FLUORIDEALERT.ORG
Fluoride Action Network

Submission to NAS on the revised draft NTP Monograph on the Systematic Review of Fluoride Exposure and Neurodevelopmental and Cognitive Health Effects

October 19, 2020

Chris Neurath
Research Director
Fluoride Action Network (FAN)

1. Introduction

As the nominators for the NTP's systematic review of fluoride neurotoxicity we are pleased to submit comments to the NAS peer-review committee on the revised NTP monograph. We nominated fluoride for review in 2015 and it is noteworthy that since then the scientific evidence has rapidly expanded with the highest quality studies all being published since 2017. In just the months following the first draft NTP monograph there have been several more high-quality studies published that have been added to the revised monograph. The studies published in the past 5 years are important for their high quality and because many found adverse neurotoxic effects at low exposure levels, including the level (0.7 mg/L) used in artificial water fluoridation.

At the open meeting of this NAS committee in October 2019 we submitted comments on the first draft of the NTP monograph. We are including those comments as Appendix 3. We found that the objective aspects of systematic review were well conducted and followed OHAT guidelines but the more subjective aspects had some serious short-comings, including lack of transparency and lack of pre-specified methods in the protocol.

We found the main conclusion, that *fluoride poses a presumed hazard of developmental neurotoxicity*, to be well supported by the body of scientific evidence. However, we found that the section titled "Generalizability to the U.S. Population" had serious problems. It was, in essence, an informal risk assessment that incorporated simplistic exposure assessments and dose-response assessments to reach a conclusion about the confidence that fluoride at exposure levels in the US are likely to cause developmental neurotoxicity. The protocol, however, contained no mention of a Generalizability assessment which represents a fundamental violation of transparency and pre-specification. The Generalizability section appeared to be *ad hoc* and tacked on to the report.

This NAS committee also criticized the inclusion of the Generalizability section and recommended that NTP eliminate it and restrict itself to a hazard assessment and state clearly that the purpose of the NTP systematic review was not to weigh in on what a safe dose might be:

> Lastly, the discussion section of the monograph provides an informal assessment of the evidence with regard to exposure and concludes that adverse health effects are observed largely in association with exposures above those associated with water fluoridation. *The basis of that conclusion is not apparent and seems to contradict the earlier assertion that nearly all the studies are positive, including ones that evaluated groups exposed to lower concentrations*. More important, as noted above, *this discussion gives a false impression that NTP conducted a formal dose–response assessment. NTP should be clear that the monograph cannot be used to assess what concentrations of fluoride are safe.* [*emphasis added*; NAS 2020, p 5]

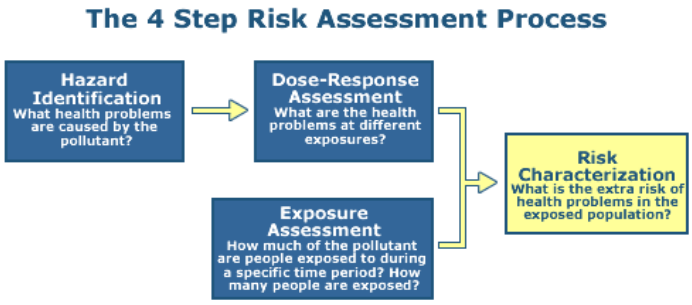
While we agree with the NAS's criticism of the Generalizability section we also believe the NTP's systematic review found sufficient evidence to answer the key question:

Is artificial water fluoridation in the US likely to be causing harm from developmental neurotoxicity?

The NAS committee acknowledged this was a question of central interest.

To answer that question the NTP would have to use rigorous exposure assessment and dose-response assessment methods rather than *ad hoc* informal methods. The US EPA provides the following diagram (Figure 3) to clarify the distinction between the four steps of a risk assessment. It shows that to determine whether there is a risk at the exposure level due to artificial fluoridation both Dose-Response Assessment and Exposure Assessment are required. The EPA has extensive guidance for conducting valid assessments.

Figure 1.



Step 4: Risk Characterization is the last step of a human health risk assessment.
from https://www.epa.gov/risk/conducting-human-health-risk-assessment

Given that the revised monograph has addressed most of the suggestions for improvement offered by NAS, we believe there is now even stronger support for the presumed hazard conclusion. We urge NAS to now take a broader view and compare the fluoride neurotoxicity monograph to NTP's other monographs, as a way to check whether NTP has applied consistent standards for evaluating different chemicals, a primary goal of the OHAT systematic review process.

We also remind the NAS committee and the NTP that neither should be attempting to balance purported dental benefits against neurotoxic risks. Your role should be confined to objectively evaluating the scientific evidence of neurotoxicity. Any balancing of risks and benefits gets into the realm of risk management and policy formulation, which should be outside the scope of NTP and this NAS committee.

For those who might persist in type to balance dental benefits against neurotoxic risks, we would point out the irony that there has never been a Randomized Controlled Trial (RCT) of fluoridated water or fluoridated salt. This was confirmed by both the York Review and a Cochrane Review [York 2003, York 2000, Cochrane 2015]. Both these authoritative reviews concluded the evidence for fluoride's effectiveness at preventing dental caries is surprisingly weak, especially under current conditions where fluoride toothpaste is widely used. In contrast, a Cochrane Review identified 70 RCTs for fluoridated toothpaste and concluded there was strong support for its effectiveness [Cochrane 2003].

The irony is thus that if the OHAT methodology were applied to the question of fluoridation's effectiveness at reducing caries, the evidence would not warrant a confidence conclusion better than "suspected effective".

STRENGTHS

2. How has the revised NTP monograph addressed our comments and the recommendations of the NAS?

We believe the revised monograph has strengthened support for the conclusion that fluoride poses a presumed hazard of developmental neurotoxicity. The revised monograph clarifies and refines the NTP's systematic review. The updated literature search has also provided increased confidence in its conclusions because of the addition of several high-quality studies finding adverse effects. The number of studies rated as lower Risk of Bias (RoB) by NTP has increased from 18 to 29. This increase is partly from the newly published studies and partly from a revised RoB scoring protocol that more appropriately scored studies on key domains like confounding.

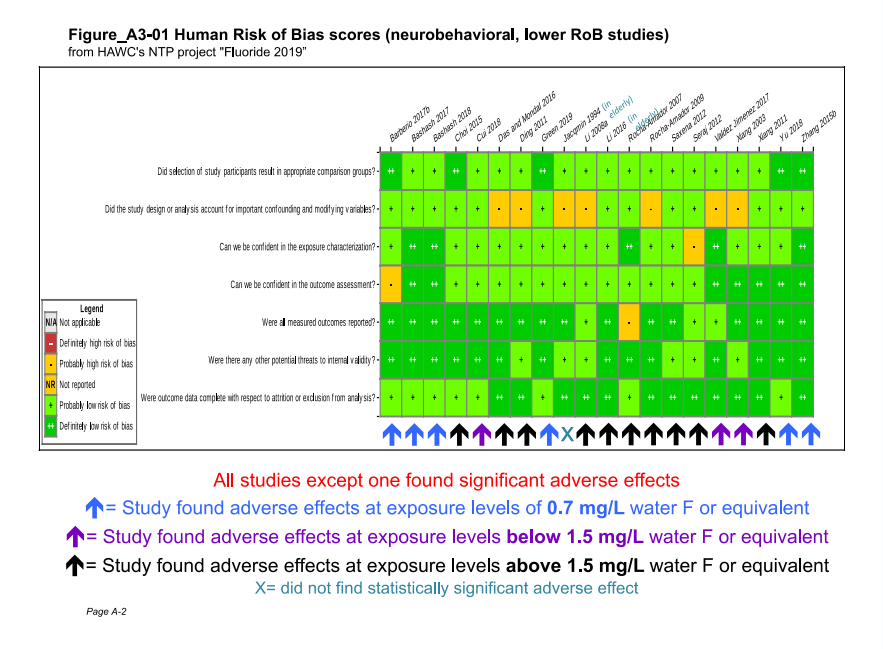
Figures 1 and 2 compare the lower RoB studies heatmap for the first draft 2019 monograph to the updated heatmap of the revised 2020 monograph. The figures also include added information on the exposure levels in each study and whether it found any statistically significant adverse effects. The figures concisely summarize both the consistency of the evidence and its relevance to exposures in the US from fluoridated water. They illustrate the substantial increase in higher quality evidence in the revised monograph. They also illustrate a doubling of the number of high-quality studies finding adverse effects at exposures relevant to artificial water fluoridation (0.7 mg/L) from 6 to 13 studies.

