April 28, 2023

Kathleen M. Gray, Ph.D.
Chair, NTP Board of Scientific Counselors
c/o Office of Policy, Review, and Outreach
Division of Translational Toxicology
National Institute of Environmental Health Sciences
P.O. Box 12233
Research Triangle Park, NC 27709

Re: NTP Monograph on the State of the Science Concerning Fluoride Exposure and Neurodevelopmental and Cognitive Health Effects: A Systematic Review

Dear Dr. Gray:

On behalf of our 159,000 members, we would like to share our recommendations for improving the scientific integrity, clarity, transparency, and timeliness of the National Toxicology Program’s third (and purportedly final) draft report, titled Draft NTP Monograph on the State of the Science Concerning Fluoride Exposure and Neurodevelopmental and Cognitive Health Effects: A Systematic Review.¹ We offer these comments in response to your Federal Register notice of March 31, 2023 (88 FR 19315).

This systematic review, which has been underway for several years, is intended to summarize the literature about a relationship (if any) between fluoride exposure and neurodevelopmental and cognitive health. The original report has been revised several times, which is a common practice for peer-reviewed papers.

NTP Director Rick Woychik has tasked the NTP Board of Scientific Counselors to determine whether NTP adequately addressed outside questions and criticisms of its methods, conclusions, clarity, and transparency, including those raised by the National Academies of Sciences, Engineering and Medicine.² ³ NASEM is the acknowledged gold standard for providing independent and objective advice on complex scientific issues.

On May 4, the BSC will hold a virtual meeting to discuss whether and how the draft report should move forward, based on the findings of a BSC Working Group⁴ and comments from the public.

We urge the BSC to adopt the following recommendations to improve the report’s scientific integrity, clarity, transparency, and timeliness—and to support the lay public’s scientific literacy and that of local elected officials who determine community water fluoridation policies. Our reasons, which are detailed in several attachments, are consistent with those expressed by the NASEM peer review committee.*

* NTP took the unusual step of abandoning its course of peer review with NASEM after the acknowledged gold standard peer panel twice reported that major revisions were needed for the report to survive scientific scrutiny. Instead, NTP hand-picked its own panel to peer review the third (and purportedly final) draft now being discussed. Changing peer reviewers is not a standard practice.
1. NTP should complete the full course of peer review with its original independent peer review panel, NASEM. (NASEM is the acknowledged gold standard for peer review.)

2. NTP should update and publish its meta-analysis of epidemiological studies, but only after it survives peer review by NASEM or been accepted for publication in a reputable journal, as the BSC Working Group recommended.

3. NTP should provide clear context for statements about low-level fluoride exposures, as NASEM recommended.2,3

4. NTP should include a stand-alone disclaimer indicating that the report should not be construed as an indictment of low-level fluoride exposures, as NASEM recommended.2,3

5. NTP should revise its risk of bias rating for several studies, based on NASEM’s concerns2,3 and the enclosed analysis.

We call special attention to NTP’s apparent aversion to adding a clear, strongly worded disclaimer about the report’s treatment of low-level fluoride exposures (<1.5 mg/L), including concentrations recommended for community water fluoridation (0.7 mg/L).

On two occasions, the NASEM committee urged NTP to “state clearly”, “[reiterate] at the end”, and “make it clear that the monograph cannot be used to draw any conclusions regarding low fluoride exposure concentrations…” The acknowledged gold standard peer review panel also made clear that statements about low-level fluoride exposures were “outside its scope” and that “comments or inferences that are not based on rigorous analyses should be avoided…”2,3

NTP ignored that recommendation. Instead, the latest version is full of non-contextualized statements about “potential associations”1 and the evidence being “unclear.”1 In one area, NTP even states, “[L]ower concentrations of fluoride may support reduced IQ in humans” without offering any data or context to support its claim.

Statements suggesting “more studies are needed”1 are technically accurate. But without context, the lay reader might conclude the lack of evidence justifies a precautionary approach to community water fluoridation, which the report does not validate.

These non-contextualized statements can easily be misconstrued. In fact, they may be indicative of a desire to retain in some form the blanket hazard assessment that appeared in the first two drafts and was eventually removed after the NASEM committee (2021) determined, “[T]he monograph falls short of providing a clear and convincing argument that supports its assessment.”3

We strongly urge that all references to “hazard conclusions found in previous draft monographs”1 be accompanied by a clear follow-up statement indicating why NASEM recommended that those hazard conclusions be withdrawn.

The BSC is now in the fortunate position of knowing how current version of this report will be used. For example, on March 15—the day the current draft was made public—anti-fluoridation activists5 issued a press release claiming that NTP “could not detect any safe exposure, including at levels common from drinking artificially fluoridated water.”6 The press release further claimed, “There is now little question that a large body of scientific evidence supports a
conclusion that fluoride can lower child’s IQ, including at exposure levels from fluoridated water.”

The non-contextualized language that inspired this rhetoric is exactly why the NASEM committee (2021) urged NTP to “make it clear that the monograph cannot be used to draw any conclusions regarding low fluoride exposure concentrations…”3

In the interest of supporting the scientific literacy of both the lay public and local elected officials who determine community water fluoridation policies, it is critical that the state of the science report, the meta-analysis, and any corresponding press releases or public statements heed NASEM’s recommendation to include a disclaimer about low-level fluoride concentrations (<1.5 mg/L). A disclaimer akin to the following would provide context for the report and help prevent the findings from being misconstrued or mischaracterized.

This state of the science report should not be construed as an indictment of consistent low-level fluoride exposures (<1.5 mg/L), including concentrations recommended for community water fluoridation (0.7 mg/L). Community water fluoridation is the purposeful upward adjustment of the naturally occurring fluoride content in water to levels recommended by the United States Public Health Service (0.7 mg/L) to prevent tooth decay.7

The report should also not be used to draw conclusions about the fluoride content of toothpaste, fluoride supplements, or any other dental treatments.

An examination of the literature on low-level fluoride exposures did not validate the hypothesis that consistent exposure to low levels of fluoride (<1.5 mg/L) poses a risk to neurodevelopmental and cognitive health. Additional research may inform that point.

We strongly urge that a disclaimer of this kind be include in the science report, the meta-analysis, and any corresponding press releases or public statements.

A clear, strongly worded disclaimer of this kind would address the NASEM committee’s criticism about the report’s lack of context2,3 and “lack of details in several places and the lack of clarity on several substantive issues.”3 It would also comport with the recommendations of the White House Task Force on Scientific Integrity, which called for federal agencies to adopt better methods of communicating scientific findings to ensure lay audiences have an accurate understanding of science.8,9

At a time when the public’s trust in federal research is declining,10 failing to include a disclaimer about low-level fluoride concentrations (and address other shortcomings) could determine whether the public’s health will be driven by science…or by unanswered rhetoric.

The Centers for Disease Control and Prevention hailed community water fluoridation as one of ten great public health achievements of the 20th century.11,12 It is a safe and inexpensive way to reduce tooth decay by at least 25 percent in the population.13 It would be a shame to distract from over 75 years of public health success over a simple matter of communicating the science, which is often more nuanced than a sound bite can convey.

Thank you for providing us the opportunity to comment. If you have any questions, please contact Mr. Robert J. Burns.
References


ATTACHMENT A
ANALYSIS AND RECOMMENDATIONS

NTP MONOGRAPH ON THE STATE OF THE SCIENCE CONCERNING FLUORIDE EXPOSURE AND NEURODEVELOPMENTAL AND COGNITIVE HEALTH EFFECTS: A SYSTEMATIC REVIEW

NTP MONOGRAPH 08
April 28, 2023

In its third (and purportedly final) draft state-of-the-science report, NTP states, “This review finds, with moderate confidence, that higher fluoride exposure (e.g., represented by populations whose total fluoride exposure approximates or exceeds the World Health Organization Guidelines for Drinking-water Quality of 1.5 mg/L of fluoride) is consistently associated with lower IQ in children.”

