

A Benchmark Dose Analysis for Maternal Pregnancy Urine-Fluoride and IQ in Children

Philippe Grandjean,^{1,2*} Howard Hu,³ Christine Till,⁴ Rivka Green,⁴ Morteza Bashash,³ David Flora,⁴ Martha Maria Tellez-Rojo,⁵ Peter X.K. Song,⁶ Bruce Lanphear,⁷ Esben Budtz-Jørgensen⁸

¹ Harvard T.H. Chan School of Public Health, Department of Environmental Health, Boston, MA 02115, USA.

² Department of Public Health, University of Southern Denmark, Odense, Denmark.

³ Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, CA, USA

⁴ Faculty of Health, York University, Ontario, Canada

⁵ Centro de Investigacion en Salud Poblacional, Instituto Nacional de Salud Publica, Cuernavaca, Morelos, Mexico

⁶ Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI, USA

⁷ Faculty of Health Sciences, Simon Fraser University, British Columbia, Canada

⁸ Department of Biostatistics, University of Copenhagen, Denmark.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/risa.13767](https://doi.org/10.1111/risa.13767).

This article is protected by copyright. All rights reserved.

* Address correspondence to Philippe Grandjean, Environmental Medicine, University of Southern Denmark,

J.B. Winslows vej 17A, 5000 Odense C, Denmark; tel: +45 6550 3769; fax: +45 6591 1458;

pgrandjean@health.sdu.dk

Abstract

As a guide to establishing a safe exposure level for fluoride exposure in pregnancy, we applied benchmark dose modeling to data from two prospective birth cohort studies. We included mother-child pairs from the Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT) cohort in Mexico and the Maternal-Infant Research on Environmental Chemicals (MIREC) cohort in Canada. Maternal urinary fluoride concentrations (U-F, in mg/L, creatinine-adjusted) were measured in urine samples obtained during pregnancy. Children were assessed for IQ at age 4 (n=211) and between 6 and 12 years (n=287) in the ELEMENT cohort, and 3-4 years (n=407) in the MIREC cohort. We calculated covariate-adjusted regression coefficients and their standard errors to assess the association of maternal U-F concentrations with children's IQ measures. Assuming a benchmark response of 1 IQ point, we derived benchmark concentrations (BMCs) and benchmark concentration levels (BMCLs). No deviation from linearity was detected in the dose-response relationships, but boys showed lower BMC values than girls. Using a linear slope for the joint cohort data, the BMC for maternal U-F associated with a 1-point decrease in IQ scores was 0.31 mg/L (BMCL, 0.19 mg/L) for the youngest boys and girls in the two cohorts, and 0.33 mg/L (BMCL, 0.20 mg/L) for the MIREC cohort and the older ELEMENT children. Thus, the joint data show a BMCL in terms of the adjusted U-F concentrations in the pregnant women of approximately 0.2 mg/L. These results can be used to guide decisions on preventing excess fluoride exposure in pregnant women.

Keywords: benchmark dose; cognitive deficits; fluoride; neurotoxicity; prenatal exposure; pregnancy.

1. INTRODUCTION

This article is protected by copyright. All rights reserved.

The EPA's Maximum Contaminant Level Goal (MCLG) of 4.0 mg/L for fluoride in drinking water was first set in 1985 to protect against chronic fluoride toxicity in the form of crippling skeletal fluorosis (U.S. Environmental Protection Agency, 1985). In 2006, the U.S. National Research Council (NRC) concluded that fluoride may adversely affect the brain (National Research Council, 2006). Since then, a substantial number of cross-sectional studies, mostly in communities with chronic fluoride exposure, have shown lower cognitive performance in children growing up in areas with higher fluoride concentrations in drinking water, as summarized in meta-analyses (Choi et al., 2015; Duan, Jiao, Chen, & Wang, 2018; Tang, Du, Ma, Jiang, & Zhou, 2008). Support for fluoride neurotoxicity has also emerged from experimental studies (Bartos et al., 2018; Mullenix, Denbesten, Schunior, & Kernan, 1995; National Toxicology Program, 2020). Despite the existence of recent prospective birth cohort studies (Bashash et al., 2017; R. Green et al., 2019; Valdez Jimenez et al., 2017), no meta-analysis has so far focused on prenatal fluoride exposure.

Fluoride is found in many minerals, in soil and thus also in groundwater (National Research Council, 2006). Since the mid-1940s, fluoride has been added to many drinking water supplies in order to prevent tooth decay (U.S. Environmental Protection Agency, 1985). Community water fluoridation is practiced in the U.S., Canada, and several other countries, whereas some, like Mexico, add fluoride to table salt. Fluoridated water accounts for about 40-70 percent of daily fluoride intake in adolescents and adults living in these communities (U.S. Environmental Protection Agency, 2010). The fluoride concentration in drinking water roughly equals the fluoride concentration in urine (National Research Council, 2006), as also recently shown in the Canadian cohort of pregnant women (Till et al., 2018). In addition to fluoridation, some types of tea, such as black tea, constitute an additional source of exposure (Krishnankutty et al., 2021; Rodríguez et al., 2020; Waugh, Godfrey, Limeback, & Potter, 2017).

Fluoride is readily distributed throughout the body, with bones and teeth as storage depots. During pregnancy, fluoride crosses the placenta and reaches the fetus (National Research Council, 2006; World Health Organization, 2006). As fluoride is rapidly eliminated via urine, the adjusted urine-fluoride (U-F) concentration mainly represents recent absorption (Ekstrand & Ehrnebo, 1983; World Health Organization, 2006). Pregnant women may show lower U-F concentrations than non-pregnant controls, perhaps due to fetal uptake and storage in hard tissues (Opydo-Symaczek & Borysewicz-Lewicka, 2005).

For the purpose of identifying safe exposure levels, regulatory agencies routinely use benchmark dose (BMD) calculations (European Food Safety Authority, 2009; U.S. Environmental Protection Agency, 2012). As long recognized (National Research Council, 1989), fluoride is not an essential nutrient, and dose-dependent toxicity can therefore be considered monotonic. As with lead (Budtz-Jørgensen, Bellinger, Lanphear, Grandjean, & International Pooled Lead Study Investigators, 2013), BMD results can be generated from regression coefficients and their standard errors for the association between maternal U-F concentrations and the child's IQ score (Grandjean, 2019). The BMD is the dose leading to a specific change (denoted BMR) in the response (in this case, an IQ loss), compared with unexposed children. A decrease of 1 IQ point is an appropriate BMR, as specified by the European Food Safety Authority and also recognized by the U.S. EPA (Budtz-Jørgensen, et al., 2013; European Food Safety Authority, 2010; Gould, 2009; Reuben et al., 2017). The present study uses data from two prospective birth cohort studies (Bashash, et al., 2017; R. Green, et al., 2019) to calculate the benchmark concentration of urine-fluoride (U-F) associated with a 1-point decrement in Full Scale IQ (FSIQ).

