Developmental fluoride neurotoxicity: A systematic review and meta-analysis
Choi, Anna; Sun, Guifan; Zhang, Ying; Grandjean, Philippe

Published in:
Environmental Health Perspectives

Publication date:
2012

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Developmental Fluoride Neurotoxicity: A Systematic Review and Meta-Analysis

Anna L. Choi,1 Guifan Sun,2 Ying Zhang,3 and Philippe Grandjean1,4
1Department of Environmental Health, Harvard School of Public Health, Boston, Massachusetts, USA; 2School of Public Health, China Medical University, Shenyang, China; 3School of Stomatology, China Medical University, Shenyang, China; 4Institute of Public Health, University of Southern Denmark, Odense, Denmark

BACKGROUND: Although fluoride may cause neurotoxicity in animal models and acute fluoride poisoning causes neurotoxicity in adults, very little is known of its effects on children’s neurodevelopment.

OBJECTIVE: We performed a systematic review and meta-analysis of published studies to investigate the effects of increased fluoride exposure and delayed neurobehavioral development.

METHODS: We searched the MEDLINE, EMBASE, Water Resources Abstracts, and TOXNET databases through 2011 for eligible studies. We also searched the China National Knowledge Infrastructure (CNKI) database, because many studies on fluoride neurotoxicity have been published in Chinese journals only. In total, we identified 27 eligible epidemiological studies with high and reference exposures, end points of IQ scores, or related cognitive function measures with means and variances for the two exposure groups. Using random-effects models, we estimated the standardized mean difference between exposed and reference groups across all studies. We conducted sensitivity analyses restricted to studies using the same outcome assessment and having drinking-water fluoride as the only exposure. We performed the Cochran test for heterogeneity between studies, Begg's funnel plot, and Egger test to assess publication bias, and conducted meta-regressions to explore sources of variation in mean differences among the studies.

RESULTS: The standardized weighted mean difference in IQ score between exposed and reference populations was −0.45 (95% confidence interval: −0.56, −0.35) using a random-effects model. Thus, children in high-fluoride areas had significantly lower IQ scores than those who lived in low-fluoride areas. Subgroup and sensitivity analyses also indicated inverse associations, although the substantial heterogeneity did not appear to decrease.

CONCLUSIONS: The results support the possibility of an adverse effect of high fluoride exposure on children’s neurodevelopment. Future research should include detailed individual-level information on prenatal exposure, neurobehavioral performance, and covariates for adjustment.

words included combinations of “fluoride” or “drinking water fluoride,” “children,” “neurodevelopment” or “neurolologic” or “intelligence” or “IQ.” We also used references cited in the articles identified. We searched records for 1980–2011. Our literature search identified 39 studies, among which 36 (92.3%) were studies with high and reference exposure groups, and 3 (7.7%) studies were based on individual-level measure of exposures. The latter showed that dose-related deficits were found, but the studies were excluded because our meta-analysis focused on studies with the high- and low-exposure groups only. In addition, two studies were published twice, and the duplicates were excluded.

Inclusion criteria and data extraction. The criteria for inclusion of studies included studies with high and reference fluoride exposures, end points of IQ scores or other related cognitive function measures, presentation of a mean outcome measure, and associated measurement of variance (95% confidence intervals [CIs] or SEs and numbers of participants). Interpretations of statistical significance are based on an alpha level of 0.05. Information included for each study also included the first author, location of the study, year of publication, and numbers of participants in high-fluoride and low-fluoride areas. We noted and recorded the information on age and sex of children, and parental education and income if available.

Statistical analysis. We used STATA (version 11.0; StataCorp, College Station, TX, USA) and available commands (Stern 2009) for the meta-analyses. A standardized weighted mean difference (SMD) was computed using both fixed-effects and random-effects models. The fixed-effects model uses the Mantel–Haenszel method assuming homogeneity among the studies, whereas the random-effects model uses the DerSimonian and Laird method, incorporating both a within-study and an additive between-studies component of variance when there is between-study heterogeneity (Egger et al. 2001). The estimate of the between-study variation is incorporated into both the SE of the estimate of the common effect and the weight of individual studies, which was calculated as the inverse sum of the within and between study variance. We evaluated heterogeneity among studies using the $I^2$ statistic, which represents the percentage of total variation across all studies due to between-study heterogeneity (Higgins and Thompson 2002). We evaluated the potential for publication bias using Begg and Egger tests and visual inspection of a Begg funnel plot (Begg and Mazumdar 1994; Egger et al. 1997). We also conducted independent meta-regressions to estimate the contribution of study characteristics (mean age in years from the age range and year of publication in each study) to heterogeneity among the studies. The scoring standard for the Combined Raven’s Test—The Rural edition in China (CRT-RC) test classifies scores of $\leq 69$ and 70–79 as low and marginal intelligence, respectively (Wang D et al. 1989). We also used the random-effects models to estimate risk ratios for the association between fluoride exposure and a low/marginal versus normal Raven’s test score among children in studies that used the CRT-RC test (Wang D et al. 1989). Scores indicating low and marginal intelligence ($\leq 69$ and 70–79, respectively) were combined as a single outcome due to small numbers of children in each outcome subgroup.

