Submission to the NAS and NTP

Dose-Response assessment of fluoride neurotoxicity studies

Chris Neurath
Research Director
Fluoride Action Network

November 30, 2020

TABLE of CONTENTS

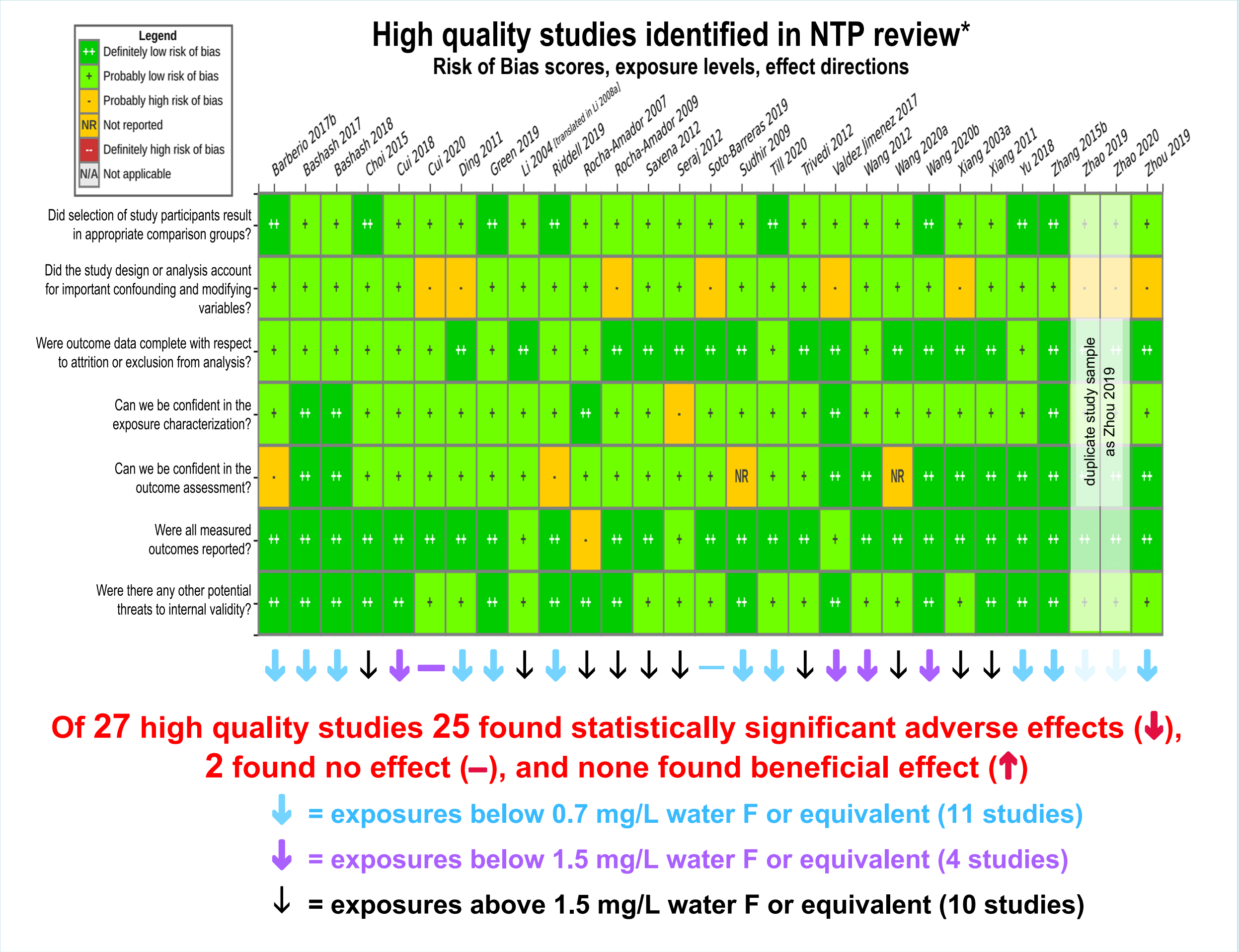
1. Summary of dose-response assessment of studies scored by NTP as higher quality

We conducted three stages of dose-response assessment. The results of each stage can be summarized in three figures:

Figure 1. Risk of Bias heat map with added indicators of exposure level and effect direction.
Figure 2. Forest plot, subgroup meta-analysis by exposure level.
Figure 3. Meta-regression plot, effect size by exposure level.

Following these three summary figures are sections describing methods; the justification for conducting dose-response assessments; and additional analyses.

1.1. Main results



Legend

++	Definitely low risk of bias
+	Probably low risk of bias
-	Probably high risk of bias
NR	Not reported
--	Definitely high risk of bias
N/A	Not applicable

High quality studies identified in NTP review*

Risk of Bias scores, exposure levels, effect directions

	Barberio 2017b	Bashash 2017	Bashash 2018	Choi 2015	Cui 2018	Cui 2020	Ding 2011	Green 2019	Li 2004 [translated in Li 2008a]	Riddell 2019	Rocha-Amador 2007	Rocha-Amador 2009	Saxena 2012	Seraj 2012	Soto-Barreras 2019	Sudhir 2009	Till 2020	Trivedi 2012	Valdez Jimenez 2017	Wang 2012	Wang 2020a	Wang 2020b	Xiang 2003a	Xiang 2011	Yu 2018	Zhang 2015b	Zhao 2019	Zhao 2020	Zhou 2019
Did selection of study participants result in appropriate comparison groups?	++	+	+	++	+	+	+	++	+	++	+	+	+	+	+	+	++	+	+	+	+	++	+	+	++	++	+	+	+
Did the study design or analysis account for important confounding and modifying variables?	+	+	+	+	+	-	-	+	+	+	+	-	+	+	-	+	+	+	-	+	+	+	-	+	+	+	-	-	-
Were outcome data complete with respect to attrition or exclusion from analysis?	+	+	+	+	+	+	++	+	++	+	+	++	++	++	++	++	+	++	++	+	++	++	++	++	+	++	+	++	++
Can we be confident in the exposure characterization?	+	++	++	+	+	+	+	+	+	+	++	+	+	-	+	+	+	+	++	+	+	+	+	+	+	++	+	+	+
Can we be confident in the outcome assessment?	-	++	++	+	+	+	+	+	+	-	+	+	+	+	+	NR	+	+	++	++	NR	++	++	++	++	++	+	++	++
Were all measured outcomes reported?	++	++	++	++	++	++	++	++	+	++	-	++	++	+	++	++	++	++	+	++	++	++	++	++	++	++	++	++	++
Were there any other potential threats to internal validity?	++	++	++	++	++	+	+	++	+	++	++	++	+	+	+	++	+	+	++	+	+	++	+	++	++	++	+	+	+
	↓	↓	↓	↓	↓	—	↓	↓	↓	↓	↓	↓	↓	↓	—	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓

duplicate study sample as Zhou 2019

Of 27 high quality studies 25 found statistically significant adverse effects (↓), 2 found no effect (–), and none found beneficial effect (↑)

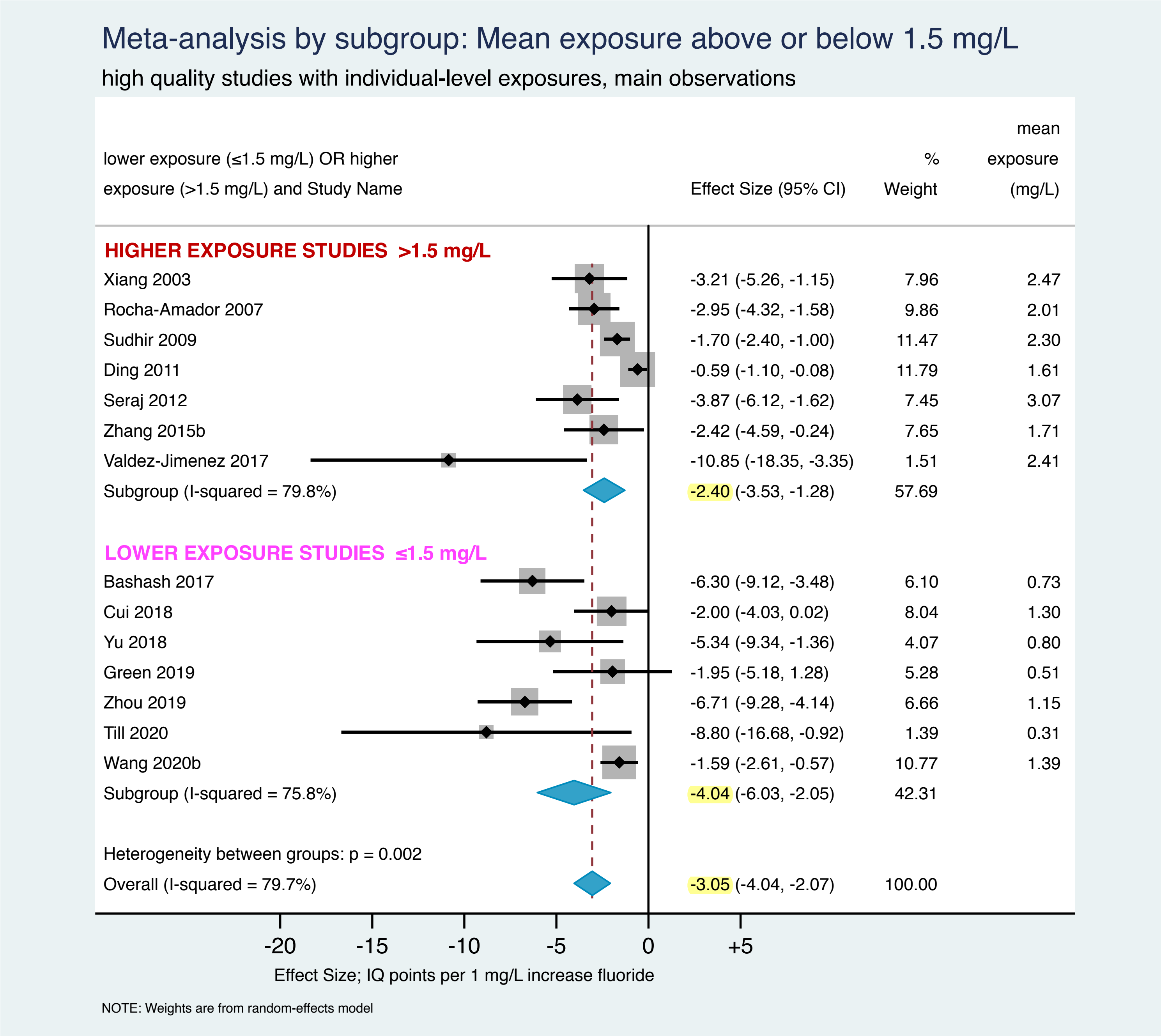
↓ = exposures below 0.7 mg/L water F or equivalent (11 studies)

↓ = exposures below 1.5 mg/L water F or equivalent (4 studies)

↓ = exposures above 1.5 mg/L water F or equivalent (10 studies)

Figure 1. Risk of Bias (RoB) heat map with added exposure level and effect direction indicators. Heat map is for the lower RoB studies identified by NTP. Green color for RoB domains indicates lower RoB. Criteria for assigning studies to exposure level categories are described in FAN submission to NAS of November 5, 2020. Original figure from ISEE 2020 Conference poster [Neurath et al 2020].