NTP’s claim for a potential association is based on 19 publications it judged to have a low risk-of-bias. The National Academies of Sciences, Engineering and Medicine peer review committee (2021) expressed serious concerns about whether NTP’s risk-of-bias methodology was sound and had been consistently applied. There is no indication that those concerns have been resolved to NASEM’s satisfaction.

At a time when the public’s trust in federal research is declining, the BSC might consider whether the evaluation methods, clarity, transparency, and timeliness of NTP’s report—and the piecemeal way NTP plans to release it—will support the lay public’s scientific literacy and that of local elected officials who determine community water fluoridation policies. The answer could determine whether public policy will be driven by science…or by unanswered rhetoric.

We therefore ask the BSC to issue the following recommendations to improve the report’s scientific integrity, clarity, transparency, and timeliness.

1. **NTP should complete the full course of peer review with its original independent peer review panel, NASEM. (NASEM is the acknowledged gold standard for peer review.)**

2. **NTP should update and publish its meta-analysis of epidemiological studies, but only after it survives peer review by NASEM or been accepted for publication in a reputable journal, as the BSG Working Group recommended.**

3. **NTP should provide clear context for statements about low-level fluoride exposures, as NASEM recommended.**

4. **NTP should include a stand-alone disclaimer indicating that the report should not be construed as an indictment of low-level fluoride exposures, as NASEM recommended.**

---

* NTP took the unusual step of abandoning its course of peer review with the NASEM committee after the acknowledged gold standard peer reviewer twice reported that major revisions were needed for the report to survive scientific scrutiny. Instead, NTP hand-picked its own panel to peer review the third (and purportedly final) draft now being discussed. Changing peer reviewers is not a standard practice.
5. **NTP should revise its risk-of-bias ratings for several studies, based on NASEM’s concerns and the enclosed analysis.**

1. **NTP should complete the full course of peer review with its original independent peer review panel, NASEM. (NASEM is the acknowledged gold standard for peer review.)**

At a time when the public’s trust in federal research is declining, the public deserves to know that its research agencies are not arbitrarily changing peer reviewers when the results are not to their liking.

NTP began its work by asking NASEM to serve as the report’s independent peer reviewer. NASEM is the acknowledged gold standard for providing independent and objective advice on complex scientific issues.

NTP then took the unusual step of abandoning its course of peer review with NASEM after the acknowledged gold standard peer review organization twice reported that major revisions were needed for the report to survive scientific scrutiny. Instead, NTP hand-picked its own panel to peer review the third (and purportedly final) draft now being discussed. Arbitrarily changing peer reviewers when the results are not to the authors’ liking is not a standard practice.

In the third (and purportedly final) draft, NTP implies it has resolved NASEM’s major concerns by stating, “NTP has responded to the NASEM committee’s comments on the revised draft (September 16, 2020) in a separate document (placeholder for URL) and revised relevant sections of this monograph.” We question NTP’s credibility to presume its response to the NASEM committee is adequate.

Replacing a gold standard peer review committee with a hand-picked group of reviewers is not a standard practice. It is not consistent with the spirit of an independent peer review and the action has not been adequately explained in the report. It raises legitimate questions about the report’s scientific integrity, as well as NTP’s motivations.

**NTP should complete the full course of peer review with its original peer review panel, NASEM.**

*If NTP is intent on proceeding without the full course of peer review without NASEM, NTP needs to immediately release or include a summary of its “separate document” to NASEM in the report itself, which would include an explanation for the “revised relevant sections of this monograph”. It is a relatively simple task that would prevent the reader from having to go back and forth between documents to determine what was done and why. Moreover, it would add a level of transparency that is lacking.*

Again, the public deserves to know that its research agencies are not arbitrarily changing peer reviewers when the results are not to their liking. The key question is whether an agency’s desire to publish an outdated report quickly outweighs the public’s need for a report whose evaluation methods, clarity, transparency, and timeliness are beyond reproach.
2. **NTP should update and publish its meta-analysis of epidemiological studies, but only after it survives peer review by NASEM or been accepted for publication in a reputable journal.**

At a time when the public’s trust in federal research is declining, are we simply to take NTP’s word that its meta-analysis—which has yet to survive peer review or even be accepted for publication—shows “there was no need to downgrade for publication bias”?  

NTP has provided no context for its proposal to publish its meta-analysis separately, or explained why the data are not already compiled in a statistically meaningful manner.  

A meta-analysis, which is used to detect publication bias, is essential to a report of this kind. In its first peer review, the NASEM committee (2020) criticized NTP for not performing a meta-analysis, stating, “Given that meta-analysis is a useful tool for aggregating and summarizing data and analyzing comparable studies, the committee strongly recommends that NTP reconsider its decision not to perform one.”  

In its second peer review, the NASEM committee (2021) expressed serious concerns about the meta-analysis that NTP eventually performed, questioning whether its risk-of-bias methodology was sound and had been consistently applied. The gold standard peer review organization used the term “worrisome remaining inconsistencies” to describe NTP’s meta-analysis, noting in its second peer review:

> Intrinsics remain in the application of risk-of-bias criteria to individual studies, particularly in NTP’s evaluation of how various studies handled major confounders, co-exposures, and outcomes...For example, Broadbent et al. 2015 and Cui et al. 2020 were deemed high risk for bias for confounding, whereas Trivedi et al. 2012 and others were not...The committee also identified several studies whose classification changed in revisions in the draft monograph without any justification provided (Sudhir et al. 2009; Trivedi et al. 2012; Das and Modal 2016).

The NASEM committee (2021) further reported the need for major revision before NTP’s meta-analysis would survive scientific scrutiny, noting:

> The committee had difficulty in following various aspects of the reported methods, identified a few worrisome remaining inconsistencies, was not able to find some key data used in the meta-analysis, and had concern about the wording of some conclusions.

> The revised monograph states that it addressed the independence issue, but the exact process used for selection of a single publication remains unclear, and in the meta-analysis, two reports on the same population are inappropriately included...It would be useful for the monograph to identify clearly which publications were derived from which study to minimize concerns about potential selection bias; doing so would also help to define the publications selected for the meta-analysis.

NTP should examine the studies included in the meta-analysis in greater depth to determine whether each study properly accounted for its design because not doing so could invalidate the meta-analysis results.
It would be useful for the monograph to identify clearly which publications were derived from which study to minimize concerns about potential selection bias; doing so would also help to define the publications selected for the meta-analysis.

NTP should review all its analyses to ensure that overlapping publications are not included in any single meta-analysis. That exercise is especially important given that the issue of “double counting” was a substantive concern of the committee in its first review.

NTP should... [provide] more information on each study result, including the actual result used from each study (SMDs, regression coefficients, and CIs), any data that NTP might have used to calculate the results (for example, means, standard deviations, and sample sizes), and other key information (for example, exposure concentrations of the high- and low-fluoride groups, the method used to assess exposure and outcome, which populations overlap, and information obtained from study authors).

NTP should also include subgroup or sensitivity analyses that respond to the committee’s concerns about blinding, complex sampling designs, and statistical analyses that account for clustered study designs.

Information on the meta-analysis protocols and information on the meta-analysis results are presented in several places. That approach forces the reader to go back and forth between sections and between documents to determine what was done or to obtain a clear picture of the meta-analysis findings.

It is unclear whether these issues were resolved to NASEM’s satisfaction because NTP removed the meta-analysis from its third draft. Instead, NTP stated, “[The] meta-analysis conducted in association with this systematic review further informs this issue and will be published separately.”