Results

Six of the 34 studies identified were excluded because of missing information on the number of subjects or the mean and variance of the outcome (see Figure 1 for a study selection flow chart and Supplemental Material, Table S1 (http://dx.doi.org/10.1289/ehp.1104912) for additional information on studies that were excluded from the analysis). Another study (Trivedi et al. 2007) was excluded because SDs reported for the outcome parameter were questionably small (1.13 for the high-fluoride group, and 1.23 for the low-fluoride group) and the SMD ($–10.8$; 95% CI: $–11.9$, $–9.6$) was $>10$ times lower than the second smallest SMD ($–0.95$; 95% CI: $–1.16$, $–0.75$) and 150 times lower than the largest SMD (0.07; 95% CI: $–0.083$, 0.22) reported for the other studies, which had relatively consistent SMD estimates. Inclusion of this study in the meta-analysis resulted with a much smaller pooled random-effects SMD estimate and a much larger $F$ ($–0.63$; 95% CI: $–0.83$, $–0.44$, $F^2$ 94.1%) compared with the estimates that excluded this study ($–0.45$; 95% CI: $–0.56$, $–0.34$, $F^2$ 80%) (see Supplemental Material, Figure S1). Characteristics of the 27 studies included are shown in Table 1 (An et al. 1992; Chen et al. 1991; Fan et al. 2007; Guo et al. 1991; Hong et al. 2001; Li FH et al. 2009; Li XH et al. 2010; Li XS 1995; Li Y et al. 1994; Li Y et al. 2003; Lin et al. 1991; Lu et al. 2000; Poursealami et al. 2011; Ren et al. 1989; Seraj et al. 2006; Sun et al. 1991; Wang G et al. 1996; Wang SH et al. 2001; Wang SX et al. 2007; Wang ZH et al. 2006; Xiang et al. 2003; Xu et al. 1994; Yang et al. 1994; Yao et al. 1996, 1997; Zhang JW et al. 1998; Zhao et al. 1996). Two of the studies included in the analysis were conducted in Iran (Poursealami et al. 2011; Seraj et al. 2006); the other study cohorts were populations from China. Two cohorts were exposed to fluoride from coal burning (Guo et al. 1991; Li XH et al. 2010); otherwise populations were exposed to fluoride through drinking water. The CRT-RC test was used to measure the children’s intelligence in 16 studies. Other intelligence measures included the Wechsler Intelligence tests (3 studies; An et al. 1992; Ren et al. 1989; Wang ZH et al. 1996), Binet IQ test (2 studies; Guo et al. 1991; Xu et al. 1994), Raven’s test (2 studies; Poursealami et al. 2011; Seraj et al. 2006), Japan IQ test (2 studies; Sun et al. 1991; Zhang JW et al. 1998), Chinese comparative intelligence test (1 study; Yang et al. 1994), and the mental work capacity index (1 study; Li Y et al. 1994). Because each of the intelligence tests used is designed to measure general intelligence, we used data from all eligible studies to estimate the possible effects of fluoride exposure on general intelligence.

In addition, we conducted a sensitivity analysis restricted to studies that used similar tests to measure the outcome (specifically, the CRT-RC, Wechsler Intelligence test, Binet IQ test, or Raven’s test), and an analysis restricted to studies that used the CRT-RC. We also performed an analysis that excluded studies with co-exposures including iodine and arsenic, or with non-drinking-water fluoride exposure from coal burning.