Summary of findings shown in Figure 1. When considering the 27 studies identified by NTP as higher quality, there was great consistency in finding statistically significant adverse neurotoxic effects at exposure levels below 1.5 mg/L fluoride in water or equivalent in urine. Over 90% of such studies found adverse effects. Thus, the studies identified and scored as higher quality by NTP contradict NTP's conclusion that the evidence below 1.5 mg/L is "inconsistent" and "unclear". The evidence below 1.5 mg/L is remarkably consistent. Furthermore, when examining just those studies finding adverse effects below 0.7 mg/L, there is equally high degree of consistency, with 11 out of 12 such studies finding an adverse effect.



lower exposure (≤1.5 mg/L) OR higher exposure (>1.5 mg/L) and Study Name	Effect Size (95% CI)	% Weight	mean exposure (mg/L)
HIGHER EXPOSURE STUDIES >1.5 mg/L			
Xiang 2003	-3.21 (-5.26, -1.15)	7.96	2.47
Rocha-Amador 2007	-2.95 (-4.32, -1.58)	9.86	2.01
Sudhir 2009	-1.70 (-2.40, -1.00)	11.47	2.30
Ding 2011	-0.59 (-1.10, -0.08)	11.79	1.61
Seraj 2012	-3.87 (-6.12, -1.62)	7.45	3.07
Zhang 2015b	-2.42 (-4.59, -0.24)	7.65	1.71
Valdez-Jimenez 2017	-10.85 (-18.35, -3.35)	1.51	2.41
Subgroup (I-squared = 79.8%)	-2.40 (-3.53, -1.28)	57.69	
LOWER EXPOSURE STUDIES ≤1.5 mg/L			
Bashash 2017	-6.30 (-9.12, -3.48)	6.10	0.73
Cui 2018	-2.00 (-4.03, 0.02)	8.04	1.30
Yu 2018	-5.34 (-9.34, -1.36)	4.07	0.80
Green 2019	-1.95 (-5.18, 1.28)	5.28	0.51
Zhou 2019	-6.71 (-9.28, -4.14)	6.66	1.15
Till 2020	-8.80 (-16.68, -0.92)	1.39	0.31
Wang 2020b	-1.59 (-2.61, -0.57)	10.77	1.39
Subgroup (I-squared = 75.8%)	-4.04 (-6.03, -2.05)	42.31	
Heterogeneity between groups: p = 0.002			
Overall (I-squared = 79.7%)	-3.05 (-4.04, -2.07)	100.00	

Figure 2. Forest plot showing results of subgroup meta-analysis for exposures above 1.5 mg/L compared to those with exposures below 1.5 mg/L. Random-effects meta-analysis.

Summary of findings shown in Figure 2. The remarkable consistency of statistically significant adverse effects, both above and below 1.5 mg/L, is further illustrated with this meta-analysis forest plot of 14 higher quality studies with individual-level exposures. In the subgroup of studies with mean exposures below 1.5 mg/L all 7 studies found adverse effects, with 6 of the 7 being statistically significant. The same numbers of studies above 1.5 mg/L had adverse effects and statistically significant adverse effects. The pooled effect size estimates were highly significant in both exposure subgroups, but was larger in those with mean exposures below 1.5 mg/L compared to the subgroup of studies with exposures above 1.5 mg/L (-4.04 compared to -2.40 IQ points per 1 mg/L increase in fluoride exposure). The consistency and strength of evidence below 1.5 mg/L is at least as great as that above 1.5 mg/L. The larger effect at lower doses is similar to what has been found with childhood lead exposure [Lanphear et al 2005]. In the Lanphear et al 2005 pooled analysis, they reported an estimated IQ decrement in the lowest blood lead dose range from 2.4 to 10 µg/dL of 3.9 IQ points while for an increase over the higher dose range 10 to 20 µg/dL the loss was only 1.9 IQ points. Note that these values of IQ loss from childhood lead exposure can genuinely be described as "on par" with that from fluoride. The similar magnitude of effects of lead and fluoride are discussed further below.

While there is heterogeneity amongst the studies, it is explainable by differences in study design, exposure measures, and outcome measures. All of these reasons for difference in effect sizes amongst studies are considered under OHAT guidance to not constitute grounds for considering the body of evidence to be inconsistent [NTP 2019, OHAT Handbook p52-53]. A section describing additional analyses, below, further discusses factors explaining heterogeneity.

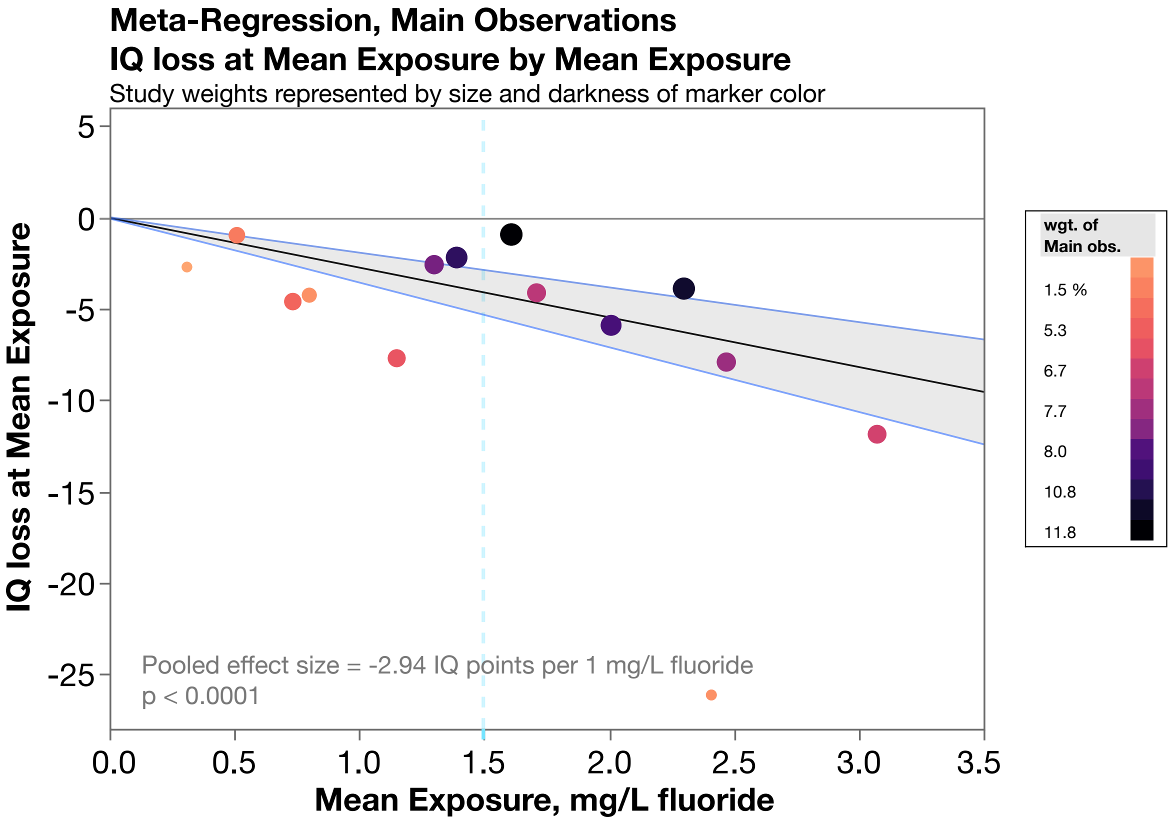


Figure 3. Meta-regression plot; predicted effect size at mean exposure level by mean exposure level. Individual study observations superimposed on regression plot with color-coded dots. Darker colors and larger size of dots represent greater weight. Weighted least-squares linear regression model with specified no intercept. Gray shaded area represents 95%CI.

Summary of findings shown in Figure 3. This refinement of the dose-response assessment goes beyond stratifying exposures into two groups (above 1.5 mg/L or below 1.5 mg/L) and shows the relationship between exposure treated as a continuous variable and effect size. It demonstrates a highly significant linear dose-response relationship with a substantial effect size, including at levels below 1.5 mg/L. Non-linear meta-regression models were also considered but did not fit the data better than a linear fit.

1.2. Discussion

Each of the three approaches to dose-response assessment supports a conclusion of clear and robust consistency of neurodevelopmental harm from fluoride at exposures below 1.5 mg/L. This exposure level is considered relevant to the US population, including to those drinking artificially fluoridated water at 0.7 mg/L. Linear dose-response meta-regression models support a conclusion that there is no identified threshold down to exposure levels at least as low as 0.7 mg/L.

While there is heterogeneity of effect magnitude, there is no inconsistency in effect direction, with all 27 of the higher quality studies finding adverse effects, 25 of which were statistically significant. The same high degree of consistency of effect direction was found in studies with exposures below 1.5 mg/L as in studies above 1.5 mg/L. Furthermore, the variations in effect magnitude may all be explainable by differences in study designs, exposure measures, and outcome measures. All of these explanations are considered by NTP's systematic review guidance [OHAT 2019 handbook] to be acceptable explanations precluding any downgrading of conclusions about overall confidence in the evidence.

We explored several factors which explained heterogeneity of effect magnitude. Most obviously, the dose (exposure to fluoride) explained variation in effect magnitude. The subgroup meta-analysis comparing studies above 1.5 mg/L to those below 1.5 mg/L demonstrates this dose effect (Figure 2) as does the dose-response meta-regression (Figure 3). Another study design factor which explained variation in effect magnitude was the Exposure Period. When stratified into "early life" and "later childhood" the studies that measured early life exposures found twice as great loss of IQ per 1 mg/L fluoride as the studies that measured exposures in later childhood (Figure 8). The related study design factor "age at outcome assessment" was also found to explain some of the variation in effect magnitude. A meta-regression model that included both dose (Mean exposure) and mean age at outcome assessment as covariates, found statistically significant linear effects for both factors (Figures 11 and 12, Table 3). Predicted IQ decreased with increasing dose and increased with increasing age of outcome assessment. There is some correlation between age at outcome assessment and exposure period, since many of the studies were cross-sectional and measured fluoride exposure at the same time as outcome assessment, rather than prospectively.

One of the studies seems to stand out as an "outlier" with much greater effect magnitude than the others. That is the Valdez-Jimenez 2017 study. However, there are several study design differences specific to this study which may explain this unusually large effect size. First, it assessed outcome at a much younger age than any of the other studies. Using the Bayley Scales of infant development, it assessed toddlers at a mean age of 0.69 years, while all the other studies assessed outcomes at age 3 years or older, with a pooled mean age of 8 years. Also, it was the only study to use the Bayley Scales for outcome assessment. Most of the other studies used versions of the Ravens or Wechsler Scales. Second, the study population had low Socioeconomic Status as measured by a marginalization index. None of the mothers had any college education and only 12% completed high school. Third, there was a high rate of premature birth in the cohort (34%) suggestive of relatively poor overall health in the sample population. Each of these factors may explain why fluoride exposure was associated with a larger effect size in this study than in any of the other studies.

1.3. Conclusion

Rigorous dose-response assessment methods, applied to the studies NTP identified as higher quality, strongly support a conclusion that there is great consistency of adverse neurodevelopmental effects – in particular reduced IQ – at fluoride exposure levels relevant to the US population from artificial fluoridation. The NTP's own dose-response assessment has numerous weaknesses which led it to an unsupportable conclusion that there is "inconsistency" for exposures below 1.5 mg/L. The NTP monograph should be revised to include a rigorous and valid dose-response assessment or any discussion dependent on dose-response assessment should be eliminated, as previously recommended by the NAS committee.

> *Rigorous dose-response assessment methods ... strongly support a conclusion that there is great consistency of adverse neurodevelopmental effects ... at fluoride exposure levels relevant to the US population from artificial fluoridation.*

2. Methods for dose-response assessment

2.1. General methods

1. We focused on the human epidemiological studies the NTP scored as lower Risk of Bias (RoB).

2. We gave special attention to the studies with individual-level exposure information, because they are the most informative for dose-response assessment. They also tended to be the highest quality studies and many were in the lower dose ranges. Also, the NTP planned to conduct dose-response meta-analyses with these studies but then chose not to without good justification [FAN 2020-Oct-19 submission to NAS on revised NTP monograph].