No context was given for its decision to publish the meta-analysis separately, or why the data were not already compiled in a statistically meaningful manner. NTP also did not give a timeline to ensure the dates of the literature search would be consistent with the state of the science report.

Considering NTP’s initial decision to forgo a meta-analysis, we question whether NTP is bound to pursue its publication. We also question whether a journal will accept it for publication and whether it would survive peer review. (It is already rumored that one journal, JAMA Pediatrics, did not accept it for publication.)

**NTP should be required to publish its meta-analysis, but only after it survives peer review by NASEM or been accepted for publication in a reputable journal.** That approach would help satisfy the concerns of those who question why NTP abandoned its peer review with NASEM, which was highly critical of NTP’s meta-analysis. It would also add a level of transparency that is lacking.

**NTP also needs to update its literature search.** This would be consistent with NTP’s already stated intent to add at least one more study to the meta-analysis that was not available during the original study period.† As the BSC Working Group noted, “[A] journal would likely ask the

† The NTP report states, “NTP is aware that this study was published after April 2021 (Ibarluzea et al. 2021) and, therefore, is not included in this monograph because it is beyond the dates of the literature search... The study will be examined as part of the NTP meta-analysis, which is being prepared as a separate report for publication.”
NTP authors to update the literature search."5

NTP’s finding is based on 19 studies. At least eight more have been published since the study period ended in 2020. One—a meta-analysis by Veneri et al. 2023—was published just four months ago in the journal *Environmental Research*.6

Unlike NTP’s meta-analysis, Veneri et al. 2023 survived the full course of peer review by its original peer review panel. Veneri et al. found:6

>[T]he limitations of most studies…with particular reference to the risk of residual confounding, raise uncertainties about both the causal nature of such relation and the exact thresholds of exposure involved. Such key issues can only be confirmed by additional, high-quality longitudinal studies.

Kumar et al. 2023, another meta-analysis published just this month, reached similar conclusions:7

These meta-analyses show that fluoride exposure at the concentration used in CWF is not associated with lower IQ scores. However, the reported association observed at higher fluoride levels in endemic areas requires further investigation. Uncritical acceptance of fluoride-IQ studies, including non-probability sampling, inadequate attention to accurate measurement of exposure, covariates and outcomes, and inappropriate statistical procedures has hindered methodological progress. Therefore, the authors urge a more scientifically robust effort to develop valid prenatal and postnatal exposure measures and to use interventional studies to investigate the fluoride-IQ hypothesis in populations with high fluoride (endemic) exposure.

In other words, the current state of the science does not validate the hypothesis that fluoride exposure is consistently associated with lower IQ in children. **We therefore support the BSC Working Group’s recommendation**6 **for NTP to update the study period and publish the report and meta-analysis separately and proximate to one another.**

**For clarification purposes, the revised report should survive the full course of peer review with NASEM, and the meta-analysis should survive the full course of peer review with either NASEM or a reputable peer reviewed journal.**

In addition to Veneri et al. 2023 and Kumar et al. 2023, NTP should include the following studies:


Guth, S., Hüser, S., Roth, A. et al. Toxicity of fluoride: critical evaluation of evidence for human developmental neurotoxicity in epidemiological studies, animal experiments and...


Again, at a time when the public’s trust in federal research is declining,² are we simply to take NTP’s word that its meta-analysis—which has yet to survive peer review or even be accepted for publication—shows “there was no need to downgrade for publication bias”? We question whether an agency’s desire to publish an outdated report quickly should outweigh the public’s need for a report whose methods, conclusions, clarity, transparency, and timeliness are beyond reproach.

3. **NTP should provide clear context for statements about low-level fluoride exposures, as NASEM recommended.**

The third (and purportedly final) draft report is full of non-contextualized statements about “potential associations”¹ between fluoride exposure and IQ, and the evidence being “unclear.”¹ In one area, NTP even states, “[L]ower concentrations of fluoride may support reduced IQ in humans” without offering any data or context to support its claim.¹

Statements suggesting “more studies are needed”¹ are technically accurate. Without context, however, the lay reader might conclude the lack of evidence justifies a precautionary approach to community water fluoridation.

For example:

**Associations** between lower total fluoride exposure [e.g., represented by populations whose total fluoride exposure was lower than the WHO Guidelines for Drinking-water Quality of 1.5 mg/L of fluoride (WHO 2017)] and children’s IQ remain unclear.

**More studies are needed to fully understand the potential** for lower fluoride exposure to affect children’s IQ.

**More studies at lower exposure levels are needed to fully understand potential associations** in ranges typically found in the United States (i.e., <1.5 mg/L in water). However, it should be noted that, as of April 2020, CWS supplying water with ≥1.5 mg/L naturally occurring fluoride served 0.59% of the U.S. population (~1.9 million people) (CDC Division of Oral Health 2020).

Although any effects in the brain or neurological tissue at lower concentrations of fluoride may support reduced IQ in humans, it may be difficult to distinguish the potential effects of
fluoride on learning and memory functions from other neurological or general health outcomes.

These non-contextualized statements can easily be misconstrued. In fact, they may be indicative of a desire to retain in some form the blanket hazard assessment that was removed from the third (and purportedly final) draft after the NASEM committee (2021) determined, “[T]he monograph falls short of providing a clear and convincing argument that supports its assessment.”

**We strongly urge that all references to “hazard conclusions found in previous draft monographs” be accompanied by a clear follow-up statement indicating why NASEM recommended that those hazard conclusions be withdrawn.**

Further, the NASEM committee (2021) made clear that statements about low level fluoride exposures were outside the monograph’s purview and inferences about “potential associations”—alongside claims suggesting the evidence is “unclear” or “may support” associations—should be avoided. In its second peer review, the NASEM committee wrote:

> Little or no conclusive information can be garnered from the revised monograph about the effects of fluoride at low exposure concentrations (less than 1.5 mg/L). **NTP therefore should make it clear that the monograph cannot be used to draw any conclusions regarding low fluoride exposure concentrations**, including those typically associated with drinking-water fluoridation. Drawing conclusions about the effects of low fluoride exposures (less than 1.5 mg/L) would require a full dose-response assessment, which would include at a minimum more detailed analyses of dose-response patterns, models, and model fit; full evaluations of the evidence for supporting or refuting threshold effects; assessment of the differences in exposure metrics and intake rates; more detailed analyses of statistical power and uncertainty; evaluation of differences in susceptibility; and detailed quantitative analyses of effects of bias and confounding of small effect sizes. Those analyses fall outside the scope of the NTP monograph, which focuses on hazard identification and not dose-response assessment. **Given the substantial concern regarding health implications of various fluoride exposures, comments or inferences that are not based on rigorous analyses should be avoided…**

**NTP needs to provide context for or otherwise remove comments and inferences about low-level fluoride exposures, as the NASEM committee recommended.** Doing so would help prevent the lay reader from misconstruing the state of the literature on low-level fluoride exposures, particularly given the most recent studies identified in Recommendation #2.

4. **NTP should include a stand-alone disclaimer indicating that the report should not be construed as an indictment of low-level fluoride exposures, as NASEM recommended.**

We question why NTP has been so averse to adding a disclaimer clarifying that its literature review did not validate the hypothesis that consistent exposure to low levels of fluoride impact neurodevelopment and cognition. The lay reader would have trouble knowing the report’s findings are limited to fluoride exposures that are more than double what the USPHS recommends for community water fluoridation. A disclaimer would help prevent the findings from being mischaracterized in debates about fluoridating local water systems.
Community water fluoridation is an inexpensive way to reduce tooth decay by at least 25 percent in the population. The CDC hailed it as one of ten great public health achievements of the 20th century.

