

Review of: "A systematic review and meta-analysis of the association between fluoride exposure and neurological disorders"

Christopher Neurath

Potential competing interests: I work for a not-for-profit organization whose activities include public education and scientific research on environmental health issues. Research and information dissemination on questions of fluoride exposure safety are a major focus of my work. I have no other competing interests.

Peer-review comments on Miranda et al (2021) systematic review of fluoride neurotoxicity

Chris Neurath

Research Director

American Environmental Health Studies Project (AEHSP)

The recently published paper by Miranda et al is a systematic review and meta-analysis of human studies of fluoride neurotoxicity [[Miranda 2021](#) (A systematic review and meta-analysis of the association between fluoride exposure and neurological disorders, *Sci Rep* 11, 22659)]. Miranda concluded that the studies they selected supported a GRADE certainty classification of “very low certainty”. However, their review has serious limitations and I suggest that if the GRADE methodology had been faithfully followed, the conclusion should properly be “high certainty” that fluoride is neurotoxic, the highest of four GRADE levels of evidence.

I identified 25 weaknesses in the Miranda review. An outline list of these weaknesses is presented below, followed by detailed point-by-point explanations for each weakness.

OUTLINE OF WEAKNESSES

1. The quality of this review compares unfavorably to that of the ongoing US National Toxicology Program (NTP) systematic review of fluoride (F) neurotoxicity. [[NTP 2020](#), [NTP 2021](#)]
The NTP’s review is more comprehensive and objective than the Miranda review.

2. However, the NTP review shares some problems with the Miranda review, most of which appear to be due to improper influence by dental interests trying to water down the conclusions. Our detailed comments on the NTP review are available here:

<https://fluoridealert.org/researchers/the-national-toxicology-program/1990-2021-submissions-to-ntp-nas/>

- 3. Miranda had flawed search and selection methods that excluded many relevant studies.**
- 4. A weak literature search missed key papers.**
- 5. Highest quality studies with individual-level exposure data were excluded.**
- 6. A systematic review that fails to identify and include many of the strongest papers on a given topic lacks objectivity.**
- 7. The Miranda review discarded important information about dose when it collapsed exposures into just two categories. It thereby reduced its ability to find associations between F and neurotoxicity.**
- 8. The Miranda review design is unable to assess dose-response relationships between F and IQ. Our own dose-response assessment found a strong relationship.**
- 9. The decision to convert all study results to odds ratios is problematic because it discards information. This may lead to an underassessment of the strength of the evidence.**
- 10. The GRADE method of evaluating a body of evidence is designed for clinical interventions intended to improve health, not observational studies of adverse environmental health effects. More appropriate methods for grading environmental health evidence are available.**
- 11. The authors' main conclusion is unsupported by their own data.**
- 12. The review contradicts its own risk of bias scoring.**
- 13. The heterogeneity found in effect size in the meta-analyses can be explained by differences amongst the studies, so is not grounds for downgrading.**
- 14. The Miranda review resorts to a subjective "narrative" for choosing its overall evidence conclusion, and thus deviates from the GRADE method.**
- 15. The Miranda review does not follow the GRADE guidance for evaluating imprecision.**
- 16. The authors also failed to follow the GRADE method that provides for upgrading**

observational studies when there is evidence of dose-response relationships.

17. The authors also failed to follow the GRADE method that provides for upgrading when the effect is large and consistent.

18. If the GRADE methodology had been faithfully followed, the starting grade of “low certainty” for observational studies would have been upgraded one level for dose-response relationship and another level for large magnitude and consistency of effect. The final grade should therefore have been “high certainty”, the highest grade, rather than the “very low certainty” chosen by the authors, the lowest possible grade.

19. The sensitivity analyses were not adequately explained or documented.

20. The authors claim there was evidence of publication bias that justified downgrading the confidence conclusion, but then acknowledge that when 3 “outlier” studies were excluded in a sensitivity analysis there is no longer evidence of publication bias in the remaining 7 studies.