3. When possible, we used data extracted by NTP for the results and the exposures. For some studies that had multiple analyses using different subgroups or different exposure or outcome measures, the NTP chose to use analyses with smaller effect sizes, thereby reducing the pooled effect sizes, their statistical significance, and their consistency. When a study had multiple analyses, we instead chose the analysis with the largest statistically significant effect, with the additional criterion that it has the most adjustment. This approach is consistent with standard risk assessment methodology, such as that of the EPA [EPA 2020, website on risk assessment].

4. For studies with available individual-level exposure data, or those having at least three different exposure-level groups, we applied Benchmark Dose (BMD) assessment methods to estimate the dose expected to produce an average 1 IQ point loss. The BMD method is the preferred method of dose-response assessment by the EPA. We chose -1 IQ points as the Benchmark Response (BMR) because the EPA has offered guidance that for developmental neurotoxicity, an average predicted loss of 1 to 2 IQ points in any subpopulation is an unacceptable risk. For studies without suitable data to conduct BMD assessments, the mean exposure of the higher of the two exposure groups was used in the dose-response meta-analyses described below. For those with individual-level continuous exposure data the mean exposure was used in dose-response meta-analyses described below.

5. Our first stage of dose-response assessment was to count and determine the proportion of the studies finding statistically significant adverse effects at exposures below 1.5 mg/L fluoride in water (or its equivalent in urine). For the studies where we were able to conduct BMD assessments, the calculated BMD at -1 IQ point was used to classify the study into one of three exposures levels: ≤0.7, >0.7 to ≤1.5, and >1.5 mg/L. We used the calculated BMD rather than the calculated BMDL (Benchmark Dose Lower Bound), leading to less protective results. In this respect we deviated from standard risk-assessment and dose-response assessment methodology, which focuses on the more protective BMDL. We compared the number and proportion of studies finding adverse effects below 1.5 mg/L to those finding adverse effects above 1.5 mg/L. This was the simplest way of evaluating consistency for lower doses compared to higher doses.

6. In a second stage of dose-response assessment, we used meta-analyses, including subgroup meta-analyses by exposure level, to estimate pooled effects and examine heterogeneity.

7. In a third stage of dose-response assessment, we use meta-regression, with exposure level being the independent variable and predicted absolute IQ loss at mean exposure the dependent variable. We examined the shape of the pooled dose-response relationship and whether it was statistically significant when including studies over the full range of exposures, and also for the restricted set of studies with exposures below 1.5 mg/L.

8. Much of the data extraction phase of our dose-response assessment, along with the first stage results, was already reported and submitted to the NAS and NTP in our submissions of November 5, 2020 [FAN 2020-Nov-5 Exposure classification protocol and individual study details for dose-response assessment].

9. To make the dose-response assessment applicable to the population of the US, an exposure assessment is required. We adopted the US EPA's exposure assessment for drinking water [EPA 2019 Exposure Factors Handbook]. The EPA found that high-end consumers of tap water (the top 95th percentile) drink more than twice as much water as the average consumer. Thus, 5% of the US population will receive at least as great an internal dose when drinking water with a concentration of 0.7 mg/L as will the average population of those drinking water with a 1.5 mg/L concentration of fluoride. We therefore conclude that studies at an average concentration of 1.5 mg/L, or lower, are relevant to drinking water concentrations of 0.7 mg/L in the US, the target level for most artificially fluoridated water systems in the US. This conclusion is consistent with that of the NTP in the revised monograph, but for a different reason. The NTP reasoned that a certain portion of the US population had water with natural fluoride levels of about 1.5 mg/L. Our reasoning is that there is a large subpopulation in the US who will consume at least twice as much water at 0.7 mg/L as the average and thus receive the equivalent internal dose as the average consumer drinking 1.5 mg/L water.

2.2. Detailed methods of meta-analyses

Meta-analyses were done with Stata 15 software and the **admetan** extension. Full details are available in the Appendix which includes a Stata .do file.

Meta-analysis pooling of aggregate data used the random-effects inverse-variance model with DerSimonian-Laird estimate of tau^2.

Data was either obtained from NTP-extracted data available in HAWC, or extracted directly from original published papers. The full data set is available in an attached Excel file.

When a study had more than one result (observation), the "Critical observation" was defined as the result with the largest effect size that was statistically significant. This was typically a subsample of the full sample or a specific exposure measure when more than one exposure measure was assessed in the study. The "Main observation" was defined as the result using the full sample that had the most adjustment.

Exposure measures were the concentration of fluoride (F) in drinking water or in urine. Urine fluoride measures in units of mg F/$g_{creatinine}$ were converted to mg/L F. Urine F adjusted for creatine was converted to urine F adjusted for specific gravity. Following NTP guidance, urine F concentrations were considered to have a 1:1 equivalence to water F concentrations for the purposes of pooling observations.

Mean exposures and Standard Deviations were extracted from the NTP's HAWC database or from original published papers, if reported. If not reported, then medians or midpoints of ranges were used, following the revised NTP monograph guidance.

Effect sizes were regression coefficients (Beta values) with units of "IQ change for each 1 mg/L increase in fluoride concentration". All except Valdez-Jimenez 2017 were for linear regression models. The effect size for Valdez-Jimenez 2017 was reported as a log-linear model coefficient. For use in the meta-analyses we calculated the difference in IQ for the 1 mg/L increase centered on the mean exposure level for the entire sample.

All outcomes were neurocognitive function tests in units of IQ score (or equivalent) with mean and SD of approximately 100±15. Since similar scales were used for all study outcomes, no standardization was performed.

If regression coefficient 95% confidence intervals were not reported in original published papers, then the methods of the Cochrane Handbook for conducting meta-analyses were used to calculate 95%CIs from Standard Deviations, Standard Errors, or *p*-values [Cochrane 2020 Handbook for Systematic Reviews].

2.3. Detailed methods of meta-regressions

Summary

Meta-regressions were done with JMP 14 software, using the study weights generated in meta-analyses using the **admetan** extension in Stata 15 [Fisher 2015].

For most analyses, ordinary least squares regression models were used. For spline models, a 3-knot cubic regression-spline model was used, based on the Stone and Koo method [JMP help 2020, Stone and Koo 1985].

The independent variable in dose-response meta-regressions was the Mean exposure for the study observation, as described above. The dependent variable was the effect size as absolute predicted IQ loss at the Mean exposure.

Additional meta-regressions exploring the reason for heterogeneity had the independent variable Mean Age at Outcome Assessment for each study. When a study did not report mean age, but instead the range of ages, the midpoint of the range was used.

The meta-regressions are based on the methods used by Vlaanderen et al [2010]. Briefly, the Mean Exposure for each study was extracted from the NTP's HAWC database, if available, otherwise from the original published paper. The Mean Effect Size was similarly extracted along with 95%CI or ±SD or ±SE. Most studies reported effect sizes as the linear regression model coefficient and we standardized them to the difference in IQ for a 1 mg/L increase in fluoride. All studies found adverse effects of fluoride on IQ so all found lower IQ with increased fluoride exposure. The predicted loss of IQ at the Mean Exposure was then calculated and this was used as the outcome in the meta-regressions. A total of 14 studies had suitable data.

Weights were calculated for each study using a random effects meta-analysis in the **admetan** extension in Stata 15. Forest plots from the meta-analyses, with weights, are reported separately (see Figures 2 and 4-8).

To force the meta-regressions to zero effect size at zero exposure "no intercept" was specified for the regression models. The weights obtained in the meta-analyses were applied in the regression models. A variety of models with different dose-response curves were examined but none gave a better fit to the data than linear least squares models so we focused on results for linear models.

Four primary subgroupings of studies were examined to investigate possible sources of heterogeneity and to better understand the dose-response relationship:

1. Main observations, all exposures
2. Critical observations, all exposures
3. Main observations, restricted to studies with mean exposures below 1.5 mg/L fluoride
4. Critical observations, restricted to studies with mean exposures below 1.5 mg/L fluoride

The criteria for Critical observations was applied to studies with multiple observations in several subsamples or using several exposure or outcome measures. The Critical observation was the result with the largest statistically significant effect size, amongst adjusted estimates. Conversely, the Main observation was for the entire sample but was also for the largest statistically significant effect amongst results for the entire sample. For some studies the Critical observation was also the Main observation.

Results of meta-regressions are presented in graphs with the regression line and 95% confidence interval plotted along with superimposed individual study data points. The data points indicate weight by size of the circles and also by the shade and darkness of color from yellow to black having greater weight. Additional graphs display the study names to label the data points.

As described previously urine fluoride concentrations were considered equivalent to drinking water fluoride concentrations for the purposes of the meta-analyses and meta-regressions.

3. Reasons for conducting dose-response assessment

The Fluoride Action Network (FAN) has previously submitted comments on both drafts of the NTP's monograph on fluoride neurotoxicity [FAN 2019-Nov-6, FAN 2020-Oct-19]. While we have applauded the NTP literature search, data extraction, and evidence synthesis, we have raised concerns that the NTP has stepped outside their study question with what amounts to a dose-response assessment. The NTP monograph includes conclusions about the strength of evidence that fluoride poses a hazard to developing brains at exposure levels occurring in the US.

Despite the NAS recommendation to avoid such a discussion, the revised NTP monograph retains its section on "Generalizability to the US Population" and has even gone further with dose-response assessment by adding dose-response meta-analyses.

If NTP had done a valid and credible dose-response assessment, FAN would not object to its inclusion. The question of the safety of artificial water fluoridation is the "elephant in the room" and should not be side-stepped. We believe the NTP has gathered more than sufficient evidence to address the question head-on today. To delay addressing this question risks harming the developing brains of millions of children born every year with exposures from water fluoridation.

Sir Austin Bradford-Hill, at the conclusion of his article setting out his classic principles for evaluating causation in epidemiological studies, posed a warning that is relevant to the current state of evidence of fluoride's developmental neurotoxicity:

> All scientific work is incomplete – whether it be observational or experimental. All scientific work is liable to be upset or modified by advancing knowledge. That does not confer upon us a freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a given time. [Bradford-Hill 1965]

To put off the question of whether exposures in the US are likely to be causing neurodevelopmental harm is to risk causing population-wide loss of IQ at a similar level as occurred from leaded gasoline before it was banned.

In just the past year, independent experts have publicly expressed concern that fluoridated water may be posing a risk on par with that of lead (Pb). They include the recently retired director of the NTP, Dr. Linda Birnbaum [Lanphear, Till, & Birnbaum 2020]. They include Dr. Bruce Lanphear, who played a major role in developing the evidence upon which NTP concluded even low levels of Pb were harmful [Lanphear 2005, NTP 2012]. Dr. Phillipe Grandjean, a recognized leader in uncovering the evidence of mercury's harm to the developing brain, and more recently of perfluorinated compounds' immunotoxic effects at extremely low levels, has made the same recommendation [Grandjean 2019]. So have Dr. Howard Hu and Dr. Christine Till, principal investigators of the NIH-funded studies of fluoride neurotoxicity [Hu 2020, Lanphear, Till, & Birnbaum 2020]. So have the editors of JAMA Pediatrics, after an exhaustive editorial- and peer-review of the Green 2019 paper [Christakis & Rivara 2019 podcast].

Given the opinions of these authorities, who have taken a careful look at the current evidence, FAN believes the evidence of fluoride neurotoxicity at exposures common in the US must be addressed and not put off to another day. FAN thus offers the present rigorous dose-response assessment, based on the human studies NTP identified as higher quality.