2. METHODOLOGY

2.1. Study Cohorts

In the Early Life Exposures in Mexico to Environmental Toxicants (ELEMENT) project, mother-child pairs were successively enrolled in longitudinal birth cohort studies from the same three hospitals in

Mexico City which serve low to moderate income populations. A full description of the cohorts and associated methods is provided in a recent “Cohort Profile” article (Perng et al., 2019). Urinary samples were collected from pregnant women between 1997 and 1999 (cohort 2A, n=327) and between 2001 and 2003 (cohort 3 with calcium intervention and placebo arms, n=670). Cohort 2A was designed as an observational birth cohort of lead toxicodynamics during pregnancy, while Cohort 3 was designed as a randomized double-blind placebo-controlled trial of calcium supplements. Women were included in the current study if they had at least one biobanked urine sample for fluoride analysis, a urinary creatinine concentration, complete data of adjusted covariates, and their child underwent cognitive testing at age 4 years (n=287) and/or between ages 6 and 12 years (n=211). Of the 287 participants with data on General Cognitive Index (GCI) outcomes and other variables, 110 were from Cohort 2A, 93 were from the Cohort 3 Calcium intervention arm and 84 were from the Cohort 3 Placebo arm. Among participant in the GCI outcome, U-F data were available for all three trimesters (n=25), two trimesters (n=121), or one trimester (n=141). Of the 211 participants with data on IQ outcomes, 78 were recruited from Cohort 2A, 75 from the Cohort 3 Calcium intervention arm and 58 from the Placebo arm; U-F data for IQ outcome were available for all three trimesters (n=10), two trimesters (n=82), or one trimester (n=119).

In the Maternal-Infant Research on Environmental Chemicals (MIREC) program, 2,001 pregnant women were recruited between 2008 and 2011 from ten cities across Canada. Women were recruited from prenatal clinics if they were at least 18 years old, less than 14 weeks of gestation, and spoke English or French. Exclusion criteria included fetal abnormalities, medical complications, and illicit drug use during pregnancy; further details have been previously described (Arbuckle et al., 2013). A subset of children (n = 601) in the MIREC Study was evaluated for the developmental phase of the study (MIREC-Child Development Plus) at 3-4 years of age from six of

the ten cities included in the original cohort, half of which were fluoridated. Of the 601 children who completed the neurodevelopmental testing in entirety, 526 (87.5%) mother-child pairs had all three U-F samples; of these, 512 (85.2%) had specific gravity measures, while 407 (67.7%) had creatinine data, as well as complete covariate data; 75 (12.5%) women were missing one or more trimester U-F samples, and 14 women (2.3%) were missing one or more covariates.

2.2. Exposure Assessment

All urine samples from the two studies were analyzed by the same laboratory at the Indiana University School of Dentistry using a modification of the hexamethyldisiloxane (Sigma Chemical Co., USA) microdiffusion method with the ion-selective electrode (Martinez-Mier et al., 2011).

In the ELEMENT study, spot (second morning void) urine samples were collected during the first trimester (mean \pm SD: 13.7 \pm 3.5 weeks for cohort 2A and 13.6 \pm 2.1 weeks for cohort 3), second trimester (24.4 \pm 2.9 weeks for cohort 2A and 25.1 \pm 2.3 weeks for cohort 3), and third trimester (35.0 \pm 1.8 weeks for cohort 2A and 33.9 \pm 2.2 weeks for cohort 3). The samples were collected into fluoride-free containers and immediately frozen at the field site and shipped and stored at -20°C at the Harvard School of Public Health, and then at -80°C at the University of Michigan School of Public Health. To account for variations in urinary dilution at time of measurement, the maternal U-F concentration was adjusted for urinary creatinine, as previously described (Thomas et al., 2016). An average of all available creatinine-adjusted U-F concentrations during pregnancy (up to a maximum of three samples) was computed and used as the exposure parameter.

In the MIREC study, urine spot samples were collected at each trimester, i.e., first trimester at 11.6 \pm 1.6 (mean \pm SD) weeks of gestation, second trimester at 19.1 \pm 2.4 weeks, and third trimester at 33.1 \pm 1.5 weeks. Maternal U-F concentrations at each trimester were adjusted for both creatinine and

specific gravity, as described previously (Till et al., 2020). For this joint analysis, however, we elected to use the U-F concentrations adjusted for creatinine to keep the urine dilution factor consistent with the adjustment procedure in ELEMENT. For each woman, the average maternal U-F concentration was derived only if a valid U-F value was available for each trimester.

2.3. Assessment of Intelligence

The ELEMENT study (Bashash, et al., 2017) used the McCarthy Scales of Children's Abilities (MSCA) Spanish version to measure cognitive abilities at age 4 years and derive a General Cognitive Index (GCI) as a standardized composite score. The MSCA was administered by trained psychometrists or psychologists who were supervised by an experienced clinical child psychologist. For children aged 6-12 years, a Spanish-version of the Wechsler Abbreviated Scale of Intelligence (WASI) was administered to derive FSIQ as a measure of global intellectual functioning. In the MIREC study, children's intellectual abilities (R. Green, et al., 2019) were assessed at age 3-4 years using the FSIQ from the Wechsler Preschool and Primary Scale of Intelligence, Third Edition (WPPSI-III). A trained research assistant who was supervised by a psychologist administered the WPPSI-III in either English or French. In both studies, examiners were blinded to the children's fluoride exposure. All raw scores were standardized for age.

The GCI shows concurrent validity with intelligence tests, including the Stanford-Binet IQ ($r=.81$) and FSIQ ($r=.71$) from the Wechsler Preschool and Primary Scale of Intelligence (WPPSI) (Kaplan & Sacuzzo, 2010). Likewise, the FSIQ of the WASI (ELEMENT cohort) and WISC-III (MIREC cohort) is strong ($r=.81$) (Wechsler, 1991). The high covariance between the various measures of intellectual ability provides justification for pooling IQ scores across the two cohorts.

2.4. Covariate Adjustment

This article is protected by copyright. All rights reserved.

For the ELEMENT study, data were collected from each subject by questionnaire on relevant parameters, gestational age was estimated by registered nurses, and maternal IQ was estimated using subtests of the Wechsler scale standardized for Mexican adults. Covariates included gestational age (weeks), birth weight, sex, age at outcome measurement and the following maternal characteristics: parity (being first child), smoking history (ever smoked vs non-smoker), marital status (married vs other), age at delivery, IQ, education (years of education), and sub-cohort (Cohort 2A, Cohort 3 calcium intervention or placebo).