21. Errors in references.

22. There is insufficient transparency and incomplete reporting.

23. There are unstated conflicts of interest with the dental industry and fluoridation-promoting organizations.

24. Conflicts are not just direct financial conflicts, and can be reputational, and indirect, such as the ability to obtain future research funding from private or public sources.

25. Given all the problems in the Miranda systematic review, It does not appear to be an objective assessment of the current body of evidence for fluoride’s neurotoxicity. It appears instead, that the authors are trying to defend the use of fluoride in dentistry. In contrast, independent scientists from outside the dental field have concluded that the existing evidence is sufficient to warn that pregnant women, infants, and children should avoid fluoridated water and limit other sources of fluoride exposure to reduce the risk of lost IQ.

DETAILED EXPLANATIONS OF WEAKNESSES

1. The quality of this review compares unfavorably to that of the ongoing US National Toxicology Program (NTP) systematic review of fluoride (F) neurotoxicity. [NTP 2020, NTP 2021]

The NTP's review is more comprehensive and objective than the Miranda review. This systematic review and meta-analysis is worth comparing to the existing ongoing rigorous systematic review and meta-analysis by the US National Toxicology Program (NTP), a division of the National Institutes of Health / National Institute of Environmental Health Sciences (NIH/NIEHS). The NTP review is over 300 pages long with hundreds of pages of additional online posted supporting data. It has undergone extensive public and independent peer-review and has been ongoing for 5 years with over 50 contributors named. The authors of the NTP review are not associated with the dental community and have expertise in toxicology, environmental epidemiology, and conducting systematic reviews.

2. However, the NTP review shares some problems with the Miranda review, most of which appear to be due to improper influence by dental interests trying to water down the conclusions.

We have made detailed critiques of the NTP review and have conducted our own dose-response assessment based on the studies identified by NTP:

List of submissions from the Fluoride Action Network (FAN) on NTP review:

<https://fluoridealert.org/researchers/the-national-toxicology-program/1990-2021-submissions-to-ntp-nas/>

Most recent critique of NTP review from FAN:

<https://fluoridealert.org/wp-content/uploads/ntp.neurath.submission-to-nas-on-revised-ntp-monograph.10-19-20.pdf>

Dose-response assessment using the high quality studies identified by NTP; summary:

<https://fluoridealert.org/wp-content/uploads/neurath-2020-isee-eposter-ver11.pdf>

Dose-response assessment; full report [Neurath 2020a]:

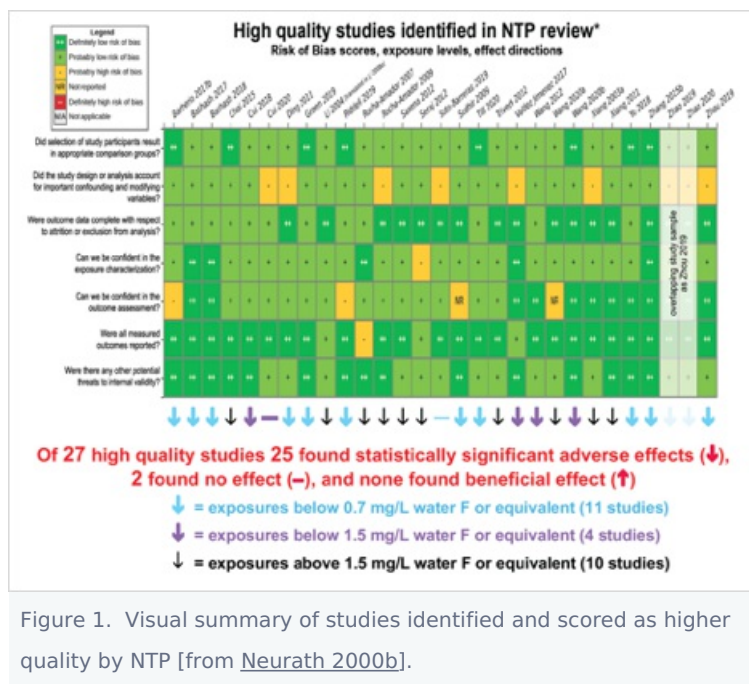
<https://fluoridealert.org/wp-content/uploads/neurath-2020.fan-dose-response-assessment-of-ntp-studies-with-corrections.pdf>