Pooled SMD estimates. Among the 27 studies, all but one study showed random-effect SMD estimates that indicated an inverse association, ranging from $–0.95$ (95% CI: $–1.16$, $–0.75$) to $–0.10$ (95% CI: $–0.63$, $–0.44$, $F^2$ 94.1%). We also performed an analysis that excluded studies with co-exposures including iodine and arsenic, or with non-drinking-water fluoride exposure from coal burning.

Figure 1. Flow diagram of the meta-analysis.
Table 1. Characteristics of epidemiological studies of fluoride exposure and children's cognitive outcomes.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study location</th>
<th>Location</th>
<th>No. of high-exposure group</th>
<th>No. of reference group</th>
<th>Age range (years)</th>
<th>Fluoride exposure</th>
<th>Outcome measure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ren et al. 1989</td>
<td>Shandong, China</td>
<td>160</td>
<td>169</td>
<td>8–14</td>
<td>High-/low-fluoride villages</td>
<td>Not specified</td>
<td>Wechsler Intelligence test²</td>
<td>Children in high-fluoride region had lower IQ scores</td>
</tr>
<tr>
<td>Chen et al. 1991</td>
<td>Shanxi, China</td>
<td>320</td>
<td>320</td>
<td>7–14</td>
<td>Drinking water</td>
<td>4.55 mg/L (high); 0.99 mg/L (reference)</td>
<td>CRT-RC</td>
<td>The average IQ of children from high-fluoride area was lower than that of the reference area</td>
</tr>
<tr>
<td>Guo et al. 1991</td>
<td>Hunan, China</td>
<td>60</td>
<td>61</td>
<td>7–13</td>
<td>Fluoride in coal burning</td>
<td>118.1–1391.7 mg/kg (coal burning area); Control area used wood</td>
<td>Chinese Binet²</td>
<td>Average IQ in fluoride coal-burning area was lower than that in the reference area</td>
</tr>
<tr>
<td>Lin et al. 1991</td>
<td>Xinjiang, China</td>
<td>33</td>
<td>86</td>
<td>7–14</td>
<td>Drinking water</td>
<td>0.08 mg/L (high); 0.34 mg/L (reference)</td>
<td>CRT-RC</td>
<td>Children in the high-fluoride area had lower IQ scores compared with the children from the reference area</td>
</tr>
<tr>
<td>Sun et al. 1991</td>
<td>Guizhou, China</td>
<td>196</td>
<td>224</td>
<td>6.5–12</td>
<td>Rate of fluorosis</td>
<td>Fluorosis: 98.36% (high); not specified (reference)</td>
<td>Japan IQ test²</td>
<td>Mean IQ was lower in all age groups except ≤7 years in the area with high fluoride and aluminum (limited to high-fluoride population only)</td>
</tr>
<tr>
<td>An et al. 1992</td>
<td>Inner Mongolia, China</td>
<td>121</td>
<td>121</td>
<td>7–16</td>
<td>Drinking water</td>
<td>2.1–7.6 mg/L (high); 0.6–1.0 mg/L (reference)</td>
<td>Wechsler Intelligence test²</td>
<td>IQ scores of children in high-fluoride areas were significantly lower than those of children living in reference fluoride area</td>
</tr>
<tr>
<td>Li Y et al. 1994</td>
<td>Sichuan, China</td>
<td>106</td>
<td>49</td>
<td>12–13</td>
<td>Burning of high-fluoride coal to cook grain in high-fluoride area</td>
<td>4.7–31.6 mg/kg (high); 0.5 mg/kg (reference)</td>
<td>Child mental work capacity</td>
<td>Early, prolonged high fluoride intake causes a decrease in the child’s mental work capacity</td>
</tr>
<tr>
<td>Xu et al. 1994</td>
<td>Shandong, China</td>
<td>97</td>
<td>32</td>
<td>8–14</td>
<td>Drinking water</td>
<td>1.8 mg/L (high); 0.8 mg/L (reference)</td>
<td>Binet-Simon²</td>
<td>Children had lower IQ scores in high-fluoride area than those who lived in the reference area.</td>
</tr>
<tr>
<td>Yang et al. 1994</td>
<td>Shandong, China</td>
<td>30</td>
<td>30</td>
<td>8–14</td>
<td>Drinking water</td>
<td>2.97 mg/L (high); 0.5 mg/L (reference)</td>
<td>Chinese comparative intelligence test</td>
<td>The average IQ score was lower in children from high-fluoride and -iodine area than those from the reference area, but the results were not significant</td>
</tr>
<tr>
<td>Li XS et al. 1995</td>
<td>Guizhou, China</td>
<td>681</td>
<td>226</td>
<td>8–13</td>
<td>Urine, Dental Fluorosis Index</td>
<td>0.81–2.69 mg/L (high); 1.02 mg/L (reference); DFI 0.8–3.2 (high); DFI &lt; 0.4 (reference)</td>
<td>CRT-RC</td>
<td>Children living in fluorosis areas had lower IQ scores than children living in non-fluorosis areas</td>
</tr>
<tr>
<td>Wang G et al. 1996</td>
<td>Xinjiang, China</td>
<td>147</td>
<td>83</td>
<td>4–7</td>
<td>Drinking water</td>
<td>&gt;1.0–8.65 mg/L (high); 0.58–1.0 mg/L (reference)</td>
<td>Wechsler Intelligence test²</td>
<td>Average IQ score was lower in children in the high-fluoride group than those in the reference group</td>
</tr>
<tr>
<td>Yao et al. 1996</td>
<td>Liaoning, China</td>
<td>266</td>
<td>270</td>
<td>8–12</td>
<td>Drinking water</td>
<td>2.8–11.5 mg/L (high); 1 mg/L (reference)</td>
<td>CRT-RC</td>
<td>Average IQ scores of children residing in exposed fluoride areas were lower than those in the reference area</td>
</tr>
<tr>
<td>Zhao et al. 1996</td>
<td>Shanxi, China</td>
<td>160</td>
<td>160</td>
<td>7–14</td>
<td>Drinking water</td>
<td>4.12 mg/L (high); 0.91 mg/L (reference)</td>
<td>CRT-RC</td>
<td>Children living in high-fluoride and arsenic area had significantly lower IQ scores than those living in the reference fluoride (and no arsenic) area</td>
</tr>
<tr>
<td>Yao et al. 1997</td>
<td>Liaoning, China</td>
<td>188</td>
<td>314</td>
<td>7–14</td>
<td>Drinking water</td>
<td>2 mg/L (exposed); 0.4 mg/L (reference)</td>
<td>CRT-RC</td>
<td>IQ scores of children in the high-fluoride area were lower than those of children in the reference area</td>
</tr>
</tbody>
</table>