Detailed comparison of effect magnitude of fluoride to lead

A direct comparison can be made between the F-IQ dose-response and that for Pb-IQ, as well as between exposures to Pb and F common in the US population at various times. The pooled studies dose-response regression model for Pb and IQ is summarized in the figure, reproduced below, from the Lanphear et al [2005] study:

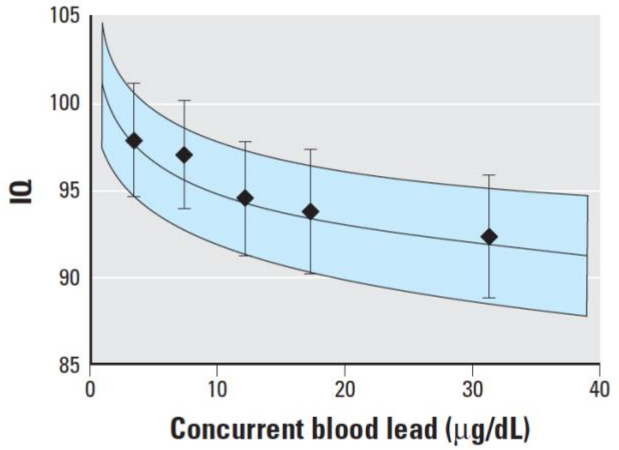


Figure 3. Log-linear model (95% CIs shaded) for concurrent blood lead concentration, adjusted for HOME score, maternal education, maternal IQ, and birth weight. The mean IQ (95% CI) for the intervals < 5 µg/dL, 5–10 µg/dL, 10–15 µg/dL, 15–20 µg/dL, and > 20 µg/dL are shown.

[from Lanphear et al 2005]

The dose-response relationship plotted above can be compared to the dose-response meta-regression plots for fluoride (Figures 3, 9, 10, and 11)

At the time most of the studies in this pooled analysis for Pb were conducted, in the 1990s, the mean blood Pb level in the US in children age 0-5y was about 3.6 µg/dL (NHANES III 1988-1994). In the figure above, the lowest exposure interval (black diamonds) has a mean blood Pb level of 3.4 µg/dL. In 2015-2016, the most recent NHANES survey found a mean level of 1.0 µg/dL and the proportion of children with blood Pb > 10 µg/dL had decreased from 9.0% to 0.2% from the 1990s NHANES III survey.

The pooled estimate of the Effect Size (Beta) in the Lanphear 2005 meta-analysis was a loss of 6.2 IQ points for an increase in blood Pb from <1 to 10 µg/dL. According to the latest NHANES data, with only 0.2% of children exceeding 10 µg/dL, only 0.2% of children will have a loss of IQ as large as 6.2 IQ points. The loss of IQ to the child with average exposure of 1.0 µg/dL will be virtually undetectable. The average loss of IQ over the entire population back in the 1990s would have been about 1/3rd of 6 IQ points because the average exposure was 3.6 µg/dL at that time. Thus, the average loss would have been about 2 IQ points. This is considered to be a serious adverse effect on a population-wide basis and was the reason the CDC lowered its level of concern from 10 µg/dL to 5 µg/dL [CDC 2020].

When experts have said the effect of F on IQ is “on par” with that of Pb, they are referring to Pb exposures common about 30 years ago, not for exposures common today. In the 1990s, about 10% of children would have been expected to have lost about 6 IQ points, on average, because of Pb, and the overall population loss would have amounted to roughly 2 IQ points.

The mean maternal urine F level today in fluoridated parts of the US is about 0.7 mg/L [Uyghurturk et al 2020]. The lowest observed levels are about 0.1 mg/L or lower. The average predicted loss of IQ for a 1 mg/L increase in F is about 4 IQ points, based on our dose-response meta-regression (Figure 2). This pooled dose-response effect appears to have no lower threshold and is based on some studies with exposures as low as about 0.1 mg/L (ELEMENT and MIREC cohort studies), so it is not an extrapolation.

The average loss of IQ in fluoridated areas is 4 * (0.7 - 0.1 mg/L) or 2.4 IQ points. This is at least as large as the average loss of IQ in the 1990s from Pb, which is why experts have said the adverse effects on IQ from F is "on par" with that from Pb. But they are comparing today's F effects to the effects of Pb exposures about 30 years ago when Pb levels were about 3 times greater than today.

4. Comments on NTP's dose-response meta-analyses

The only quantitative dose-response type assessment conducted by NTP was dose-response meta-analyses. However, the NTP used inappropriate methods in these meta-analyses, and misinterpreted some of the results. Specifically:

1. Although the revised protocol planned to do dose-response meta-analyses on group-level exposure studies and on individual-level exposure studies, the NTP monograph abandoned dose-response meta-analyses of individual-level studies. This is a serious limitation, because the studies with individual-level exposure information tended to be the highest quality studies and tended to be done at lower exposures than the group-level studies.

2. In the meta-analyses and presumably the group-level dose-response meta-analyses, the NTP made inappropriate choices for which observations to use for several of the studies. The NTP's choices, in all cases, led to underestimating adverse effect sizes and underestimating the consistency of adverse effects. The Green 2019 study is the best example, where an inappropriate observation was chosen for the group-level meta-analysis (comparison of mean IQ for urine F ≤0.8 mg/L compared to >0.8 mg/L which was never intended as an analysis and has very low statistical power), and an inappropriate observation was chosen for the individual level meta-analyses (IQ at age 6-12 years instead of GCI at age 4 years). It is not clear what observation from Green 2019 was used in the group-level meta-analyses since no details were reported for which observations were included in each of these.

3. The NTP's group-level dose-response meta-analyses inappropriately stratified by urine F and water F, thereby reducing their power compared to if these had been combined.

4. The NTP's group-level dose-response meta-analyses inappropriately focused on a subset of observations with mean exposures below 1.5 mg/L instead of using studies with observations at all levels of exposure. This again reduced the statistical power of the dose-response meta-analysis, and subverts the basic principle of using all available information across a range of doses to estimate the dose-response relationship.

5. The NTP used only the mean exposure to assign an exposure level to a study, even though most studies had exposure groups with a wide range of exposures, including exposures much lower than the mean. A more appropriate method of dose-response assessment is to use all the available information, such as with Benchmark Dose (BMD) methods. For many studies, BMD methods demonstrate that the dose-response curve predicts substantial adverse effects even at doses much lower than the mean exposures. When we applied BMD methods to studies with suitable data where there were at least 3 exposures groups, most of those studies support a prediction of adverse effects at exposure levels below the mean exposure levels.

Our dose-response assessment corrected these and other errors or inappropriate methods used by NTP. This document describes our methods and results in detail, including where they differ from those of the NTP.

To address the first problem listed above in NTP's dose-response assessment (failure to do dose-response meta-analysis with individual-level studies), an explanation may be that the NTP's protocol described a method that is not suitable for individual-level studies, only group-level studies. The protocol specified that NTP would use the **drmeta** extension with Stata statistical software to conduct dose-response meta-analyses. However, the authors of the **drmeta** extension [Filippini et al 2019] state this method is not helpful for identifying deviations from a linear dose-response relationship:

> We also performed a dose-response meta-analysis to understand the shape of the curve relating air pollution and disease risk. To do that, we used methodology developed by Greenland and Longnecker (1992) and Orsini et al. (2012) that has been applied in other contexts (Crippa et al. 2018b; Vinceti et al. 2016) to estimate the trend from the RRs across categories of pollutant exposure levels and their approximate pointwise 95% confidence intervals (CIs) based on asymptotic normality. ... *We excluded from this dose-response meta-analysis the studies not reporting any exposure category cut points for the investigated pollutant and the studies providing only RR estimates based on 1-unit increment in exposure based on a linear model because these studies could not contribute to the assessment of departure from linearity*. [Filippini et al 2019; *emphasis added*]

However, amongst the studies of fluoride and IQ, most individual-level studies that examined a variety of dose-response relationships besides linear, found, just as we did, that a linear dose-response model fit as well as any non-linear models considered. There is little evidence for a threshold within the range of observations in most of the studies, which are as low as 0.2 mg/L.

To allow the higher quality individual-level studies to be analyzed in a dose-response meta-analysis, we used meta-regression with mean exposure as the independent variable. We followed the method used by Vlaanderen et al [2010] in their exploration applying different meta-regression models to observational epidemiology studies of the association between occupational benzene exposures and leukemia risk.

5. Additional meta-analyses, meta-regressions, and dose-response analyses

All our meta-analyses and meta-regressions were restricted to the studies NTP rated as higher quality, or lower Risk of Bias (RoB). We also restricted to just those studies with individual-level exposure measures. These tended to be the highest quality of the higher quality studies, and also tended to be studies with relatively low exposure levels. Thus, they are especially pertinent to assessing the dose-response relationship at exposure levels relevant to the United States, and to artificial fluoridation.

Consistent with standard risk assessment methods, we identified observations from each study that were considered the "critical observations", which are the observations with the largest effect size. In risk assessment, the goal is to protect all members of the population, including those in the most sensitive subpopulations or the subpopulations with the highest exposures. All the subpopulations we identified for critical observations were substantial in size, and would constitute millions of Americans.

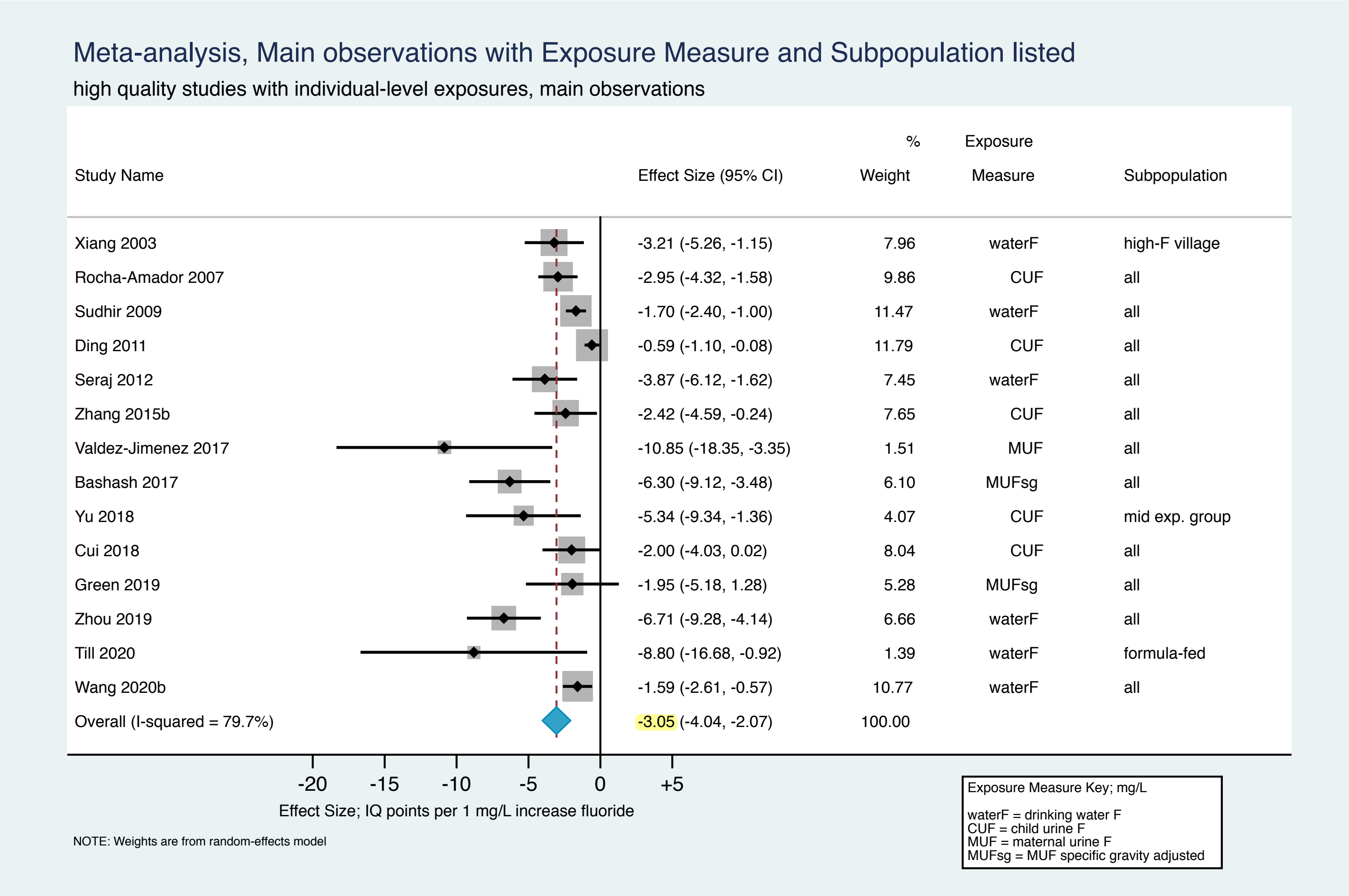
Many studies had multiple observations, examining different subpopulations or using different exposure measures or outcome measures. In those studies with multiple observations, our criteria for determining the critical observation was that it have at least as much adjustment as other observations, and that it have the largest statistically significant effect size. In specific studies, this critical observation was in a particular subpopulation by gender (boys in Green 2019), or in a subpopulation by genetic variant (val/val variant in COMT gene in Zhang 2015, TT variant in DRD2 gene in Cui 2018). In one study, the critical observation was for the outcome measured at a certain age (age 4 years for Bashash 2017).