Video in-depth discussion of the NTP systematic review methods and our dose-response assessment based on NTP review:

<https://fluoridealert.org/articles/ntp-neurath-summary/>

Figure 1 is a visual summary of the higher quality studies identified by NTP, showing the NTP's Risk of Bias scores assigned for each domain of each study, with additional effect direction and mean dose information that demonstrates the remarkable consistency of adverse neurotoxic outcomes associated with F,

including in studies with mean F exposures below 0.7 mg/L.



3. Miranda had flawed search and selection methods that excluded many relevant studies. The Miranda review excluded, or failed to even identify, some of the highest quality studies of F and IQ identified in the NTP review. Examples of high-quality papers that seem to have been missed in their search include [Jimenez-Valdez 2017](#), [Bashash 2017](#), [Bashash 2018](#), [Yu 2018](#), [Cui 2018](#), [Green 2019](#), [Riddell 2019](#), [Zhao 2021](#), [Yu 2021](#), [Cantoral 2021](#), and [Wang 2021](#). Several of these are prospective longitudinal cohort studies, which is often the strongest available design for observational studies of adverse effects. Exclusion of studies with individual-level exposure measures seems to have been built-in to the review from the outset, because the PECO statement apparently allows only studies with two exposure groups, a “high” and a “low”.

4. A weak literature search missed key papers. An example of the flawed search methods can be seen in the Miranda Supplement 1 in their choice of search terms in Google Scholar. Their sole search string required three terms to *all* be found in a paper:

Fluoride+"Neurological manifestations"+Intelligence

This search string returns only about 150 results, with very few of the results meeting the authors other inclusion criteria. If a less restrictive search string had been used, such as:

Fluoride+Intelligence

Google Scholar returns about 24,000 results, with many of the top results being clearly applicable.

The authors' use of such a restrictive search may explain why they apparently did not identify some of the most important and prominent epidemiological studies of fluoride neurotoxicity.

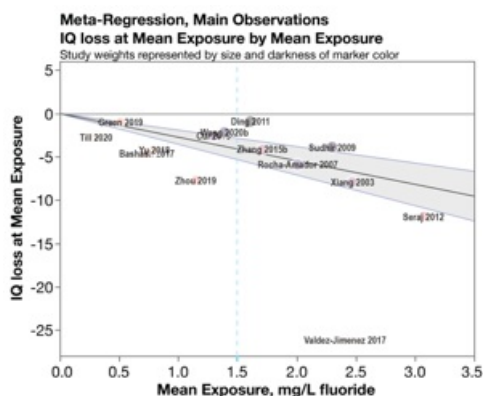
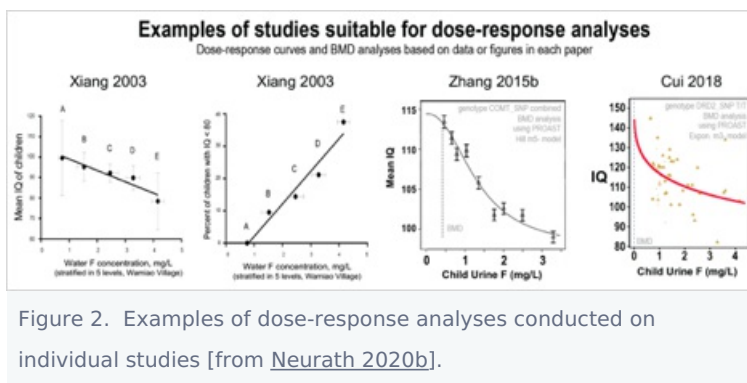
5. Highest quality studies with individual-level exposure data were excluded. The Miranda review restricted itself to only studies with group-level exposure data and thereby excluded many of the highest quality studies which often have individual-level exposure data. The large majority of the 23 excluded studies listed in Supplement 3 are excluded because of their "Absence of IQ comparison considering high and low fluoride exposure regions". Unmentioned is that many of these studies have individual-level fluoride exposures which cover a continuous range, thereby allowing more accurate estimates of fluoride exposure as well as evaluation of dose-response relationships. These excluded studies are often the highest quality studies, and many have exposures well below the 1.0 mg/L water concentration so are in the range chosen by the authors as relevant to artificial water fluoridation.