Passions run deep about the population-based public health practice of adjusting the fluoride concentration of public water supplies to the levels recommended by the U.S. Public Health Service to prevent tooth decay (0.7 mg/L). In fact, opposition to community water fluoridation has been the driving force for this report.

For over the last 75 years, opponents have argued that fluoride is toxic and causes numerous harmful health effects, fluoride does not prevent tooth decay, fluoridation is costly, and fluoridation interferes with freedom of choice and infringes on individual rights. It was even called a Communist plot in the 1950s and a conspiracy between the U.S. government, the dental-medical establishment, and industry in the 1970s.

NASEM understood the public health ramifications of the report’s limited findings being mischaracterized in debates about whether to fluoridate community water systems. In its first peer review, the NASEM committee (2020) wrote:

The committee found some issues associated with data presentation and communication of various aspects of the process that are discussed further in the context of the evaluation of the animal and human evidence. One particular aspect of communication needs to be emphasized here. Many people are interested in whether water fluoridation to prevent tooth decay poses a threat to human neurodevelopment and cognition. Although the monograph provides some discussion of dose–response relationships, NTP did not conduct a formal dose–response assessment and needs to state clearly that the monograph is not designed to be informative regarding decisions about fluoride concentrations for water fluoridation.

NTP did not conduct a formal dose-response assessment that could inform a discussion on water fluoridation. NTP needs to state clearly that the monograph is not designed to be informative with respect to decisions about the concentrations of fluoride that are used for water fluoridation. That point should be reiterated at the end of the monograph with some indication that its evaluation of the literature is focused on hazard identification of fluoride and that it does not draw any conclusions regarding drinking-water fluoridation or other fluoride sources, such as toothpaste or other dental treatments. Although NTP does not explicitly claim that it has done something other than hazard identification, the context into which the monograph falls calls for much more carefully developed and articulated communication on this issue.

The NASEM committee (2021) reiterated in its second review:

The report must present its methods clearly, document the results transparently, and provide the rationale for conclusions in such a way that even those who disagree with them will appreciate that the process by which they were derived is clear and was implemented without error. The question is not whether this committee or the multiple audiences come to the same conclusions but rather whether the methods and analysis documented in the monograph support NTP’s conclusions...

Little or no conclusive information can be garnered from the revised monograph about the effects of fluoride at low exposure concentrations (less than 1.5 mg/L). NTP therefore should make it clear that the monograph cannot be used to draw any conclusions regarding low
flouride exposure concentrations, including those typically associated with drinking-water fluoridation. Drawing conclusions about the effects of low fluoride exposures (less than 1.5 mg/L) would require a full dose-response assessment, which would include at a minimum more detailed analyses of dose-response patterns, models, and model fit; full evaluations of the evidence for supporting or refuting threshold effects; assessment of the differences in exposure metrics and intake rates; more detailed analyses of statistical power and uncertainty; evaluation of differences in susceptibility; and detailed quantitative analyses of effects of bias and confounding of small effect sizes. Those analyses fall outside the scope of the NTP monograph, which focuses on hazard identification and not dose-response assessment. Given the substantial concern regarding health implications of various fluoride exposures, comments or inferences that are not based on rigorous analyses should be avoided…

[I]t is extremely important for it to be able to withstand scientific scrutiny by those who have vastly different opinions on the risks and benefits associated with fluoride exposure. The committee strongly recommends that NTP improve the revised monograph by seriously considering the suggestions that are provided in this letter report to improve its clarity and transparency.

NTP ignored NASEM’s recommendation. Instead, the report is full of non-contextualized statements about “potential associations” between fluoride exposure and IQ, and the evidence being “unclear.” In one area, NTP even states, “[L]ower concentrations of fluoride may support reduced IQ in humans” without offering any data or context to support its claim.

Statements suggesting “more studies are needed” are technically accurate. Without context, however, the lay reader might conclude the lack of evidence justifies a precautionary approach to community water fluoridation.

These non-contextualized statements can easily be misconstrued. In fact, they may be indicative of a desire to retain in some form the blanket hazard assessment that appeared in the first two drafts and was eventually removed after the NASEM committee (2021) determined, “[T]he monograph falls short of providing a clear and convincing argument that supports its assessment.”

We strongly urge that all references to “hazard conclusions found in previous draft monographs” be accompanied by a clear follow-up statement indicating why NASEM recommended that those hazard conclusions be withdrawn.

NTP’s poorly worded language has already had consequences. For example, on March 15—the day the third (and purportedly final) draft was made public—anti-fluoridation activists issued a press release claiming NTP “could not detect any safe exposure, including at levels common from drinking artificially fluoridated water.” The press release further claimed, “There is now little question that a large body of scientific evidence supports a conclusion that fluoride can lower child’s IQ, including at exposure levels from fluoridated water.”

The press release is consistent with a 2020 editorial from the NTP director who commissioned the report, suggesting that the unpublished, non-peer reviewed second draft was justification enough to end community water fluoridation nationwide. It is highly unusual for a researcher to comment on work that has not survived peer review.
The BSC is now in the fortunate position of **knowing** how the current version of this report will be used. **NTP should therefore adopt NASEM’s recommendation to add a disclaimer about low-level fluoride concentrations to the final report.** A disclaimer akin to the following would address legitimate concerns about the findings being misconstrued or mischaracterized in debates about fluoridating community water systems.

**DISCLAIMER**

*This state-of-the-science report should not be construed as an indictment of consistent low-level fluoride exposures (<1.5 mg/L), including concentrations recommended for community water fluoridation (0.7 mg/L). Community water fluoridation is the purposeful upward adjustment of the naturally occurring fluoride content in water to levels recommended by the United States Public Health Service (0.7 mg/L) to prevent tooth decay.*

*The report should also not be used to draw conclusions about fluoride content of toothpaste, fluoride supplements, or any other dental treatments.*

*An examination of the literature on low-level fluoride exposures did not validate the hypothesis that consistent exposure to low levels of fluoride (<1.5 mg/L) poses a risk to neurodevelopmental and cognitive health. Additional research may inform that point.*

A clear, strongly worded disclaimer of this kind would address the NASEM committee’s criticism about the report’s lack of context[^3] and “lack of details in several places and the lack of clarity on several substantive issues.”[^4] It would also comport with the recommendations of the White House Task Force on Scientific Integrity, which called for better methods of communicating scientific findings to ensure lay audiences have an accurate understanding of science.[^20][^21]

As the NASEM committee (2021) observed, “[i]t is extremely important for [the monograph] to be able to withstand scientific scrutiny by those who have vastly different opinions on the risks and benefits associated with fluoride exposure.”[^4] The gold standard peer reviewer therefore urged NTP to “make it clear that the monograph cannot be used to draw any conclusions regarding low fluoride exposure concentrations…”[^4]

**Crafty language**, the **former NTP director’s editorial**, and other actions—such as **abandoning the course of peer review with NASEM**, **removing the criticized meta-analysis** (which NTP initially declined to perform), and **asking the public to take on faith that there is no need to downgrade a number of risk-of-bias determinations**—is consistent with a pattern of behavior that suggests the line between research and activism may have been blurred. We question whether this is typical of how NTP carries-out its work and whether further oversight is needed.

At a time when the public’s trust in federal research is declining[^5], the BSC might consider whether the lack of clarity in NTP’s report—and the piecemeal way that NTP plans to release it—will contribute to (or detract from) the lay public’s understanding of science and that of local elected officials who determine community water fluoridation policies. The answer could determine whether the public’s health will be driven by science…or by **unanswered rhetoric**.

**5. NTP should revise its risk of bias rating for several studies, based on NASEM’s concerns and the enclosed analysis.**
At a time when the public's trust in federal research is declining, are we simply to take NTP's word that its risk-of-bias ratings—which have yet to survive peer review—show "there was no need to downgrade for publication bias"?