The MIREC study selected similar covariates from a set of established predictors of fluoride metabolism and cognitive development, including sex, city of residence, HOME score, maternal education (dichotomized as bachelor's degree or higher: yes/no), and maternal race/ethnicity (dichotomized as white: yes/no). Covariates included in the original studies (Bashash, et al., 2017; R. Green, et al., 2019) were retained in the statistical calculations in the present study. Due to a growing body of epidemiologic studies showing sex-specific effects associated with neurotoxic exposures (Levin, Dow-Edwards, & Patisaul, 2021), including fluoride (R. Green, et al., 2019; R. Green, Rubenstein, Popoli, Capulong, & Till, 2020), interactions between sex and U-F exposure were examined.

2.5 Benchmark Concentration (BMC) Calculations

The BMC is the U-F concentration that reduces the outcome by a prespecified level (known as the benchmark response, BMR) compared to an unexposed control with the same covariate profile (Budtz-Jørgensen, Keiding, & Grandjean, 2001; Crump, 1995). We based the benchmark calculations on regression models with p covariates in the following form:

$$IQ = \alpha_0 + \alpha_1 \times \text{covariate}_1 + \dots + \alpha_p \times \text{covariate}_p + f(c) + \epsilon$$

where c is the urine-fluoride concentration and f is the concentration-response function, and ϵ is a normally distributed error term with a mean of 0 (and a variance of σ^2). To assess the linearity of the concentration-response relationship, several models were considered. In addition to the standard linear model, where $f(c) = \beta c$, we estimated a squared effect, where $f(c) = \beta c^2$, and two piecewise-linear models (or broken-stick) with breakpoints at 0.5 and 0.75 mg/L. Piecewise-linear models are useful in benchmark calculations because the slope of the concentration-response function is allowed to change linearity at the breakpoint, and in such models, benchmark calculations are less sensitive to exposure-associated effects occurring only at high concentration levels. Furthermore, to allow for the possibility of different exposure effects in boys and girls, each concentration-response model was also fitted with the inclusion of an interaction with sex.

Models were fitted separately in the two cohorts yielding analyses that were similar to those presented in the original publications (Bashash, et al., 2017; R. Green, et al., 2019) based on the original raw data and with the covariate adjustments as originally justified. Sensitivity analyses were carried out using the MIREC specific gravity-adjusted U-F values joint with the ELEMENT creatinine-adjusted U-F values as well. The Mexico study controlled for maternal bone lead stores (the primary source of prenatal lead exposure in this cohort) and blood-mercury during pregnancy, although the sample size was reduced by about one-third; the effect estimates for fluoride on child IQ increased and remained statistically significant ($p < 0.01$) (Bashash, et al., 2017). Likewise, controlling for lead, mercury, PFOA, arsenic, and manganese in the MIREC study did not result in any appreciable change of the U-F estimates (R. Green, et al., 2019). Thus, these other neurotoxicants were not included as covariates in the present calculations. Using the regression coefficients, we first calculated BMC results for each cohort and then derived joint BMCs by combining regression coefficients from the two cohorts.

Given that the BMC reduces the outcome by the BMR, a smaller BMR will result in lower BMC and benchmark concentration level (BMCL) results. For the child IQ as the outcome variable, the BMR is 1 IQ point. In our regression model, the IQ difference between unexposed subjects and subjects at the BMC is given by $f(0) - f(\text{BMC})$, and therefore the BMC satisfies the equation $f(0) - f(\text{BMC}) = \text{BMR}$. We use concentration-response functions with $f(0) = 0$, and therefore the BMC is given by

$$\text{BMC} = f^{-1}(-\text{BMR})$$

In a regression model with a linear concentration-response function [$f(c) = \beta c$], we get $\text{BMC} = -\text{BMR}/\beta$. If the estimated concentration-response is increasing (indicating a beneficial effect), the BMC is not defined, and the BMC is then indicated by ∞ .

The main result of the BMC analysis is the BMCL, which is defined as a lower one-sided 95% confidence limit of the BMC (Crump, 1995). In the linear model,

$$\text{BMCL} = -\text{BMR}/\beta_{\text{lower}}$$

where β_{lower} is the one-sided lower 95% confidence limit for β (Budtz-Jørgensen, et al., 2013). In the other models considered, we calculated the BMCL by first identifying a lower confidence limit for $f(c)$ and then finding the concentration (c) where confidence limit is equal to -BMR.

Finally, we derived two sets of joint benchmark concentrations: The MIREC results (FSIQ score) were combined with ELEMENT outcomes using either GCI or FSIQ scores for all subjects where the creatinine-adjusted U-F was available. Joint benchmark concentration results were obtained under the hypothesis that the concentration-response functions were identical in the two studies. Under this hypothesis, the concentration-response function [$f(c)$] was estimated by combining the regression coefficients describing $f(c)$. Again, using the linear model as an example, we estimated the joint regression coefficient by weighing together cohort-specific coefficients. Here we used optimal weights proportional to the inverse of the squared standard error. In a Wald test, we tested whether the exposure effects in the two cohorts were equal. We calculated sex-dependent BMC results from regression models that included interaction terms between sex and $f(c)$. The fit of the regression models was compared by twice the negative log-likelihood [$-2 \log L$] as supplemented by the Akaike Information Criterion (AIC); the latter is provided in the tables. For both measures, a lower value indicates a better fit, but AIC-based differences below 4 are not considered important. For sex-dependent results, the AIC value for both boys and girls represents the fit of a model that includes an interaction between sex and exposure. As the linear model is nested in the piecewise linear model, the fit of these two models can be directly compared. Thus, we calculated the p-value for the hypothesis that the concentration-response is linear in a test where the alternative was the

piecewise linear model. Here a low p-value indicates that the linear model has a poorer fit. As specific-gravity adjusted U-F values were available for an additional 105 MIREC subjects, we carried out sensitivity analyses using these data jointly with ELEMENT's creatinine-adjusted data.

3. Results

Table 1 shows the regression coefficients obtained from the two outcomes (GCI and IQ score) in the ELEMENT study and the IQ score in the MIREC study. As previously reported (Bashash, et al., 2017; R. Green, et al., 2019), maternal U-F exposure predicts significantly lower IQ scores in boys and girls in the ELEMENT cohort, while it does not show a statistically significant association for boys and girls combined in the MIREC cohort. However, for the linear association, the difference between the two studies is not statistically significant and the combined data show highly significant U-F regression coefficients (Table 1). A sensitivity analysis using the larger number of observations with specific-gravity adjusted U-F did not show significant differences between the two cohort studies and yielded joint U-F effects that were significant.