6. A systematic review that fails to identify and include many of the strongest papers on a given topic lacks objectivity.

7. The Miranda review discarded important information about dose when it collapsed exposures into just two categories. It thereby reduced its ability to find associations between F and neurotoxicity. The choice of concentrations to be classified as "low/adequate" and "high" is unrelated to the exposure levels found in the actual papers. It was instead chosen based on WHO criteria pertaining to effectiveness at reducing dental caries. The WHO criteria were developed with no consideration of possible adverse neurocognitive effects. The decision to lump the exposure levels in all studies into just two categories also makes this review incapable of ascertaining a dose-response relationship. Therefore, the design of this review precluded it from upgrading the GRADE certainty when a dose-response relationship is found. In our own dose-response analyses of F-IQ studies, in which we conducted quantitative meta-analyses and meta-regressions, we found a strong dose-response relationship amongst the multiple studies identified in the NTP review [[Neurath 2020a](#)]. Furthermore, even when examined one-by-one, many of the higher quality studies with individual-level exposure data also found a clear dose-response relationship [[Xiang 2003](#), [Ding 2011](#), [Valdez-Jimenez 2017](#), [Bashash 2017](#), [Bashash 2018](#), [Green 2019](#), [Till 2020](#)]. Indeed, a recent pooled analysis that combines individual-level data from the two strongest studies in a Benchmark Dose analysis found very strong support for a dose-response relationship with no threshold. It predicted an adverse effect on IQ at exposures at least as low as 0.3 mg/L [[Grandjean 2021](#)]. Also, see our dose-response analysis for full details on all the individual studies that found dose-response relationships [[Neurath 2020a](#)]. Figure 2 below shows examples of dose-response

analyses of several individual studies. Figure 3 shows the dose-response from meta-regression of all the studies scored higher quality by NTP that had individual-level exposure data.

Given this strong evidence of a clear dose-response relationship from individual studies and from pooled results, the GRADE method allows an upgrading of the overall certainty conclusion.



8. The Miranda review design is unable to assess dose-response relationships between F and IQ. Our own dose-response assessment found a strong relationship. Details on the exposure levels and ranges in all studies identified in the NTP review are available in our dose-response assessment [Neurath 2020a]. The NTP also has details on all identified studies, but their extracted information on exposure levels is not as complete as ours, and is not as easily accessible. The NTP data is posted at the HAWC website:

<https://hawcproject.org/assessment/405/>

9. The decision to convert all study results to odds ratios is problematic because it discards information. This may lead to an underassessment of the strength of the evidence. Few of the original studies reported results by odds ratios. Most reported mean IQ by exposure group or difference in mean IQ between exposure groups. Many higher quality studies used individual-level exposures and outcomes and continuous variables and therefore reported results as regression coefficients (Beta values). Converting any of these types of results to odds of having “low IQ” discards information by collapsing continuous scale exposures and/or outcomes into a dichotomous scale outcome of the odds ratio between children having either “low IQ” or “not low IQ”.

The loss of information stemming from this decision may impair the power of the analyses and their ability to detect associations. Consequently, this decision may have led to an underestimate of the strength of the overall body of evidence.

Indeed, the apparent primary reference used by Miranda for meta-analysis methods recommends that the type of effect size used in the meta-analysis should typically be based on the type of effect size used in the primary studies [[Borenstein 2009 book](#)]. Use of odds ratios would only have been appropriate if most of the primary studies reported dichotomous outcomes, but most did not.