The NASEM committee twice expressed serious concerns about the methodology NTP used to determine study bias and questioned whether it had been consistently applied. The gold standard peer review organization used the term "worrisome remaining inconsistencies", noting in its second peer review:

![Inconsistencies remain in the application of risk-of-bias criteria to individual studies, particularly in NTP’s evaluation of how various studies handled major confounders, co-exposures, and outcomes…For example, Broadbent et al. 2015 and Cui et al. 2020 were deemed high risk for bias for confounding, whereas Trivedi et al. 2012 and others were not…The committee also identified several studies whose classification changed in revisions in the draft monograph without any justification provided (Sudhir et al. 2009; Trivedi et al. 2012; Das and Modal 2016).

A more recent meta-analysis, Veneri et al. 2023, found “noticeable differences of the estimates across categories of overall study quality, with a general trend towards weaker or null associations in the most carefully conducted studies.” The authors further noted:

![The serious adverse effect found in lower quality studies according to [risk-of-bias], could be at least in part due to the methodological limitations of those studies, thus increasing the uncertainty about the actual association between fluoride exposure and children’s cognitive neurodevelopment and reaffirming the strong need for properly designed and higher quality research on this topic.

Kumar et al. 2023, another meta-analysis published just this month, reached similar conclusions:

These meta-analyses show that fluoride exposure at the concentration used in CWF is not associated with lower IQ scores. However, the reported association observed at higher fluoride levels in endemic areas requires further investigation. Uncritical acceptance of fluoride-IQ studies, including non-probability sampling, inadequate attention to accurate measurement of exposure, covariates and outcomes, and inappropriate statistical procedures has hindered methodological progress. Therefore, the authors urge a more scientifically robust effort to develop valid prenatal and postnatal exposure measures and to use interventional studies to investigate the fluoride-IQ hypothesis in populations with high fluoride (endemic) exposure.

In terms of NTP's low and moderate risk-of-bias ratings, we note that in a pilot study of 51 children Choi et al. 2015—the sample size is too small to warrant a low risk-of-bias rating.

Two prospective secondary analysis studies—Bashash et al. 2017 and Green et al. 2019—rely on an invalid biomarker (spot maternal urinary fluoride) as a proxy for measuring fetal fluoride exposure. Thomas et al. 2016 reported lack of association between spot maternal urinary fluoride and maternal plasma fluoride in their multiple regression analysis. The Spearman coefficient was 0.29 in first trimester and -0.24 in third trimester. In fact, Thomas et al. 2016 found maternal plasma fluoride levels to be some 40 times lower than urinary fluoride levels.
Using an invalid biomarker alone warrants a high risk-of-bias rating.

Eight papers (based on 11 separate publications) are cross-sectional evaluations of endemic fluorosis areas. Cross-sectional study design cannot rule out reverse causality in endemic fluorosis areas. As stated in Guth et al. 2021:

> It is possible that parents with higher IQ read or inform themselves about the possible health hazards to children, and therefore avoid fluoride exposure. In this case, high maternal/parental intelligence [which is correlated with children’s IQ] would be causally linked to lower fluoride exposure rather than high fluoride exposure causing lower intelligence in children.

Except for Yu et al. 2018, these papers also rely on questionable methods, such as non-probability convenience (or purposive) sampling of endemic fluorosis areas and statistical operations that rely on randomness for their validity (e.g., hypothesis testing or linear regression).

Additionally, the authors of the eight papers made no effort to validate the data in the 11 foundational publications, or to scrutinize the analytical methods used for the findings.

For example, the initial confidence is based on comparison of the groups used. Many studies from high fluoride areas do not provide sufficient data to support this key criterion. Notwithstanding other limitations, these eight publications are too sufficiently flawed to warrant a moderate confidence in the body of evidence.

**Risk-of-bias.** NTP graded a number of studies as having a low risk-of-bias despite acknowledging individual critical elements had a high risk-of-bias. Examples are Ding et al. 2011 (high risk-of-bias for confounding); Rocha-Amador et al. 2007 (high risk-of-bias for selective reporting); Seraj et al. 2012 (high risk-of-bias for exposure assessment); Trivedi et al. 2012 (high risk-of-bias for statistical analysis), etc.

Out of ten publications, nine used a non-probability convenience sample. Only one study (Yu et al. 2018) sampled more than ten villages/towns/cities, which should decrease confidence in the body of evidence. The community-level effect was not adequately addressed in any of the studies. Often, the exposure measure is one or two samples of spot urinary fluoride with or without adjustment for urinary dilution. This is not a valid measure of long-term fluoride exposure. Only one study (Bashash et al. 2017) adjusted for maternal IQ.

**Unexplained inconsistencies.** Previous meta-analyses have shown substantial unexplained heterogeneity. Duan et al. 2018 conducted a meta-analysis of standardized mean difference in IQ scores between higher water fluoride communities (Mean F=3.7 mg/L) and normal fluoride communities (Mean F=0.6 mg/L). The summary results indicated high water fluoride exposure was associated with lower intelligence levels (standardized mean difference: -0.52; 95% CI: -0.62 to -0.42; P < 0.001). However, there was substantial heterogeneity (I² =69.1%; P < 0.001). The authors were unable to explain the source of heterogeneity. Studies conducted after 2014 show that the effect sizes are smaller and not statistically significant, including Green et al. 2019 and Bashash et al. 2017.

There are also inconsistencies within the same study. For example, Green et al. 2019 (the Maternal-Infant Research on Environmental Chemicals study, or MIREC) reported a differential effect such that the association between maternal urinary fluoride (MUF) and IQ was found only
in boys. However, Till et al. 2020 stated that MUF was not statistically significant either in boys or girls once postnatal fluoride was added to the model. Farmus et al. 2021 reported that fluoride exposures (during any trimester, average across all trimesters, infancy, and childhood) was not significantly associated with IQ outcomes after city was controlled and correction for multiple testing was applied. While Bashash et al. 2017 reported a threshold at 0.8 mg/L MUF (ages 6-12 years), Thomas et al. 2014 found no evidence of a detectable adverse outcome on offspring (ages 1-3 years) neurobehavioral development associated with maternal fluoride exposure during pregnancy.

**Imprecision.** As NASEM observed in its second review, standard errors are underestimated. Of most concern are the studies that used fluoride concentration measured at the community level as the exposure—see, for example, Seraj et al. 2012, Till et al. 2020, Trivedi et al. 2012, and Wang et al. 2012. When everyone in a community is subject to the same exposure, the standard error of the difference in means between high-exposure and low-exposure groups increases multiplicatively by the square root of a variance inflation factor (VIF) equal to \( [1 + (n - 1)r] \), where n is the number of persons in each community and r is the correlation in outcomes (such as IQ score) between members of the same community (Murray 1998; Donner and Klar 2000; Feng et al. 2001). The same phenomenon occurs in randomized control trials that assign treatment to groups of persons. Thus, unless within-community clustering is accounted for in the analysis—for example, through a random effects model—standard-error estimates will be too small and confidence intervals (CIs) too narrow...For individual-level exposures, such as urinary fluoride concentration, the VIF is probably smaller than one would see for community level exposures because some communities might contain people in multiple exposure groups. However, it is still important to account for clustering in the analysis because one would expect most people in a community to be in the same exposure group.

Note that when the average cluster size is large (e.g., n=66 in Green et al. 2019), even an interclass correlation coefficient of 0.2 will greatly impact VIF.