Table 2 shows the BMC results obtained from the regression coefficients for each sex and for both sexes. The BMC and BMCL are presented for the MIREC study, the ELEMENT (GCI and IQ) study, and combined across the two cohorts. The AIC results did not reveal any important differences between the model fits, except that the linear slope appeared superior to the squared for the joint results that included the Mexican GCI data. For the linear models, the joint BMCL in terms of U-F (creatinine-adjusted) is approximately equal for the MIREC-ELEMENT IQ model (0.20 mg/L) and MIREC-ELEMENT GCI model (0.19 mg/L). Likewise, for the squared models, the joint BMCL in terms of U-F is approximately equal for the MIREC-ELEMENT IQ model (0.77 mg/L) and MIREC-ELEMENT GCI model (0.81 mg/L). When using the larger number of specific gravity-adjusted U-F results from

the MIREC cohort, the joint analysis with the ELEMENT data yielded results that were very close to those shown in Table 2, i.e., with BMC values of about 0.19 mg/L for the linear model and about 0.63 mg/L for the squared model (data not shown).

Linear models allowing for sex-dependent effects showed a slightly better fit in the AIC mainly due to the significant interaction terms in the MIREC cohort. Although the BMCL in the MIREC cohort is clearly higher in girls than boys (0.61 vs. 0.13 mg/L), the overall BMCL for both sexes in the MIREC cohort (0.23 mg/L) is closer to the one for boys than the one for girls (Table 2). Sex-linked differences were not significant in the ELEMENT study.

Table 3 shows results using piecewise linear functions, with one breakpoint at 0.75 mg/L and one at 0.5 mg/L. A piecewise linear model is more flexible than a linear model, but AIC results showed that the joint piecewise linear models in Table 3 did not fit better than the standard linear models in Table 2. Thus, the hypothesis of a linear concentration-response relation could not be rejected: for the joint MIREC-ELEMENT IQ model, p-values for likelihood testing were $p=0.18$ and $p=0.15$ when the linear model was tested against models using breakpoints of 0.5 and 0.75 mg/L, respectively. For the joint MIREC-ELEMENT GCI model, the corresponding p-values were $p=0.83$ and $p=0.48$.

The shapes of the linear, the squared, and one piecewise concentration-response curves are shown in Figure 1. In accordance with the BMC values, the figure shows that the squared model has a weaker slope at low concentrations, while the low-concentration slope for the piece-wise association is steeper.

4. DISCUSSION

Experimental and cross-sectional epidemiology studies have provided evidence of fluoride neurotoxicity, especially when the exposure occurs during early brain development (Grandjean, 2019). As early as 2006, sufficient evidence was available to warrant further consideration of the possible brain toxicity of fluoride exposure with an emphasis on vulnerable populations (National Research Council, 2006). We now have thorough prospective epidemiology evidence on populations exposed to fluoridated water (about 0.7 mg/L) or comparable exposure from fluoridated salt and other sources. The present study is based on data from two prospective birth cohort studies (Bashash, et al., 2017; R. Green, et al., 2019) that include detailed assessment of child IQ and urinary fluoride concentrations during pregnancy. In these two studies, the mean U-F concentration (creatinine-adjusted) was similar among pregnant women living in Mexico City (0.89 mg/L) and the pregnant women living in fluoridated cities in Canada (0.84 mg/L).

Due to the brain's continued vulnerability across early development (Grandjean, 2013), early infancy may also be a vulnerable period of exposure for adverse effects from fluoride, especially among bottle-fed infants who receive formula reconstituted with fluoridated water (Till et al., 2019). Still, the effects of fetal exposure (i.e., U-F in pregnancy) in the MIREC Study remained significant when adjusting for exposure occurring in infancy. Likewise, in the ELEMENT study, the effect of maternal U-F was only marginally reduced after controlling for child U-F; fluoride exposure in school-age children showed a weaker and non-statistically significant association with child IQ (Bashash, et al., 2017). Taken together, these findings suggest that fetal brain development is highly vulnerable to fluoride exposure.

The magnitude of the fluoride-associated IQ losses is in accordance with findings in cross-sectional studies carried out in communities where the children examined had likely been exposed to chronic water-fluoride concentrations throughout development (Choi, Sun, Zhang, & Grandjean, 2012). More recent studies have shown similar results (Wang et al., 2020; Yu et al., 2018), and benchmark dose calculations (Hirzy, Connett, Xiang, Spittle, & Kennedy, 2016) relying on a large cross-sectional study (Xiang et al., 2003) showed results on the linear association similar to the ones obtained in the current analysis. These findings provide additional evidence that fluoride is a developmental neurotoxicant (i.e., causing adverse effects on brain development in early life). Given the ubiquity of fluoride exposure, the population impact of adverse effects from fluoride may be even greater than for other toxic elements like lead, mercury, and arsenic (Nilsen et al. 2020). Adverse effects of the latter trace elements are associated with blood concentrations that are about 100-fold lower than the serum-fluoride concentration that corresponds to the benchmark concentration (Grandjean, 2019).

A few retrospective studies have been carried out in communities with elevated fluoride exposure, though with imprecise exposure assessment that mostly relied on proxy variables, and without prenatal fluoride measurements (Aggeborn & Ohman, 2017; Broadbent et al., 2015). In addition to IQ outcome studies, the ELEMENT cohort found that elevated maternal U-F concentrations were associated with higher scores on inattention on the Conners' Rating Scale, an indication of Attention-Deficit/Hyperactivity Disorder (ADHD) behaviors (Bashash et al., 2018). Other studies on attention outcomes found an association between water fluoridation and diagnosis of ADHD in Canada, although data on child U-F did not replicate this association (Riddell, Malin, Flora, McCague, & Till, 2019), which is consistent with the ELEMENT study of child U-F and IQ (Bashash, et al., 2017). Likewise, increased risk of ADHD was reported to be associated with water fluoridation at the state level in the U.S. (Malin & Till, 2015), although inclusion of mean elevation at the residence as a covariate made the association non-significant (Perrott, 2018).

Individual vulnerability may play a role in fluoride neurotoxicity. In the original MIREC study, boys were more vulnerable to prenatal fluoride neurotoxicity than girls (R. Green, et al., 2019) suggesting that sex-dependent endocrine disruption may play a role (Bergman et al., 2013), among other sex-differential possibilities. Genetic predisposition to fluoride neurotoxicity may also exist (Cui et al., 2018; Zhang et al., 2015), but has so far not been verified. Other predisposing factors, such as iodine deficiency (Malin, Riddell, McCague, & Till, 2018) may contribute. For such reasons, regulatory agencies routinely use an uncertainty factor to derive safe exposure levels that are lower than the BMCL.

