that the point of departure falls out side of the range of observation be used as a criterion for deciding to modify or not use the result of the benchmark calculation.

No

Expert 3

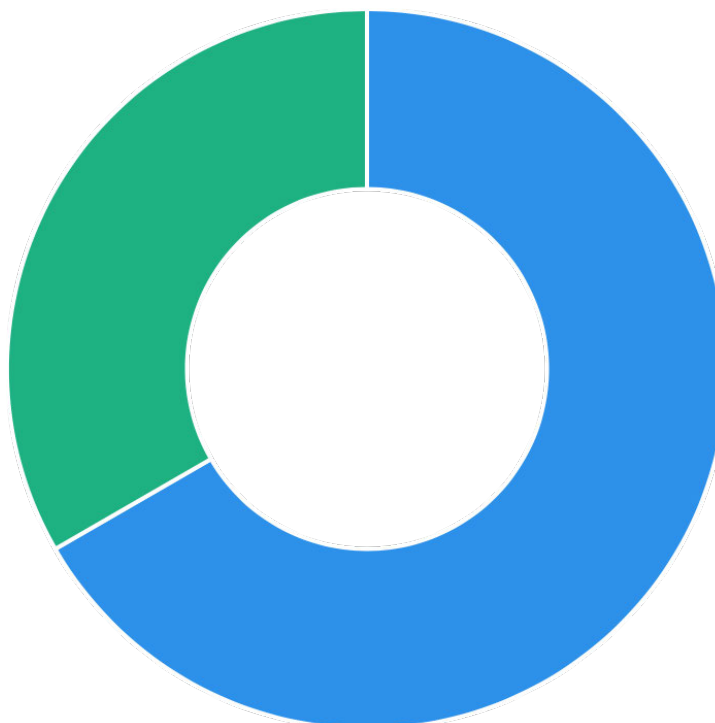
Extrapolation outside the range of observations is speculative. (A well-validated causal biological model might support such extrapolation, but no such model is used in the Budtz-Jorgensen and Grafdean (2018) study.)

I cannot answer

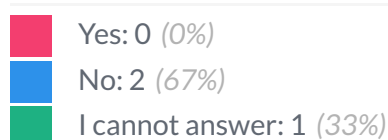
Expert 2

I don't feel I have a solid enough understanding of BMD methodology to answer this question

Are the results of this study generalizable to other populations? If not, what are the limitations of this study to other populations and what improvements to the study design could be made to address these limitations?



Legend



answers: 3

skips: 0

Answer Explanations

I cannot answer

Expert 1

I do agree with the other responders that perhaps the somewhat unique environment of the study population and particularly their exposure circumstances may limit the generalization of results to other study populations, although unless particular important limitations are shown to exist, I would not recommend that this be used as a reason not to use the results of the study in setting human exposure advisories. I don't know that anything could be changed in the study design to address potential concerns, given the data that are available.

No

Expert 3

The external validity of conclusions has not been addressed in the Budtz-Jorgensen and Grafdean (2018) study (except for a statement that "Because the PFAS exposure levels are similar to those of other populations [7], our results should be applicable beyond the Faroes." This reasoning does not appear to be sound, as it does not consider

whether conditions for transportability, external validity, and generalization of results are met.)

No

Expert 2

In my view, this is a very unique population in terms of its characteristics, and sources of exposure to PFASs. I would be uncomfortable generalizing these findings to other populations.