**Publication bias.** There is also evidence of publication bias. For example, the Thomas et al. 2014 thesis that showed a beneficial effect of fluoride exposure in the Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT) was not published. Another example is that Green et al. 2019 do not discuss the lack of effect of MUF on FSIQ in their paper. There was also the sudden removal of a critical sentence from the final pre-print version of Farmus et al. 2021: “However, exposures do not significantly associate with IQ outcomes once city is controlled and FDR is applied.”

Again, at a time when the public’s trust in federal research is declining, are we simply to take NTP’s word that its risk-of-bias ratings—which have yet to survive peer review—show “there was no need to downgrade for publication bias”? We question is whether an agency’s desire to publish an outdated report quickly should outweigh the public’s need for a report whose evaluation methods, clarity, transparency, and timeliness are beyond reproach.

**References**

1 National Toxicology Program. May 2022. [Third] Draft NTP Monograph on the State of the Science Concerning Fluoride Exposure and Neurodevelopmental and Cognitive Health Effects: A Systematic...
Attachment A

Analysis and Recommendations

Review. Office of Health Assessment and Translation, Division of the NTP, National Institute of Environmental Health Sciences, National Institutes of Health, U.S. Department of Health and Human Services. doi:10.22427/NTP-MGRAPH-8


12 Connett, Michael, Executive Director, Fluoride Action Network, to Thayer, Kristina, Director, Office of Health Assessment and Translation, National Toxicology Program, Durham, NC. November 6, 2015. Available at: fluoridealert.org/wp-content/uploads/connett-ntp.11-6-15.pdf (accessed April 28, 2023)


A meta-analysis, which is used to detect publication bias, is essential to a report of this kind. The NASEM peer review committee (2021) expressed serious concerns about NTP’s meta-analysis, questioning whether its risk-of-bias methodology was sound and had been consistently applied. The gold standard peer review organization used the term “worrisome remaining inconsistencies” to describe NTP’s meta-analysis, noting in its second peer review:

"Inconsistencies remain in the application of risk-of-bias criteria to individual studies, particularly in NTP’s evaluation of how various studies handled major confounders, co-exposures, and outcomes...For example, Broadbent et al. 2015 and Cui et al. 2020 were deemed high risk for bias for confounding, whereas Trivedi et al. 2012 and others were not...The committee also identified several studies whose classification changed in revisions in the draft monograph without any justification provided (Sudhir et al. 2009; Trivedi et al. 2012; Das and Modal 2016)."

We urge you to recommend that NTP revisit its meta-analysis to account for more recent literature. This would be consistent with NTP’s stated intent to add at least one more study to the meta-analysis that was not available during the original study period, and the BSC Working Group’s observation that “[A] journal would likely ask the NTP authors to update the literature search.” It would also allow for a more recent meta-analysis, published earlier this year, to be included:


We also urge you to recommend that NTP revisit its risk-of-bias ratings for the studies herein to account for the mitigating factors identified below. These mitigating factors warrant an adjustment to NTP’s risk-of-bias determinations.

---


2 The NTP report of May 2022 states, “NTP is aware that this study was published after April 2021 (Ibarluzea et al. 2021) and, therefore, is not included in this monograph because it is beyond the dates of the literature search...The study will be examined as part of the NTP meta-analysis, which is being prepared as a separate report for publication.”