Both prospective studies adjusted for a substantial number of cofactors. Prenatal and early postnatal lead exposure did not influence the ELEMENT fluoride-associated IQ deficits (Bashash, et al., 2017). Adjustment for other neurotoxicants or risk factors, such as arsenic and lead exposure, did not appreciably change the estimates in the MIREC study (R. Green, et al., 2019). While BMC results were calculated for the creatinine-adjusted U-F available from both studies, U-F results adjusted for specific-gravity were available for an additional 105 MIREC women; if using the latter U-F data, slightly lower BMC results were obtained, as compared to those based on creatinine-adjusted data only. Higher results were obtained for the squared, and lower for the broken linear slopes, but neither showed a superior fit to the data when compared to the linear relationship between maternal U-F and child IQ.

The increased precision using the average maternal U-F concentration as an indicator of prenatal fluoride exposure results in stronger statistical evidence of fluoride-associated deficits, compared with using cross-sectional or retrospective studies. Still, the amount of fluoride that reaches the brain during early brain development is unknown, and even the maternal U-F concentration measurements may be considered somewhat imprecise as dose indicators. Such

imprecision, likely occurring at random, will tend to underestimate fluoride neurotoxicity (Grandjean & Budtz-Jørgensen, 2010).

Author Manuscript

The prospective studies offer strong evidence of prenatal neurotoxicity, and the benchmark results should inspire a revision of water-fluoride recommendations aimed at protecting pregnant women and young children. While systemic fluoride exposure has been linked to dental health benefits in early studies (Iheozor-Ejiro et al., 2015), these benefits occur in the oral cavity after teeth have erupted (Featherstone, 2000), thus suggesting that use of fluoridated toothpaste and other topical treatment should be considered for alternative caries prevention.

5. CONCLUSIONS

Two prospective studies examined concentration-dependent cognitive deficits associated with the maternal U-F during pregnancy; one of the studies (Bashash, et al., 2017) measured child IQ at two ages and found similar results, whereas the other study (R. Green, et al., 2019) found a fluoride-IQ effect only in boys. We explored the shape of the concentration-response curve by using a standard linear shape and compared with a squared exposure and a piecewise linear function that allowed a change in steepness at two points within the range of exposures. Comparisons between the models suggest that the standard linear function is a reasonable approximation. All of these estimates have a certain degree of uncertainty, and emphasis should therefore be placed on the joint BMC results from the two studies and involving both sexes. These findings, using a linear concentration dependence, suggest an overall BMCL for fluoride concentrations in urine of approximately 0.2 mg/L. The results of this benchmark analysis should be incorporated when developing strategies to facilitate lowering fluoride exposure among pregnant women.

COMPETING INTEREST STATEMENT

PG has served as an expert on the hazards of environmental chemicals on behalf of the plaintiffs in *Food & Water Watch v. US EPA*. HH and BL served as non-retained expert witnesses (uncompensated) for the same trial, in which they offered testimony regarding the studies their respective teams on fluoride exposure and neurobehavioral outcomes. All other authors have no interest to declare.

FUNDING

The ELEMENT study was supported by U.S. NIH R01ES021446, NIH R01-ES007821, NIEHS/EPA P01ES022844, NIEHS P42-ES05947, NIEHS Center Grant P30ES017885 and the National Institute of Public Health/Ministry of Health of Mexico. The MIREC study was supported by the Chemicals Management Plan at Health Canada, the Ontario Ministry of the Environment, and the Canadian Institutes for Health Research (grant # MOP-81285). PG is supported by the NIEHS Superfund Research Program (P42ES027706). CT is supported by the NIEHS (grants R21ES027044; R01ES030365-01).