The exposure level classifications in Figures 1 and 2 are based on a methodology documented in Appendix 2. Briefly, if data was available for Benchmark Dose assessment, the calculated Benchmark Dose (BMD) was considered to be the lowest dose having an adverse effect as long as the predicted dose-response curve did not suggest a threshold below 0.7 or 1.5 mg/L. For studies with only two exposure level groups, the mean or median of the higher exposure group was compared to the cut-points to determine classification.

Our methodology of exposure level classification for each study was refined between the 2019 and the 2020 monographs. This produced changes in classification of a few of the studies. Details of our classification method and their application to each study are provided in Appendix 2. A summary of our risk assessment for the US population is available in Appendix 1 as a conference poster from the International Society for Environmental Epidemiology (ISEE) 2020 Conference.

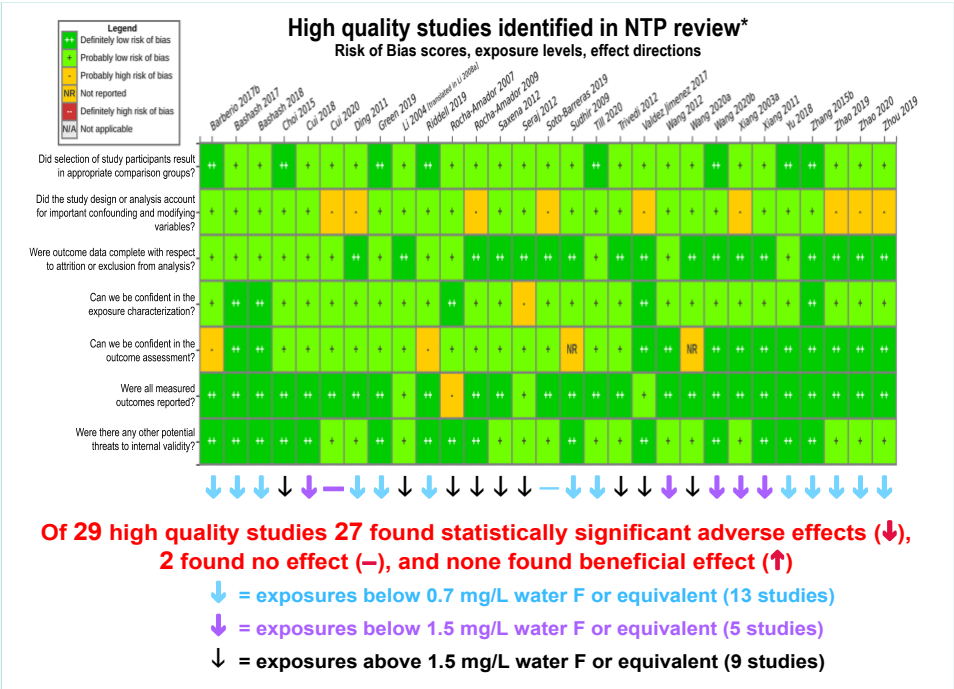
Figures 1 and 2 also illustrate the great consistency of findings. For the revised 2020 monograph, 27 of 29 higher quality studies (93%) found statistically significant adverse effects. Of the 20 studies with findings relevant to exposures below 1.5 mg/L, 18 found statistically significant adverse effects (90%). Of the 14 studies with findings relevant to exposures of 0.7 mg/L, 13 found statistically significant adverse effects (93%).

Figure 2. RoB scores heatmap of NTP 2019 monograph higher quality studies.



(Note: Figure 1 used different exposure levels symbology with up-pointing arrows indicating studies finding statistically significant adverse effects whereas Figure 2 has down-pointing arrows indicating statistically significant adverse effects. Two studies are in elderly so are not considered developmental neurotoxicity.)

Figure 3. RoB scores heatmap of revised NTP 2020 monograph higher quality studies.



(Note: Figure 1 used different exposure levels symbology with up-pointing arrows indicating studies finding statistically significant adverse effects whereas Figure 2 has down-pointing arrows indicating statistically significant adverse effects.)

3. Improvements to RoB scoring validity in revised NTP monograph

An example of how the revised monograph improved its Risk of Bias (RoB) scoring criteria can be found in the RoB domain for confounding.

The NAS emphasized that on the issue of potential confounding, that magnitude and direction of bias from confounding were important but not addressed. NAS also emphasized that it is systematic bias over the whole body of evidence that is of most concern:

> **"Confounding**
> NTP developed a reasonable list of the factors most likely to cause confounding in the literature as a whole (NTP 2019, p. 29); in several cases, it provided thoughtful discussions of the likelihood of confounding by some of the factors." [NAS 2020]

The best example of improved RoB scoring in the revised monograph is for arsenic, which in some places in the world has been found to have a moderate positive association with natural F levels in groundwater. Lead, on the other hand, is not known to have any association with natural levels of F in drinking water so would be of less concern, especially when evaluating the totality of the evidence rather than individual studies which might have some unusual local situation where Pb was associated with F. The revised NTP monograph has modified and clarified their RoB confounding criteria related to arsenic and Pb. For arsenic the NTP now assesses the likelihood of confounding by arsenic in those studies that did not explicitly address arsenic by referring to a world map that shows were groundwater with elevated arsenic is most likely to occur. For studies conducted in areas without likely elevated arsenic, the NTP considers the study to probably not be affected by confounding by arsenic. For Pb, the NTP now acknowledges that Pb is not known to be associated with natural water F levels, so it is unlikely to be an actual confounder of most studies. Therefore, NTP no longer requires Pb to be addressed for a study to achieve a score of probably low RoB in the confounding domain.

We are aware of only one study that might plausibly have actual confounding by Pb [Broadbent 2015]. The direction of bias would be to attenuate an adverse effect. The Broadbent 2015 study location had the unique situation in which many of the non-fluoridated children lived in an area with extremely corrosive drinking water because of unusual hydrogeological conditions. The large majority of fluoridated children, on the other hand, had water that was not corrosive [Osmunson 2016].

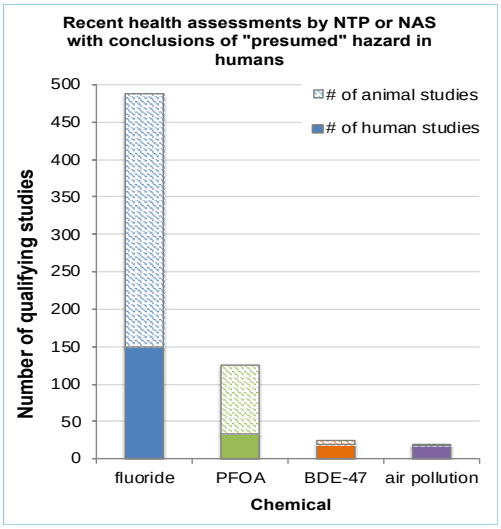
We note that these revisions in RoB scoring did not result in many substantive changes in overall RoB ratings of studies, but they have improved the validity and clarity of the scoring, so they have strengthened the confidence in the conclusions.

4. Comparisons to NTP systematic reviews of other chemicals

Judgements about what constitutes “consistency” of evidence are largely subjective. Neither OHAT guidance nor the NTP protocol offer criteria for what will be deemed “high consistency”, “low consistency” or “excessive inconsistency”. Nevertheless, we believe when there are substantial numbers of high-quality human studies and over 90% of them consistently find statistically significant adverse effects, the conclusion must be that the evidence shows high consistency. That conclusion applies to the subgroup of studies relevant to exposures below 1.5 mg/L and those below 0.7 mg/L as well.

Without explicit OHAT or protocol guidance for what constitutes sufficient consistency for a conclusion of presumed hazard, it is informative to compare the NTP fluoride monograph to NTP monographs for other chemicals. What amount, quality and consistency of evidence in humans and/or animals was sufficient for OHAT to conclude other chemicals were presumed hazards? Figure 4 shows the quantity of studies available in other recent systematic reviews performed using OHAT methodology.

Figure 4.



The NTP fluoride neurotoxicity systematic review identified many more studies than the other chemicals deemed presumed hazards. The large majority of the fluoride studies showed statistically significant adverse effects.