Continued.

The average IQ of children from high-fluoride area was lower than that of the reference area.

The average IQ of children from high-fluoride area was lower than that of the reference area.

The overall test of between-study variance was explained by the two covariates, and only 47.3% of the residual variation was attributable to heterogeneity. The overall test of both covariates in the model remained significant (p = 0.0053).

On further restriction of the model to exclude the 7 studies with arsenic and iodine as co-exposures and fluoride originating from coal burning (thus including only the 9 with fluoride exposure from drinking water), neither age nor year of publication was a significant predictor, and the overall test of covariates was less important (p = 0.062), in accordance with the similarity of intelligence test outcomes and the source of exposure in the studies included.
Although official reports of lead concentrations in the study villages in China were not available, some studies reported high percentage (95–100%) of low lead exposure (less than the standard of 0.01 mg/L) in drinking-water samples in villages from several study provinces (Bi et al. 2010; Peng et al. 2008; Sun 2010). 

**Publication bias.** A Begg’s funnel plot with the SE of SMD from each study plotted against its corresponding SMD did not show clear evidence of asymmetry, although two studies with a large SE also reported relatively large effect estimates, which may be consistent with publication bias or heterogeneity (Figure 3). The plot appears symmetrical for studies with larger SE, but with substantial variation in SMD among the more precise studies, consistent with the heterogeneity observed among the studies included in the analysis. Begg (p = 0.22) and Egger (p = 0.11) tests did not indicate significant (p < 0.05) departures from symmetry.

**Pooled risk ratios.** The relative risk (RR) of a low/marginal score on the CRT-RC test (≤80) among children with high fluoride exposure compared with those with low exposure (16 studies total) was 1.93 (95% CI: 1.46, 2.55; I² 58.5%). When the model was restricted to 9 studies that used the CRT-RC and included only drinking-water fluoride exposure (Chen et al. 1991; Fan et al. 2007; Li XH et al. 2010; Li XS et al. 1995; Li Y et al. 2003; Lu et al. 2000; Wang ZH et al. 2006; Yao et al. 1996, 1997), the estimate was similar (RR = 1.75; 95% CI: 1.16, 2.65; I² 70.6%). Although fluoride exposure showed inverse associations with test scores, the available exposure information did not allow a formal dose–response analysis. However, dose-related differences in test scores occurred at a wide range of water-fluoride concentrations.