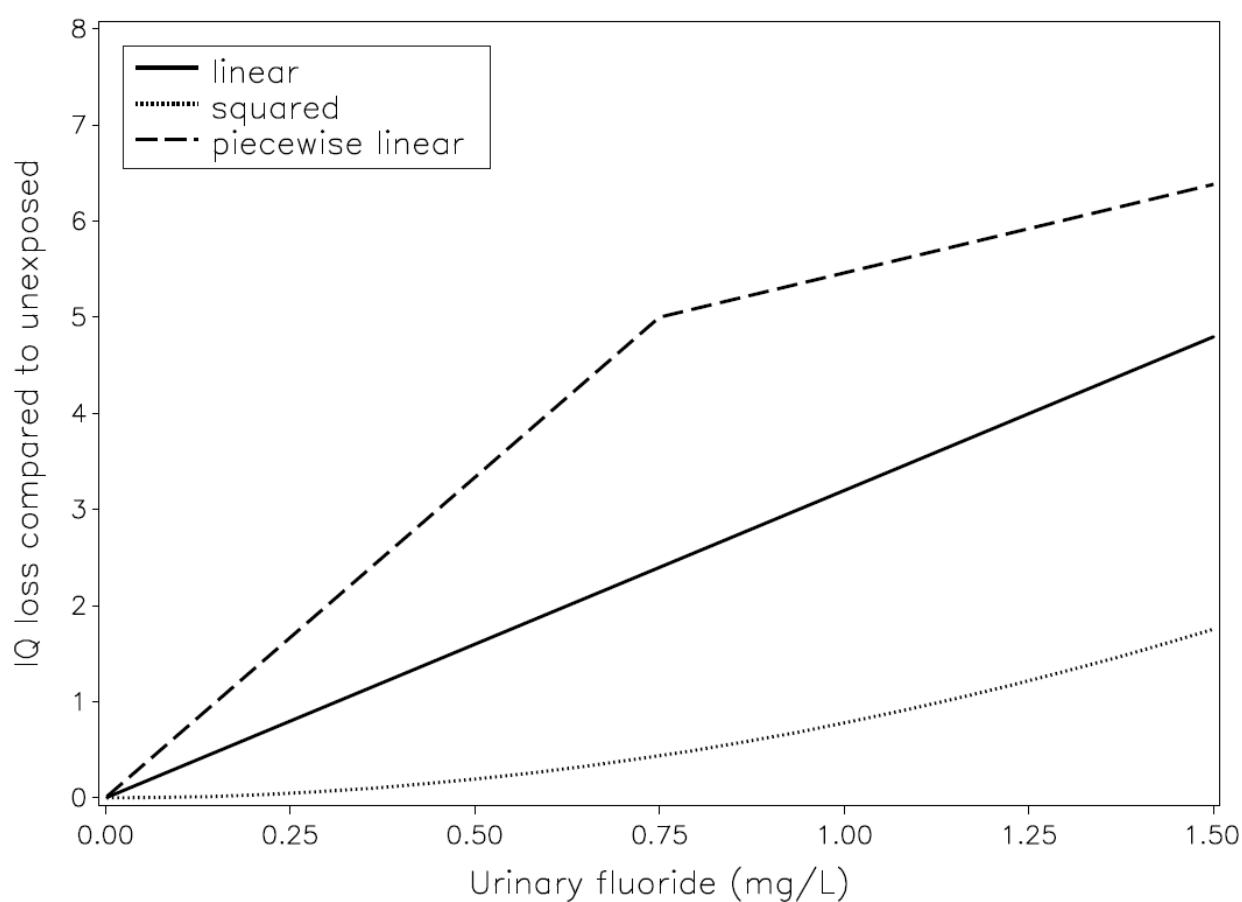
ACKNOWLEDGMENTS

The authors gratefully acknowledge: Nicole Lupien, Stéphanie Bastien, and Romy-Leigh McMaster and the MIREC Study Coordinating Staff for their administrative support, Dr. Jillian Ashley-Martin for providing feedback on the manuscript, as well as the MIREC study group of investigators and site investigators, Alain Leblanc from the INSPQ for measuring the urinary creatinine; Dr. Angeles Martinez-Mier, Christine Buckley, Dr. Frank Lippert and Prithvi Chandrappa for their analysis of urinary fluoride at the Indiana University School of Dentistry; Linda Farmus for her assistance with statistical modeling. This MIREC Biobank study was funded by a grant from the National Institute of Environmental Health Science (NIEHS) (grant #R21ES027044). The MIREC Study was supported by

the Chemicals Management Plan at Health Canada, the Ontario Ministry of the Environment, and the Canadian Institutes for Health Research (grant # MOP-81285).

Figure legend

Fig. 1. Association between creatinine-adjusted maternal urinary-fluoride (U-F) concentration in pregnancy and child IQ loss for the larger number of children (joint for GCI in ELEMENT and MIREC). Covariate-adjusted models are shown for the linear (solid), squared (dotted), and piecewise (dashed) linear curve with breakpoint 0.75 mg/L. The BMC is the U-F concentration that corresponds to an IQ loss of 1 (numbers shown in Tables 2 and 3).



A

Table 1. Regression coefficients adjusted for confounders for the change in the outcome, for boys and girls combined, at an increase by 1 mg/L in creatinine-adjusted maternal urine fluoride concentration for IQ in the MIREC study, GCI (upper rows) and IQ (lower rows) in the ELEMENT study, and a joint calculation. The column to the right (p_{diff}) shows the p-value for a hypothesis of identical regressions in the two studies. Two concentration-response models are used, a linear and one with the squared exposure variable.

	MIREC		ELEMENT		Joint MIREC-ELEMENT		
model	beta	p	beta	p	beta	p	p_{diff}
	FSIQ (n=407)		GCI (n=287)				
Linear	-2.01	0.16	-6.29	0.007	-3.20	0.008	0.12
Squared	-0.419	0.40	-2.68	0.02	-0.780	0.09	0.07
	FSIQ (n=407)		IQ (n=211)				
Linear	-2.01	0.16	-5.00	0.01	-3.07	0.01	0.22
Squared	-0.419	0.40	-2.65	0.002	-0.998	0.023	0.025

Table 2. Benchmark concentration results (mg/L urinary fluoride, creatinine-adjusted) for a BMR of 1 IQ point obtained from the MIREC study and the two cognitive assessments from the ELEMENT study as well as the joint results. Two concentration-response models are used, a linear and one with the squared exposure variable. For both models, sex-specific and joint benchmark results are provided. The fit of the regression models was compared by the AIC (where lower values indicate a better fit).

Study	Sex	MIREC (n=407)		ELEMENT IQ (n=211)		ELEMENT GCI (n=287)		MIREC and ELEMENT IQ (n=618)		AIC
		BMC	BMCL	BMC	BMCL	BMC	BMCL	BMC	BMCL	
Linear	Both	0.497	0.228	0.200	0.122	0.159	0.099	0.326	0.201	470
Linear	Boys	0.201	0.125	0.275	0.130	0.148	0.084	0.222	0.144	470
Linear	Girls	∞	0.609	0.160	0.091	0.169	0.087	1.098	0.275	470
Squared	Both	1.545	0.896	0.614	0.496	0.611	0.467	1.008	0.768	470
Squared	Boys	0.840	0.622	0.684	0.496	0.581	0.435	0.787	0.619	470
Squared	Girls	∞	1.262	0.576	0.449	0.642	0.434	1.637	0.866	470

Abbreviations: AIC, Akaike Information Criterion; BMC, benchmark concentration; BMCL, benchmark concentration level; BMR, benchmark response; GCI, Global Cognitive Index; IQ, Intelligence Quotient.

Table 3. Benchmark concentration (BMC) results (mg/L urinary fluoride, creatinine-adjusted) for a BMR of 1 IQ point obtained from the MIREC study and the two cognitive assessments from the ELEMENT study as well as the joint results. Two piecewise linear concentration-response models (with urinary fluoride breakpoints at 0.5 and 0.75 mg/L) are used. For both models, sex-dependent

and joint benchmark results are provided. The fit of the regression models was compared by the AIC (where lower values indicate a better fit).

Study	Sex	MIREC (n=407)		ELEMENT IQ (n=211)		ELEMENT GCI (n=287)		MIREC and ELEMENT IQ (n=618)		
		BMC	BMCL	BMC	BMCL	BMC	BMCL	BMC	BMCL	AIC
Breakpoint 0.5	Both	1.751	0.092	2.688	0.431	1.004	0.042	1.073	0.139	4770.6
Breakpoint 0.5	Boys	0.086	0.040	2.953	0.135	0.725	0.011	0.156	0.053	4766.7
Breakpoint 0.5	Girls	∞	0.309	2.363	0.024	1.144	0.046	2.913	0.428	4766.7
Breakpoint 0.75	Both	0.166	0.081	1.283	0.149	0.115	0.050	0.284	0.112	4769.8
Breakpoint 0.75	Boys	0.082	0.049	1.379	0.121	0.127	0.035	0.136	0.070	4769.4
Breakpoint 0.75	Girls	∞	0.125	1.155	0.052	0.109	0.044	1.365	0.140	4769.4

Abbreviations: AIC, Akaike Information Criterion; BMC, benchmark concentration; BMCL, benchmark concentration level; BMR, benchmark response; GCI, Global Cognitive Index; IQ, Intelligence Quotient.