The "main observations" were the observations in a study that included the full study sample and that were considered to be the primary results by the study authors. When there were multiple results considered to be "primary" by the study authors, we chose that with the largest statistically significant effect size. For example, in Bashash 2017, the results for all children at the age 4 outcome assessment had a larger effect size than at the age 6-12 assessment. Both results were considered "primary" results by the study authors, so we chose the age 4 results. In another example, the Till 2020 study reported results for formula-fed infants separately from those for exclusively breast-fed infants. The result for formula-fed infants was larger than for breast-fed and was chosen as the main observation. Since the metabolism of fluoride causes extremely low levels to occur in breast milk even when the mother is drinking high levels, the exposures in breast-fed infants were not considered relevant for estimating effects in formula-fed infants who received much higher direct exposures. Therefore, combining breast-fed and formula-fed was not considered appropriate.

In some studies the critical observation was the same as the main observation, usually where the sample was the entire sample, or "all".

5.1. Additional meta-analyses

Additional detail for each study is provided in the forest plot in Figure 4, where additional columns list the Exposure Measure and Subpopulation for each study.



Meta-analysis, Main observations with Exposure Measure and Subpopulation listed

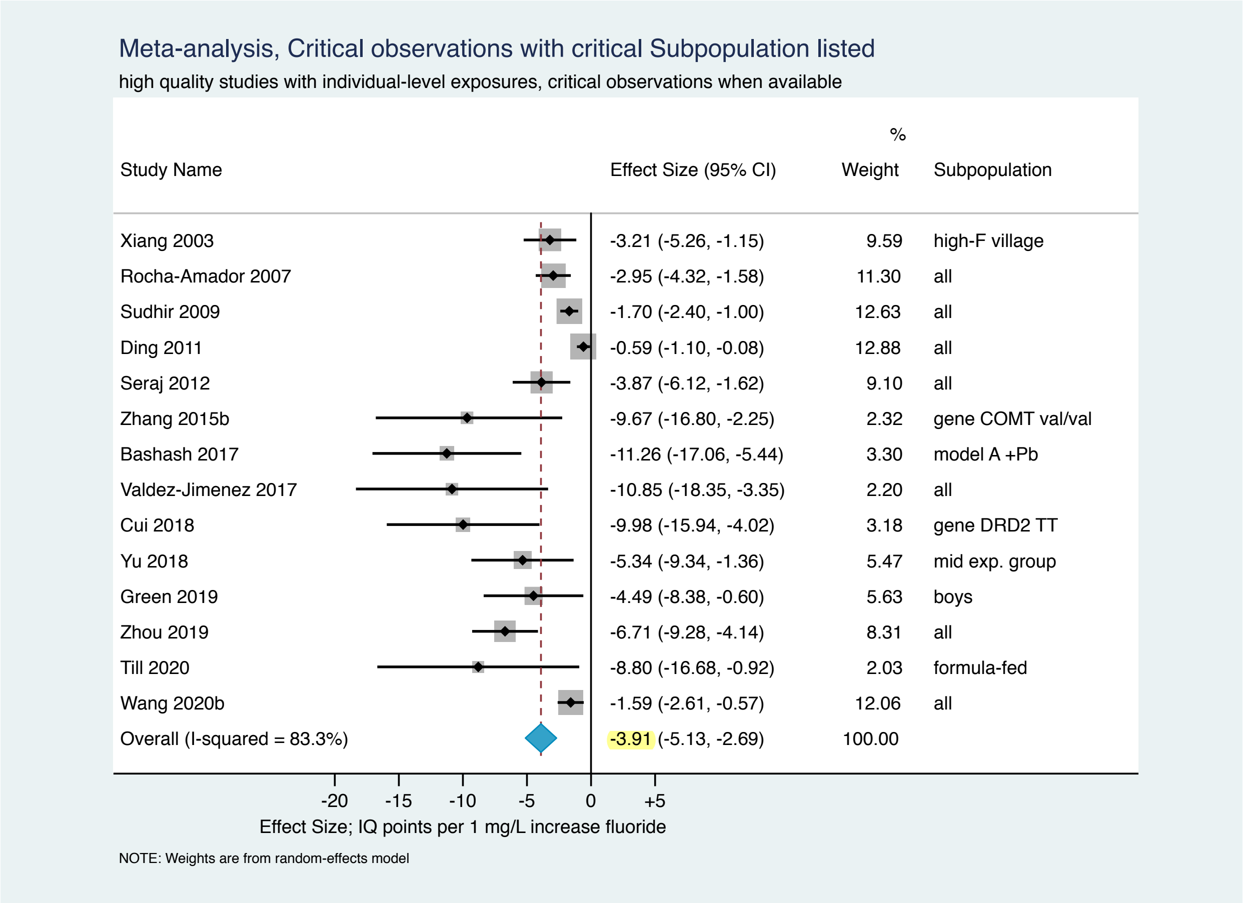
high quality studies with individual-level exposures, main observations

Study Name	Effect Size (95% CI)	% Weight	Exposure Measure	Subpopulation
Xiang 2003	-3.21 (-5.26, -1.15)	7.96	waterF	high-F village
Rocha-Amador 2007	-2.95 (-4.32, -1.58)	9.86	CUF	all
Sudhir 2009	-1.70 (-2.40, -1.00)	11.47	waterF	all
Ding 2011	-0.59 (-1.10, -0.08)	11.79	CUF	all
Seraj 2012	-3.87 (-6.12, -1.62)	7.45	waterF	all
Zhang 2015b	-2.42 (-4.59, -0.24)	7.65	CUF	all
Valdez-Jimenez 2017	-10.85 (-18.35, -3.35)	1.51	MUF	all
Bashash 2017	-6.30 (-9.12, -3.48)	6.10	MUFsg	all
Yu 2018	-5.34 (-9.34, -1.36)	4.07	CUF	mid exp. group
Cui 2018	-2.00 (-4.03, 0.02)	8.04	CUF	all
Green 2019	-1.95 (-5.18, 1.28)	5.28	MUFsg	all
Zhou 2019	-6.71 (-9.28, -4.14)	6.66	waterF	all
Till 2020	-8.80 (-16.68, -0.92)	1.39	waterF	formula-fed
Wang 2020b	-1.59 (-2.61, -0.57)	10.77	waterF	all
Overall (I-squared = 79.7%)	-3.05 (-4.04, -2.07)	100.00		

-20
-15
-10
-5
0
+5
Effect Size; IQ points per 1 mg/L increase fluoride
NOTE: Weights are from random-effects model
Exposure Measure Key; mg/L
waterF = drinking water F
CUF = child urine F
MUF = maternal urine F
MUFsg = MUF specific gravity adjusted

Figure 4. Forest plot of meta-analysis of Main observations, with additional details listed: Exposure Measure and Subpopulation.

Figure 5 is a forest plot of the meta-analysis of Critical observations. It includes an additional column of information describing the subpopulation that determined the critical observation.



Meta-analysis, Critical observations with critical Subpopulation listed

high quality studies with individual-level exposures, critical observations when available

Study Name	Effect Size (95% CI)	% Weight	Subpopulation
Xiang 2003	-3.21 (-5.26, -1.15)	9.59	high-F village
Rocha-Amador 2007	-2.95 (-4.32, -1.58)	11.30	all
Sudhir 2009	-1.70 (-2.40, -1.00)	12.63	all
Ding 2011	-0.59 (-1.10, -0.08)	12.88	all
Seraj 2012	-3.87 (-6.12, -1.62)	9.10	all
Zhang 2015b	-9.67 (-16.80, -2.25)	2.32	gene COMT val/val
Bashash 2017	-11.26 (-17.06, -5.44)	3.30	model A +Pb
Valdez-Jimenez 2017	-10.85 (-18.35, -3.35)	2.20	all
Cui 2018	-9.98 (-15.94, -4.02)	3.18	gene DRD2 TT
Yu 2018	-5.34 (-9.34, -1.36)	5.47	mid exp. group
Green 2019	-4.49 (-8.38, -0.60)	5.63	boys
Zhou 2019	-6.71 (-9.28, -4.14)	8.31	all
Till 2020	-8.80 (-16.68, -0.92)	2.03	formula-fed
Wang 2020b	-1.59 (-2.61, -0.57)	12.06	all
Overall (I-squared = 83.3%)	-3.91 (-5.13, -2.69)	100.00	

NOTE: Weights are from random-effects model

Figure 5. Forest plot of meta-regression of Critical observations, with listing of Subpopulation that defines the Critical observation.

When comparing meta-analyses using Main observations to meta-analyses using Critical observations, the consistency, pooled effect sizes, and pooled statistical significance were all greater in the Critical observations. This is illustrated in Figure 6 below where Main observations had a pooled effect of -3.05 IQ points per 1 mg/L fluoride, and Critical observations had a pooled effect of -3.91. Also, all 14 of the Critical observations had statistically significant adverse effects, while 13 of the 14 Main observations had statistically significant adverse effects.