### Discussion

Findings from our meta-analyses of 27 studies published over 22 years suggest an inverse association between high fluoride exposure and children’s intelligence. Children who lived in areas with high fluoride exposure had lower IQ scores than those who lived in low-exposure or control areas. Our findings are consistent with an earlier review (Tang et al. 2008), although ours more systematically addressed study selection and exclusion information, and was more comprehensive in:

- **a** including 9 additional studies,
- **b** performing meta-regression to estimate the contribution of study characteristics as sources of heterogeneity, and
- **c** estimating pooled risk ratios for the association between fluoride exposure and a low/marginal Raven’s test score.

As noted by the NRC committee (NRC 2006), assessments of fluoride safety have relied on incomplete information on potential

---

**Table 1. Continued.**

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study location</th>
<th>No. in high-exposure group</th>
<th>No. in reference group</th>
<th>Age range (years)</th>
<th>Fluoride exposure</th>
<th>Outcome measure</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zhang JW et al. 1998</td>
<td>Xinjiang, China</td>
<td>51</td>
<td>52</td>
<td>4–10</td>
<td>Drinking water</td>
<td>Japanese IQ Test&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Average IQ scores of children residing in high-fluoride and arsenic area were lower than those who resided in the reference area</td>
</tr>
<tr>
<td>Lu et al. 2000</td>
<td>Tianjin, China</td>
<td>60</td>
<td>58</td>
<td>10–12</td>
<td>Drinking water</td>
<td>CRT-RC&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Children in the high-fluoride area scored significantly lower IQ scores than those in the reference area</td>
</tr>
<tr>
<td>Deng et al. 2001</td>
<td>Shandong, China</td>
<td>85</td>
<td>30</td>
<td>8–14</td>
<td>Drinking water</td>
<td>CRT-RC&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Average IQ scores were significantly lower in high-fluoride group (and -iodine) than the reference group</td>
</tr>
<tr>
<td>Wang Sh et al. 2001</td>
<td>Inner Mongolia, China</td>
<td>720</td>
<td>236</td>
<td>6–13</td>
<td>Fluorosis</td>
<td>CRT-RC&lt;sup&gt;c&lt;/sup&gt;</td>
<td>No significant difference in IQ scores of children in the high-fluoride/high-iodine and reference fluoride/low-iodine areas</td>
</tr>
<tr>
<td>Li Y et al. 2003</td>
<td>Jiangsu, China</td>
<td>222</td>
<td>290</td>
<td>8–13</td>
<td>Drinking water</td>
<td>CRT-RC&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Average IQ of children in high-fluorosis area was lower than that in the reference area</td>
</tr>
<tr>
<td>Seraj et al. 2006</td>
<td>Tehran, Iran</td>
<td>41</td>
<td>85</td>
<td>Not specified</td>
<td>Drinking water</td>
<td>Raven&lt;sup&gt;d&lt;/sup&gt;</td>
<td>The mean IQ of children in the high-fluoride area was significantly lower than that from the reference fluoride area</td>
</tr>
<tr>
<td>Wang ZH et al. 2006</td>
<td>Shanxi, China</td>
<td>202</td>
<td>166</td>
<td>8–12</td>
<td>Drinking water</td>
<td>CRT-RC&lt;sup&gt;c&lt;/sup&gt;</td>
<td>The IQ scores of children in the high-fluoride group were significantly lower than those in the reference group</td>
</tr>
<tr>
<td>Fan et al. 2007</td>
<td>Shandong, China</td>
<td>42</td>
<td>37</td>
<td>7–14</td>
<td>Drinking water</td>
<td>CRT-RC&lt;sup&gt;c&lt;/sup&gt;</td>
<td>The average IQ scores of children residing in the high-fluoride area were lower than those of children residing in the reference area</td>
</tr>
<tr>
<td>Wang SX et al. 2007</td>
<td>Shanxi, China</td>
<td>253</td>
<td>196</td>
<td>8–12</td>
<td>Drinking water and urine</td>
<td>CRT-RC&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Mean IQ scores were significantly lower in the high-fluoride group than from the reference group in the fluoride/arsenic areas</td>
</tr>
<tr>
<td>Li et al. 2009</td>
<td>Hunan, China</td>
<td>60</td>
<td>20</td>
<td>8–12</td>
<td>Coal burning</td>
<td>CRT-RC&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Mean IQ was lower in children in coal-burning areas compared to those in the reference group</td>
</tr>
<tr>
<td>Li FH et al. 2010</td>
<td>Henan, China</td>
<td>347</td>
<td>329</td>
<td>7–10</td>
<td>Drinking water</td>
<td>CRT-RC&lt;sup&gt;c&lt;/sup&gt;</td>
<td>No significant difference in IQ scores between children in the exposed and reference groups</td>
</tr>
<tr>
<td>Pourselami et al. 2011</td>
<td>Iran</td>
<td>59</td>
<td>60</td>
<td>6–9</td>
<td>Drinking water</td>
<td>Raven&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Children in the high-fluoride group scored significantly lower than those in the reference group</td>
</tr>
</tbody>
</table>