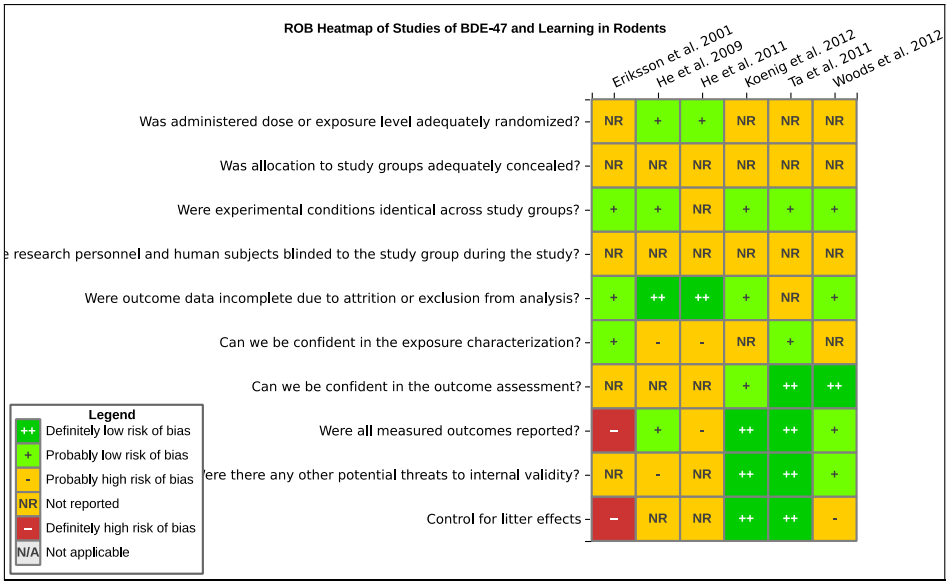
For comparison purposes, the systematic review of the brominated fire retardant chemical BDE-47 identified only 10 human epidemiological studies on neurocognitive effect and 6 animal studies on learning and memory [NAS 2017, Lam 2017]. Of the 10 human studies, 5 would be rated lower RoB using a scoring method similar to OHAT (the Navigation Guide method). With regard to consistency, none of the human studies found a statistically significant reduction in IQ from 10-fold higher BDE-47 exposure (0% consistency). Nine of the 10 did find non-significant adverse effects and 1 found non-

significant beneficial effect. Meta-analysis found a pooled estimate of -3.7 IQ points that was statistically significant.

Compare this number, quality, and consistency of human studies with that of fluoride. For fluoride 65 human studies were identified with 29 rated as lower RoB. Over 90% of these higher quality studies found not just an adverse effect, but a statistically significant adverse effect. The pooled effect magnitude in meta-analysis of the studies with group-level data was about -7 IQ points. Most of the fluoride studies had a smaller exposure difference than the 10-fold difference in the BDE-47 meta-analysis, so fluoride's effect size could be considered much greater.

The quantity and quality of animal studies of BDE-47 to support the conclusion of presumed neurotoxic in humans was limited, as seen in Figure 5 which is the RoB heatmap for BDE-47, for all the identified animal studies on learning. Of the 6 animal studies 5 found "some indication of effect on at least one measure of learning" for a consistency of 80% [NAS 2017, p 8].

Figure 5.



from NAS 2017 report: https://doi.org/10.17226/24758
HAWC visualization: https://hawcproject.org/summary/visual/353/

Comparing the quantity, quality, and consistency of the body of evidence for BDE-47 to that for fluoride demonstrates that fluoride has much stronger support for a conclusion of presumed hazard. Even when restricted to human fluoride neurotoxicity studies relevant to 0.7 mg/L exposures, there is a stronger body of evidence supporting a conclusion of presumed hazard, than for BDE-47.

Similar comparison with other chemicals having recent OHAT systematic reviews finds only Pb has a stronger body of evidence supporting a conclusion of presumed hazard.

5. Meta-analyses support conclusion of presumed hazard

An important recommendation of the NAS, that we concur with, was to conduct meta-analyses. The revised monograph followed this recommendation and has an extensive appendix with various meta-analyses. We find these meta-analyses to be helpful for assessing the body of evidence and the reasons for heterogeneity. They strongly support the conclusion that fluoride is a presumed developmental neurotoxin. Pooled effect estimates showing adverse effects are highly significant not just for the overall group of studies, but for almost all of the subgroup meta-analyses as well. The magnitude of the pooled effect estimates are consistent with the previous meta-analyses by Choi et al 2015 and Duan et al 2018. For the overall group and for most subgroups the magnitude is about -0.5 Standardized Mean Difference (SMD) or about -7 IQ points. This is a large effect and is comparable to the IQ loss from childhood lead poisoning.

While fairly high degree of heterogeneity in effect size was found, this can be explained by the differences in study design, study populations, exposure levels, exposure measures, and outcome measures. The heterogeneity does not weaken the conclusion of high consistency with regards to statistically significant adverse effects.

WEAKNESSES

6. OHAT guidance on “unexplained inconsistency” ignored

Although OHAT provides no guidance on what constitutes consistency, it does clearly state that only “unexplained inconsistency” is grounds for downgrading evidence. The revised monograph Generalizability section argues that studies at exposures relevant to the US are inconsistent. However, on closer inspection, the inconsistency is only on effect magnitude or individual study dose-response relationships. All of these inconsistencies are explainable from differences in study design, populations, gender, exposure measures, exposure levels, and outcome measures. OHAT guidance says all of these reasons for differences are acceptable explanations and do not justify downgrading.

> Inconsistency that can be explained, such as variability in study populations, would not be eligible for a downgrade. Potential sources of inconsistency across studies are explored, including consideration of population or animal model (e.g., cohort, species, strain, sex, lifestage at exposure and assessment); exposure or treatment duration, level, or timing relative to outcome; study methodology (e.g., route of administration, methodology used to measure health outcome); conflict of interest, and statistical power and risk of bias. Generally, there is no downgrade when identified sources of inconsistency can be attributed to study design features

such as differences in species, timing of exposure, or health outcome assessment. [NTP 2019 OHAT Handbook, p 53]

7. NTP used inappropriate methods for "Generalizability to the U.S."

While the evidence base and NTP's conclusion of presumed hazard has strengthened since the previous draft monograph, we also find that NTP has ignored the NAS recommendations to focus on hazard assessment and to avoid risk assessment and trying to identify a safe exposure level. The section titled "Generalizability to the U.S. Population" goes beyond Hazard Assessment into what is essentially risk assessment. Rather than omit the Generalizability section NTP has expanded it. Furthermore, we find that NTP has not used proper exposure assessment and dose-response assessment methods to underpin their revised Generalizability section. Their informal methods have tended to downgrade the evidence and understate the risk at low doses. The NTP has still not included any mention of a generalizability assessment in their protocol so there continues to be a lack of transparency and pre-specification. The only addition to their protocol is a description of a planned dose-response meta-analysis which would presumably contribute to their generalizability section. However, a dose-response meta-analysis is not in itself a dose-response assessment. Ultimately, the planned dose-response meta-analysis was not even conducted for the studies with individual-participant data, which includes most of the strongest studies. Finally, the revised protocol addition describing dose-response meta-analysis was not released until September 16, 2020 and no public comment period was provided. This is a further deficiency of transparency and openness to public comment. If there had been opportunity to comment on the revised protocol we would have raised these concerns before the revisions to the systematic review were implemented.

8. Exposure assessment is simplistic and inadequate

The extent of NTP's exposure assessment to support its Generalizability section appears to be a single footnote with a link to a CDC website [NTP 2020 monograph, p 2]. Furthermore, the linked website does not contain any of the exposure information stated in the footnote but instead gives a general description of a confidential database managed by the CDC. The database is only accessible to approved CDC staff and state oral health and drinking water officials, not the public. The confidential database is called the Water Fluoridation Reporting System or WFRS.
https://www.cdc.gov/fluoridation/data-tools/reporting-system.html

There is thus no transparency in the exposure assessment. The public summaries of the WFRS data (CDC's My Water's Fluoride website) are not sufficient to support a valid exposure assessment of the US population and have been found to contain serious limitations and errors such as reporting water systems having artificial fluoridation at 1.2 mg/L as being at 0.7 mg/L.

9. Proper exposure assessment demonstrates that the NTP's presumed hazard conclusion applies directly to doses from artificial fluoridation

Even if the exposure information in the footnote could be verified and was reliable, it is insufficient for a valid exposure assessment. It is a summary of drinking water fluoride concentrations and only for public water systems and only for naturally occurring levels. Valid exposure assessments require dose information which requires information on the amount of water consumed in addition to its concentration. As described in more details below, the EPA has found that the 95th percentile consumer of water drinks more than twice as much as the average consumer. That finding applies to all ages. Therefore, the top 5% of consumers (millions of people in the US) when drinking water with a concentration of 0.7 mg/L will receive the same doses as the average person in a study where the drinking water concentration is 1.5 mg/L. Thus, studies finding that average exposures to 1.5 mg/L cause neurotoxic harm directly support a conclusion that millions of people in the US with artificial fluoridation at a concentration 0.7 mg/L will be harmed. The NTP has made the fundamental error of conflating concentration with dose and not accounting for the wide range of doses that will occur for any given concentration of fluoride in drinking water.