REFERENCES

- Aggeborn, L., & Ohman, M. (2017). *The Effects of Fluoride in Drinking Water*. Uppsala, Sweden: Institute for Evaluation of Labour Market and Education Policy.
- Arbuckle, T. E., Fraser, W. D., Fisher, M., Davis, K., Liang, C. L., Lupien, N., et al. (2013). Cohort profile: the maternal-infant research on environmental chemicals research platform. *Paediatr Perinat Epidemiol*, 27(4), 415-425.
- Bartos, M., Gumilar, F., Gallegos, C. E., Bras, C., Dominguez, S., Monaco, N., et al. (2018). Alterations in the memory of rat offspring exposed to low levels of fluoride during gestation and lactation: Involvement of the alpha7 nicotinic receptor and oxidative stress. *Reprod Toxicol*, 81, 108-114.

- Bashash, M., Marchand, M., Hu, H., Till, C., Martinez-Mier, E. A., Sanchez, B. N., et al. (2018). Prenatal fluoride exposure and attention deficit hyperactivity disorder (ADHD) symptoms in children at 6-12 years of age in Mexico City. *Environ Int*, 121(Pt 1), 658-666.
- Bashash, M., Thomas, D., Hu, H., Martinez-Mier, E. A., Sanchez, B. N., Basu, N., et al. (2017). Prenatal Fluoride Exposure and Cognitive Outcomes in Children at 4 and 6-12 Years of Age in Mexico. *Environ Health Perspect*, 125(9), 097017.
- Bergman, A., Heindel, J. J., Kasten, T., Kidd, K. A., Jobling, S., Neira, M., et al. (2013). The impact of endocrine disruption: a consensus statement on the state of the science. *Environ Health Perspect*, 121(4), A104-106.
- Broadbent, J. M., Thomson, W. M., Ramrakha, S., Moffitt, T. E., Zeng, J., Foster Page, L. A., et al. (2015). Community Water Fluoridation and Intelligence: Prospective Study in New Zealand. *Am J Public Health*, 105(1), 72-76.
- Budtz-Jørgensen, E., Bellinger, D., Lanphear, B., Grandjean, P., & International Pooled Lead Study Investigators. (2013). An international pooled analysis for obtaining a benchmark dose for environmental lead exposure in children. *Risk Anal*, 33(3), 450-461.
- Budtz-Jørgensen, E., Keiding, N., & Grandjean, P. (2001). Benchmark dose calculation from epidemiological data. *Biometrics*, 57(3), 698-706.
- Choi, A. L., Sun, G., Zhang, Y., & Grandjean, P. (2012). Developmental fluoride neurotoxicity: a systematic review and meta-analysis. [Research Support, Non-U.S. Gov't]. *Environ Health Perspect*, 120(10), 1362-1368.
- Choi, A. L., Zhang, Y., Sun, G., Bellinger, D. C., Wang, K., Yang, X. J., et al. (2015). Association of lifetime exposure to fluoride and cognitive functions in Chinese children: a pilot study. *Neurotoxicol Teratol*, 47, 96-101.
- Crump, K. S. (1995). Calculation of benchmark doses from continuous data. *Risk Analysis*, 15(1), 79-89.
- Cui, Y., Zhang, B., Ma, J., Wang, Y., Zhao, L., Hou, C., et al. (2018). Dopamine receptor D2 gene polymorphism, urine fluoride, and intelligence impairment of children in China: A school-based cross-sectional study. *Ecotoxicol Environ Saf*, 165, 270-277.
- Duan, Q., Jiao, J., Chen, X., & Wang, X. (2018). Association between water fluoride and the level of children's intelligence: a dose-response meta-analysis. *Public Health*, 154, 87-97.
- Ekstrand, J., & Ehrnebo, M. (1983). The relationship between plasma fluoride, urinary excretion rate and urine fluoride concentration in man. *J Occup Med*, 25(10), 745-748.

- European Food Safety Authority. (2009). Guidance of the Scientific Committee on Use of the benchmark dose approach in risk assessment. *EFSA Journal*, 1150, 1-72.
- European Food Safety Authority. (2010). EFSA Panel on Contaminants in the Food Chain (CONTAM); Scientific Opinion on Lead in Food. *EFSA Journal*, 8(4), 1570.
- Featherstone, J. D. (2000). The science and practice of caries prevention. *J Am Dent Assoc*, 131(7), 887-899.
- Gould, E. (2009). Childhood lead poisoning: conservative estimates of the social and economic benefits of lead hazard control. *Environ Health Perspect*, 117(7), 1162-1167.
- Grandjean, P. (2013). *Only One Chance: How Environmental Pollution Impairs Brain Development — and How to Protect the Brains of the Next Generation*. New York: Oxford University Press.
- Grandjean, P. (2019). Developmental fluoride neurotoxicity: an updated review. *Environ Health*, 18(1), 110.
- Grandjean, P., & Budtz-Jørgensen, E. (2010). An ignored risk factor in toxicology: The total imprecision of exposure assessment. *Pure Appl Chem*, 82(2), 383-391.
- Green, R., Lanphear, B., Hornung, R., Flora, D., Martinez-Mier, E. A., Neufeld, R., et al. (2019). Association Between Maternal Fluoride Exposure During Pregnancy and IQ Scores in Offspring in Canada. *JAMA Pediatr*, 173, (in press).
- Green, R., Rubenstein, J., Popoli, R., Capulong, R., & Till, C. (2020). Sex-specific neurotoxic effects of early-life exposure to fluoride: A review of the epidemiologic and animal literature. *Current Epidemiology Reports*, (in press).
- Hirzy, J. W., Connett, P., Xiang, Q. Y., Spittle, B. J., & Kennedy, D. C. (2016). Developmental neurotoxicity of fluoride: A quantitative risk analysis towards establishing a safe daily dose of fluoride for children. [Article]. *Fluoride*, 49(4), 379-400.
- Iheozor-Ejiofor, Z., Worthington, H. V., Walsh, T., O'Malley, L., Clarkson, J. E., Macey, R., et al. (2015). Water fluoridation for the prevention of dental caries. *Cochrane Database Syst Rev*(6), CD010856.
- Kaplan, R. M., & Saccuzzo, D. P. (2010). *Psychological Testing: Principles, Applications, & Issues, Eighth Edition*. Belmont, CA: Wadsworth.
- Krishnankutty, N., Jensen, T. S., Kjær, J., Jørgensen, J. S., Nielsen, F., & Grandjean, P. (2021). Public health risks from tea drinking: Fluoride exposure. *Scand J Public Health*, (in press).
- Levin, E. D., Dow-Edwards, D., & Patisaul, H. (2021). Introduction to sex differences in neurotoxic effects. *Neurotoxicol Teratol*, 83, 106931.