All the other systematic reviews of F neurotoxicity with meta-analyses have used difference in mean IQ directly, or if studies used different types of IQ scales they used Standardized Mean Differences (SMDs) [[Choi 2012](#), [Duan 2018](#), [NTP 2020](#)]. None of the other systematic reviews have converted individual study results to odds ratios.

10. The GRADE method of evaluating a body of evidence is designed for clinical interventions intended to improve health, not observational studies of adverse environmental health effects. More appropriate methods for grading environmental health evidence are available.

The Miranda review used the GRADE method for drawing conclusions about the overall confidence in the totality of included studies, yet the GRADE method is intended for evaluating studies of medical interventions where study designs are usually restricted to RCTs (Randomized Controlled Trials). The GRADE method has been criticized as inappropriate for evaluating observational epidemiology studies for questions of environmental epidemiology [[Woodruff 2011](#), [Woodruff 2014](#)]. A specific criticism is that GRADE gives all observational studies an initial rating of “low” confidence, but in environmental epidemiology of adverse effects observational designs are usually the only ethically feasible design. The NTP has done groundbreaking work to develop systematic review methods appropriate for observational epidemiology studies [[Rooney 2014](#), [NTP 2020 OHAT Handbook](#)]. The OHAT method has been recognized by the US National Academies of Sciences, the US EPA, and independent researchers [[NAS 2017](#), [Eick 2020](#)]. The OHAT method is more appropriate for a systematic review of fluoride neurotoxicity than GRADE,

and is being used in the NTP evaluation. Another method tailored specifically for environmental epidemiology questions is the Navigation Guide [Woodruff 2014]. Either OHAT or the Navigation Guide would have been more appropriate than GRADE. Furthermore, a comparison of several different methods of conducting systematic reviews found dramatic differences in outcomes when they were applied to the same set of studies on an environmental epidemiology topic [Eick 2020]. This highlights the subjectivity that can creep into systematic reviews, even at the outset, when choosing a method.

11. The authors' main conclusion is unsupported by their own data. The authors state:

Discussion

This systematic review and meta-analysis gathered evidence showing that, following the WHO classification of low and high levels in the drinking water, exposure to low/adequate water F levels is not associated with any neurological damage, while exposure to high levels is.

Since this review reduced every study to a comparison between just two exposure levels, classified by the reviewers as either “high” or “low/adequate”, it makes no sense to say “high” water F exposure is associated with harm but “low/adequate” exposure is not. There would need to be a comparison of the “low/adequate” exposures to an even lower exposure before any conclusion could be reached that “low/adequate” exposures did not cause harm.

12. The review contradicts its own risk of bias scoring. Despite finding the large majority of individual studies were at low risk of bias (19 of 27 studies or 70% were rated “low risk of bias”) the authors conclude the overall evidence has a “serious risk of bias”. In the wider context of toxic chemicals and developmental neurotoxins in particular, to find 19 higher quality human epidemiological studies, all of which point in the same direction of an adverse effect, represents a much stronger body of evidence than is usually available. The authors' conclusion that the overall body of evidence has “serious risk of bias” is unsupported by their own evaluations.

13. The heterogeneity found in effect size in the meta-analyses can be explained by differences amongst the studies, so is not grounds for downgrading. The authors' failure to consider the different study designs, outcome measures, exposures and exposure contrasts in different studies is another severe limitation of the Miranda review. The heterogeneity of effect magnitude in the meta-analyses may be explained by any of these differences between the different studies. Thus, the authors improperly downgraded the evidence certainty when citing the high degree of heterogeneity, without having tried to find explanations for the heterogeneity. In our own meta-analyses of F-IQ studies identified by NTP, we found that not only dose but also age at outcome evaluation explained much of the heterogeneity in effect magnitude [Neurath 2020a]. There are also other adequate explanations for

heterogeneity of effect magnitude: different study designs, different study populations, different outcome assessments, different ways of measuring exposure. Miranda failed to properly apply the GRADE method, which only allows downgrading when heterogeneity is unexplained.