10. Leading experts, historical perspective; analogy with childhood Pb poisoning

Recently, respected experts have noted the comparability of IQ loss from fluoride and from childhood lead poisoning. The most prominent amongst these experts is Dr. Linda Birnbaum, the director of NIEHS and NTP during much of the time the fluoride systematic review was being conducted. [Lanphear, Till, Birnbaum 2020; Bellinger 2019 editorial; Christakis & Rivera 2019 (podcast)]. These and other experts have also publicly stated that the evidence at exposures relevant to artificial fluoridation warrants recommending that pregnant mothers and young children avoid fluoridated water Lanphear, Till, Birnbaum 2020; Bashash & Hu 2020 commentary; Grandjean 2019 review; Christakis & Rivera 2019 (podcast)]. These expert opinions support the NTP's overall conclusion that fluoride is a presumed developmental neurotoxin but also go beyond the NTP for exposures relevant to the US.

Similarly to the history of recognizing low-level childhood lead (Pb) poisoning, the NTP appears to be trailing expert opinion. The NTP issued a monograph on low-level Pb toxicity in 2012 that relied, in large part, on a landmark study by Lanphear et al 2005 in which the authors said:

> "Lead is a confirmed neurotoxin, but questions remain about lead-associated intellectual deficits at blood lead levels < 10 μg/dL...."
>
> "We conclude that environmental lead exposure in children who have maximal blood lead levels < 7.5 μg/dL is associated with intellectual deficits."

The Lanphear 2005 paper concluded by recommending that blood Pb levels below 10 ug/dL be considered harmful and that efforts should be made to further reduce exposures to Pb [Lanphear 2005].

The earlier history of "low-level" Pb neurotoxicity research is also relevant, in particular because one of the co-authors of Lanphear 2005 was Dr. Herbert Needleman. Needleman did the groundbreaking studies finding neurotoxic harm at low levels (below what the CDC considered safe at the time) in the 1980s and 1990s. Needleman's work, and he personally, were attacked and criticized for decades before his work was fully accepted [Denworth 2018 Toxic Truth: A Scientist, a Doctor, and the Battle over Lead]. Similar arguments were used against Needleman and the Pb evidence as the NTP is now making to conclude the evidence of fluoride neurotoxicity is "inconsistent" and "unclear" at exposures relevant to the US.

Thirty years after Needleman's studies and 7 years after the Lanphear 2005 study, the NTP finally reached the same conclusion that Lanphear, Needleman and other experts had previously reached: that low-level Pb was a likely developmental neurotoxin [NTP 2012 monograph on low-level Pb].

The CDC took even longer to act on the scientific evidence, and did not lower their childhood blood Pb level of concern to 5 ug/dL until after the NTP 2012 report.

The NTP was at least 7 years behind the experts in the field of "low-level" lead neurotoxicity. The NTP appears to be dragging their feet on "low-level" fluoride as well.

We believe the existing scientific evidence, identified and rated for quality by NTP, is more than sufficient to conclude that fluoride exposures relevant to the US and to artificial fluoridation are a presumed hazard, or at the least, a suspected hazard. To support this conclusion, we have used the same data NTP has extracted from studies and have applied widely accepted methods of dose-response assessment and exposure assessment, rather than the *ad hoc* methods NTP used in their Generalizability section. Appendix 2 contains the details. We offer this as an example of what the NTP should have done to address the question of whether water fluoridation at 0.7 mg/L poses a risk of developmental neurotoxicity.

The NTP should either conduct such a valid risk assessment with attendant exposure assessment and dose-response assessment or should omit from the monograph the section on generalizability and any discussion of dose-response assessment of the overall body of evidence.

11. Specific examples of downgraded evidence in revised NTP monograph

- **Excluding largest effect in the strongest study.** For the Bashash 2017 study, which was one of the strongest studies, and at exposures relevant to artificial water

fluoridation, the NTP improperly focused on a minor secondary analysis and largely excluded consideration of the primary analyses, especially the primary analysis with the largest effect at the lowest exposure levels. The NTP focused on a comparison with the dichotomous exposure levels of <0.8 mg/L or ≥0.8 mg/L child urine F. There was only a small difference in IQ score between these two groups and it was not statistically significant. The small difference may be explained by the reduced information in the analysis and because child urine F was the exposure measure, not maternal urine F.

In contrast to this secondary analysis, the primary planned analyses of the study were largely excluded in narrative and meta-analysis portions of the NTP monograph. They were the multiple regression models between maternal urine F and GCI score for 4-year olds and WASI FSIQ score for 6-12-year olds. Both analyses found large statistically significant effects. The NTP further discriminated against the findings in GCI scores at 4 years old by treating it not as a measure of neurocognitive development but as an "other outcome". In meta-analyses and summaries of data the GCI analysis was thus excluded. Yet the GCI analysis found a linear dose-response relationship with no threshold, while the WASI FSIQ analysis found what may be a threshold at 0.8 mg/L. By excluding the GCI analysis, the NTP excluded the larger effect that occurred at lower doses. The GCI score should have been classified by NTP as a neurocognitive outcome rather than "other outcome". GCI is generally considered as a valid measure of neurocognitive development and has a strong correlation with several of the tests NTP did classify as tests of neurocognitive development. The abbreviation GCI stands for General Cognitive Index, which in itself should have helped NTP recognize it as a test of neurocognitive development.

• **Excluding strongest low-dose studies from dose-response meta-analysis.** The closest the revised NTP monograph gets to a proper dose-response assessment to support its generalizability section is a dose-response meta-analysis. There are several problems with it, however, the most serious being that they excluded the 10 studies with individual-participant data and instead relied on lower quality evidence from studies with only group-level analyses. The 10 individual-participant data studies included all of the highest quality studies and many of the studies at low doses, so this exclusion is especially problematical for conducting a valid and balanced dose-response assessment. It is important to note that this was a planned analysis in the revised protocol, so the decision to not conduct it is troubling. All other planned meta-analyses were conducted. Furthermore, the very brief reasons given for not conducting the analysis are unjustified and represent a double-standard. Here is what the NTP monograph said about the individual-participant dose-response meta-analysis:

> "A dose-response meta-analysis using the effect estimates reported in studies with individual-level exposure was considered. However, because of the small number of studies (n = 10), the various types of exposure metrics, and the different types of reported effect estimates that could not be combined, a dose-response meta-analysis of these studies could not be conducted." [NTP 2020 revised monograph p 253]

Taking each of the three stated reasons separately:

1.) “small number of studies (n = 10)” The claim that 10 studies are insufficient is contradicted by the NTP’s own actions elsewhere in the monograph. The NTP conducted dose-response meta-analyses on studies without individual-level data when there were as few as 4 studies and meta-analyses on subgroups with as few as 2 studies.

2.) “various types of exposure metrics” Just as with “small number of studies”, in other dose-response meta-analyses and meta-analyses the NTP has combined various types of exposure metrics. Elsewhere in the NTP monograph this issue is discussed and NTP concludes that for comparison purposes urine F levels can be considered equivalent to water F levels on a 1-to-1 ratio [NTP 2020 monograph p 72].

3.) “different types of reported effect estimates” Again, in other meta-analyses and dose-response meta-analyses the NTP combined studies with different effect measures. The revised NTP protocol considered a wide range of tests to be classifiable under the general domain “Leaning, Memory, Intelligence, Cognitive Development” [NTP 2020 protocol, Table 6, p 20]. Ten different specific tests were listed as examples that fit within this domain of outcomes.

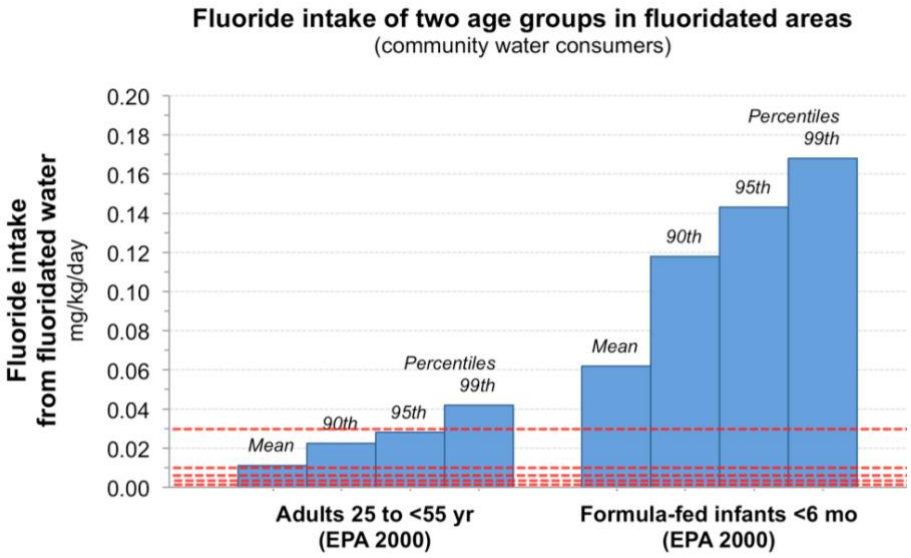
NTP has exhibited a clear double-standard when claiming they were unable to conduct a dose-response meta-analysis on the 10 studies with individual-level exposure data. Also, this represents a failure of the NTP to follow their revised protocol.