risks. In regard to developmental neurotoxicity, much information has in fact been published, although mainly as short reports in Chinese that have not been available to most expert committees. We carried out an extensive review that includes epidemiological studies carried out in China. Although most reports were fairly brief and complete information on covariates was not available, the results tended to support the potential for fluoride-mediated developmental neurotoxicity at relatively high levels of exposure in some studies. We did not find conclusive evidence of publication bias, although there was substantial heterogeneity among studies. Drinking water may contain other neurotoxicants, such as arsenic, but exclusion of studies including arsenic and iodine as co-exposures in a sensitivity analysis resulted in a lower estimate, although the difference was not significant. The exposed groups had access to drinking water with fluoride concentrations up to 11.5 mg/L (Wang SX et al. 2007); thus, in many cases concentrations were above the levels recommended (0.7–1.2 mg/L; DHHS) or allowed in public drinking water (4.0 mg/L; U.S. EPA) in the United States (U.S. EPA 2011). A recent cross-sectional study based on individual-level measure of exposures suggested that low levels of water fluoride (range, 0.24–2.84 mg/L) had significant negative associations with children’s intelligence (Ding et al. 2011). This study was not included in our meta-analysis, which focused only on studies with exposed and reference groups, thereby precluding estimation of dose-related effects.

The results suggest that fluoride may be a developmental neurotoxicant that affects brain development at exposures much below those that can cause toxicity in adults (Grandjean 1982). For neurotoxins such as lead and methylmercury, adverse effects are associated with blood concentrations as low as 10 nmol/L. Serum fluoride concentrations associated with high intakes from drinking water may exceed 1 mg/L, or 50 µmol/L—more than 1,000 times the levels of some other neurotoxins that cause neurodevelopmental damage. Supporting the plausibility of our findings, rats exposed to 1 ppm (50 µmol/L) of water fluoride for 1 year showed morphological alterations in the brain and increased levels of aluminum in brain tissue compared with controls (Vaner et al. 1998).

The estimated decrease in average IQ associated with fluoride exposure based on our analysis may seem small and may be within the measurement error of IQ testing. However, as research on other neurotoxics has shown, a shift to the left of IQ distributions in a population will have substantial impacts, especially among those in the high and low ranges of the IQ distribution (Bellinger 2007).

Our review cannot be used to derive an exposure limit, because the actual exposures of the individual children are not known. Misclassification of children in both high- and low-exposure groups may have occurred if the children were drinking water from other sources (e.g., at school or in the field).

The published reports clearly represent independent studies and are not the result of duplicate publication of the same studies (we removed two duplicates). Several studies (Hong et al. 2001; Lin et al. 1991; Wang SH et al. 2001; Wang SX et al. 2007; Xiang et al. 2003; Zhao et al. 1996) report other exposures, such as iodine and arsenic, a neurotoxicant, but our sensitivity analyses showed similar associations between high fluoride exposure and the outcomes even after these studies were excluded. Large tracts of China

---

### Table 2: Sensitivity analyses of pooled random-effects standardized mean difference (SMD) estimates of child’s intelligence score with high exposure of fluoride.