14. The Miranda review resorts to a subjective “narrative” for choosing its overall evidence conclusion, and thus deviates from the GRADE method. The authors jettison the GRADE method and seem to substitute their own undocumented and unsupported conclusions. In particular, in Table 3, which summarizes how they arrived at their overall certainty conclusion, the only factor that is not satisfactory is “imprecision”, which the authors list as “serious”. Yet they do not offer even a narrative explanation for why they believe there is “serious imprecision” in the overall evidence. If by “imprecision”, they mean heterogeneity of effect magnitude, then as described above, GRADE does not allow downgrading when the heterogeneity is explainable. If by “imprecision” they mean there are excessively wide confidence intervals in many of the individual studies or in the pooled effect estimate, then they are contradicting their own findings. Table 3 states that 24 of 26 studies found statistically significant differences of IQ when comparing the higher exposure group to the lower exposure group. By definition, a study or a pooled result that is statistically significant is precise enough to conclude that there is likely a true difference in IQ associated with exposure. Furthermore, 24 of 26 studies finding statistically significant adverse effects is a noteworthy degree of consistency that is rarely found in systematic reviews of observational human studies, even of well-recognized neurotoxins such as lead and mercury.

15. The Miranda review does not follow the GRADE guidance for evaluating imprecision. The GRADE method for deciding about imprecision is described in detail in Ryan et al 2016, which is the reference cited by Miranda for how it applied the GRADE method [[Ryan 2016](#)]. Miranda does not seem to have followed that GRADE guidance because it would not have allowed a conclusion of severe imprecision. The guidance starts by specifying a minimum sample size in order to calculate sufficiently precise effect estimates. The pooled data from the Miranda meta-analysis clearly has a more-than-adequate sample size (Miranda Figure 2 shows almost 3,000 children with more than 1,100 “events” of reduced IQ while the GRADE guidance states that 300 events indicates a sufficient sample size). The GRADE guidance then asks whether the upper and lower 95% confidence limits are both *consistent* in the sense of both being considered harmful effects, or are *inconsistent* with one showing no effect or a beneficial effect and the other a harmful effect. Only if they are inconsistent might there be grounds for saying there is so much imprecision that the confidence conclusion should be downgraded. In the Miranda meta-analysis there is complete consistency with the lower 95% CL odds ratio being 2.41 and the upper 95% CL odds ratio being 6.23, and the mean effect estimate OR=3.88. All of these ORs should clearly be considered serious adverse effects, with even the smallest magnitude being a 2.41 increase in odds of low IQ in the children with higher F exposure. There is absolutely no inconsistency so there are no grounds in the GRADE method for rating this finding as having severe imprecision, or even a minor amount of

imprecision.

The Miranda authors seem to have ignored the GRADE method and substituted their own subjective decision when they say there is severe imprecision in the effect estimates, and improperly downgraded the initial overall evidence confidence level from “low” to “very low”.

16. The authors also failed to follow the GRADE method that provides for upgrading observational studies when there is evidence of dose-response relationships. As described above, if they had designed their review to be capable of detecting dose-response relationships, they would have found clear evidence of such, which would warrant upgrading by one level when following the GRADE method.

17. The authors also failed to follow the GRADE method that provides for upgrading when the effect is large and consistent. The other grounds for upgrading observational studies in GRADE is when there is a large effect size. As stated above, the average estimate from the pooled meta-analysis was an OR of 3.88, and even the lower 95% CL was a very large OR of 2.41.

The GRADE guidance specifically states that a Relative Risk greater than 2 ($RR > 2$) that is consistent in two or more studies that are of higher quality is sufficient to upgrade by one level [[Ryan 2016, p 15](#)]. The Miranda meta-analysis found a pooled OR of 3.88, which is much greater than 2, and that was based on ten higher quality studies, almost all of which found a statistically significant adverse effect, and none of which found a beneficial effect.

By the GRADE methodology, the certainty rating clearly deserves an upgrade by one level for the large magnitude and consistency of effect.