- **Unnecessary division of studies lowers power in dose-response meta-analyses.** The NTP’s dose-response meta-analysis of *group-level studies* unnecessarily stratified by whether exposure was measured in urine F or water F [NTP 2020 monograph Table A5-3, p 254]. This stratification reduced the statistical power and produced lower confidence in pooled results for each subgroup of studies.

- **Simplistic exposure assessment underestimates hazard at doses relevant to US.** The NTP’s simplistic exposure assessment assumed that only water F or urine F concentrations below 1.5 mg/L were applicable to the US. They further assumed that only water F or urine F concentrations of 0.7 mg/L are applicable to artificial fluoridation in the US. This perpetuates a fundamental error made by many fluoridation proponents that concentration is equivalent to dose. The US EPA has conducted rigorous exposure assessments of F from drinking water. They find that the 95th percentile of water consumers, on a mL per kg body weight basis, consume about twice as much water and fluoride as the average consumer (see Figure 6). Therefore, the 95th percentile consumers drinking water with a concentration of 0.7 mg/L are receiving double the dose of the average water consumer. They thus receive the same dose as the average water consumer drinking water with a concentration of 1.5 mg/L. The consequence is that studies finding harm at water concentrations of 1.5 mg/L are relevant to the top 95th percentile water consumers. This subpopulation represents millions of people in the US

and must be considered when generalizing from the results of epidemiological studies to the actual exposures in the US. This realistic exposure assessment alone is sufficient for the NTP to conclude that artificially fluoridated water at 0.7 mg/L is a presumed developmental neurotoxin for the 5% of the US population who consume the most water. This realistic exposure assessment also greatly expands the number of studies which should be considered relevant to exposures in the US and from artificial water fluoridation. Instead of 1.5 mg/L as the cut-off for relevance, the level should be 3.0 mg/L. Using this more appropriate cut-off, the NTP's dose-response meta-analysis shows that the group-level studies with water F as the exposure measure already show a pooled estimate that is statistically significant and large (SMD -0.27 equivalent to -4 IQ points for studies with mean water F below 2 mg/L) [NTP 2020 monograph Table A5-3, p 254].

Figure 6. Distribution of fluoride intake from fluoridated water, USA



Intakes based on EPA 2000 Water Intake Estimates in US, p. IV-3.

• **NTP's simplistic dose-response meta-analysis methods underestimated effects at low doses because they used the mean exposure while most studies had individual-level exposures that ranged well below the mean.** Furthermore, NTP dichotomized studies by whether the mean exposure was above or below a cut-off of 1.5 mg/L [NTP 2020 monograph Table A5-3, p 254]. The loss of information in taking the mean and then dichotomizing by the mean value is contrary to standard dose-response assessment methodology. For example, the EPA currently prefers Benchmark Dose (BMD) methods be used for dose-response assessment. BMD methods account for the totality of the data and provide estimates of the dose likely to

cause a specified degree of harm. That dose, called the BMD, is frequently lower than the mean dose.

To illustrate the difference between the information available in the complete individual-level data and just the mean or the dichotomized mean, we use simulated data from a hypothetical study, first plotted as a scattergram with all the data points (Figure 7) and then as a single data point showing the mean dose and mean response (Figure 8). These illustrations are then followed by the results of a BMD analysis of the same individual-level simulated data (Figure 9).

The mean exposure in this hypothetical study is 1.7 mg/L. That puts it over NTP's cut-off for relevance to exposures in the US. Yet the full exposure distribution and dose-response relationship as shown in the scattergram clearly shows it is relevant to exposures below 1.5 mg/L. This illustrates why NTP's informal dose-response analysis and generalizability discussion are invalid and will underestimate the confidence that exposures in the US and from artificial fluoridation will produce harm.

Figure 7. Hypothetical individual-level data:

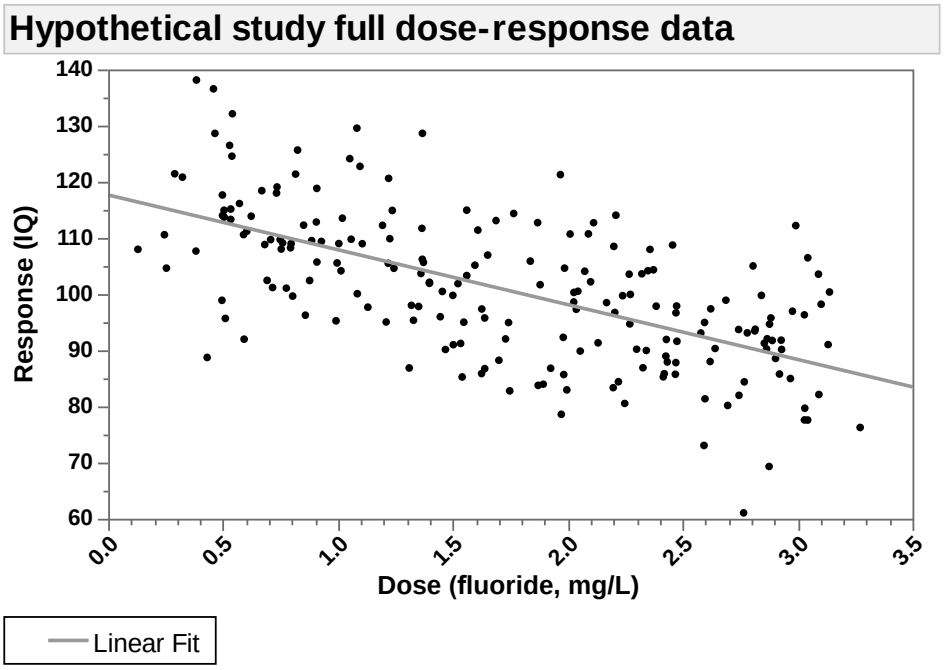


Figure 8. Here is what NTP has reduced it to:

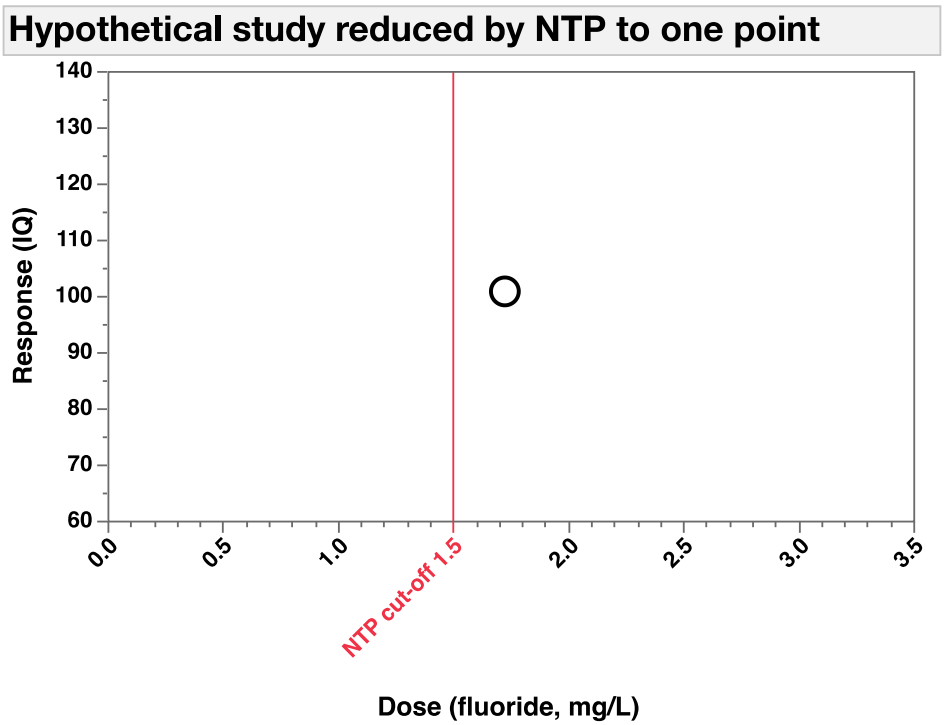
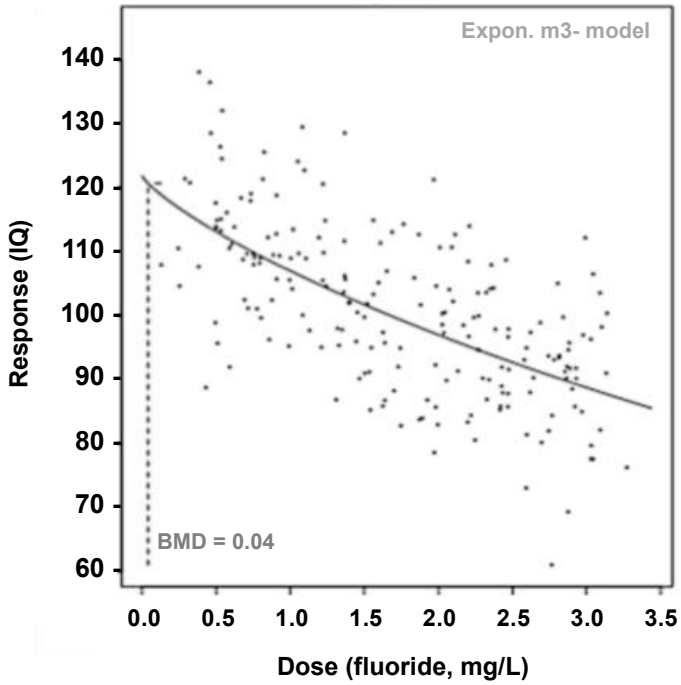