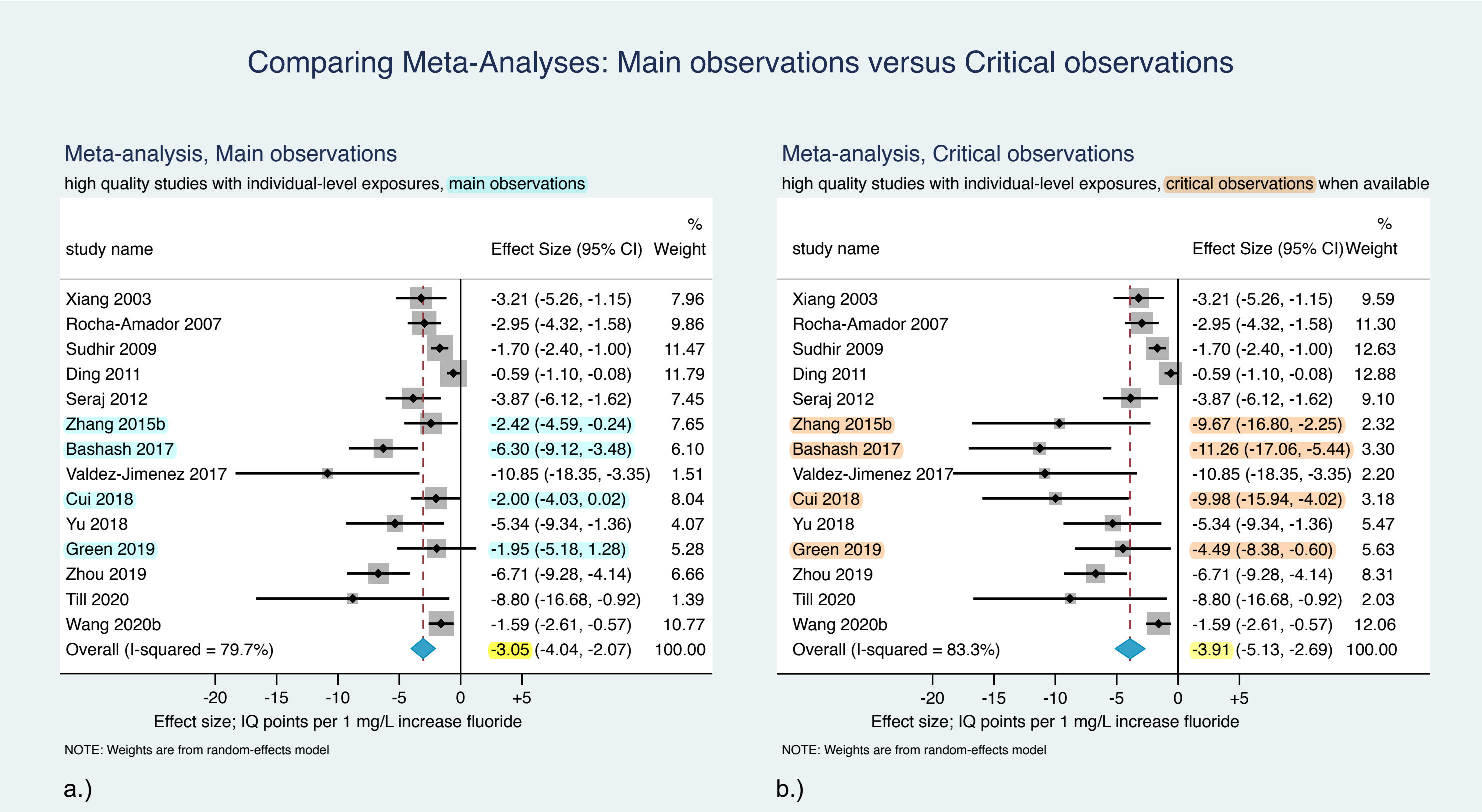


Figure 6. Comparison of meta-analyses, a.) Main observations versus b.) Critical observations. Studies in which the Critical observation is different from the Main observation and has greater effect are highlighted.

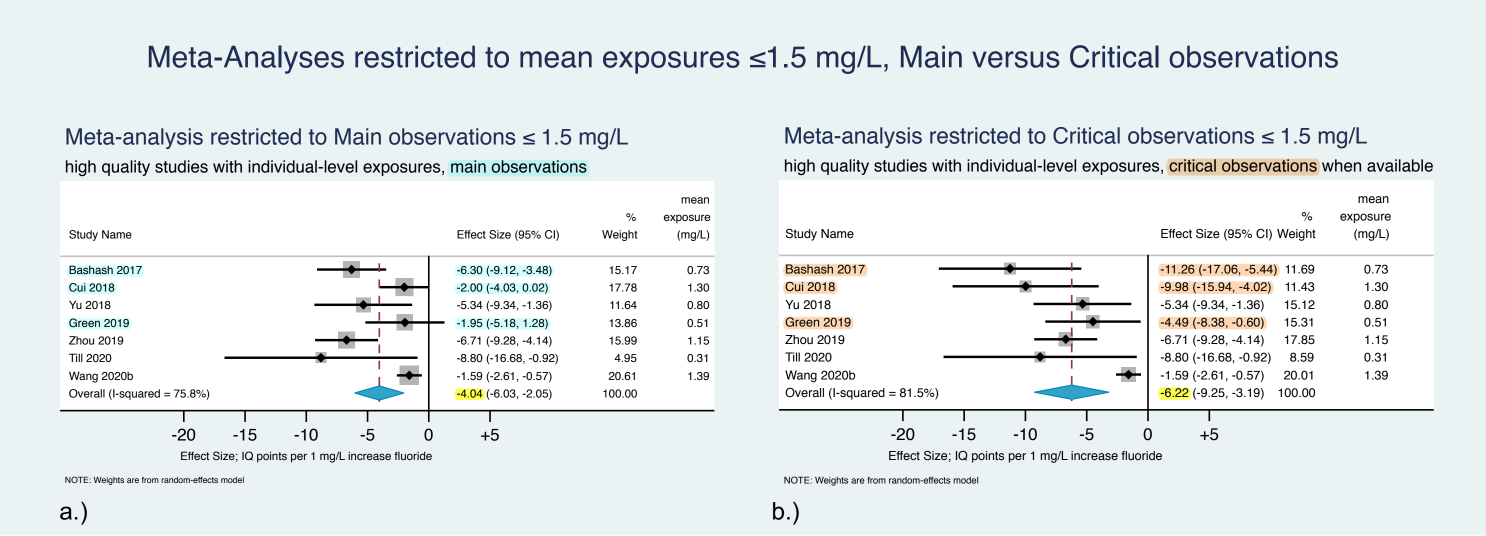


Figure 7. Comparison of meta-analyses for studies restricted to those having mean exposures below 1.5 mg/L, a.) Main observations versus b.) Critical observations. Studies in which the Critical observation is different from the Main observation and has greater effect are highlighted.

In a further refinement of meta-analyses, we restricted studies to just those with mean exposures below 1.5 mg/L. All of these are indisputably relevant to exposures common in the United States and no extrapolation of dose-response is required from observed higher exposure levels to unobserved lower exposure levels. We identified 7 such studies, and present a forest plot for the Main observations that includes an added column showing the mean exposure (Figure 7a). For comparison purposes, we also did a meta-analysis of these 7 studies but using the Critical observations (Figure 7b). The consistency of these studies is very high, with all studies finding reductions in IQ with increasing F exposure, with all but one observation being statistically significant. The pooled effect size of the Critical observations is -6.22 IQ points per 1 mg/L increase in fluoride, which rivals the predicted IQ loss from lead poisoning in children with highly elevated blood lead levels. This provides direct evidence that sizable US subpopulations who are more sensitive to fluoride or receive high internal doses may be experiencing substantial average drops in IQ. The genetic variants typically occur in 5% or more of the population. The gender subpopulation (males) occurs in 50% of the population. The formula-fed subpopulation represents about 50% of US infants. In addition to these heightened sensitivity groups, about 5% of the overall population at all ages consumes more than twice as much municipal water as the average consumer, and will thus receive at least twice the internal dose of fluoride as the average water consumer.

Exposure period explains some heterogeneity

An *a priori* important source of heterogeneity amongst the studies was expected to be exposure period. We stratified measured exposure period into two periods: early life exposure (prenatal through infancy) and childhood (classified as age 1 years and older). The biological rationale for this expectation is that the more rapidly developing brain is considered more sensitive to neurotoxic effects than when development is slower at older ages, and because the blood-brain barrier is not well developed in the fetus and through the first 6 months of life. A given external exposure may thus produce higher internal exposures in the target tissue, the brain, during this early life period. Another reason to expect greater effects for a given water or urine F concentration in the earlier life period is because infants consume substantially more water per body weight than older children.

Figure 8 shows a forest plot of the subgroup meta-analysis by exposure period. It shows a highly statistically significant effect in both periods, but the magnitude is more than twice as large for prenatal and infant exposures as for childhood exposures (-5.90 IQ points per 1 mg/L fluoride, compared to -2.58). Thus, exposure period explains some of the heterogeneity amongst the studies, and the larger effect from early life exposure supports the *a priori* expectation that this period would be more sensitive.

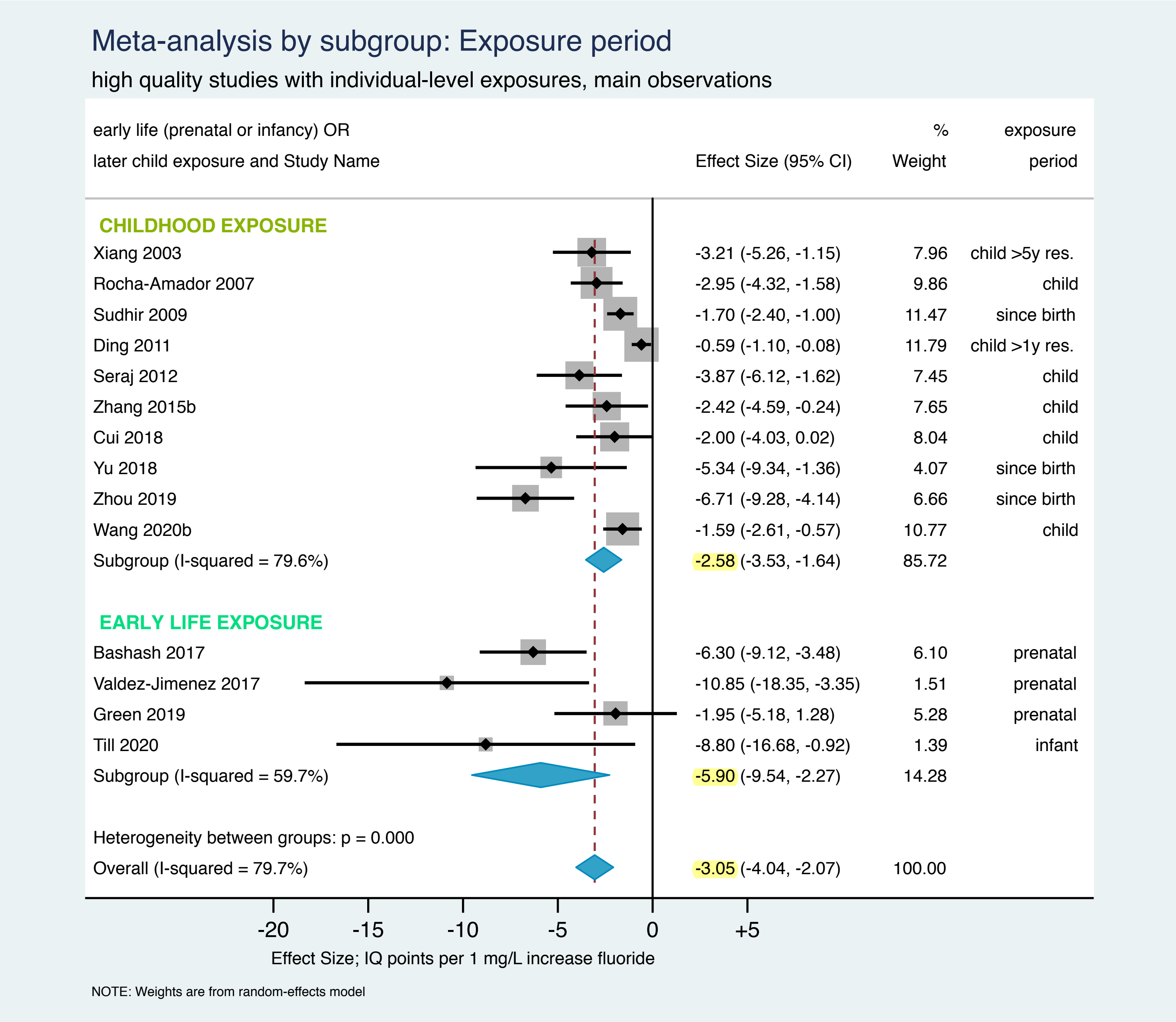


Figure 8. Forest plot of subgroup meta-analysis with exposure period stratified into early life exposure and childhood exposure. Main observations.