<table>
<thead>
<tr>
<th>Study</th>
<th>Rationale</th>
</tr>
</thead>
</table>
| Bashash (2017) | NTP rated Bashash et al. (2017) as having a low risk-of-bias. However, the study relies on convenience sampling of spot maternal urinary fluoride—an invalid biomarker—as a proxy for measuring fetal fluoride exposure. Using an invalid biomarker alone warrants a definitely high risk-of-bias rating.  
This cohort study is based on a convenience sample drawn from multiple hospitals (clusters) in Mexico. Study results based on convenience sampling cannot be used to draw inferences.  
Spot maternal urinary fluoride is the proxy for fetal exposure. However, Thomas et al. (2016) showed a weak correlation between urinary F and plasma F. There was no association between urinary fluoride and plasma fluoride in a multiple regression analysis. Not a valid biomarker.  
The source of F is salt. Therefore, the higher fluoride exposure is confounded by higher salt intake, which is associated unhealthy diet and poor pregnancy outcomes. The authors did not assess whether the lower IQ is due to an unhealthy diet. Cantoral et al. (2021) cited this as a limitation in their analysis of the ELEMENT data.  
The authors did not account for clustering resulting from samples drawn from hospitals.  
The study is not compliant with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) research methodology, which is a best practice for studies of this kind. To avoid bias, or even the perception of bias, an independent, STROBE-compliant analysis of the MIREC and ELEMENT data is warranted.  
Selective reporting: This study should receive a definitely high risk-of-bias rating for selective reporting because it excluded the positive findings associated with fluoride exposure from the Thomas et al. (2014) dissertation that analyzed the same ELEMENT cohorts. |
| Choi (2015)  | NTP rated this study as having a low risk-of-bias. However, not only is the sample size too small, but the study also relies on unrelated exposure variables. A high risk-of-bias rating is warranted.  
This is a pilot study of 51 students in China. The authors also used dental fluorosis as an exposure variable, which is a postnatal phenomenon. Dental fluorosis of primary teeth is extremely rare, even in endemic fluorosis areas. |
<table>
<thead>
<tr>
<th>Study</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cui (2018)</td>
<td>NTP rated Cui et al. (2018), Cui et al. (2020), Wang et al. (2020), Yu et al. (2018), and Zhang et al. (2015) as having a low risk-of-bias. However, the study designs (cross-sectional) do not account for reverse causality, which is likely in these cases. A high risk-of-bias determination is warranted. These publications are from a more extensive study of 2886 resident children, aged 7 to 13 years, randomly from endemic and non-endemic fluorosis areas in Tianjin, China (Yu et al. (2018)). Used a complex survey design (stratified sampling of clusters from endemic and non-endemic fluorosis areas). Cui et al. (2018), Cui et al. (2020), and Zhang et al. (2015) selected a subset of schools based on IQ scores and F levels, leading to selection bias. However, the authors did not account for the complex survey design and the standard errors are therefore underestimated. There is a possibility of Type 1 error. Exposure measure is a spot urinary fluoride of unproven validity. NTP highlighted only the statistically significant results but left out the results that did not show statistically significant results. Yu et al. (2018) showed a threshold effect such that there is no effect of fluoride on IQ below 3.4 mg/L fluoride in water ($B=-0.04 (-0.33, 0.24)$) or below 1.6 mg/L urinary F ($B=0.36 (-0.29, 1.01)$).</td>
</tr>
<tr>
<td>Cui (2020)</td>
<td></td>
</tr>
<tr>
<td>Wang (2020)</td>
<td></td>
</tr>
<tr>
<td>Yu (2018)</td>
<td></td>
</tr>
<tr>
<td>Zhang (2015)</td>
<td></td>
</tr>
<tr>
<td>Ding (2011)</td>
<td>NTP rated Ding et al. (2011) as having a low risk-of-bias despite finding a high risk-of-bias for confounding. The cross-sectional study design also does not account for reverse causality, which is likely in this case. A high risk-of-bias rating is therefore warranted. The authors selected schools from 4 sites in Inner Mongolia, China. All four sites were in endemic and nonendemic fluorosis areas. The authors did not account for the cluster sampling design. The standard errors are therefore underestimated. Additionally, the regression equation included only age.</td>
</tr>
<tr>
<td>Green (2019)</td>
<td>NTP rated Green et al. (2019) and Till et al. (2020) as having a low risk-of-bias. However, both studies rely on spot maternal urinary fluoride—an invalid biomarker—as a proxy for measuring fetal fluoride exposure. An invalid biomarker alone justifies a high risk-of-bias rating. NTP also lists these as prospective cohort studies; however, there is only one IQ measurement. Had Green et al. (2019) and Till et al. (2020) assessed the validity of this biomarker, they would have found that Thomas et al. (2016) reported lack of association between spot maternal urinary fluoride and maternal plasma</td>
</tr>
<tr>
<td>Study</td>
<td>Rationale</td>
</tr>
<tr>
<td>-------</td>
<td>-----------</td>
</tr>
<tr>
<td></td>
<td>fluoride in their multiple regression analysis. The Spearman coefficient was 0.29 in first trimester and -0.24 in third trimester. In fact, Thomas et al. (2016) found maternal plasma fluoride levels to be some 40 times lower than urinary fluoride levels.</td>
</tr>
<tr>
<td></td>
<td>With respect to Green et al. (2019), the convenience sample was drawn from seven hospitals in six cities (clusters) in Canada, creating a hierarchical data structure. The statistical analysis did not adequately account for the city-level effect. IQ varied by as much as 8 points between the non-fluoridated cities of Vancouver and Kingston (page 30, Green 2018 Master’s Thesis).</td>
</tr>
<tr>
<td></td>
<td>A single staff person from each study site administered in-person IQ assessments. Thus, the assessor was matched to the city. This would be considered a fatal flaw in any RCT or case-control study.</td>
</tr>
<tr>
<td></td>
<td>Further, Green et al. (2019) is not compliant with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) research methodology, which is a best practice for studies of this kind. To avoid bias, or even the perception of bias, an independent, STROBE-compliant analysis of the MIREC and ELEMENT data is warranted.</td>
</tr>
<tr>
<td></td>
<td>In its peer review of NTP’s second draft, the NASEM (2021) committee reported¹:</td>
</tr>
<tr>
<td></td>
<td>In the case of Green et al. (2019), NTP learned from the investigators that accounting for city-level clustering via a random-effects model “showed similar results to the main model.” More details should be provided regarding the similarity of results because although overall conclusions might not have changed, the results of the meta-analysis could be affected by incorrect exposure-effect or standard-error estimates.</td>
</tr>
<tr>
<td></td>
<td>The Canadian Agency for Drug and Technologies in Health also analyzed Green et al. (2019) study and determined a high risk-of-bias in the study²:</td>
</tr>
<tr>
<td></td>
<td>The study by Green et al., 2019 concluded that “maternal exposure to higher levels of fluoride during pregnancy was associated with lower IQ scores in children aged 3 to 4 years.” (p. E1) This conclusion was not supported by the data... Between nonfluoridated and fluoridated maternal exposure (assessed by MUFSG or daily fluoride intake), the difference in mean FSIQ in total children (108.07 ± 13.31 versus 108.21 ± 13.72) was minimal. The average FSIQ in boys in the non-fluoridated and fluoridated groups were 106.31 ± 13.60 and 104.78 ± 14.71, respectively, and in girls were 109.86 ± 12.83 and 111.47 ± 11.89, respectively. According to the WPPSI test scoring, these numbers were considered as normal, as a score of 90 to 109 represents average intelligence.</td>
</tr>
<tr>
<td>Study</td>
<td>Rationale</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Study (Rocha-Amador 2007)</td>
<td>NTP rated Rocha-Amador et al. (2007) as having a low risk-of-bias overall despite finding a high risk-of-bias for selective reporting. The cross-sectional study design does not account for reverse causality, which is likely in this case. A high risk-of-bias rating is therefore warranted.</td>
</tr>
<tr>
<td></td>
<td>This cross-sectional study of 132 children of age 6-10 was conducted in areas of Brazil where mean levels of Arsenic in water were 17 and 19 times higher than WHO limits in Salitrail (mean F level in water 5.3 mg/L) and 5 de Febrero (9.4 mg/L F), respectively. However, it was not included in the regression model.</td>
</tr>
<tr>
<td>Study (Green et al. 2019)</td>
<td>The authors identified limitations in the study and where possible implemented measures to reduce their impact. However, a number of uncertainties remain (e.g., estimation of prenatal fluoride exposure, other unmeasured factors affecting child IQ) which limit this study’s ability to confirm a causal relationship between prenatal fluoride and deficits in child IQ.</td>
</tr>
<tr>
<td></td>
<td>Health Canada and CADTH evaluations are included as Attachments C and D, respectively.</td>
</tr>
<tr>
<td>Study (Till et al. 2020)</td>
<td>With respect to Till et al. (2020), the authors reported that after postnatal exposure was introduced into the model, maternal urinary fluoride was not associated with FSIQ in boys or girls. The authors found two outliers in the same cohort, and the association became non-significant when two outliers were removed.</td>
</tr>
<tr>
<td></td>
<td>NTP reports correspondence with Till about assessors lack of knowledge of fluoridation status; however, Till neglected to mention that a single staff person from each study site administered in-person IQ assessments of 3 and 4 year-olds. Thus, the assessor was matched to the city with no attempt to assess inter-rater reliability. This would be considered a fatal flaw in any RCT or case-control study.</td>
</tr>
</tbody>
</table>

Given that these values were available during data collection period, it was unclear about the authors’ rationale to further explore the associations between maternal fluoride exposure and children’s IQ. Indeed, adjusted estimates with a limited set of covariates showed no statistically significant association between an increase of 1 mg/L in MUFSG and FSIQ, PIQ or VIQ in all children. These were not discussed or considered when formulating the conclusion.

Given that these values were available during data collection period, it was unclear about the authors’ rationale to further explore the associations between maternal fluoride exposure and children’s IQ. Indeed, adjusted estimates with a limited set of covariates showed no statistically significant association between an increase of 1 mg/L in MUFSG and FSIQ, PIQ or VIQ in all children. These were not discussed or considered when formulating the conclusion.

Health Canada and CADTH evaluations are included as Attachments C and D, respectively.
<table>
<thead>
<tr>
<th>Study</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>While height for age was included in the model, age was not. Mothers’ education levels differed among the three areas with low fluoride community with the highest level of education. This community-level effect was not controlled. Therefore, NTP noted that the results might still be biased.</td>
</tr>
<tr>
<td>Saxena (2012)</td>
<td>NTP rated Saxena et al. (2012) as having a low risk-of-bias. However, the cross-sectional study design does not account for reverse causality, which is likely in this case. A high risk-of-bias grading is therefore warranted.</td>
</tr>
<tr>
<td></td>
<td>This is a cross-sectional study of 120 children in India from 3 endemic areas, and 50 children from 1 non-endemic area were included in the analysis. The mean urinary fluoride level in the non-endemic areas was 2.25 mg F/L which is about three times higher compared to a fluoridated area.</td>
</tr>
<tr>
<td></td>
<td>NTP correctly noted that the author’s use of linear regression for an ordinal IQ outcome with five levels was inappropriate. Similarly, the authors used ANOVA for socioeconomic status and other variables measured with an ordinal scale. This alone should have received a high risk-of-bias rating.</td>
</tr>
<tr>
<td>Seraj (2012)</td>
<td>NTP rated Seraj et al. (2012) as having a low risk-of-bias overall despite finding a high risk-of-bias for exposure assessment. However, the cross-sectional study design does not account for reverse causality, which is likely in this case. A high risk-of-bias rating is therefore warranted.</td>
</tr>
<tr>
<td></td>
<td>This cross-sectional study of 293 6- to 11-year-old children in Iran from five selected rural areas. The authors state that these areas were similar in their general demographic and geographic characteristics, with the inhabitants having a comparable socioeconomic status and similar occupations. However, there is no data to support the comparability of areas. NTP somehow found indirect evidence of comparability.</td>
</tr>
<tr>
<td></td>
<td>NTP rated probably high risk-of-bias for exposure assessment. The authors did not provide data to indicate that the mean was representative of the fluoride levels over 12 years and throughout the village.</td>
</tr>
<tr>
<td></td>
<td>The statistical analysis is also difficult to comprehend.</td>
</tr>
<tr>
<td>Soto-Barreras (2019)</td>
<td>NTP rated Soto-Barreras et al. (2019) as having a low risk-of-bias despite finding a high risk-of-bias for confounding and water fluoride exposure. Also, the cross-sectional study design does not account for reverse causality, which is likely in this case. A high risk-of-bias rating is therefore warranted.</td>
</tr>
</tbody>
</table>
### Study Rationale