Figure 9. Hypothetical study Benchmark Dose (BMD) analysis:



Best fit model Exponential m3- using PROAST website of EFSA; BMR = -1 IQ point.

Benchmark Dose (BMD) analysis uses all available data from a study to estimate at what dose an adverse effect is predicted to occur. It was developed as an improvement over other methods of dose-response assessment and is EPA's preferred method when suitable data is available. This BMD analysis considered non-linear dose-response models.

This hypothetical study example, that NTP would have classified as too high a dose to be relevant to exposures below 1.5 mg/L, is found to provide clear evidence of an adverse effect well below 1.5 mg/L when analyzed with the BMD method.

12. Additional weaknesses of the revised NTP monograph

• **Improperly downgraded the animal evidence to "inadequate" despite the NTP 2016 review of the animal evidence concluding it was "low to moderate".** The

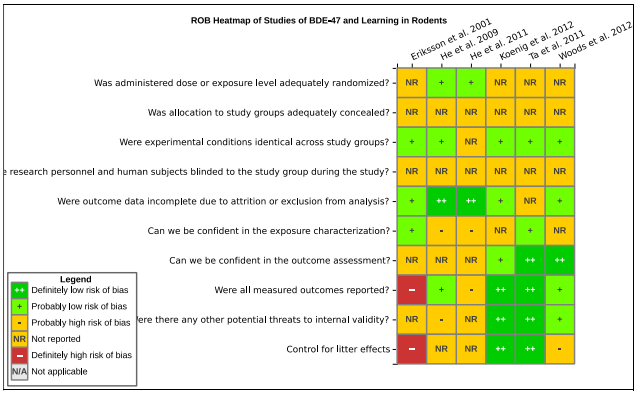
NAS committee specifically chastised NTP for improperly downgrading the evidence based on the claim that sensory/motor effects might have played a role in the deficits in the learning and memory tests. Despite this clear rejection of that argument the NTP persisted and continues to use it in the revised monograph to downgrade the animal evidence. Furthermore, additional animal studies were identified since the NTP 2016 review, several of which were scored high quality. Therefore, it is difficult to understand how the revised NTP monograph can downgrade the overall body of animal evidence to "inadequate".

The NAS also raised concern in the other direction about some of the animal studies which the NAS suggested should have higher RoB scores. The concerns were mostly because of deficiencies in reporting, such as not reporting whether researchers were blinded to exposure status and whether litter effects had been controlled.

However, just as comparison to other OHAT systematic reviews for other chemicals provides perspective for the strength of evidence necessary to reach overall confidence conclusions, it is appropriate to consider other OHAT reviews and how they scored individual animal studies for RoB and how they assessed the overall animal evidence for confidence. The same NAS review used as an example above, for BDE-47, and employing OHAT methodology, demonstrates that the OHAT methodology applied to much weaker animal evidence than is available for fluoride was sufficient to give an animal evidence confidence rating of "moderate".

Figure 10 shows the total extent of animal studies of BDE-47 upon which a "moderate" confidence rating was concluded for effects on learning. Of the 6 animal studies, 5 found "some indication of effect on at least one measure of learning" for a consistency of 83% [NAS 2017, p 8]. But the quality of all 6 studies was low.

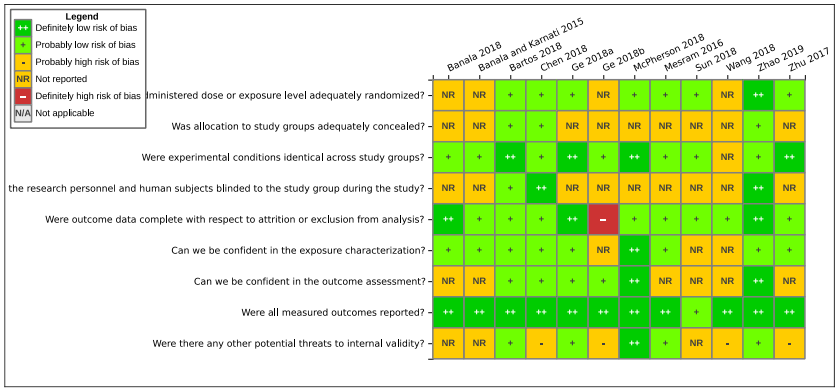
Figure 10. BDE-47 RoB heatmap, all animal studies



from NAS 2017 report: https://doi.org/10.17226/24758
HAWC visualization: https://hawcproject.org/summary/visual/353/

For comparison, Figure 11 is the RoB heatmap for just the newer fluoride animal studies.

Figure 11. Fluoride RoB heatmap, recent animal studies

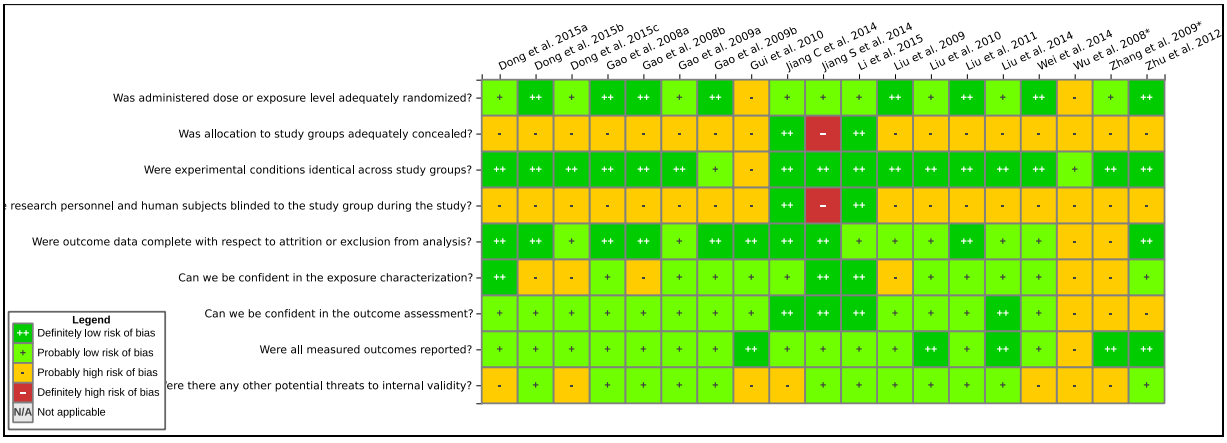


HAWC visualization: https://hawcproject.org/summary/visual/530/

Fluoride has 12 studies to just 6 for BDE-47. Five of the 12 fluoride studies would be rated lower RoB while none of the BDE-47 studies would rate lower RoB. The NTP protocol requires that to be rated "lower RoB" no more than two of the key RoB domains be rated "yellow" or "red".

These 12 animal studies are just those identified since the NTP 2016 systematic review of fluoride neurotoxicity in animals. The NTP 2016 review identified 19 additional earlier studies that used Morris Water Maze tests, considered the most applicable to learning and memory. Figure 11 is the RoB heatmap for these earlier studies.

Figure 12. Fluoride RoB heatmap, animal studies from NTP 2016



HAWC visualization: https://hawcproject.org/summary/visual/33/

Of the 19 additional animal studies, 8 would be rated as lower RoB using NTP's criteria. For fluoride the total number of animal studies is 31 with 13 rated lower RoB. The total number of BDE-47 is just 6 with none rated lower RoB.

Thus, the fluoride animal evidence is substantially greater in quantity and quality than BDE-47, yet NTP has rated confidence in fluoride "inadequate" while BDE-47 is rated "moderate".