- Malin, A. J., Riddell, J., McCague, H., & Till, C. (2018). Fluoride exposure and thyroid function among adults living in Canada: Effect modification by iodine status. *Environ Int*, 121(Pt 1), 667-674.
- Malin, A. J., & Till, C. (2015). Exposure to fluoridated water and attention deficit hyperactivity disorder prevalence among children and adolescents in the United States: an ecological association. *Environ Health*, 14, 17.
- Martinez-Mier, E. A., Cury, J. A., Heilman, J. R., Katz, B. P., Levy, S. M., Li, Y., et al. (2011). Development of gold standard ion-selective electrode-based methods for fluoride analysis. *Caries Res*, 45(1), 3-12.
- Mullenix, P. J., Denbesten, P. K., Schunior, A., & Kernan, W. J. (1995). Neurotoxicity of sodium fluoride in rats. *Neurotoxicol Teratol*, 17(2), 169-177.
- National Research Council. (1989). *Recommended Dietary Allowances* (10 ed.). Washington, DC: National Academy Press.
- National Research Council. (2006). *Fluoride in Drinking Water: A Scientific Review of EPA's Standards*. Washington, D.C.: National Academy Press.
- National Toxicology Program. (2020). *Revised draft NTP monograph on the systematic review of fluoride exposure and neurodevelopmental and cognitive health effects*. Research Triangle Park, NC: National Institute of Environmental Health Sciences.
- Opydo-Symaczek, J., & Borysewicz-Lewicka, M. (2005). Urinary fluoride levels for assessment of fluoride exposure of pregnant women in Poznan, Poland. *Fluoride*, 38(4), 312-317.
- Perng, W., Tamayo-Ortiz, M., Tang, L., Sanchez, B. N., Cantoral, A., Meeker, J. D., et al. (2019). Early Life Exposure in Mexico to ENvironmental Toxicants (ELEMENT) Project. *BMJ Open*, 9(8), e030427.
- Perrott, K. W. (2018). Fluoridation and attention deficit hyperactivity disorder - a critique of Malin and Till (2015). *Br Dent J*, 223(11), 819-822.
- Reuben, A., Caspi, A., Belsky, D. W., Broadbent, J., Harrington, H., Sugden, K., et al. (2017). Association of Childhood Blood Lead Levels With Cognitive Function and Socioeconomic Status at Age 38 Years and With IQ Change and Socioeconomic Mobility Between Childhood and Adulthood. *JAMA*, 317(12), 1244-1251.
- Riddell, J. K., Malin, A. J., Flora, D., McCague, H., & Till, C. (2019). Association of water fluoride and urinary fluoride concentrations with attention deficit hyperactivity disorder in Canadian youth. *Environ Int*, 133(Pt B), 105190.

- Rodríguez, I., Burgos, A., Rubio, C., Gutiérrez, A. J., Paz, S., Rodrigues da Silva Júnior, F. M., et al. (2020). Human exposure to fluoride from tea (*Camellia sinensis*) in a volcanic region-Canary Islands, Spain. *Environ Sci Pollut Res Int*.
- Tang, Q., Du, J., Ma, H., Jiang, S., & Zhou, X. (2008). Fluoride and children's intelligence: a meta-analysis. *Bio Trace Elem Res*, 126, 115-120.
- Thomas, D. B., Basu, N., Martinez-Mier, E. A., Sanchez, B. N., Zhang, Z., Liu, Y., et al. (2016). Urinary and plasma fluoride levels in pregnant women from Mexico City. *Environ Res*, 150, 489-495.
- Till, C., Green, R., Flora, D., Hornung, R., Martinez-Mier, E. A., Blazer, M., et al. (2019). Fluoride exposure from infant formula and child IQ in a Canadian birth cohort. *Environ Int*, 134, 105315.
- Till, C., Green, R., Flora, D., Hornung, R., Martinez-Mier, E. A., Blazer, M., et al. (2020). Fluoride exposure from infant formula and child IQ in a Canadian birth cohort. *Environ Int*, 134, 105315.
- Till, C., Green, R., Grundy, J. G., Hornung, R., Neufeld, R., Martinez-Mier, E. A., et al. (2018). Community Water Fluoridation and Urinary Fluoride Concentrations in a National Sample of Pregnant Women in Canada. *Environ Health Perspect*, 126(10), 107001.
- U.S. Environmental Protection Agency. (1985). *National Primary Drinking Water Regulations: Fluoride Final Rule and Proposed Rule*.
- U.S. Environmental Protection Agency. (2010). *Fluoride: Exposure and Relative Source Contribution Analysis*. Washington, DC: Health and Ecological Criteria Division, Office of Water, U.S. EPA.
- U.S. Environmental Protection Agency. (2012). *Benchmark dose technical guidance*. Washington, DC: Risk Assessment Forum, U.S. EPA.
- Valdez Jimenez, L., Lopez Guzman, O. D., Cervantes Flores, M., Costilla-Salazar, R., Calderon Hernandez, J., Alcaraz Contreras, Y., et al. (2017). In utero exposure to fluoride and cognitive development delay in infants. *Neurotoxicology*, 59, 65-70.
- Wang, M., Liu, L., Li, H., Li, Y., Liu, H., Hou, C., et al. (2020). Thyroid function, intelligence, and low-moderate fluoride exposure among Chinese school-age children. *Environ Int*, 134, 105229.
- Waugh, D. T., Godfrey, M., Limeback, H., & Potter, W. (2017). Black Tea Source, Production, and Consumption: Assessment of Health Risks of Fluoride Intake in New Zealand. *J Environ Public Health*, 2017, 5120504.
- World Health Organization. (2006). *Fluoride in drinking-water*. London, UK: IWA Publishing.

Xiang, Q., Liang, Y., Chen, L., Wang, C., Chen, B., Chen, X., et al. (2003). Effect of fluoride in drinking water on children's intelligence. *Fluoride*, 36(2), 84-94.

Yu, X., Chen, J., Li, Y., Liu, H., Hou, C., Zeng, Q., et al. (2018). Threshold effects of moderately excessive fluoride exposure on children's health: A potential association between dental fluorosis and loss of excellent intelligence. *Environ Int*, 118, 116-124.

Zhang, S., Zhang, X., Liu, H., Qu, W., Guan, Z., Zeng, Q., et al. (2015). Modifying effect of COMT gene polymorphism and a predictive role for proteomics analysis in children's intelligence in endemic fluorosis area in Tianjin, China. *Toxicol Sci*, 144(2), 238-245.

Author Manuscript