18. If the GRADE methodology had been faithfully followed, the starting grade of “low certainty” for observational studies would have been upgraded one level for dose-response relationship and another level for large magnitude and consistency of effect. The final grade should therefore have been “high certainty”, the highest grade, rather than the “very low certainty” chosen by the authors, the lowest possible grade. The authors failed to follow their protocol and the explicit guidance of the GRADE method and reached the opposite conclusion from what the scientific evidence supports.

19. The sensitivity analyses were not adequately explained or documented. The reference for how the sensitivity assessment was conducted does not appear to be correct (reference number 15 is to Higgins 2011 which is a Cochrane Collaborative method for assigning RoB scores to RCT studies and

contains no mention of sensitivity analyses, heterogeneity, or publication bias). Lacking a reference with any details on Miranda's method for sensitivity assessment, the only information is this short description by Miranda:

This evaluation was performed by manually omitting one study at time, one by one, and verifying its impact in the final results¹⁵

But no explanation is given as to what criteria was used to decide which studies to exclude in the sensitivity analyses reported in Supplement 5. In Supplement 5, the meta-analysis and funnel plots have excluded 3 out of the 10 higher quality studies used in the main analyses. But it is unclear why these 3 studies were excluded. Two of them (Kahn 2015 and Razdan 2017) had very large effect sizes compared to the remaining 7 studies, and this might have been the reason for excluding them, but the third excluded study (Qin 2008) has no obvious characteristics that would suggest it should be excluded.

20. The authors claim there was evidence of publication bias that justified downgrading the confidence conclusion, but then acknowledge that when 3 apparent "outlier" studies were excluded in a sensitivity analysis there is no longer evidence of publication bias in the remaining 7 studies. The results of the meta-analysis of the 7 remaining studies, presented in Supplement 5, shows the same consistency of statistically significant adverse effects, and almost as large a pooled effect size, as before the exclusion of the 3 studies (pooled result: OR = 3.23 [95%CI 2.18, 4.78]). Since the authors detected no publication bias amongst the 7 studies, they were not justified in downgrading according to GRADE guidance [[Ryan 2016, p 14](#)].

The NTP systematic review, which conducted a much more thorough literature search than Miranda 2021, and also had a wide range of meta-analyses, was able to rule out publication bias as an explanation for the findings of significant adverse effects of F, especially in the higher quality studies [[NTP 2020](#)].

21. Errors in references. Only a few references were checked but two errors were found. The large majority of references were not checked nor was the large majority of data reported in the study. A single original paper was checked against the extracted information reported in Table 1 and an error or incomplete reporting was found (see item 22 below).

- Reference 15 appears to be incorrect and is unrelated to the statement in the article that refers to it.
- Reference 16 is to a 2 page book review of a 456 page book rather than the book itself.

22. There is insufficient transparency and incomplete reporting. For example, from the information made available it is impossible to determine why the eleven high quality studies mentioned in item 4 above were never identified in any searches, or if they were identified, why they were not selected

for further evaluation to decide whether they met the inclusion criteria. The studies are [Jimenez-Valdez 2017](#), [Bashash 2017](#), [Bashash 2018](#), [Yu 2018](#), [Cui 2018](#), [Green 2019](#), [Riddell 2019](#), [Zhao 2021](#), [Yu 2021](#), [Cantoral 2021](#), and [Wang 2021](#).

Another example of possibly incomplete reporting arises in Table 1. The Razdan 2015 paper high exposure level is listed as 2.99 ppm but in the Razdan paper itself there are two places where the high exposure level is given as 4.99 ppm and one place where it is given as 2.99 ppm. Miranda either failed to notice this discrepancy in the original paper or possibly Miranda was able to contact the Razdan authors and determine the correct value was 2.99 ppm. But if Miranda had contacted the Razdan authors and received clarification, they should have noted this so that readers could be assured that Miranda had not mistakenly chosen 2.99 ppm.