With regards to consistency, almost all of the fluoride studies of learning and memory found statistically significant adverse effects. Even the much-touted McPherson 2018 study found a statistically significant adverse effect that was not acknowledged in its summary or abstract. Therefore, on consistency, fluoride is stronger than BDE-47 as well.

Finally, the doses (as measured in body tissues) found to cause adverse effects in the BDE-47 animal studies were hundreds of times higher than commonly occur in humans [Staskal 2007, EPA 2008]. In contrast, the doses of fluoride found to cause adverse effects in most of the animal studies were less than 20 times higher than commonly occur in humans, when taking into account pharmacokinetic differences. The OHAT Handbook requires conversion between external and internal dosimetry in assessing relevance of animal studies to human exposures:

> "*Exposure*
> • *Route of administration in animal studies:* External dose comparisons used to reach level of concern conclusions need to consider internal dosimetry in animal models, which can vary based on route of administration, species, age, diet, and other cofactors." [NTP 2019 OHAT Handbook, p 58]

There is no justification for downgrading the fluoride animal data because of claimed irrelevant doses to humans, as the revised NTP monograph does [NTP 2020 monograph, p 58].

In conclusion, both NTP and the NAS are applying an extreme double-standard, with fluoride having to meet a much higher bar than BDE-47. Despite much greater quantity, quality, and consistency of the animal evidence for fluoride, NTP has rated it "inadequate" and BDE-47 "moderate". This violates an overarching goal of OHAT systematic reviews, which is to consistently apply the same standards for judging the hazard of different chemicals. The OHAT handbook says an objective of the OHAT evaluation process is to "ensure consistency across evaluations" [NTP 2019 OHAT Handbook, p 1]. BDE-47 was used as an example here, but similar comparisons can be made to many other chemicals the NTP concluded had "moderate" animal evidence.

• **The NTP monograph deviated from the OHAT guidelines in its section "Generalizability to the U.S. Population".** While OHAT methodology has little guidance on dose-response analysis, the revised monograph did not even follow what is

available. The Generalizability section is essentially a dose-response and risk assessment evaluation although it did not follow OHAT methodology for such. OHAT methodology for dose-response and risk assessment is termed “Level of Concern Conclusions” (LoC) and is considered as a second type of conclusion beyond the initial “Hazard Identification Conclusion” [NTP 2019 OHAT handbook, p 3]. But the protocol for the NTP review does not include any Level of Concern assessment so there was no pre-specification of the Generalizability section. As stated above, the Generalizability section in the NTP monograph is not adequately documented or justified, even as a *post hoc* addition to the monograph. We believe this is a serious failure to follow the principles of transparency and pre-specification.

> “The National Toxicology Program (NTP) ... conducts evaluations to assess the evidence that environmental chemicals, physical substances, or mixtures (collectively referred to as "substances") cause adverse health effects and provides opinions on whether these substances may be of concern, given what is known about current human exposure levels. The opinions are referred to as NTP Level of Concern (LoC) conclusions.” [NTP 2019 OHAT Handbook, p 1]

OHAT says LoC comes after Hazard Identification and requires an exposure assessment:

> “*Exposure*
> • *Human studies:* ... In OHAT’s process, the applicability of a given exposure scenario for reaching a “level of concern” for a certain subpopulation is considered after hazard identification. For that subpopulation the health effect is interpreted in the context of what is known regarding the extent and nature of human exposure (Twombly 1998, Medlin 2003, Jahnke *et al.* 2005, Shelby 2005).” [NTP 2019 OHAT Handbook, p 58]

While OHAT offers little guidance on how NTP will conduct LoC determinations, it makes clear that an exposure assessment is required. The OHAT Handbook says NTP will “update” its LoC framework “to ensure integrated consideration of relevant and reliable evidence and to enhance transparency”. The update is projected for completion in 2016-2017 but apparently has not yet been issued. Here is the currently available extent of guidance on LoC determinations:

> • ***Level of Concern (LoC) Conclusions*** – For LoC conclusions OHAT integrates two categories of evidence: (1) health-outcome data from human, animal, and mechanistic studies to reach hazard identification conclusions and (2) information on the extent of exposure and pharmacokinetics. LoC conclusions are narrative (i.e., non-quantitative) conclusions that use a 5-point scale ranging from “negligible” to “serious” concern for exposure. As part of implementing systematic reviews the NTP will update its LoC framework to ensure integrated consideration of relevant and reliable evidence and to enhance transparency in describing how these conclusions are reached. These strategies will improve the LoC framework as a risk communication tool (expected completion in 2016-2017). The updated LoC framework will be included in a future version of the OHAT handbook.” [NTP 2019 OHAT Handbook, p 3]

The OHAT Handbook specifies that the decision for whether NTP will conduct just a Hazard Assessment or also a Level of Concern determination should be made at the problem formulation stage, before a protocol is even written [NTP 2019 OHAT Handbook, p 10]. Neither the problem formulation nor protocol for NTP's review of fluoride neurotoxicity have ever mentioned a LoC determination as an objective.

- **Meta-analyses have inadequate documentation.** The meta-analyses are not adequately documented, especially the dose-response meta-analyses. The specific studies included at each dose should be provided in a table. Forest Plots should be provided for all dose-response meta-analyses. Bubble plots showing the dose-response curve with 95%CI for the dose-response meta-analyses should be provided.

No data underlying the meta-analyses are available at the HAWC project website, nor are any visualizations like Forest Plots available through HAWC. There is a downloadable Excel file named "meta-analysis data" but it is only column headings with no data. All data used in meta-analyses and dose-response meta-analyses should be provided in HAWC in downloadable data files as well as tables and visualizations.

13. NTP should consider publication bias due to institutional conflicts of interest

A primary recommendation the NAS committee made regarding the NTP literature search concerned the possibility of publication bias introduced by use of English translations of foreign-language papers supplied by FAN:

> "The committee acknowledges FAN's efforts in providing several studies that appear to be relevant for the review. However, the process by which FAN identified and selected studies is not clear. FAN identified a number of studies published in Chinese language journals—some of which are not in PubMed or other commonly used databases—and translated them into English. That process might have led to a biased selection of studies and raises the question of whether it is possible that there are a number of other articles in the Chinese literature that FAN did not translate and about which NTP is unaware. ... The committee emphasizes that its comments regarding FAN are aimed only at evaluating bias; they are not intended to discourage stakeholder input into the systematic-review process, and the committee acknowledges and encourages the important contributions of FAN and other stakeholder organizations in this process." [NAS 2019 p 33]

FAN welcomed the NAS recommendations for assessing this potential form of bias and assisted NTP in identifying sources for the original Chinese-language papers and describing FAN's literature search process. The revised NTP monograph implemented the NAS recommendations for assessing potential publication bias and concluded there was no evidence of such bias. The revised monograph went beyond the NAS recommendations and also assessed the accuracy of the translations provided by FAN by commissioning their own translations of several of the papers. The NTP found the FAN translations to be accurate.

Given the concern about potential bias from FAN's input to the NTP systematic review, it should be pointed out that there is an "elephant in the room" that neither the NAS nor the NTP have acknowledged. That is the conflict of interest represented by institutions that promote artificial fluoridation. The US Public Health Service and the CDC Oral Health Division have long-standing policies promoting fluoridation, as do equivalent agencies in other countries with artificial fluoridation. Alongside government agencies, the other institution that promotes artificial fluoridation is dentistry, including through such organizations as the American Dental Association (ADA) and the International Association of Dental Research (IADR), both of which have long-standing policies promoting fluoridation. There may also be financial conflicts of interest from industry, although they are likely not through direct funding of neurotoxicity studies. Marketers of fluoride dental products, however, often fund dental schools and dental research.

Because these conflicts are usually not direct financial conflicts, they are unlikely to be identified through publication conflict of interest (CoI) declarations. Yet they are potentially very powerful. They can be described as reputational conflicts of interest. For toxic chemicals for which there are direct financial conflicts of interest there is a long record of industry manipulation of the science by funding academic researchers, published reviews, and original research which downplay or sow doubt about the risks of chemicals [Michaels 2008, book: Doubt is Their Product]. Non-governmental environmental health organizations such as FAN do not have the resources to influence the scientific record in these ways, but institutions with conflicts of interest do have substantial resources and power.

These institutional conflicts of interest over fluoride neurotoxicity are exemplified by the emphasis put on the Broadbent 2015 human study by many promoters of fluoridation. The Broadbent 2015 study is one of the few that have not found an adverse effect. The NTP monograph has rated this study at higher RoB, we believe justifiably, but because of its study design, not because of any acknowledgement of its authors' reputational conflicts of interest.