<table>
<thead>
<tr>
<th>Study</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sudhir (2019)</td>
<td>NTP rated Sudhir et al. (2019) as having a low risk-of-bias overall despite finding a high risk-of-bias for lack of blinding and because no information was provided to indicate that the methods to assess IQ outcomes were reliable and valid in this study population. We observe also that the cross-sectional study design does not account for reverse causality, which is likely in this case. A high risk-of-bias rating is therefore warranted.</td>
</tr>
<tr>
<td>Trivedi (2012)</td>
<td>NTP rated Trivedi et al. (2012) as having a low risk-of-bias overall despite finding a high risk for statistical analysis. However, the cross-sectional study design does not account for reverse causality, which is likely in this case. A high risk-of-bias rating is therefore warranted. This is a cross-sectional study of exactly 1000 children of 13 to 15 years of age from Nalgonda district (Andhra Pradesh), India. Clustering of children within the four areas was not accounted for in the analysis. About 70% of children in the low exposure group were in the below-average intelligence grade. The authors did not consider a multivariate analysis.</td>
</tr>
<tr>
<td>Xiang (2003)</td>
<td>NTP rated Xiang et al. (2003), Xiang et al. (2011), and Wang et al. (2012) as having a low risk-of-bias. However, the cross-sectional study design does not account for reverse causality, which is likely in this case. A high risk-of-bias grading is therefore warranted.</td>
</tr>
<tr>
<td>Xiang (2011)</td>
<td></td>
</tr>
<tr>
<td>Wang (2012)</td>
<td></td>
</tr>
</tbody>
</table>

This is a cross-sectional study of 161 children aged 9 to 10 years of age from Chihuahua, Mexico. There was no adjustment for clustering at the school level or the sampling design; however, the authors report that they did not find a relationship between fluoride exposure and IQ.
<table>
<thead>
<tr>
<th>Study</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>This is a cross-sectional study of 512 children aged 8-13 years from Wamiao (severe endemic fluorosis) and Xinhua (non-endemic) villages in Sihong County, Jiangsu Province, China. According to the authors, these “villages are situated in isolated low-income areas with less economic development and a relative lack of communication with the outside world, resulting in poor living conditions for the majority of the residents, especially the elderly and children.” The two villages are not comparable concerning the education of parents. The proportion of parents with senior high school education was 13.5% in Wamiao and 41.7% in Xinhua. NTP noted that a potential concern raised by the NASEM (2020) committee’s review was the lack of accounting for relationships in exposure between persons from the same village. Given only two villages were included and the analyses consisted of village-level comparisons (no use of individual-level covariate data), it is likely that the standard error of the difference in mean IQ between fluoride in water exposure groups will be biased, making differences appear stronger than they actually are. Without controlling for village effects and given the large differences in fluoride concentrations and IQ levels between villages, the apparent dose-response relationship could be due to a village effect in addition to a fluoride effect.</td>
<td></td>
</tr>
<tr>
<td>Broadbent (2015)</td>
<td>NTP rated Broadbent et al. (2015) as having high risk-of-bias, primarily because the authors did not account for other sources of fluoride in non-fluoridated areas. The high risk-of-bias rating is questionable, however, because NTP did not seek clarification from the authors in the same manner they did with the authors of other studies. The study should be given full and fair consideration. In response, the authors filed their own response in a letter to the editor of the American Journal of Public Health, noting:“ In the Dunedin [New Zealand] Study cohort, the majority of children who took fluoride tablet supplements did so intermittently and for only a short period of time. We have now estimated average total fluoride intake in our cohort up to age five years, including tablets, toothpastes, and dietary sources. We identified no differences in IQ in childhood or adulthood by total fluoride intake, but we did identify significantly fewer dental caries in both childhood and adulthood among those with higher estimated fluoride intake up to age five years.</td>
</tr>
</tbody>
</table>
References


Overview of York University Fluoride Study

Prepared by the Water and Air Quality Bureau, Health Canada

Health Canada’s Guidelines for Drinking Water Quality in Canada provides parameters to provinces, territories and federal Government Departments for water systems across the country. The Drinking Water Guideline for fluoride establishes a maximum acceptable concentration (MAC) for fluoride at 1.5 mg/L that factors in all sources of exposure to fluoride. The Guideline, published in 2010, was informed by published peer reviewed studies and the recommendations of an expert panel that included the Chief Dental Officer. The expert panel examined both potential adverse health effects of fluoride and the public health benefits of adding fluoride to drinking water through community water fluoridation to prevent dental caries. Since the Drinking Water Guideline for fluoride was established in 2010, Health Canada has regularly reviewed the state of the science on the health effects of fluoride and has concluded the current available science indicates that fluoride at levels below this guideline does not pose a health concern.

A York University study, “Association Between Maternal Fluoride Exposure During Pregnancy and IQ Scores in Offspring in Canada”, linking maternal fluoridation exposure during pregnancy to lower IQ scores in children aged 3 to 4 was published in JAMA Pediatrics on August 19, 2019. As is the case with all new science, Health Canada has reviewed this study and has considered it in weight of evidence-based decision-making to protect the health and safety of Canadians. It is important to note that when assessing the health risk, Health Canada looks at the available body of science—not one single study—in order to determine whether there is enough evidence to warrant a change in position.

In reviewing this study, Health Canada notes that from analysis of data and banked maternal urine (for fluoride) from the Maternal-Infant Research on Environmental Chemicals (MIREC) Study, the authors conclude that “… maternal exposure to higher levels of fluoride during pregnancy was associated with lower IQ scores in children aged 3 to 4 years.” The key element of this study is that it is an observational study, which found an association between higher levels of two different measures of fluoride exposure during pregnancy and small decreases in child IQ at 3-4 years of age. This one study is not able to prove that prenatal fluoride exposure causes deficits in child IQ, only that there was an observation of such an association. The study was well designed and analysed. The authors identified limitations in the study and where possible implemented measures to reduce their impact. However, a number of uncertainties remain (e.g., estimation of prenatal fluoride exposure, other unmeasured factors affecting child IQ) which limit this study’s ability to confirm a causal relationship between prenatal fluoride and deficits in child IQ.

This study is one of the first linking fluoride and neurological effects, and Health Canada will continue to monitor and evaluate studies as they are published. Based on the current weight of evidence, Health Canada continues to support the existing Drinking Water Guideline for fluoride. As Health Canada continues to keep abreast of scientific developments, the Department will collaborate with the Office of the Chief Dental Officer, provinces and territories and other interested stakeholders.
ATTACHMENT E

BIBLIOGRAPHY


Connett, Michael, Executive Director, Fluoride Action Network, to Thayer, Kristina, Director, Office of Health Assessment and Translation, National Toxicology Program, Durham, NC. November 6, 2015. Available at: fluoridealert.org/wp-content/uploads/connett-ntp.11-6-15.pdf (accessed April 28, 2023)


Green R. Prenatal Fluoride Exposure and Neurodevelopmental Outcomes in a National Birth Cohort. A Thesis Submitted to the Faculty of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Master of Arts, Graduate Program in Psychology. York University, Toronto, Ontario, July 2018. yorkspace.library.yorku.ca/xmlui/handle/10315/35502