5.2. Additional meta-regressions

We conducted additional meta-regressions of the studies rated higher quality by NTP that had individual-level exposure measures. These meta-regressions were conducted to further examine the dose-response relationship and to explore reasons for heterogeneity.

The first additional meta-regression was restricted to the 7 studies with direct observations at exposures below 1.5 mg/L. The results are very similar to the meta-regression that includes the 14 studies at all exposure levels. Figure 9b shows that the predicted pooled effect size was still statistically significant, and slightly larger, for those studies with mean exposures below 1.5 mg/L compared to that for studies at all exposures (-3.32 IQ points per 1 mg/L increase in fluoride, compared to -2.94). Thus, studies with direct exposures relevant to the US provide as strong evidence of substantial loss of IQ as do studies with higher exposures. Figure 9 compares the meta-regression with studies at all exposures (Figure 9a) to those with exposures below 1.5 mg/L (Figure 9b). Also shown are duplicate plots with points labeled with the study names rather than marked by dots, to allow linking specific studies to specific data points (Figures 9c and 9d).

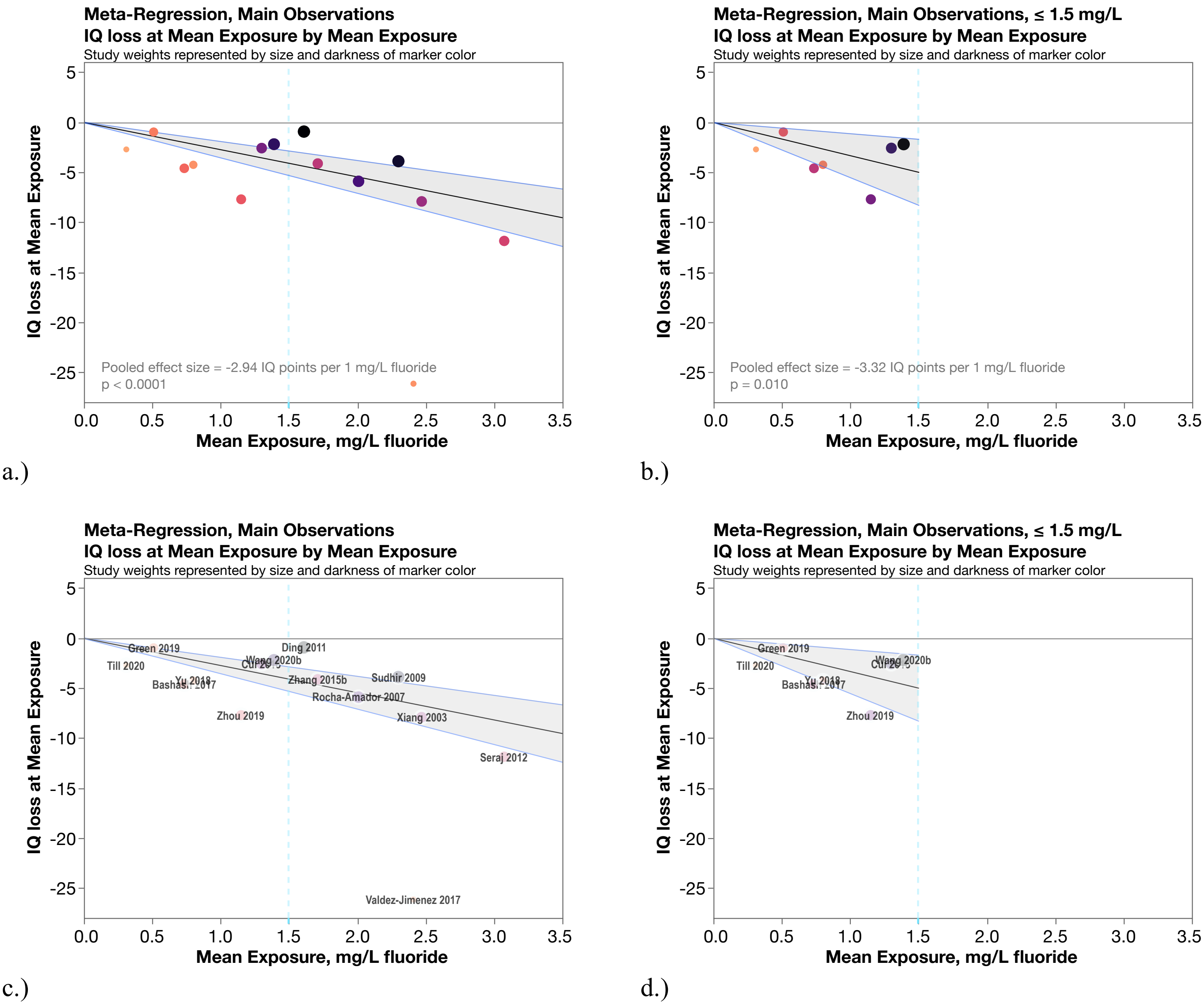
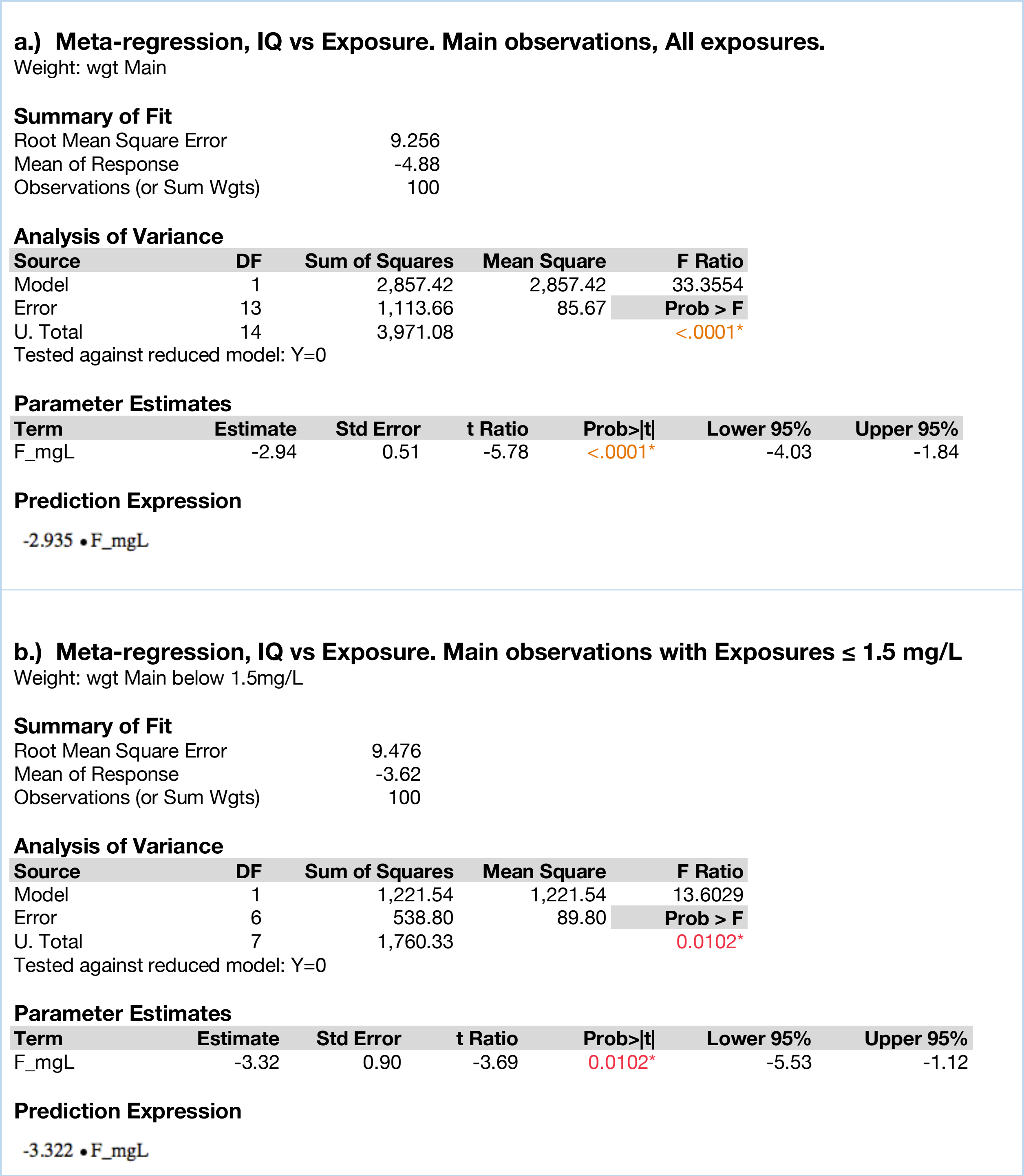


Figure 9. Meta-regression plots; a.) All exposure levels; b.) Exposures ≤1.5 mg/L; c.) & d.) with study name labels.

Statistics for the meta-regressions plotted in Figure 9a and 9b are provided in Table 1.

Table 1. Statistics for linear meta-regression models: a.) IQ vs Exposure for All exposures; b.) IQ vs Exposure for Exposures ≤1.5 mg/L.



a.) Meta-regression, IQ vs Exposure. Main observations, All exposures.
Weight: wgt Main

Summary of Fit

Root Mean Square Error	9.256
Mean of Response	-4.88
Observations (or Sum Wgts)	100

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	2,857.42	2,857.42	33.3554
Error	13	1,113.66	85.67	**Prob > F**
U. Total	14	3,971.08		<.0001*

Tested against reduced model: Y=0

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|	Lower 95%	Upper 95%
F_mgL	-2.94	0.51	-5.78	<.0001*	-4.03	-1.84

Prediction Expression

-2.935 • F_mgL

b.) Meta-regression, IQ vs Exposure. Main observations with Exposures ≤ 1.5 mg/L
Weight: wgt Main below 1.5mg/L

Summary of Fit

Root Mean Square Error	9.476
Mean of Response	-3.62
Observations (or Sum Wgts)	100

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	1,221.54	1,221.54	13.6029
Error	6	538.80	89.80	**Prob > F**
U. Total	7	1,760.33		0.0102*

Tested against reduced model: Y=0

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|	Lower 95%	Upper 95%
F_mgL	-3.32	0.90	-3.69	0.0102*	-5.53	-1.12

Prediction Expression

-3.322 • F_mgL

As mentioned above, several non-linear meta-regression models were also examined but were not found to fit the data better than the linear model. These included logarithmic, quadratic, and flexible spline models. Figure 10 shows the plot of the spline model, with superimposed dots for the individual study data. This is a 3-knot cubic regression-spline model, using the Stone and Koo method [JMP 2020, Stone and Koo 1985].