We propose that NTP should evaluate the possibility of publication bias from reputational conflicts of interest by considering author affiliation to fluoridation promoting institutions, especially the profession of dentistry. We propose an explicit criteria to identify studies with such potential conflicts of interest: those studies where the first author or the last author (or the senior author if explicitly designated), or at least 50% of authors have dental degrees or are affiliated with a dental institution.

The Broadbent 2015 paper meets this criterion because the lead author was associated with a dental school as were 3 out of a total of 7 authors.

Closer examination of the Broadbent 2015 study reveals clear evidence of author bias. The opening paragraph of the paper is an advocacy statement defending fluoridation and asserting that it has been proven safe:

> "Community water fluoridation (CWF) is a cost-effective,[1,2] safe,[3] and environmentally friendly[4] means of reducing dental caries rates[3] and social inequalities.[5] However, CWF has recently been criticized as a cause of IQ deficits among children,[6] despite a lack of evidence to support that claim. This claim was considered pivotal in the recent rejection of CWF by voters in Portland, Oregon,[7] and by local government politicians in Hamilton, New Zealand. It is likely that such claims may continue to be lobbied against CWF worldwide." [Broadbent 2015]

The paper goes on to discuss details of political debates about fluoridation, taking sides with those defending fluoridation.

Focusing on just the studies rated as higher quality by NTP, which excludes Broadbent 2015, there are 4 out of 29 studies that meet the criteria of having potential reputational conflicts of interest due to author affiliation with dental institutions:

> Saxena 2012 - 2 out of 3 total authors have dental school affiliation, including first & last authors; India.
>
> Seraj 2012 - 4 out of 7 total authors have dental school affiliation, including first and last; Iran.
>
> Soto-Barreras 2018 - 6 out of 7 total authors have dental school affiliation, including the first author; Mexico.
>
> Sudhir 2009 - 2 out of 4 total authors have dental school affiliation, including last author who is probably senior author; India.

None of the other higher quality studies meet the criteria for having potential dental conflict of interest.

Of these 4 higher quality studies with potential institutional dental conflict of interest, 1 found "no effect" and the other 3 found statistically significant adverse effects for a 25% rate of finding "no effect". Amongst the 25 higher quality studies with no potential institutional conflict of interest only 1 found "no effect", a rate of 4%. Amongst the higher quality studies, there was thus a 6-fold greater rate of finding "no effect" for those with potential dental conflicts of interest.

Another refinement of institutional conflict of interest assessment is possible by distinguishing those papers with dental authors from countries that have national policies of promoting artificial fluoridation versus countries that do not. Of the 4 higher quality studies, Soto-Barreras 2018 authors are in Mexico which has a policy of promoting salt fluoridation. The other three papers with dental authors are from India and Iran, countries without such a policy. Only the Soto-Barreras 2018 paper found "no effect", for a 100% rate amongst those with dentist authors in countries with a policy of promoting fluoridation.

It should be pointed out that the Broadbent 2015 dental authors are also in a country with a national policy of promoting fluoridation, New Zealand.

We believe the NTP monograph should incorporate recognition of the potential conflict of interest posed by dental authors, especially from countries with national policies of promoting fluoridation. There appears to be a clear risk of publication bias amongst these studies. Without such recognition, the NTP would be engaged in a double-standard because it has only assessed FAN's translations as a potential source of publication bias.

14. "Presumed hazard" is *de facto* highest confidence rating possible under OHAT method

Of all chemicals assessed by NTP under OHAT guidelines [https://ntp.niehs.nih.gov/publications/monographs/index.html], only one was given a confidence rating of "known hazard". That was the chemical warfare agent sarin for the hazard of neurotoxicity [NTP 2019 sarin monograph]. The NTP noted that reaching the conclusion of "known hazard" was based largely on the existence of controlled human trials. Aside from the surprise that this incredibly deadly chemical should even need an NTP review to be declared a "known hazard" to humans, this also indicates the near impossibility for most environmental chemicals to reach a "known hazard" rating from NTP, no matter how strong the human observational study evidence. The NTP monograph should acknowledge that "presumed hazard" is the highest confidence rating that has been achieved by any chemical without controlled human trials evidence.

At last year's NAS committee open session, committee members asked NTP whether it was possible for a chemical to achieve "known hazard" rating without human experimental studies and the NTP said it was possible. However, in practice it appears to require an unreasonably high bar.

15. Literature search update missed important recent study on adolescents

A study by Malin et al 2019 finding that fluoride exposure in adolescents was associated with disruptions in sleep patterns was not identified in the literature update [Malin 2019]. Apparently NTP's search criteria did not recognize sleep disruption as a form of developmental neurotoxicity. However, sleep is largely mediated by neurological functions and can impact neurological and psychological wellbeing in ways that may not be measured by intelligence tests. The Malin 2019 study is relevant because it was done with a NHANES sample of children age 16-19 years, that is nationally representative of the US. Exposure was measured through individual-level tap water samples. The study found a statistically significant doubling of odds of sleep-apnea symptoms for an increase of 0.5 mg/L in water F concentration. The authors suggested that fluoride may affect the pineal gland and melatonin production for which there is

some animal study evidence. As a neuroendocrine organ in the brain, adverse effects on the pineal gland should be considered neurotoxic effects.

While this is the first study to ever examine sleep patterns in relationship to fluoride, it opens the possibility that fluoride neurotoxic effects might extend beyond prenatal and earlier childhood to adolescence. This could enlarge the portion of the population subject to neurotoxic harm from fluoride.

16. Availability of NHANES data for fluoride studies in US: urine samples unanalyzed and dental fluorosis data withheld

The Malin 2019 study is also important because it is one of the first studies to investigate fluoride neurotoxicity in the US, and the first to use NHANES data. For future work, the NTP monograph should note that for survey cycles 2013-2014 and 2015-2016 tap water fluoride and plasma fluoride concentrations of children were measured and are publicly available.

However, CDC is apparently not aware that NHANES also collected urine samples for fluoride determination in its 2015-2016 cycle [TSCA lawsuit deposition of CDC representative; NHANES 2016 procedures manual]. This is the first time urine F samples have been systematically collected in the US despite 75 years of water fluoridation. NHANES has not yet released any data on the fluoride levels in the urine samples and it appears there may not be any plans to actually do fluoride analyses on these samples. They were sent to a CDC long-term sample storage facility rather than an analytical laboratory. The NTP monograph should recommend that CDC measure the fluoride concentrations in the samples and release the data.

NHANES also collected dental fluorosis data using a validated fluorosis imaging system method in the 2013-2014 and 2015-2016 cycles [NHANES 2013 Oral Health Examiners Manual, Dye 2018]. NHANES has not released this data publicly. NHANES has released clinically assessed fluorosis data but then issued a caution that they were concerned about its validity. Therefore, the more objectively determined fluorosis imaging system data is especially important. Dental fluorosis is a biomarker of early life fluoride exposure that could allow more accurate estimates of exposures during early life periods than contemporary water F, plasma F, or urine F measures. Evidence points to the early life period as most vulnerable to fluoride developmental neurotoxicity. The NTP monograph should recommend that CDC release the dental fluorosis data.

17. Conclusion

The scientific evidence is more than sufficient to justify a conclusion that fluoride is a presumed developmental neurotoxin in children. Compared to other chemicals

reviewed by NTP and given a “presumed hazard” rating, there is greater quantity, quality, and consistency in the fluoride human studies.

It is extremely unlikely there could be any unidentified studies that could alter this conclusion. Likewise, the quantity and consistency of evidence mean it is extremely unlikely that any new studies could weaken this conclusion.

NTP’s revisions have addressed NAS’s recommendations and the revised monograph is substantially strengthened and more transparent as a result.

However, the Generalization section is weak and should be removed as recommended by NAS or redone using valid risk assessment methods. A valid risk assessment requires a valid exposure assessment and valid dose-response assessment. FAN has offered a risk assessment using methodologically rigorous and appropriate methods following EPA guidance. Our risk assessment finds that exposures to levels below 1.5 mg/L and below 0.7 mg/L both pose a high likelihood of neurotoxic harm to at least some proportion of children in the US population. Even without a formal risk assessment, recognition that about 5% of the population will receive twice the average dose because of greater than average water consumption provides sufficient support for this conclusion, when coupled with the NTP conclusion of presumed hazard above 1.5 mg/L.

APPENDIX #1

ISEE 2020 Conference poster and abstract [Neurath et al 2020].

APPENDIX #2

Rigorous risk assessment including exposure assessment and dose-response assessment applicable to doses due to artificial water fluoridation in the US. The risk assessment will include proper dose-response meta-analysis of both group-level exposure and individual-level exposure studies.

APPENDIX #3

FAN comments to first NAS committee meeting October 2019

1. Summary followed by additional details
2. Figures to accompany FAN comments