<table>
<thead>
<tr>
<th>Model</th>
<th>Available studies for analysis</th>
<th>SMD (95% CI)</th>
<th>ρ-Value test of heterogeneity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Exclude nonstandardized tests</td>
<td>23</td>
<td>-0.44 (-0.54, -0.33)</td>
<td>77.6% &lt; 0.001</td>
</tr>
<tr>
<td>2. Exclude non-CRT-RC Tests</td>
<td>16</td>
<td>-0.26 (-0.48, -0.05)</td>
<td>82.5% &lt; 0.001</td>
</tr>
<tr>
<td>3. Exclude studies with other exposures (iodine, arsenic)</td>
<td>9</td>
<td>-0.29 (-0.44, -0.14)</td>
<td>81.8% &lt; 0.001</td>
</tr>
<tr>
<td>or non-drinking-water fluoride exposure</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Mental work capacity (Li Y et al. 1994); Japan IQ (Sun et al. 1991; Zhang JW et al. 1998); Chinese comparative scale of intelligence test (Yang et al. 1994); *4Wechsler intelligence test (An et al. 1992; Ren et al. 1989; Wang G et al. 1996); Chinese Binet IQ (Guo et al. 1991); Raven (Poureslami et al. 2011; Seraj et al. 2006); Binet-Simon (Xu et al. 1994). *5Iodine (Hong et al. 2001; Lin et al. 1991; Wang SH et al. 2001); arsenic (Wang SX et al. 2007); Xiang et al. 2003; Zhao et al. 1996; (Zhang JW et al. 1998 was already excluded; see notes a and b). *6Fluoride from coal burning (Li FH et al. 2009) (Guo et al. 1991 and Li Y et al. 1994 were already excluded; see notes a and b).

---

### Figure 2: Random-effect standardized mean difference (SMD) estimates and 95% CIs of child’s intelligence score associated with high exposure to fluoride. SMDs for individual studies are shown as solid diamonds (●), and the pooled SMD is shown as an open diamond (○). Horizontal lines represent 95% CIs for the study-specific SMDs.
have superficial fluoride-rich minerals with little, if any, likelihood of contamination by other neurotoxins that would be associated with fluoride concentrations in drinking water. From the geographic distribution of the studies, it seems unlikely that fluoride- attributed neurotoxicity could be attributable to other water contaminants.

Still, each of the articles reviewed had deficiencies, in some cases rather serious ones, that limit the conclusions that can be drawn. However, most deficiencies relate to the reporting of where key information was missing. The fact that some aspects of the study were not reported limits the extent to which the available reports allow a firm conclusion. Some methodological limitations were also noted. Most studies were cross-sectional, but this study design would seem appropriate in a stable population where water supplies and fluoride concentrations have remained unchanged for many years. The current water fluoride level likely also reflects past developmental exposures. In regard to the outcomes, the inverse association persisted between studies using different intelligence tests, although most studies did not report age adjustment of the cognitive test scores.

Fluoride has received much attention in China, where widespread dental fluorosis indicates the prevalence of high exposures. In 2008, the Ministry of Health reported that fluorosis was found in 28 provinces with 92 million residents (China News 2008). Although microbiologically safe, water supplies from small springs or mountain sources created pockets of increased exposure near or within areas of low exposures, thus representing exposure settings close to the ideal, because only the fluoride exposure would differ between nearby neighborhoods. Chinese researchers took advantage of this fact and published their findings, though mainly in Chinese journals and according to the standards of science at the time. This research dates back to the 1980s, but has not been widely cited at least in part because of limited access to Chinese journals.

In its review of fluoride, the NRC (2006) noted that the safety and the risks of fluoride at concentrations of 2–4 mg/L were incompletely documented. Our comprehensive review substantially extends the scope of research available for evaluation and analysis. Although the studies were generally of insufficient quality, the consistency of their findings adds support to existing evidence of fluoride-associated cognitive deficits, and suggests that potential developmental neurotoxicity of fluoride should be a high research priority. Although reports from the World Health Organization and national agencies have generally focused on beneficial effects of fluoride (Centers for Disease Control and Prevention 1999; Petersen and Lennon 2004), the NRC report examined the potential adverse effects of fluoride at 2–4 mg/L in drinking water and not the benefits or potential risks that may occur when fluoride is added to public water supplies at lower concentrations (0.7–1.2 mg/L) (NRC 2006).

In conclusion, our results support the possibility of adverse effects of fluoride exposures on children’s neurodevelopment. Future research should formally evaluate dose–response relations based on individual-level measures of exposure over time, including more precise prenatal exposure assessment and more extensive standardized measures of neurobehavioral performance, in addition to improving assessment and control of potential confounders.

References


Teratol 17:169–177.


Environmental Health Perspectives | VOLUME 120 | NUMBER 10 | October 2012

1367