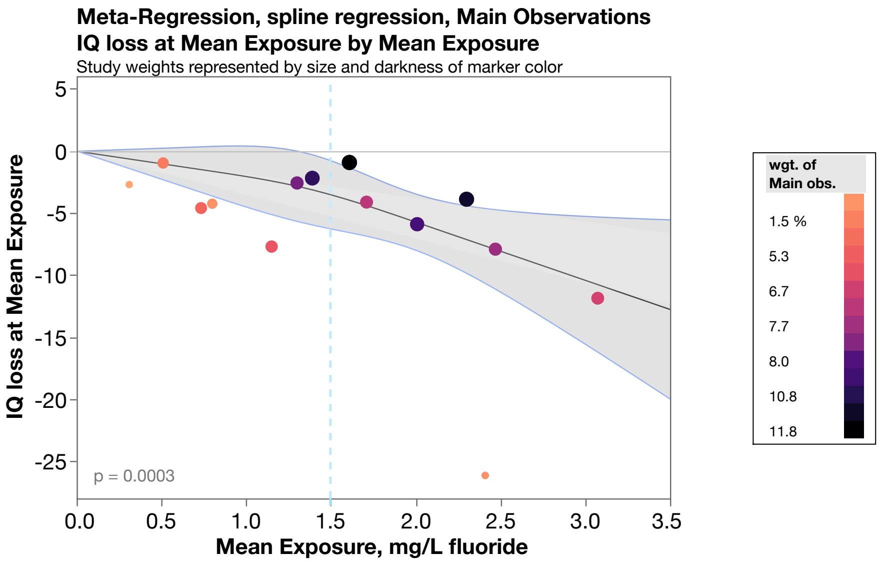


Figure 10. Spline meta-regression, Main observations, All exposure levels.

The dose-response curve is close to linear, with just a slight flattening below 1.5 mg/L. The regression model is highly statistically significant and the predicted average IQ loss at exposures relevant to the US is about 2 to 3 IQ points.

Table 2. Statistics for regression-spline meta-regression of IQ loss by Mean Exposure.

Summary of Fit

Root Mean Square Error	9.332
Mean of Response	-4.88
Observations (or Sum Wgts)	100

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	2	2,926.09	1,463.05	16.8008
Error	12	1,044.98	87.08	**Prob > F**
U. Total	14	3,971.08		0.0003*

Tested against reduced model: Y=0

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|	Lower 95%	Upper 95%
F_mgL&Knotted	-2.019	1.152	-1.75	0.1052	-4.529	0.492
F_mgL&Knotted@0.734	-1.096	1.234	-0.89	0.3920	-3.785	1.593

Prediction Expression

$$\left(\begin{array}{l} -2.019 \bullet \text{F\_mgL} \\ + -1.096 \bullet \left(\begin{array}{l} \left(\text{Maximum}\left(\text{F\_mgL} - 0.734, 0\right)^{3} - \text{Maximum}\left(\text{F\_mgL} - 1.3713, 0\right)^{3} \bullet 2 \right) \\ + \text{Maximum}\left(\text{F\_mgL} - 2.0085, 0\right)^{3} \end{array} \right) \end{array} \right)$$

To examine reasons for heterogeneity amongst studies, meta-regression models were tried for various factors that had continuous variables. Besides dose (Mean exposure, mg/L fluoride), the only factor that had a statistically significant association with effect size was age at outcome assessment, when entered in a multivariable model that also had dose. Both dose and age were statistically significant in the model. The younger the mean age at outcome assessment, the larger the IQ loss. The relationship between effect size and these two independent variables is shown a marginal effect plots (or margin plots), where the non-plotted variable is held at its mean value for each plot (Figure 11).

Margin plots for multivariable meta-regression: Predicted IQ loss by Mean Exposure and Mean Age.
Each covariate set at its mean value for the margin plots.

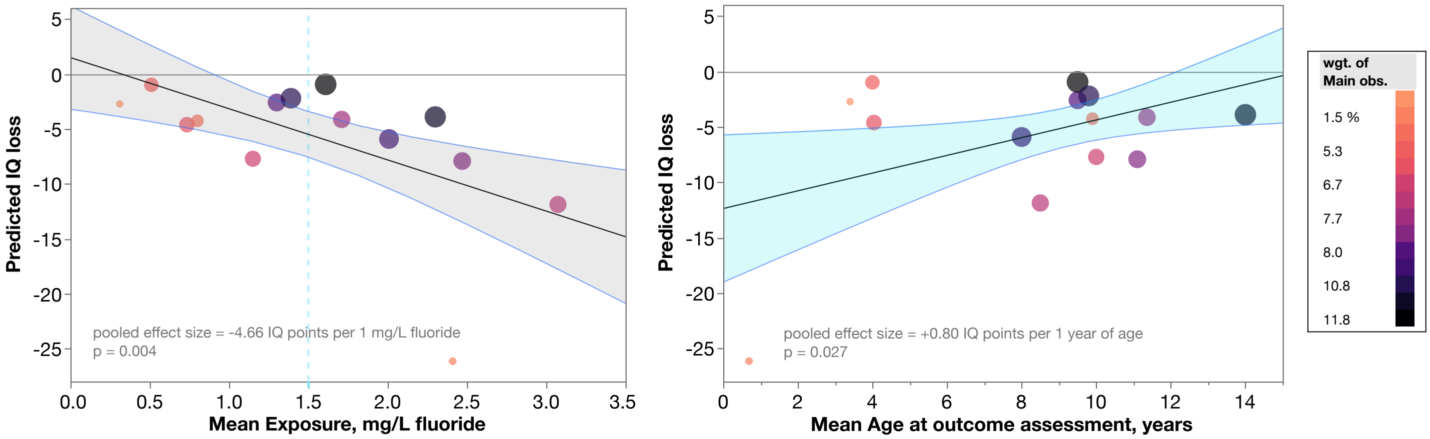


Figure 11. Marginal effects plots for the model: Predicted IQ loss by Mean exposure and Mean Age at outcome assessment.

Table 3. Summary statistics for multivariable meta-regression model.

Summary of Fit

RSquare	0.562
RSquare Adj	0.482
Root Mean Square Error	7.957
Mean of Response	-4.88
Observations (or Sum Wgts)	100

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	2	894.05	447.03	7.0603
Error	11	696.47	63.32	**Prob > F**
C. Total	13	1,590.52		0.0107*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-4.473	2.835	-1.58	0.1429
Mean Exposure, F_mg/L	-4.659	1.288	-3.62	0.0041*
Mean Age at outcome, y	0.800	0.312	2.56	0.0266*

Prediction Expression

$$-4.473 + -4.659 \bullet \text{F\_mgL} + 0.7995 \bullet \text{Mean\_Age\_y}$$

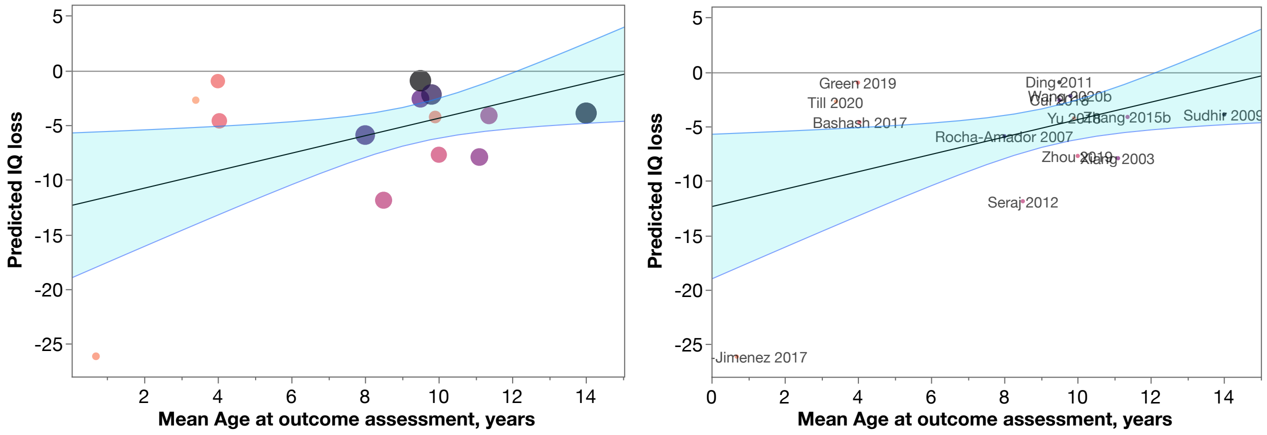


Figure 12. Margins plot for Mean Age at Outcome Assessment with superimposed bubble plot in left pane and superimposed study names in right pane.

A possible explanation for the greater effect size as Mean Age at Outcome Assessment decreases may be that the studies at younger ages are also capturing effects from earlier life exposure periods, whereas studies with outcome assessment at older ages, were capturing exposures across older age periods. In a previous subgroup meta-analysis (Figure 8 above) we found that early life exposure periods were associated with larger effect sizes. This may be due to greater sensitivity of the fetal and infant brain to disruption from fluoride. It might also reflect effect attenuation from bias toward the null due to greater random exposure measurement error when exposure is measured at older ages. Exposure estimated at a single time in childhood may not as accurately reflect lifetime average exposure as when exposure is measured in early life.

The multivariable meta-regression model that includes both dose and outcome age provides greater support to a conclusion that there is a robust dose-response relationship between fluoride exposure and loss of IQ. With outcome age accounted for in the model, the fluoride exposure effect becomes larger in magnitude and statistical significance. The effect size (Beta) is -4.7 IQ points per 1 mg/L fluoride with p=0.004. In the bivariate meta-regression the effect size is only -3.2 IQ points per 1 mg/L fluoride with p=0.04 (see Figure 13 and Table 4 below).

Bivariate meta-regression: Predicted IQ loss by Mean Exposure.

Weight: wgt Main

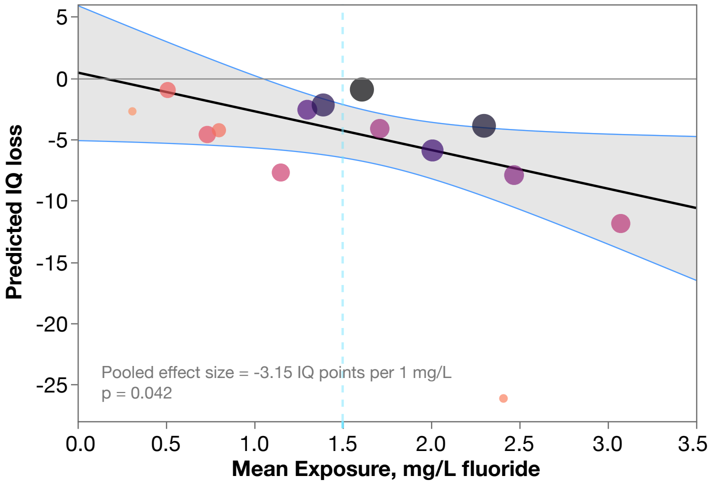


Figure 13. Bivariate meta-regression model: IQ loss at Mean Exposure vs. Mean Exposure.

Table 4. Statistics for meta-regression bivariate model: IQ loss at Mean Exposure vs. Mean Exposure.

Summary Statistics, All exposure level observations

Weight: wgt Main

	Value	Lower 95%	Upper 95%	Signif. Prob
Correlation	-0.549	-0.836	-0.026	0.0420*
Covariance	-11.69			
Count	14			

Variable	Mean	Std Dev
F_mgL	1.684	1.925
predicted IQ loss at mean Exposure	-4.879	11.061

Linear Fit

predicted IQ loss at mean Exposure = 0.4339273 - 3.1554817*F_mgL

Summary of Fit

RSquare	0.302
RSquare Adj	0.243
Root Mean Square Error	9.622
Mean of Response	-4.88
Observations (or Sum Wgts)	100

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	1	479.60	479.60	5.1806
Error	12	1,110.92	92.58	**Prob > F**
C. Total	13	1,590.52		0.0420*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|	Lower 95%	Upper 95%
Intercept	0.434	2.525	0.17	0.8664	-5.067	5.935
F_mgL	-3.155	1.386	-2.28	0.0420*	-6.176	-0.135