23. There are unstated conflicts of interest with the dental industry and fluoridation-promoting organizations. The authors claim they have no conflicts of interest, yet several are dentists working in dental schools and have received funding from dental industry sources such as the major dental pharmaceutical company Colgate [IADR/Colgate grants to [Pessan in 2009, 2013, 2014, & 2020](#)]; and Borrow Foundation funding for Buzalaf [[Buzalaf 2018](#), [Buzalaf 2011](#), [Zohoori 2019](#), [Cassiano 2016](#), [Pessan 2015](#), [Rodrigues 2009](#), etc.] The Colgate company has a clear financial interest because it markets fluoride dental products, and the Borrow Foundation has a goal of promoting the use of fluoride for dental health [[Borrow Foundation website](#)]. Dentists and the dental industry are the main promoters and defenders of fluoridation worldwide. For example, the American Dental Association (ADA) vigorously promotes and defends fluoridation [[ADA policy statement](#)]. The International Association of Dental Research (IADR) has a long-standing policy promoting fluoridation and claiming unequivocally it is “safe and effective” [[IADR 1999](#), [IADR 2018](#)].

Fluoridation and other fluoride modalities have been the most-promoted caries-preventative approaches of dentistry for over 50 years, so any scientific evidence that raises concerns about fluoride’s safety poses serious reputational risks to the dental profession and dental industry, not to mention financial risks. Dental interests have frequently tried to deny or downplay the evidence of fluoride’s neurotoxicity.

24. Conflicts are not just direct financial conflicts, and can be reputational, and indirect, such as the ability to obtain future research funding from private or public sources. Nature journals updated their policy on competing interests in 2018 specifically adding non-financial types of conflicts:

“What makes a conflict of interest in science? Definitions differ, but broadly agree on one thing: an influence that can cloud a researcher’s objectivity. For some people, that influence can be money. But there are other influences that can interfere, such as institutional loyalty, personal beliefs and ambition.”

<https://www.nature.com/articles/d41586-018-01420-8>

Nature's full policy on competing interests is stated here:

<https://www.nature.com/nature/editorial-policies/competing-interests>

25. Given all the problems in the Miranda systematic review, It does not appear to be an objective assessment of the current body of evidence for fluoride's neurotoxicity. It appears instead, that the authors are trying to defend the use of fluoride in dentistry. In contrast, independent scientists from outside the dental field have concluded that the existing evidence is sufficient to warn that pregnant women, infants, and children should avoid fluoridated water and limit other sources of fluoride exposure to reduce the risk of lost IQ.

Leading scientists in the field of environmental epidemiology, including those specializing in neurotoxicity, along with medical experts from outside the dental field, have concluded the current evidence for fluoride's developmental neurotoxicity warrants a warning to pregnant women, infants and children to avoid exposure to fluoride, including artificially fluoridated water. The experts include: [Grandjean 2019](#), [Bashash & Hu 2020](#), [Lanphear & Till & Birnbaum 2020](#), [Till & Green 2021](#), [Till & Lanphear 2021](#); and editors of the Journal of the American Medical Association (JAMA) [Christakis & Rivara 2019](#). Dr. Linda Birnbaum is not only a leading researcher in environmental health, she is also the former director of NIEHS and NTP [[Lanphear & Till & Birnbaum 2020](#)]. In contrast to these experts in the field of environmental health, it is mostly dentists and dental organizations that have downplayed or denied the evidence of fluoride's neurotoxicity [[American Dental Association \(ADA\) 2021](#), [International Association of Dental Research \(IADR\) 1999](#), [IADR 2018](#), [American Association of Public Health Dentistry 2010](#), [CDC Oral Health Division 2021](#)].

In conclusion, the Miranda systematic review of fluoride neurotoxicity has numerous weaknesses which raise concerns about whether it is an objective assessment of the evidence and whether it has improperly downgraded its evidence conclusion. If the GRADE method were properly applied to the overall body of evidence, the confidence that F is neurotoxic would be rated "high certainty" rather than the rating assigned by Miranda of "very low certainty".